KU Leuven
Group Biomedical Sciences
Faculty of Medicine
Dep. Neurosciences
Lab. Exp. ORL

KU LEUVEN
DOCTORAL SCHOOL
BIOMEDICAL SCIENCES

IDENTIFICATION OF THREE NATURAL VOICE GROUPS BY PHONETOGRAPHY

A data driven approach

Hugo LYCKE

Leuven, June 4th 2013

Doctoral thesis in Biomedical Sciences

KU Leuven
Group Biomedical Sciences
Faculty of Medicine
Dep. Neurosciences
Lab. Exp. ORL

IDENTIFICATION OF THREE NATURAL VOICE GROUPS BY PHONETOGRAPHY

A data driven approach

Hugo LYCKE

Promotor:	Prof. Dr. Felix de Jong
Co-promotores:	Prof. Dr. Wivine Decoster
	Prof. Dr. Marc Van Hulle
Chair:	Prof. Dr. Mark Jorissen
Jury members:	Prof. Dr. Frans Debruyne
	Prof. Dr. Michael Fuchs (Universität Leipzig, Germany)
	Prof. Dr. Louis Heylen (Universiteit Gent)
	Prof. Dr. Harm Schutte (RU Groningen, the Netherlands)
	Prof. Dr. Jan Wouters

Leuven, June 4th 2013

Doctoral thesis in Biomedical Sciences

Short description

From the literature and our own inquiries among singing teachers, students and professional singers it became clear that there exists no generally accepted method for objective voice classification. Frequency range is generally considered to be an important factor in voice classification, but our own exploratory studies indicated that this is not a panacea for voice classification by itself. At this point, one may question the existence of three basic female and male voice types by nature. In an attempt to break out of the controversy, a new perspective is adopted in this study by letting the data speak for itself (i.e, a data-driven approach). In a pilot study consisting data of 327 female singing students we investigated if basic groups (clusters) can de distinguished by generally applied and easily clinically understandable frequency-intensity related parameters, derived from the voice range profile (VRP), and how much there are. The Ward's minimum variance method was used. Alas, the result was negative. Subsequently, more "intelligent" parameters derived from the VRP (combinations, or even non-linear ones) were used. The data of 206 female conservatory singing students and 256 male subjects, consisting of 9 young singing students, 17 professional singers, 61 professional choir singers and 169 with and without singing experience, was analyzed. The same statistical technique now indicated that there were potentially two cluster solutions: 3 and 4 clusters. In order to resolve the latter, we applied a clustering technique in combination with a parameter selection procedure (to select the most discriminative parameters) and a test for consistency of the found cluster solutions. Now, the three-cluster solution turned out to be the most consistent one in both gender. The parameter that led to the best three-cluster separation in females was the ratio of the perimeter length of the chest voice part of the VRP versus the total perimeter length. In males this was the frequency of the register dip. The results of this study demonstrate that different parameters of the VRP are able to yield a clear separation into three voice clusters for each gender. Such a result is remarkable, since this may not be expected from biological proxies. One can wonder if ancient composers of vocal music had an innate feeling about the existence of three natural basic human voice categories. Further studies are necessary to link the three statistically obtained clusters to the traditional three basic female and male voice classes. A second salient finding of our study is that the parameters that have lead to the three cluster separation in both gender have to do with register transition. The results of the pertinent study may provide a basis for settling the issue of objective voice classification.

Korte beschrijving

Klinische ervaring en exploratieve studies tonen aan dat er geen algemeen aanvaarde objectieve methode voor stemclassificatie bestaat. De frequentiespanne van de stem wordt beschouwd als een belangrijke factor bij stemclassificatie, doch een van onze exploratieve studies toonde aan dat louter de frequentiespanne geen universele oplossing biedt voor stemclassificatie. Men kan zich vragen stellen bij het bestaan in de natuur van drie basis stemsoorten voor man en vrouw. In een poging om uit deze controverse te geraken, werd in deze studie uitgegaan van een nieuw perspectief door de data voor zichzelf te laten spreken, zonder veronderstellingen vooraf (een zogenaamde "data driven approach"). In een pilootstudie werd bij 327 vrouwelijke zangstudenten nagegaan of door algemeen gebruikte en klinisch gemakkelijk interpreteerbare frekwentie-intensiteit gerelateerde parameters, afgeleid uit het fonetogram, een onderscheid gemaakt kan worden tussen basisgroepen (clusters), en hoeveel dit er zijn. Dit gebeurde met de Ward's minimum variance methode. Er werden echter geen natuurlijke stemgroepen gevonden. Vervolgens werden meer "intelligente" parameters (en combinaties ervan; zelfs niet-lineaire) gebruikt, ook weer afgeleid uit het fonetogram. De gegevens van 206 vrouwelijke conservatorium zangstudenten en 256 mannelijke personen, bestaande uit 9 jonge zangstudenten, 17 professionele zangers, 61 professionele koorzangers en 169 personen met en zonder zangervaring, werden geanalyseerd. Er werden clusters gevonden met de Ward's methode, maar dit aantal kan 3 of 4 zijn. Door te bepalen of en in welke mate de personen met verschillende parameters in andere clusters terechtkomen (migratie) werd het beste aantal cluster bepaald. Zowel bij de vrouwen, alsook bij de mannen waren dit er drie. De parameter die tot de beste clusterscheiding leidde bij de vrouwen was de ratio van de lengte van de omtrek van het borststem gedeelte in het fonetogram t.o.v. de totale lengte van de omvang van het fonetogram. Bij de mannen was dit de frequentie van de registerdip. De resultaten van deze studie tonen aan dat verschillende parameters, afgeleid van het fonetogram in staat zijn een duidelijke scheiding te maken tussen drie stemclusters bij elk geslacht. Dit is een opmerkelijk resultaat, dat niet direct verwacht wordt bij biologische variabelen. Men kan zich afvragen of in het verleden componisten van vocale muziek een aanvoelen hadden omtrent het bestaan van drie natuurlijke basis stemsoorten bij de mens. Een tweede opvallende vaststelling van onze studie is dat de parameters die aanleiding gaven tot de drie clusterscheiding voor beide geslachten te maken hebben met de registerovergang. De resultaten van deze studie kunnen een basis vormen voor en een bijdrage leveren aan een objectieve methode voor stemclassificatie.

Table of contents

List of abbreviations

Freq_Dip	frequency of the register dip
dB	decibel
F_0	fundamental frequency
Freq_Range	frequency range of the VRP
Hz	Herz
Max_Freq	maximum frequency in the VRP
mean_freq_TrZone	mean frequency of the register transition zone
Perim_Chest	length of the perimeter of the chest voice part of the VRP
Perim_Head	length of the perimeter of the head voice part of the VRP
Pr	probability
R10	ratio of the surface area of the head voice part of the VRP divided by the surface area of the chest voice part of the VRP
R2	ratio of the surface area of the head voice part of the VRP divided by the total VRP surface area
R4	ratio of the perimeter of the chest voice part of the VRP divided by the total perimeter of the VRP
R9	ratio perimeter transition zone divided by the perimeter chest voice part of the VRP
RS	R-squared
SAS	Statistical Analysis System
SD	standard deviation
SPL	sound pressure level
SPR	semipartial R-squared
SPSS	statistical package for social sciences
UEP	Union of European Phoniatricians
VRP	voice range profile

Chapter 1

General introduction

The human voice is an amazing phenomenon, comprising many psychological, sociological, artistic and biological aspects. Vocal possibilities and limits are based on individual biological properties. Individual vocal qualities can give rise to optimal and even high-class artistic vocal performances, while vocal restraints can cause functional and organic voice disorders. Professional voice users, using their voice as a primary tool, are especially prone to voice problems. Vocal nodules, for instance, most often based on overload of the voice, e.g. due to an incorrect voice classification, are well-known in clinical practice. Therefore, it is important in voice and in singing education to know the physiological limits of the voice and to carefully watch them.

The influence of singing education on voice characteristics

Today's life is immersed in music: willing or not, in almost every situation, day or night, music is in the air. Especially young people have a leaning to listen to contemporary music, songs based on all kinds of regularly changing hit lists, and promoted by the market and mass media. Successful voices of the moment - good or bad ones - are imitated by young people, often without concern of the quality of the sound produced by their idols or by themselves.

The influence of voice training on vocal capabilities is well known from clinical experience but the fact that "a comparable acoustic product can be generated using different physiological strategies" is widely ignored in the voice research.[1-2] Phyland et al. point at the "wide variability among singers as one of the major confounders in the estimation of the incidence or prevalence of voice disorders among singers".[3] Variables such as the degree and sort of singing education and singing experience, the type of singing demands, the performance environments and the singing style, are potential causes of voice problems.[3]

The relationship between the singing teacher and the singing student is particularly interesting in this regard. The singing student often chooses a singing teacher with a particular voice type, which he or she likes and wants to imitate, while the singing teacher in his turn may be inclined to reinforce that attitude. However, frequently changing from one singing teacher to another during one's singing education and one's professional singing career is common practice. This means that the singer is flooded by different advice over the years, including many comments on his "real" voice type. At the same time, during the singing education, the singer becomes aware of the changing features of his or her voice, for the better or worse. Tarneaud already explained that pitch and timbre not only depend on constitutional and physiological factors but also on educational mimesis, acquired in various surroundings, family, school, and profession.[4] Each singer has not only one characteristic timbre, but a set of timbres, or a timbre transformation.[5] Voice quality depends highly on vocal techniques, thus on voice education. Many singers have developed a functional adaptation of their vocal organs, which is not always in correspondence with their anatomical and physiological abilities.[4]

Voice classification

The need for voice classification

The vocal capacity must be estimated to know the possibilities and impossibilities of the voice in order to avoid damage of the voice and to optimize vocal performance.[4,6-11] Voice classification is a method to estimate the vocal capacity and composers of vocal music wrote and write repertoires that fit to the possibilities of the voice, in classical music indicated as voice classes. The biographies of famous and less famous singers frequently mention examples of the pernicious outcomes for their voices and for their careers caused by incorrect voice classification and recent studies show the great prevalence of voice disorders by incorrect voice classification among singing students, singers and singing teachers.[12-15]

Garde made a remarkable statement: “Voice classification is as important as the determination of the blood group and can be seen as a biological constant”.[16] Bonet and Casan state that voice classification is necessary for all children who sing.[17] This “requires a musically and vocally well-trained examiner”.

Basically, classification of a voice means to determine the frequency and intensity voice range in which a subject can work without harming or fatiguing his voice and to which repertoire he should be assigned by the singing teacher. Traditionally, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor. Various singing teachers take also other aspects of the voice into consideration in voice classification, like timbre, expression and personality.

Voice classification has a great impact on a singer's life, but often "experts" disagree and singers question the received label(s) and stick to their own opinions. In addition to this, conductors, scientists, physicians, speech and voice pathologists, almost everybody (including the subject himself) feels entitled to express his opinion on this matter. However, a correct and preferably objective voice classification is indispensable to assess the physiological capacities of the human voice in order to develop and preserve them.

Voice classification in history

Songs for specific voice types are scrupulously elaborated by opera composers. In some cases composers even adapt their musical repertoire to the vocal capabilities and limitations of a distinct singer. This could explain why some singing teachers of commercial music don't bother about voice classification: songs are simply adapted or even rewritten to suit the assumed capabilities of the chosen singer or singing student.

In classical singing education great emphasis is put on voice classification. Additionally to the classic voice types, there are many subtypes, according to different roles ("Fach", e.g., for the soprano voice: coloratura soprano, lyric soprano, dramatic soprano, soubrette) and based on the characteristics of the voice such as pitch, timbre, mobility, vibrato, temperament, expression, and personality. According to Welch et al.[18] the physiological and acoustic base for subcategories of broader classes is unclear.

Nowadays many singing students are often in turmoil. By choosing a kind of education programme - be it classic or commercial - they enter a closely protected environment in which they are guided over many years into a particular direction. They choose a repertoire and take singing lessons which causes them to make restraint choices in connection with their assumed voice type. When the singing teacher is not sure about the exact voice classification at the beginning of the study, very often a cautious repertoire is chosen, because it is assumed that the voice of the young singer has still to mature. However, singing exercises definitely try to expand the singing range and this also influences the singing teacher and the singing student in their perception of the vocal evolution. As a result, if the singing teacher is not sure about the exact voice category of his/her student, there is a great chance that the voice of the young singer is forced into a repertoire, which can damage the voice and negatively affect his or her future career. Sataloff [19] stated: “Singers are habitually unhappy with the limitations of their voices. In many situations, voice teachers are to blame. Both singer and teacher must resist the impulse to show off the voice in works that are either too difficult for the singer’s level of training or simply not suited to the singer’s voice”.

The attitude to voice classification

Voice classification is one of the major objectives and challenges of scientific endeavour. The first requisite of any classification is its functionality, the second - and this is very important from a scientific viewpoint - is its general applicability. Great intra- and interindividual as well as intra- and interinvestigator variations are typical in all measurements of psychophysiological phenomena in humans.

According to Radionoff et al., contemporary music study and performance can be divided in two broad categories: Commercial Music and Classical Music. Classical Music represents genres including opera, Lied, oratorio etc., while Commercial Music represents a huge variety of genres including pop, rock, jazz, country, rhythm and blues, hip-hop, rap, gospel, musical theatre.[20] The number of study programs of Commercial Music in the USA and the United Kingdom surged significantly over the last few years. Radionoff et al. advocate this is due to the worldwide emergence of reality shows on commercial television.[20] On the other hand, they also observed that “Along with nomenclature disparity, a tremendous lack of consistency exists among curriculums of commercial music degrees”.

Many singing students are taking private singing lessons, which are not curriculum-bound. According to Radionoff et al., contemporary Commercial Music singers often complain that their singing teachers do not understand the vocal styles and requirements of a contemporary Commercial Music singer.[20] Little is known how the relatively new music institutions and individual singing teachers deal with voice classification and which criteria they use to classify their singing students.

Lycke et al. explored the opinion of contemporary singing teachers about the utility of voice classification (Addendum 1). He sent one questionnaire to 200 singing teachers via internet and a second questionnaire to 22 singing teachers of one Classical conservatory and two Musical Theatre conservatories. Of the 200 singing teachers, 72 responded (36%). In 61.1% voice classification was important for at least one reason, while 38.9% did not find voice classification an important issue. The results of this study indicate that there is a marked difference of attitude towards voice classification in singing teachers.

The methodology of voice classification

Many manuals on singing techniques do not mention how to classify a voice and scientific publications on voice classification are scarce.[21] Over the years, many factors have been mentioned which provide an indication for classifying a voice, according to the six basic voice types mentioned above, such as the size of the person, the dimensions of the vocal folds, the shape and the volume of the resonating cavities, the general and vocal muscular constitution, biotypological traits, tessitura, the speaking fundamental frequency, the passagio's, the voice timbre, endocrine and sexual aspects, and the neuropsychic condition. The investigation of all these anatomical and biotypological factors usually require several complex instruments and above all, highly trained experts. Voice range, timbre and register transitions are considered to be important classification criteria. However, the assessment of tessitura, speaking fundamental frequency, passagio's, and voice timbre require an experienced ear. Erickson et al. mention that no research has been conducted that examines the interrelationship of pitch, tessitura, and timbre as predictors of voice classification.[22] In another study, Ericson wrote: "Traditionally, voice classification has been based on three perceptual parameters: frequency range, timbre, and tessitura. Some research studies have focused on the acoustic correlates of one parameter, timbre. Yet even this parameter is not well understood".[23] She concluded that "When presented with stimuli comprising conflicting upper and lower formants, experienced listeners find it difficult to perceive voice categories". Regarding voice classification by timbre alone, Bloothooft and Plomp even raise the question of "whether perceptual voice classification, based on timbre, has a phonetical basis (formant frequency detection) or a psychoacoustical basis (sharpness detection)".[24] They even assert that the psychoacoustical basis could lead to an incorrect judgment of voice class altogether.

In the above mentioned study (Addendum 1), Lycke et al. also explored which criteria contemporary singing teachers use in voice classification. Most frequently used acoustical parameters for voice classification were frequency range/tessitura (56.0%), voice quality/timbre (56.0%), volume (12.1%) and register transition (9.0%). The conservatory singing teachers classified their students (n = 165). In the conservatory singing teachers, voice classification was an important issue in singing education. Frequency range/tessitura, voice quality, register transition, and volume were the most frequently used criteria. However, each singing teacher reported a varying individual set of voice classification criteria, depending on the singing student and on the specialty of the department. The results of this study indicate that among singing teachers different criteria for voice classification are applied. Apparently, there is no consensus about the advisability and criteria of voice classification among the various singing teachers.
In an other study, Lycke et al. explored voice classification by conservatory singing teachers and how the singing students think about their voice classification by their teachers (Addendum 2). They sent one questionnaire to 22 singing teachers: at one Belgian classical conservatory and two Musical Theatre conservatories, one in the Netherlands and one in the United Kingdom. The singing teachers were asked to classify their students (N = 165). In an other questionnaire the classified singing students responded about their voice classification. First year students (n = 73) in the Master's degree programs in speech-language pathology at a Belgian university were used as controls. A rather substantial number of singing students, 23.4% in classical singing training and 11.9% in Musical Theatre training, reported not to know their voice category. In the control group, this figure was 83.6%. The methodology and the results of voice classification were different according to the type of conservatory. Musical Theatre students were most frequently classified in the middle voice categories, while higher voices were more present in the classical conservatory group. The results of the two questionnaires showed apparently feelings of uncertainty about voice classification among both singing teachers and

their students. The results of this study indicate the hazard of neglecting careful watch on the physiological limits of the voice.

The importance of the parameter frequency and intensity in voice classification

Voices are commonly classified by ear only. This subjective method is usually based on the appreciation of timbre and/or frequency. The human ear is considered to be an excellent instrument for rapidly detecting nuances of voice quality since not all voice quality aspects can be assessed via laboratory analyses.[25]

The importance of the parameter frequency is illustrated by the representation of the frequency span of the various voice types in many voice textbooks. [4,10-11,16]

Lycke et al. explored frequency range as a parameter for female voice classification (Addendum 3). In this study 16 singing teachers of three European conservatories classified 99 of their students. They elaborated an algorithm, which is based on the limits of the female frequency range, according to 38 authors with different backgrounds. The results of the voice classification by the singing teachers were compared to the results of voice classification by the algorithm. The results of this study demonstrated that frequency range alone proved to be not suitable as the parameter for voice classification.

Another important parameter of the voice in classification is intensity. Clinicians proved not to be able to detect small differences across complex speech tasks.[26]

According to Punt, a person with normal hearing perceives notes of equal intensity at certain frequency ranges louder than those at other frequency ranges.[27] Voices sound louder as they range from bass, baritone to tenor and contralto, mezzo-soprano to soprano, which may be explained by the difference loudness-intensity in psychoacoustics. This is partly a result of this phenomenon, and partly because the lowest notes can only be sung relatively softly for physiological reasons. According to Sundberg, the differences of intensity between singers and non-singers, perceived by the human ear do not consistently correspond to sound pressure levels (SPL) differences.[28]

Plant and Younger state: “In general, intensity increases with higher subglottic air pressures, but there are also considerable variations both between individuals and within different portions of the frequency-intensity range for a given subject”.[29]

Sundberg writes: "The phonatory dimensions pitch, vocal loudness and mode of phonation (fluctuating between the extremes of hyperfunctional and hypofunctional phonation) are relevant from the point of view of phonatory hygiene as well as from the point of view of vocal pedagogy. While a not-appropriate habitual pitch or vocal loudness can be harmful to vocal health, so does the use of an inappropriate mode of phonation."[30]

The parameter intensity is considered to be important to performers engaged in vocally demanding professions and hobbies and physiologically inefficient SPL control strategies are acknowledged as potentially injurious to vocal fold tissues.[31] Hoffman-Rudy et al. studied three categories of professional voice users/vocal performers (musical theatre, choral ensembles and street theatre): "These high-risk vocal performers produce their singing or theatre voice at their maximum vocal effort level."[32] Plant and Younger come to the same conclusions: "In general, intensity increases with higher subglottic air pressures, but there are also considerable variations both between individuals and within different portions of the frequency-intensity range for a given subject".[29]

Vocal nodules in singers and hyperfunctional voice disorders in general develop by vocal strain, forcing, speaking or singing in the wrong vocal range, and by vocalizing too loudly.[33] Vocal attrition is associated with the use of increased loudness levels for extended periods of time.[34]

The use of voice range profile analysis as a tool for objective voice classification

De Bruyne explains the pathophysiology of vocal fatigue by various forms of vocal misuse: an inappropriate pitch and loudness disturb the balanced functioning of the phonatory muscles as well as the mucosa.[35] The decrease of intensity is one of the major symptoms of vocal fatigue.[36] The relation between frequency and intensity is expressed in the voice range profile (VRP), a graph showing voice intensities at various frequencies (Fig. 1). The frequencies are represented at the abscissa and the intensities on the ordinate. VRP analysis has been used for many years in clinical practice.

Behrman et al. cite a conclusion of the International Voice Committee (1986), which considers the VRP as more useful for a within-subject measure than for a between-group measure but it was not applied to voice classification.[37]

Nevertheless, the Committee recommended that the VRP be studied further to achieve a classificatory scheme for the different kinds of upper and lower contour patterns observed. It can be assumed that different VRP types could be indicative of different voice types. Some authors, praising the many possibilities of VRP analysis, give some hints in this direction.[38-45] Some authors assume that VRP provides useful information for classifying a voice. However, no specific interpretation of VRP results regarding voice classification can be found in the scientific literature.[46-51] This has not only to do with difficulties in interpreting the VRP, but also with the general conviction that voice classification belongs to the domain of singing teachers, conductors and other people with musical training, who claim to possess a trained ear for assessing voice types. The many contradictions and discussions on this subject indicate the need of an objective methodology of voice classification. The objective measurement of voice intensity, combined with the objective measurement of the frequency vocal range as applied in VRP could be considered as an objective starting point for voice classification.

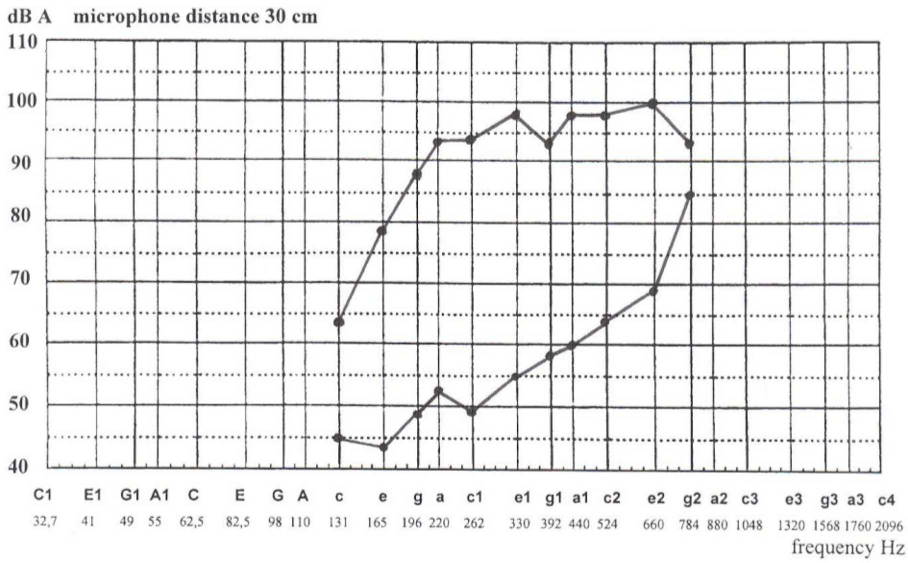


Figure 1. A Voice Range Profile of an adult woman. The frequencies are represented at the abscissa and the intensities on the ordinate.

Objectives and research questions

In traditional voice classification different parameters are considered, like the quality of the sound, the intensity with which the individual can sing, how high the chest voice can be used, and the comfort that the singer feels and the frequency of the register transition points. The weight to be assigned to these parameters is not easy to define and requires great experience. Due to the subjective nature of these parameters, combination of these parameters or incorporation into an algorithm is not readily possible. The question arises if there exist three types of female and male voices by nature. In an attempt to break out of the controversy, a new perspective is adopted in this study by letting the data speak for itself. Such an approach, called data-driven, imposes minimal assumptions on the nature of the data, what elements to use for its analysis, and in our case even the existence of natural voice groups. Voice range profiling is an objective measurement of minimum and maximum intensity at all fundamental frequencies that span the singer's range. Therefore, the voice range profile (VRP) reflects the capacity of the voice. The pertinent study uses VRP-derived parameters that are commonly applied in clinical practice. But it could very well be that the latter are insufficient to identify any natural voice groups.

Based on literature data and the results of the various exploratory studies, mentioned above (Addendum 1-3) the following research questions (RQ) are put forward:

RQ1 To verify the existence of individual- or combinations of commonly used and easily understandable VRP parameters with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

RQ2 If the answer to RQ1 is negative: to verify whether individual- or combinations of *non* commonly used and possibly non-clinically interpretable VRP parameters (more "intelligent" parameters) can partition the data into a number of clearly separated clusters.

Outline of the thesis

Chapter 1: General introduction. A frame of reference of classification of the human voice is presented. In this frame three explorative studies were carried out. In a first study, it was explored how contemporary singing teachers of Classical and Commercial Music deal with voice classification today and which criteria they use to classify their singing students (Addendum 1); in a second study, voice classification by singing teachers of two different types of conservatories (Classical singing and Musical Theatre) and how the singing students themselves think about the voice label they received by their singing teachers was explored (Addendum 2); in a third study, a survey is presented of the results of voice classification by singing teachers compared with the results of voice classification based on an algorithm deduced from the limits of frequency range/tessitura as found in the literature (Addendum 3). Objectives and research questions are formulated and an outline of the study is presented.

Chapter 2: Pilot study: Verification of the existence of individual- or combinations of commonly used and easily understandable VRP parameters with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Chapter 3: Verification of the existence of individual- or combinations of *not* commonly used and *not* easily understandable VRP parameters (more "intelligent" parameters) with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Chapter 4: Verification of the existence of individual- or combinations of *not* commonly used and *not* easily understandable VRP parameters (more "intelligent" parameters) with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic male voice categories.

Chapter 5: General discussion. The methodology and results of the various studies are discussed. A syntax is made and issues for clinical applications and future research are formulated.

References

1. Watson PJ, Hoit JD, Lansing RW, Hixon TJ. Abdominal Muscle Activity During Classical Singing. J Voice 1989;24-31.
2. Stathopoulos ET, Sapienza C. Respiratory and Laryngeal Function of Women and Men During Vocal Intensity Variation. J Speech and Hear Res 1993;64-75.
3. Phyland DJ, Oates J, Greenwood KM. Self-Reported Voice Problems Among Three Groups of Professional Singers. J Voice 1999;602-611.
4. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Paris: Librairie Maloine SA, 19612.
5. Erickson ML, Perry S, Handel S. Discrimination Functions: Can They Be Used to Classify Singing Voices? J Voice 2001;492-502.
6. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
7. Burns P. Clinical management of Country and Western singers. J Voice 1991;5:349-353.
8. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.
9. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. New York: Raven Press 1991;49-68.
10. Abitbol J. L'odyssée de la voix. Paris: Editions Robert Laffont SA; 2005.
11. Barthélémy Y. La voix libérée, Une nouvelle technique pour l'art lyrique et la rééducation vocale. Paris: Editions Robert Laffont, SA;1984.
12. Gilman M, Merati AI, Klein AM, Hapner ER, Johns MM. Performer's attitude toward seeking health care for voice issues: understanding the barriers. J Voice 2009;23: 225-228.
13. Blanchard R, de Candé R. Dieux et divas de l' opéra. Paris: Librairie Plon;1987.
14. Matheopoulos H. Les Ténors du siècle: de Caruso à Pavarotti. Paris: Editions de la Martinière; 1999.
15. Segond A. Divines divas. Paris: Gallimard; 2002.
16. Garde E. La Voix. Paris: Presses Universitaires de France;1954.
17. Bonet M, Casan P. Evaluation of Dysphonia in a Children's Choir. Folia phoniatr Logop 1994;46:27-34.
18. Welch GF, Sergeant DC, MacCurtain F. Xeroradiographic Electrolaryngographic Analysis of Male Vocal registers. J Voice 1989;3:244-256.
19. Sataloff RT. The Professional Voice: Part III. J Voice 1987;1:283-292
20. Radionoff SL, Satterfield TR, Lee E. Commercial Music: A survey of degree granting institutions in the United Kingdom. Choice for Voice 2008. Multidisciplinary Approaches to Performance, Health and Research in Voice, London: The British Voice Association 2008:1-22.

21. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. New York: Raven Press 1991:49-68.
22. Erickson ML, Perry S, Handel S. Discrimination Functions: Can They Be Used to Classify Singing Voices? J Voice 2001:492-502.
23. Erickson ML. The Interaction of Formant Frequency and Pitch in the Perception of Voice Category and Jaw Opening in Female Singers. J Voice 2004;18:24-37.
24. Bloothooft G, Plomp R. The timbre of sung vowels. J Acoust Soc Am 1988;84(3): 847-860.
25. Brewer DW. The Laryngologist's Perspective. In: Scherer RC, Ed. The Integration of Voice Science, Voice Pathology, Medicine, Public Speaking, Acting, and Singing. J Voice 1994;359-374.
26. Sapienza CM, Stathopoulos ET. Speech Task Effects on Acoustic and Aerodynamic Measures of Women with Vocal Nodules. J Voice 1995;413-418.
27. Punt NA. The Singer's and Actors's Throat. W. Heinemann Medical Books Ltd. London, 1967;2.
28. Sundberg J. Perceptual Aspects in Singing. J Voice 1994;8:106-122.
29. Plant RL, Younger RM. The Interrelationship of Subglottic Air Pressure, Fundamental Frequency, and Vocal Intensity During Speech. J Voice 2000;170-177
30. Sundberg J. Vocal Fold Vibration Patterns and Modes of Phonation. Folia Phoniatr Logop 1995;47:218-228.
31. Reich AR, Mc Henry MA. Respiratory Volumes in Cheerleaders with a History of Dysphonic Episodes. Folia phoniat 1987;39:71-77.
32. Hoffman-Ruddy B, Lehman J, Crandell C, Ingram D, Sapienza C. Laryngostrobo-scopic, Acoustic, and Environmental Characteristics of High-Risk Vocal Performers. J Voice 2001;543-552
33. Heinemann M. Stimmlippenknötchen - auch eine Definitionsfrage. Sprache - Stimme - Gehör 1981;5:27-31.
34. Sapienza CM, Crandell CC, Curtis B. Effects of Sound-Field Frequency Modulation Amplification on Reducing Teachers' Sound Pressure Level in the Classroom. J Voice 1999;375-381.
35. De Bruyne F. Stembelasting en de larynx. Stem-, Spraak-, en Taalpathologie 1996;5: 65-72.
36. Köslyk BE, Rochet AP. Laryngeal Airway Resistance in Teachers with Vocal Fatigue: A Preliminary Study. J Voice 1998;287-299.
37. Behrman A, Agresti CJ, Blumstein E, Sharma G. Meaningful Features of Voice Range Profiles from Patients with Organic Vocal Fold Pathology: A Preliminary Study. J Voice 1996;3:269-283.
38. Dejonckere Ph. Le phonétogramme, son intérêt clinique. Les cahiers d' ORL 1977;12: 865-872.
39. Gross M. Stimmfeldmessung, eine objektivere Methode der Stimmdiagnostik. Sprache - Stimme - Gehör 1980;4:100-101.
40. Hollien H, Keister E. Frequency tracking abilities of singers and non-singers. In: Lawrence VL, Ed. Transcripts of the sixt symposium “Care of the Professional Voice”. New York: The Voice Foundation 1977:42-44.
41. Klingholz F, Jolk A, Martin F. Stimmfelduntersuchungen bei Knabenstimmen (Tölzer Knabenchor). Sprache - Stimme - Gehör 1989;13:107-111.
42. Pabon JPH, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. J Speech Hear Res 1988;31:710.
43. Pabon JPH. Kwaliteitsaspecten in het fonetogram.Logopedie 1997;10:20-33.
44. Seidner W, Kröger H, Wernecke KD. Numerische Auswertung spektraler Stimmfelder. Sprache - Stimme - Gehör 1985;9:10-13.
45. Siegwart H, Scherer KR. Acoustic Concomitants of Emotional Expression in Operatic Singing: The Case of Lucia in Ardi gli incensi. J Voice 1995;249-260.
46. Cleveland TF. A Clearer View of Singing Voice Production: 25 Years of Progress. J Voice 1994;8:18-23.

47. Cleveland TF, Stone RE Jr, Sundberg J, Iwarsson J. Estimated Subglottal Pressure in Six Professional Country Singers. J Voice 1997;403-409.
48. Colton RH. Vocal Intensity in the Modal and Falsetto Registers. Folia phoniat 1973; 25:62-70.
49. Klingholz F, Martin F. Zum Entstehungsmechanismus des Vocal Fry (VF). Sprache - Stimme - Gehör 1983;7:16-21.
50. Miller DG, Schutte HK. Toward a Definition of Male 'Head' Register, Passagio, and 'Cover' in Western Operatic Singing. Folia Phoniatr Logop 1994;46:157-170.
51. Pedersen MF. Computed phonetograms in adult patients with benign voice disorders before and after treatment with a nonsedating antihistamine (Loradine). Folia Phoniatr 1991;43:60-67.

Chapter 2

Discrimination of basic female voice categories by Voice Range Profiling. A Study on the Reliability of Voice Frequency and Intensity Related Parameters

H. Lycke, W. Decoster, A. Ivanova, M.M. Van Hulle, F.I.C.R.S. De Jong

Pilot study. Unpublished results

Abstract

Aims: The purpose of the current study is to verify the existence of individual- or combinations of commonly used and easily understandable VRP parameters (i.e. the lowest- and highest frequencies the subject's voice can reach, the minimum and maximum intensity of the lowest octave, and the intensity and frequency when entering and exiting the register transition zone) with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Methods: The data from 327 females between 18 and 25 years (mean age: 20.8 years) from different conservatories were analyzed. Two complementary clustering procedures were used: Ward's minimum variance method, to determine the number of clusters, and k-means clustering, to assign the subjects to the clusters.

Results: It was found that the measure of separation was very low since no alto, mezzo and soprano clusters could be observed.

Conclusions: These results, however, do not mean that there would not exist other, more "intelligent" parameter combinations, even non-linear ones, of the voice range profile, that could lead to a clear cluster separation, and that could serve as the basis to resolve the riddle of voice classification.

Introduction

A careful watch on the physiological limits of the voice is a prerequisite to prevent functional and organic voice disorders, especially in the singing voice.[1-8] Therefore, voice specialists stress the importance of a correct voice classification.[1-3,9-15] Various singing teachers use different criteria for voice classification. Obviously, there is a need on consensus of objective criteria on which voice classification should be based.

Lycke et al. showed that all 22 singing teachers at one classical and two Musical Theatre conservatories in 3 European countries used frequency range / tessitura as a basic criterion to classify the voice of 165 of their students (see: Addendum 1 of this thesis). In an explorative study, the results of the voice classification of the singing teachers were compared to the results of voice classification that was based on the limits of voice frequency range, according to 38 authors (see: Addendum 3 of this thesis). It was found that the voice classification of the singing teachers did not agree well with the voice classification based on frequency range. Therefore, frequency range appears not to be a suitable voice parameter for classification. Voice is a multidimensional phenomenon and should be addressed as such. Voice Range Profiling is considered to be a suitable method of establishing vocal capacities.[18-20] It is therefore tempting to verify whether frequency- and intensity parameters derived from the Voice Range Profile can discriminate between the three basic female voice types, alto, mezzo-soprano, and soprano. It looked plausible to assess if the combined measurement of frequency and intensity parameters can discriminate between the three basic female voice types, alto, mezzo-soprano, and soprano.

The aim of the current study is to verify the existence of individual- or combinations of commonly used and easily understandable VRP parameters (i.e. the lowest- and highest frequencies the subject's voice can reach, the minimum and maximum intensity of the lowest octave, and the intensity and frequency when entering and exiting the register transition zone) with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Methods

The data from 327 females between 18 and 25 years (mean age: 20.8 years) from different conservatories in three European countries (Musical Theatre and Classical), was investigated. Subjects were recruited from a private practice (HL) and during auditions for Opera and Musical Theatre. Data entries that only measured part of the Voice Range Profile were removed. This reduced the data set to 268 subjects.

A voice range profile (VRP) was performed according to the UEP recommendations[16] using the ling Waves Voice Diagnostic Center, version 2.5; 2007, with a Center 322 Data Logger Sound Level Meter.

To analyze the data, 14 parameters from the VRP, listed in Table 1, and age, were taken to verify whether a natural group separation is present. For each subject the value of the 14 parameters was determined and a 14 dimensional data vector was created. Thereafter, clustering analysis was carried out in this 14 dimensional space.

As is common practice, two complementary clustering methods were used, one hierarchical and one non-hierarchical, Ward's minimum variance method[17], to determine the number of clusters, and given this number, apply k-means clustering to assign the subjects to the clusters.

The analyses were performed in SAS (release 9.2).

Table 1. The voice range profile parameters used for the assessment of natural group separation.

Parameters	Description of parameters
Highest frequency	Highest frequency (in Hz) the voice can reach
Lowest frequency	Lowest frequency (in Hz) the voice can reach
Maximum intensity	Maximum intensity (in dB(A)) of the lowest octave @ 123, 131, 147, 165 Hz
Minimum intensity	Minimum intensity (in dB(A)) of the lowest octave @ 123, 131, 147, 165 Hz
Location of register transition	The intensity (in dB(A)) and frequency (in Hz) when entering and exiting the register transition zone

Results

After applying Ward's procedure, it can be concluded that there are possibly two or three clusters in the data (Figure 1). Table 2 summarizes the clustering results for both cases.

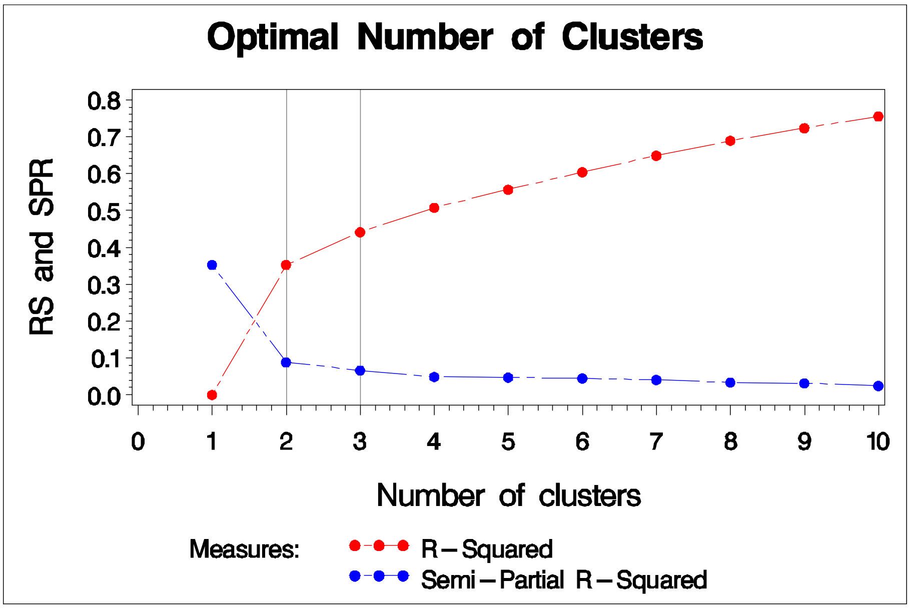


Figure 1. The results of the Ward's procedure showing possibly two or three clusters in the data.

Table 2. Results of clustering based on six frequency/intensity related parameters.

	Overall R-squared	R-squared for the best separation variable
2 clusters	0.18	Max. intensity_123, R^2=0.66
3 clusters	0.31	Max. intensity_123, R^2=0.70

The measure of separation -- overall R-squared -- is very low in the case of two or three clusters. If we look at the separation for the individual parameters, then the best case is for the maximum Intensity at 123 Hz. Note that this parameter has not always a value for each data entry (i.e., missing values). In the Ward clustering algorithm, such entries are ignored when determining the number of clusters. In the k-means clustering algorithm they are also ignored when estimating the cluster means ("seeds"), but when assigning the data entries to the clusters, the vector dimensions corresponding to the missing values are ignored (note that mean substitution could not be performed here since the data is expected to be multimodal). The downside of this procedure is that the data set effectively used for estimating the cluster number and the cluster means is reduced in size, which in turn could explain the lower overall R-squared values.

Discussion

The voice source varies along the phonatory dimensions of vocal frequency and intensity.[6] Together with other features such as timbre, this yields a multi-dimensional phenomenon. A VRP is a combined graph of frequency and intensity measurements. Voice Range Profiling can be considered as an appropriate method of assessing vocal capacities.[18-20] The choice of the set of following parameters was decided on a physiological basis.

Lowest and highest frequency

In Voice Range Profiling the melodic vocal range is measured, i.e., all producible frequencies. Tessitura was not taken into account because this is prone to a subjective interpretation. Scores for songs have always been written for specific voice types, depending on the range of frequencies (and not intensity). Therefore, frequency scoring is considered to be the most important procedure for voice classification by singing teachers. However, voice range / tessitura alone proved not to be a reliable voice classification parameter (Addendum 3).

Maximum intensity

Maximum and minimum intensity are important to a singer in order to express the many nuances of emotion in a song.[21] Although scores do not mention specific intensity requirements, it is evident that intensity plays a major part in the actual performance on stage. The results of a previous study showed that voice intensity was only used by 12.1% private singing teachers of all genres compared to 54.5% classical conservatory singing teachers as a voice classification criterion (see: Addendum 1 of this thesis). This can be explained by the fact that the human ear proves to be less susceptible for intensity differences.[22] Intensity differences of more than 10 dB (A) are required to be detectable.[23] The intensity difference that is perceived between singers and non-singers does not consistently correspond to a sound pressure level (SPL) difference.[24] In Voice Range Profiling the voice intensity is objectively measured. Plant and Younger (2000) stated “In general, intensity increases with higher subglottic air pressures, but there are also considerable variations both between individuals and within different portions of the frequency-intensity range for a given subject”. Also Wilson and Leeper (1992) and Stathopoulos and Sapienza (1993) found great variability regarding changes in laryngeal airway resistance as a function of vocal sound pressure in both male and female subjects. Keilmann et al.(1994) came to the same conclusions for children. Knowing the maximum intensity of a voice at different frequencies is also important from a clinical viewpoint. Hoffman-Rudy et al. (2001) studied three categories of professional voice users/vocal performers: musical theatre, choral ensembles, and street theatre. These “high-risk vocal performers” produced their singing or theatre voice at their maximum vocal effort level.[29] According to Mc Henry and Reich (1985), the relationship between vocal SPL control and dysphonia has received relatively little attention in research. They noted cheering intensity levels, associated with SPL´s that were close to their maximal SPL´s. Hyperfunctional voice disorders and vocal nodules in singers develop by vocal strain, forcing, singing in the false vocal range, and by vocalizing too loud.[30] According to Sapienza et al. (1999), vocal attrition is associated with the use of increased loudness levels for extended periods of time and continued use of a high intensity level can lead to further tissue changes such as vocal nodules.[31] De Bruyne (1996) explains the pathophysiology of vocal fatigue by various forms of vocal misuse: inappropriate pitch and loudness levels disturb the balanced functioning of the phonatory muscles as well as the mucosal function, and the reduction of intensity is one of the major symptoms of vocal fatigue.[32] The singers in the study of Bennett (1981) sang more loudly in the upper part of their range than in the lower, but, for each of them, the SPL increased at a different rate and reached its maximum at a different point of the range. He concluded that “not the specific values for any one pitch, but rather the behaviour of the SPL over the entire range was characteristic for each singer”.[33]

Minimum intensity

Minimum intensity is an important parameter in clinical practice.[34] When taking VRP’s, Klingholz (1990) mentions the difficulty for trained singers to phonate with low intensity. How soft a voice can be produced on a given frequency is an important diagnostic parameter in voice disorders. Especially soft phonations are difficult to produce in the highest frequencies.[35-37] According to Pabon (1997), the irregularity of the lower contour of the V.P.R. is caused by the inconstancy of the vibration of the vocal folds. The lower contour of the VRP is strongly influenced by training.[38] In his acoustic study of the voice range profile, Titze (1991,1992) predicted that the bottom curve of the Voice Range Profile should rise faster than the top curve, assuming that soft phonation produces a more sinusoidal glottal waveform than loud phonation. According to this author, the results of the minimum intensities can be very revealing.[36-37]

Location of register transition

Lycke et al. stress the great difference in appreciation of the register transition as a criterion for voice classification. Only 9.0% of private singing teachers against 60.0% of conservatory singing teachers think that register transition is an important voice classification criterion (Addendum 1). Although vocal registers are known to occupy separate portions of the total fundamental frequency range of the human voice[39], perceptual judgments alone invariably leads to discussions and controversies.[40] Manuals of singing education and voice studies mention divergent statements on shifts and boundaries of register zones, related to voice categories. VPR analysis can clearly locate a register transition. According to Klingholz et al. (1985) and Airainer and Klingholz (1993), markers of the register ranges are the transitions that are indicated “by minima in the forte contour and maxima in the piano contour, and minima of the dynamics, at specific pitches”.[41-42] Bunch (1982) stated: ”Classification of voices is made chiefly according to where the quality notes are located in the voice and where the depth and ease of sound are located in the range”.[9] Together with the (subjective) appreciation of the quality of that tone, the intensity of the produced tone is all important. What matters is the potential intensity of the tones allowing the subject to be heard on the scene, time after time, without hurting his vocal instrument. These potentialities are expected to be present in the VRP.

Taking into account the multidimensional phenomenology of the voice it may be expected that a combination of parameters is the most suitable method to describe voice characteristics, such as the voice categories alto, mezzo and soprano in the female voice. In this study, the measure of separation -- overall R-squared -- is very low in the case of two or three clusters. If the separation for individual parameters is considered, then the best case is for the maximum intensity. However, since for this parameter there were many missing values, this finding should be considered with cautiousness when deciding on its importance. A clustering analysis, based on the clinically important voice parameters, i.e. the lowest- and highest frequencies the subject's voice can reach, the minimum and maximum intensity of the lowest octave, and the intensity and frequency when entering and exiting the register transition zone, did not lead to clearly separable clusters that could be connected with the voice classes alto, mezzo and soprano.

Conclusions

Voice Range Profiling is an excellent tool for establishing vocal capacities due to its objective nature. Therefore, it looked plausible to us to rely on measurements of frequency and intensity as indicators for voice classification. The results of the cluster analysis of various frequency/intensity parameters, however, showed that this is not the case. Further studies are necessary to look for other, more "intelligent" parameters and possibly also parameter combinations to resolve the problem of voice classification.

References

1. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209-216.
2. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
3. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Paris: Librairie Maloine SA, 1961[2].
4. Burns P. Clinical management of Country and Western singers. J Voice 1991;5:349-353.
5. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.
6. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. New York: Raven Press, 1991;49-68.
7. Abitbol J. L'odyssée de la voix. Paris: Editions Robert Laffont SA; 2005.
8. Barthélémy Y. La voix libérée, Une nouvelle technique pour l'art lyrique et la rééducation vocale. Paris: Editions Robert Laffont SA;1984.

9. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F. Wyke BD, Eds. Disorders of Human Communication 6. Wien New York:Springer-Verlag; 1982.
10. Garde E. La Voix. Paris: Presses Universitaires de France;1954.
11. Dejonckere Ph. Le phonétogramme, son intérêt clinique. Les cahiers d' ORL 1977;12:865-872.
12. Gross M. Stimmfeldmessung, eine objektivere Methode der Stimmdiagnostik. Sprache - Stimme – Gehör 1980;4:100-101.
13. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimmfeld. Sprache - Stimme – Gehör 1985;9:109-111.
14. Pabon JPH, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. J Speech Hear Re 1988;31:710-722.
15. Seidner W, Kröger H, Wernecke KD. Numerische Auswertung spektraler Stimmfelder. Sprache - Stimme – Gehör 1985;9:10-13.
16. Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement / phonetography. Folia Phoniatr 1983;35:286-288.
17. Sharma S. Applied Multivariate Techniques. Wiley, New York, 1996;217.
18. Schutte HK. Over het fonetogram. Log en Fon 1975;47:82-92.
19. Klingholz F, Martin F. Die quantitative Auswertung der Stimmfeldmessung. Sprache - Stimme - Gehör 1983;7:106-110.
20. Awan SN. Phonetographic profiles and F°-SPL characteristics of untrained versus trained vocal groups. J Voice 1991;5:41-50.
21. Seidner W, Wendler J. Die Sängerstimme. Phoniatrische Grundlagen für die Gesangsausbildung. Heinrichhofen, Willemshaven, 1997.
22. Winckel F. Die Ästhetischen Komponenten der Stimmgebung. Acta Physiol Pharmacol Neerl 1956;5:56-72.
23. Brewer DW. The Laryngologist's Perspective. In: Scherer, R.C., Ed. The Integration of Voice Science, Voice Pathology, Medecine, Public Speaking, Acting, and Singing. J Voice 1994;359-374.
24. Sundberg J. Perceptual Aspects in Singing. J Voice 1994;8:106-122.
25. Plant RL, Younger RM. The Interrelationship of Subglottic Air Pressure, Fundamental Frequency, and Vocal Intensity During Speech. J Voice 2000;170-177.
26. Wilson JV, Leeper HA. Changes in Laryngeal Airway Resistance in Young Adult Men and Women as a Function of Vocal Sound Pressure Level and Syllable Context. J Voice 1992;235-245.
27. Stathopoulos ET, Sapienza C. Respiratory and Laryngeal Function of Women and Men During Vocal Intensity Variation. J Speech and Hear Res 1993;64-75.
28. Keilmann A, Bader C-A, Bergler W. Aerodynamische Aspekte der kindlichen Phonation. Sprache-Stimme-Gehör 1994;125-129.
29. Hoffman-Ruddy B, Lehman J, Crandell C, Ingram D, Sapienza C. Laryngostroboscopic, Acoustic, and Environmental Characteristics of High-Risk Vocal Performers. J Voice 2001;543-552.
30. McHenry MA, Reich AR. Effective airway resistance and vocal sound pressure level in cheerleaders with a history of dysphonic episodes. Folia Phoniatr 1985;37:223-231.
31. Sapienza CM, Crandell CC, Curtis B. Effects of Sound-Field Frequency Modulation Amplification on Reducing Teachers' Sound Pressure Level in the Classroom. J Voice 1999;375-381
32. De Bruyne F. Stembelasting en de larynx. Stem-, Spraak-, en Taalpathologie. 1996; 5:65-72.
33. Bennett G. Singing Synthesis in Electronic Music. In: Research Aspects on Singing. Autoperception. Computer Synthesis. Emotion. Health. Voice Source. Royal Swedish Academy of Music 1981;37-41.
34. Stone RE, Bell CJ, Clack TD. Minimum intensity of voice at selected levels within pitch range. Folia Phoniatr 1978;30:113-118.

35. Klingholz F. Das Stimmfeld. Eine praktische Anleitung zur Messung und Auswertung. Verlag J. Peperny, München, 1990.
36. Titze IR. Acoustic Interpretation of the Voice Range Profile (Phonetogram). NCVS Status and Progress Report-1, 1991;15-37.
37. Titze IR. Vocal Efficiency. J Voice 1992;6:135-138.
38. Pabon JPH. Kwaliteitsaspecten in het fonetogram. Logopedie 1997;10:20-33.
39. Colton RH. Vocal Intensity in the Modal and Falsetto Registers. Folia phoniat 1973;25:62-70.
40. Röhrs M, Pascher W, Ocker C. Untersuchungen über das Schwingungsverhalten der Stimmlippen in verschiedenen Registerbereichen mit unterschiedlichen stroboskopischen Techniken. Folia phoniat 1985;37:113-118.
41. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimmfeld. Sprache - Stimme - Gehör 1985;9:109-111.
42. Airainer R, Klingholz F. Quantitative Evaluation of Phonetograms in the Case of Functional Dysphonia. J Voice 1993;7:136-141.

Chapter 3

Identification of Three Basic Female Voice Groups by Voice Range Profile-Derived Parameters

Part of this work has been published in:
Discrimination of Three Basic Female Voice Types in Female Singing Students by Voice Range Profile-Derived Parameters.
H.Lycke, W. Decoster, A. Ivanova, M.M. Van Hulle, F.I.C.R.S de Jong,
Folia Phoniatr Logop 2012;64:80–86. DOI: 10.1159/000337042

Abstract

Aims: To verify the existence of individual- or combinations of *not* commonly used and *not* easily understandable clinical VRP parameters (also called "features") with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Methods: The voice range profiles of 206 female conservatory singing students were recorded, parameterized into more compact descriptions ("features"), and subjected to a cluster analysis.

Results: The 3 cluster case provided the most consistent solution, across all feature combinations. The feature that led to the best cluster separation was the ratio of the perimeter length of the chest voice part of the voice range profile versus the total perimeter length.

Conclusions: Based on a statistical analysis of the voice range profile parameters, the ratio of the perimeter length of the chest voice versus the total perimeter length was shown to yield a clear separation into three basic female voice types which in turn may serve as the basis to resolve the riddle of voice classification.

Introduction

Traditionally, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor. Many subtypes are described according to different roles and based on various characteristics, such as loudness, timbre, flexibility, vibrato, temperament, expression and personality.

Classification of a voice determines the frequency and intensity range in which a singer can work without harming or fatiguing his voice and to which repertoire he should be assigned by the singing teacher.[1-7] Correct classification of the singer's voice is indispensable in order to achieve optimum performance.[8] Incorrect voice classification can cause or increase functional and organic voice disorders.[5-6,9-13]

Despite the importance of correct voice classification, there is neither a generally accepted protocol for voice classification, nor there is consensus about what parameters should be used.

Voice range profiling is considered to be a suitable method of establishing vocal capacities.[14-15] Therefore, it may be expected that frequency and intensity parameters derived from the voice range profile (VRP) can discriminate between the three basic female voice types: alto, mezzo-soprano, and soprano.

The aim of this study was to verify the existence of individual- or combinations of *not* commonly used and *not* easily understandable clinical VRP parameters (also called "features") with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

Materials and methods

Participants and data acquisition

Voice range profiles from 206 female subjects, between 18 and 25 years old, from two Musical Theatre Conservatories, were investigated. The VRP's were performed according to the Union of European Phoniatricians (UEP) recommendations[16], using the Ling Waves Voice Diagnostic Center, version 2.5;2007, with a Center 322 Data Logger Sound Level Meter. Phonation was on the vowel /a:/ with a microphone distance of 30 cm. The minimum phonation time was 2 seconds. The data include maximum and minimum intensity measurements (in dBA) of a subject's voice taken at the fundamental frequencies that span the singer's range. At each F0, subjects were encouraged to phonate as softly and as loudly as possible, regardless of the produced vocal quality.

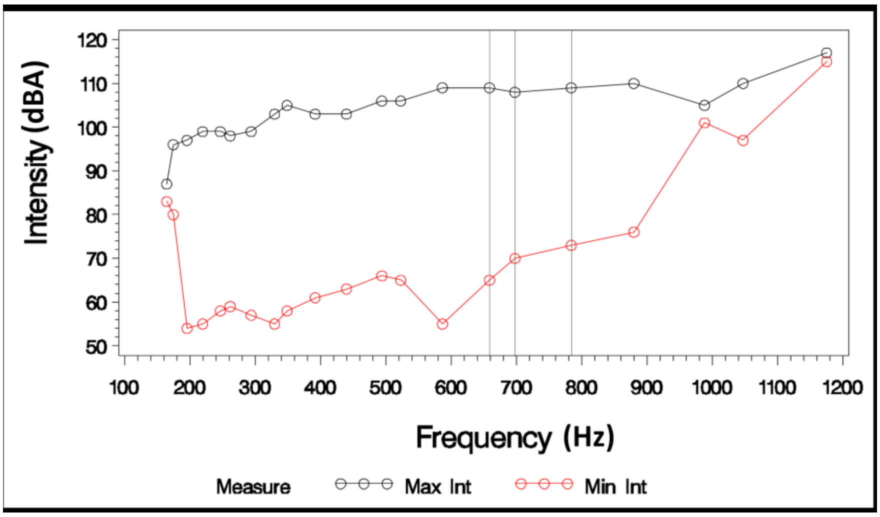


Figure 1. The voice range profile after the frequency axis was transposed to a linear scale. The two outer vertical lines define the borders of the transition zone (left and right lines) and the middle vertical line the dip of the register transition zone (at the dip in the maximum intensity).

An individual subject's VRP has, approximately, the shape of an ellipse. In order not to favor any frequency range in the analysis, the frequency axis was transformed from the non-linear note scale into a linear frequency scale (Hz) (Fig. 1). In the VRP the maximum intensity increases with increasing frequency. The register *zone* (including the marked register dip) was measured as follows: The beginning of the register zone is marked by the maximum intensity before a drop of intensity occurs. The register zone ends when the maximum intensity (at increasing frequency) exceeds the maximum intensity of the beginning of the register zone. This avoids argumentation about the "exact" register transition (point) and the possible difference when singing up or down. The marked break points reflect the more or less auditively perceived voice sound, according to the used technique.

Feature construction

Then, it was decided to represent, in a more compact manner, each VRP by a number of parameters - called features in our context - in connection with:

- the *geometry of the VRP* such as the surface area enclosed between the maximum and minimum intensity curves (Fig. 1), their frequency ranges, and their perimeter lengths;
- the *register transition zone* such as the intensity of the dip in the maximum/minimum intensity curves between the chest and head voice parts of the VRP and the frequency at which it occurs;
- the *geometry of the chest/head voice parts* of the VRP such as their surface areas and their perimeter lengths;
- the *linear characteristics of the minimum and maximum intensity curves* such as the slopes of the regression lines through the maximum and minimum intensity curves.

Finally, a number of voice frequency and intensity ratios and differences were defined based on some of the above features such as the ratio of the surface area of the chest voice to the total surface area enclosed by the maximum and minimum intensity curves. Another example is the ratio of the perimeter length of the chest voice part of the VRP to the total perimeter length.

In total, 49 features were defined. For a detailed description: see Addendum 4.

Statistical analysis and methods

Rather than looking at differences between the maximum and minimum intensity curves, another strategy was chosen. Each of the above features characterizes in a much more compact manner an individual's VRP. This opens the possibility to apply more powerful analysis techniques. Indeed, as the voice of every singer can be represented as a point in the space formed by one or more of the aforementioned features, one can apply statistical techniques to assess whether there exist natural groups in the data (called clusters) with which basic female voice categories can be discriminated. In order to answer this question the following strategy was developed:

- for every single feature and for every possible feature pair and so on, a clustering procedure is applied to the data;
- the single feature or feature combination that leads to clusters with the best cluster separating power (feature selection) is determined;
- the found clusters are related to the traditional voice frequency and intensity parameters.

Finally, based on a detailed statistical analysis, the found clusters are related to the voice classification performed by a number of singing teachers.

Clustering

Before starting the clustering analysis, the highly correlated (95% and higher) features were removed from the data set, in order to remove the mutually redundant features. Hence, the list was reduced to 34 features. The data set was standardized (equal variance along each feature dimensions) in order to eliminate the possible influence due to differences in scale (features with broadly distributed samples tend to dominate features with narrow distributions).

The following procedure was applied:

- for all single features, the Ward's minimum variance method to define the optimal number of clusters was applied (see further);
- subsequently, K-means clustering was applied on each single feature. The best single feature, leading to the best cluster separation, was retained;
- given the optimal number cluster, determined for the single features, K-means clustering was applied on all possible combinations of three features to define the best combination of three features, based on their cluster separation.

To define the optimal number of clusters, the Ward's minimum variance method was used, and the two following metrics obtained with this method were plotted as a function of the number of clusters:

- R-squared (RS) quantifies the between cluster variability vs. the total variability. The between cluster variability assesses the squared distances, per feature dimension, taken between the cluster centers (averages, centroids) and the center (average, centroid) of the whole data. The total variability is the squared distance between each data point and the center (average, centroid) of the whole data. RS varies in the interval [0, 1]. “0” means that there is no difference, “1” means that there is a maximal difference between the clusters;

- Semipartial R-squared (SPR), defines the homogeneity of each cluster.

It could be seen that the optimal number of clusters is three or four. For example, the one for feature *R4* is shown in Fig. 2.

Finally, for a given number of clusters (i.e. either 3 or 4), it was verified whether the data points of each cluster remained in the same cluster, for the case of the best single feature, as well as the best feature pair, and so on. This cluster member consistency formed the basis to opt for the three or four cluster solution (see further).

The analyses were performed with Statistical Analysis SAS/STAT® Software (release 9.2).

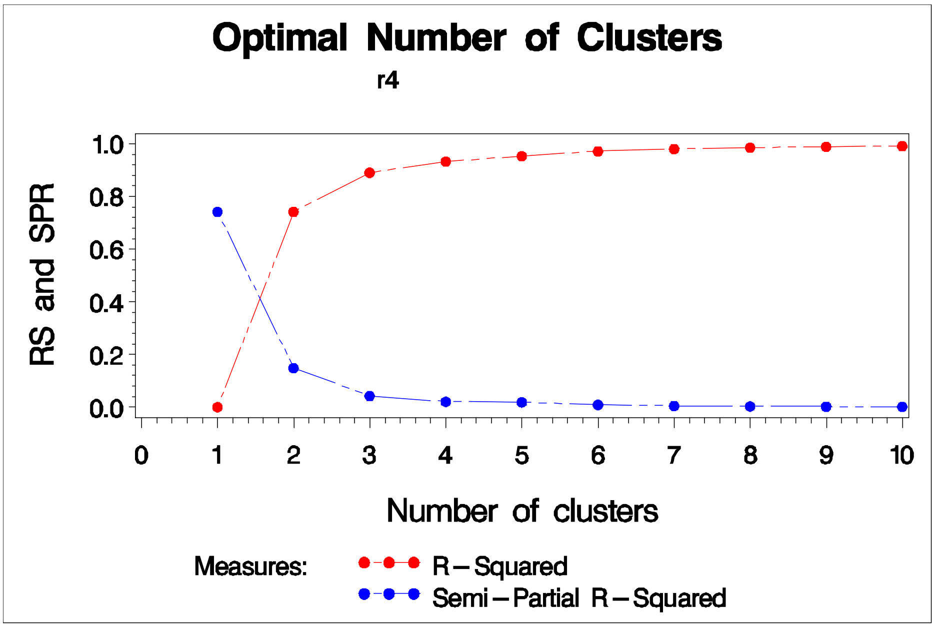


Figure 2. The results of the Ward’s procedure showing possibly two or three clusters in the data (for feature *R4*).

Results

As indicated above, when applying Ward's minimum variance method, 3 or 4 clusters turned out to be possible. We, therefore, restrict ourselves to the 3 or 4 cluster cases: when assuming 3 clusters, the single, two and three features cases are examined; the same was done when assuming 4 clusters. Note that no prior knowledge of the number of voice types is used in Ward's method or in the further evaluation of the 3 or 4 cluster options.

3 clusters solution

- *Single feature*

The best cluster separation, for three clusters, was obtained for feature *R4* (i.e. the ratio of the perimeter length of the chest voice part versus the total perimeter length – the perimeter length of both the minimum and maximum intensity curves is taken).
The overall R-squared is equal to the partial R-squared, and it is larger than 80%, which is an indication of a high cluster separation degree.
To better understand the meaning of *R4* in terms of the three clusters, the distribution of the data points along *R4* was examined. The result is that the members of cluster 1 have a low value of feature R4; those from cluster 2 have a high value of *R4* and, finally, those from cluster 3 have a middle value.

- *Combination of two features (feature pair)*

The best separation for three clusters, for the combination of two features, was obtained for *R4* and *Perim_Head* (i.e. the perimeter lengths of the maximum and minimum intensity curves corresponding to the head voice part).
The overall R squared is also larger than 80%, which is an indication of a high degree of cluster separation. The partial R squared indicates the same degree of separation.
The members of cluster 1 have a low value of feature *R4* and a high value of feature *Perim_Head*; those from cluster 2 have a high value of *R4* and a low value of *Perim_Head*; finally, those from cluster 3 have a middle value of *R4* and *Perim_Head*.

- *Combination of three features (feature triplet)*

The best separation for three clusters, for the combination of three features, was obtained for *R4, Perim_Head* and *R9* (i.e. the ratio of the perimeter length of the transition zone versus the perimeter of the chest voice part).

The overall R-squared is also larger than 80%, which is an indication of a degree of high cluster separation. The partial R-squared gives almost the same degree of separation.

The members of cluster 1 have a low value of feature *R4* and *R9* and a high value of feature *Perim_Head*; those from cluster 2 have a high value of *R4* and *R9* and a low value of *Perim_Head*; finally, those from cluster 3 have a middle value of *R4*, *Perim_Head* and *R9*.

The best single, two and three feature combinations yielded a similar cluster separation.

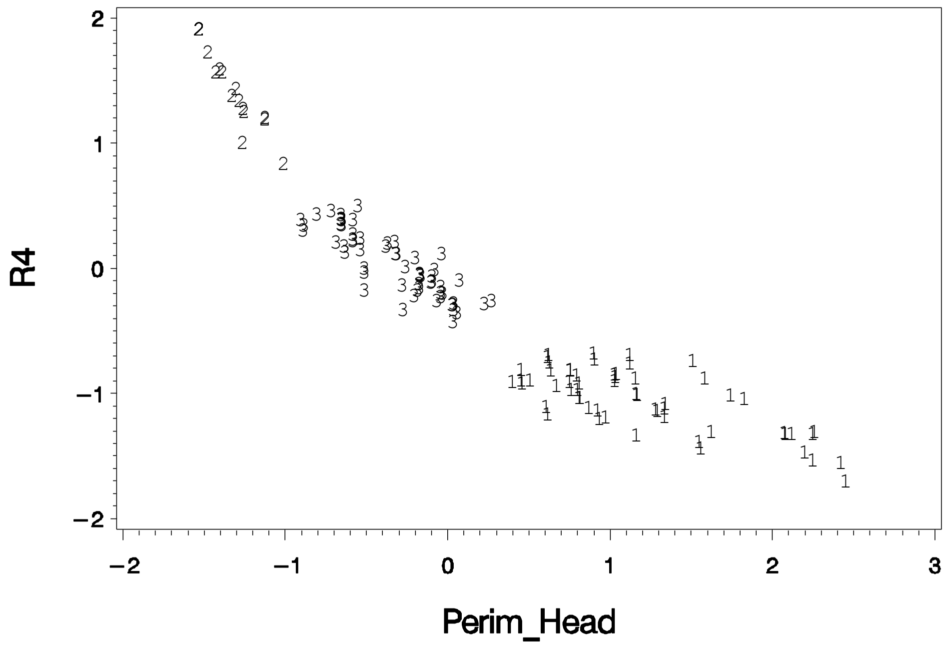
Interpretation

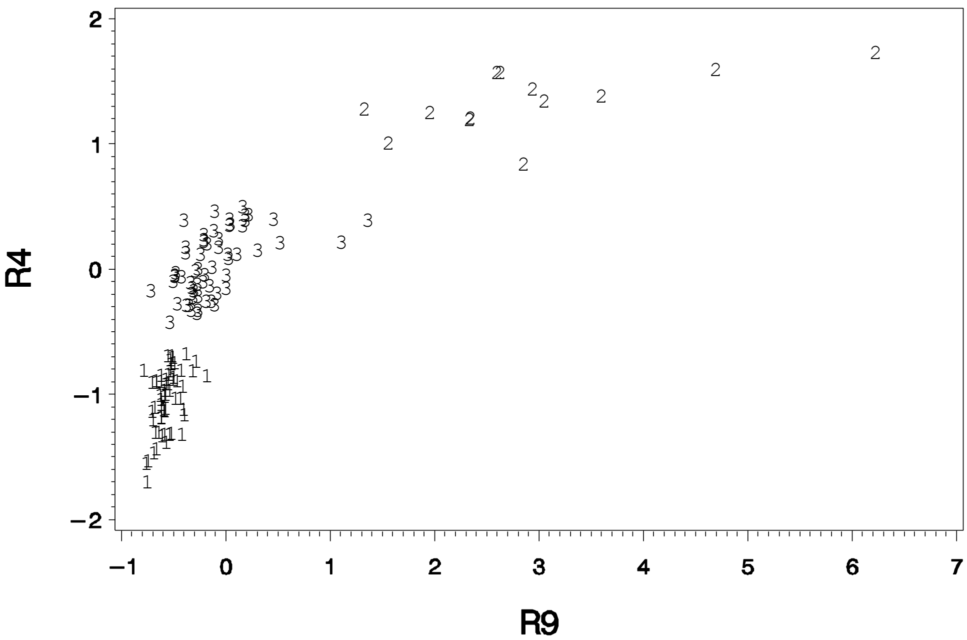
For 155 subjects out of the 206 (about 75% of data), and for each of the three retained feature combinations, the data was classified into the same cluster (Fig. 2). For the remaining 51 subjects, their data migrated from one cluster to another but only once. There is not a single subject for which the data was member of three different clusters.

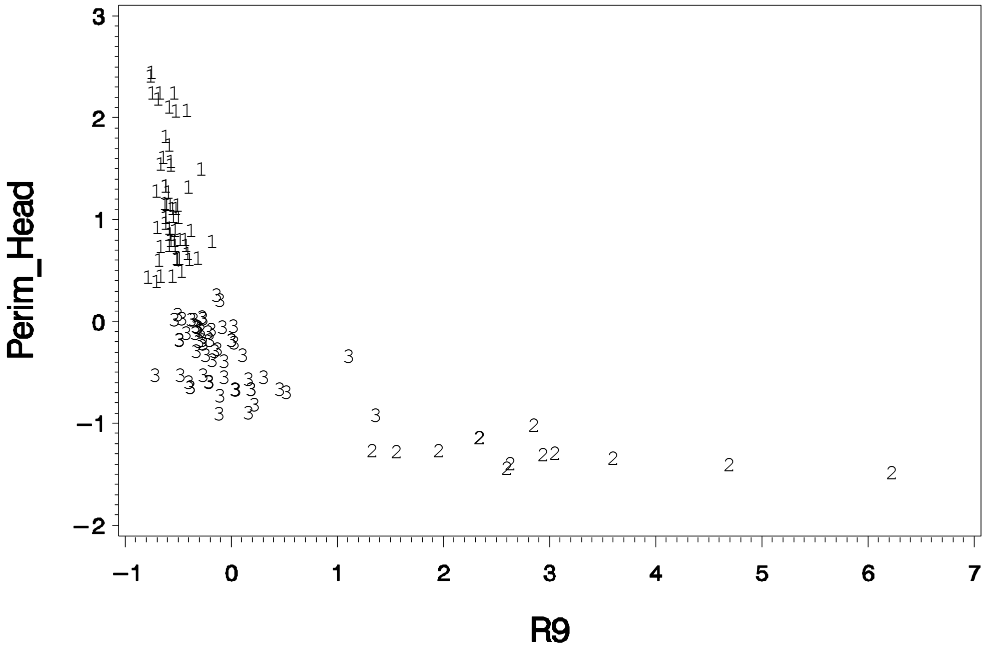
In order to attach a meaning to the three cluster solution that was obtained, a color code was added to the data points corresponding to the value they have for the more clinically understandable voice frequency and intensity parameters: *mean intensity, maximum intensity, minimum intensity, intensity range, mean frequency, and so on*.

We illustrate this only for the plot *R4* versus *Perim_Head* since the result is quite similar for the other combinations of best separating features (Fig. 3).

Figure 3 (page 48). Visualization of the cluster distribution. The three panels display the originally three-dimensional data as projections into two-dimensional subspaces formed by the three possible combinations of feature pairs: *R4* vs. *Perim_Head* (top), *R4* vs. *R9* (middle), *Perim_head* vs. *R9* (bottom). The position of each data point in each subspace follows from the respective feature scores along these axes. The data points are marked by the cluster number to which they belong (labels 1 to 3).


R4
Perim_Head
2
1
0
−1
−2
−2
−1
0
1
2
3


R4
R9
2
1
0
−1
−2
−1
0
1
2
3
4
5
6
7


Perim_Head
R9
3
2
1
0
−1
−2
−1
0
1
2
3
4
5
6
7

4 clusters solution

The same methods were used to investigate the 4 clusters solution. As a result, the best single feature is again *R4*, the best combination of two features is *R4*, *Perim_Head* and the best combination of the three features is *R4*, *Perim_Head* and *R9*. Hence, essentially the same result was obtained as in the 3 clusters solution, also with a very high degree of cluster separation (R-squared 85% and higher).

But when investigating the cluster members common to all feature combinations, it is found that only 88 subjects out of 206 (about 43 % of data) were classified into the same cluster, and that 163 subjects migrated from one cluster to another.

From the latter, we conclude that the 3 cluster case leads to the most consistent solution, across feature combinations.

Relation between 3 cluster solution and voice frequency and intensity parameters

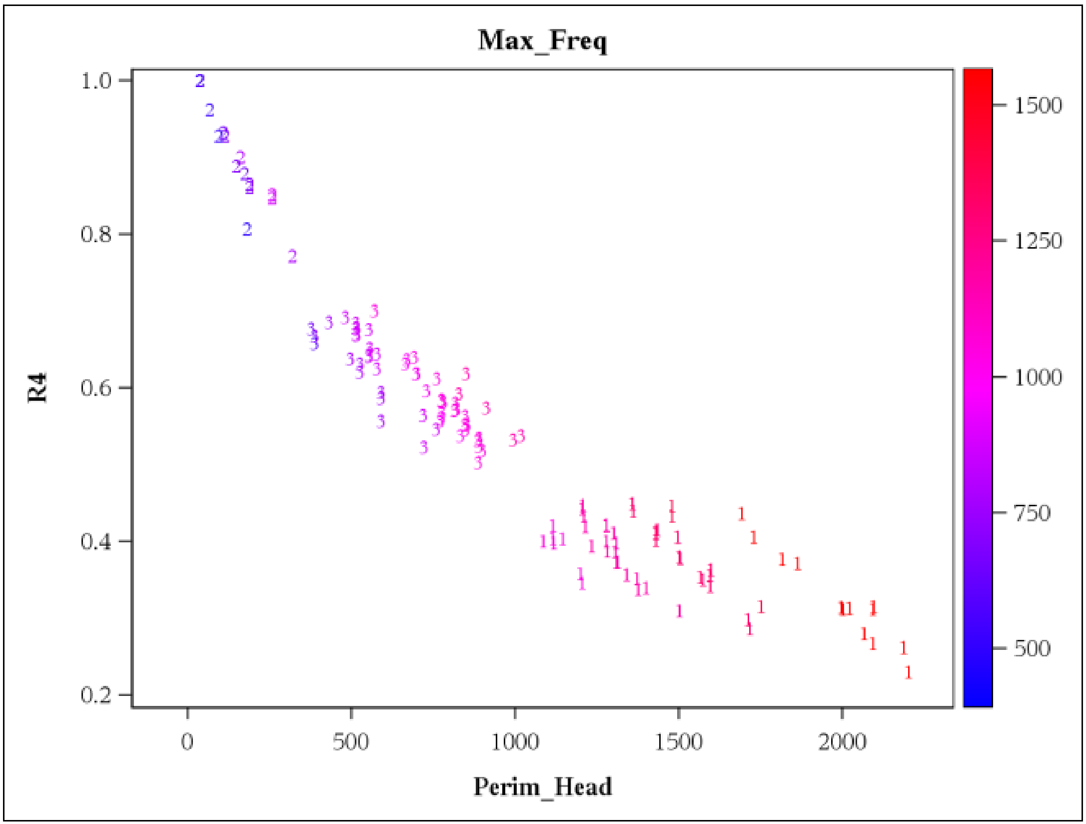
Two of the selected parameters, *maximum frequency* and *frequency range*, lead to a similar distribution over the 3 clusters (Fig. 4).

One can observe that the color distribution nearly fits a curvilinear axis connecting the clusters. This could explain why the authors were unable to observe any clusters in the space defined by the more traditional voice frequency and intensity parameters directly since in such an analysis the parameters define straight coordinate axes (results not shown).

Overall, the members of cluster 1 have a low value of feature *R4* and *R9* and a high value of feature *Perim_Head*; those from cluster 2 have a high value of *R4* and *R9* and a low value of *Perim_Head*; finally, those from cluster 3 have a middle value of *R4*, *Perim_Head*, *R9 Max_Freq*, and *Freq_Range*.

Discussion

The best cluster separation (larger than 80%), for three clusters, is obtained for feature R4: the ratio of the perimeter length of the chest voice versus the total perimeter length. This parameter may not be clinically easily understandable. However, the cluster separation is based on data points corresponding to the value they have for the more traditional voice frequency and intensity parameters



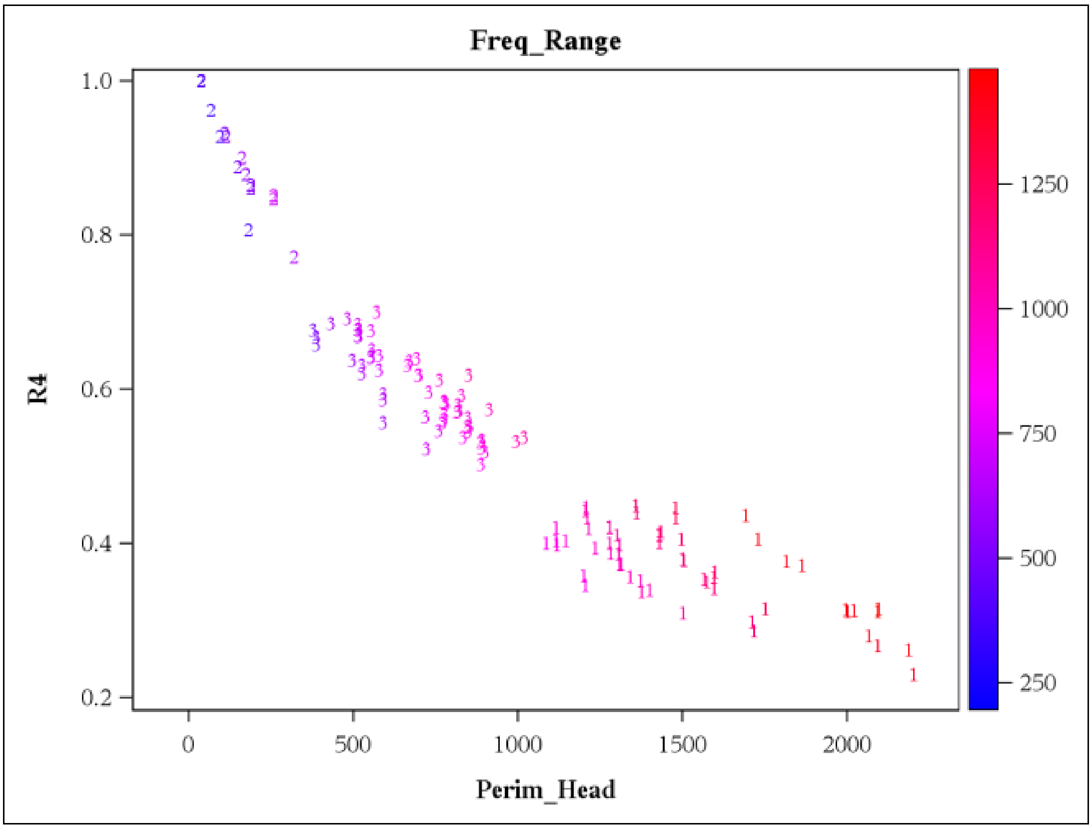


Figure 4 (previous page). The horizontal and vertical axes represent feature dimensions. The data points are labeled by the cluster to which they belong. Feature *R4* is the ratio of the perimeter of the chest voice part of the VRP and the total perimeter (dimensionless). Feature *Perim_Head* is taken as the sum of the Euclidean distances between consecutive data points on the perimeter of the head voice part of the VRP (in arbitrary units).

mentioned above. This means that the complexity of voice classification can be reduced to a more compact, yet adequate formula, easily obtainable from the VRP parameters. In fact the “perimeter total” can be seen as an expression of the total frequency range, whereas the “perimeter chest voice” marks the boundary with the head voice/falsetto. Klingholz et al. elaborated a method of analysing ellipse parameters and conclude that the register transitions, visible in the VRP, at certain frequencies, illustrate clearly the different mechanisms based on pre-phonatory larynx positions and specific muscle activities.[17-19] Pabon prefers to talk about a region where modal and falsetto registers overlap. He acknowledges that, in general, register differences are accentuated by a greater effort (louder voice production, higher pitch).[20] In previous publications automatic, computer assisted VRP registration is considered to be helpful to determine voice breaks and to indicate register contours, facilitating voice classification.[21-22] In another study, however, the author assumes that the register transition from modal to loft cannot be located with certainty from the VRP, nor by auditory perception.[20] According to Klingholz et al.[17] and Airainer and Klingholz[23], however, markers of the register ranges are the transitions which are indicated “by minima in the forte contour and maxima in the piano contour, and minima of the dynamics, at specific pitches.”
The features that lead to a clear cluster separation in this study do not take timbre into consideration. This parameter has always been evocated as decisive in voice classification by many singing teachers. Perceptual evaluation of voice quality, however, proved to be highly subjective[24] and remains controversial because of poor correlation among raters.[25-28] The pertinent study demonstrates that there exist three different female voice categories by nature. Such a result may not be expected within biological variables. However, many years ago, the French phoniatrician Garde defined voice category already as “a biological constant, as important as the determination of the blood group”.[3] whether these results also apply to male voices remains to be investigated. This remarkable clustering has to be linked to the traditionally three basic voice categories. One can only wonder if composers of vocal music had an innate feeling about the existence of three natural basic human voice categories.

Conclusions

This study demonstrates that parameter combinations of the voice range profile are able to yield a clear cluster separation to discriminate between three basic female voice categories and may serve as the basis to resolve the riddle of voice classification. However, more studies are necessary to link the results of the statistically obtained cluster separation, which discriminate between three basic voice categories, to the three basic female voice categories as commonly interpreted by most composers of vocal music and singing teachers.

References

1. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F. Wyke BD, Ed. Disorders of Human Communication 6. Wien New York, Springer-Verlag, 1982.
2. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1: 209-216.
3. Garde E. La Voix. Paris, Presses Universitaires de France, 1954.
4. Lycke H. Stemclassificatie en Fonetogram. Dynamiek tussen morfologie en functie; in de Jong F. Ed. VOX 2007: Spreken en zingen. Twee werelden apart? Nijmegen, Drukkerij Graficolor, 2007.
5. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
6. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Paris, Librairie Maloine SA, 1961[2].
7. Thurmer S. The tessiturogram. J Voice 1988;4:327-329.
8. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209-216.
9. Burns P. Clinical management of Country and Western singers. J Voice 1991;5:349-353.
10. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.
11. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. New York: Raven Press, 1991:49-68.
12. Abitbol J. L'odyssée de la voix. Paris, Editions Robert Laffont SA, 2005.
13. Barthélémy Y. La voix libérée, Une nouvelle technique pour l'art lyrique et la rééducation vocale. Paris, Editions Robert Laffont SA, 1984.
14. Schutte HK Over het fonetogram. Log en Fon 1975;47:82-92.
15. Klingholz F, Martin F. Die quantitative Auswertung der Stimmfeldmessung. Sprache - Stimme - Gehör 1983;7:106-110.
16. Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement / phonetography. Folia Phoniatr 1983;35:286-288.
17. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimmfeld. Sprache - Stimme - Gehör 1985;9:109-111.
18. Klingholz F, Martin F, Jolk A. Das dreidimensionale Stimmfeld. Laryng. Rhinol Otol 1986:65.
19. Klingholz F, Jolk A., Martin F. Stimmfelduntersuchungen bei Knabenstimmen (Tölzer Knabenchor). Sprache - Stimme - Gehör 1989;13:107-111.
20. Pabon JPH. Kwaliteitsaspecten in het fonetogram. Logopedie 1997;10:20-33.
21. Pabon JPH, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. J Speech Hear Res 1988;31:710-722.
22. Pabon JPH. Objective acoustic voice-quality parameters in the computer phonetogram. J Voice 1991;3:203-216.
23. Airainer R, Klingholz F. Quantitative Evaluation of Phonetograms in the Case of Functional Dysphonia. J Voice 1993;7:136-141.

24. Giovanni A, Robert D, Establier N, Teston B, Zanaret M, Cannoni M. Objective Evaluation of Dysphonia: Preliminary Results of a Device Allowing Simultaneous Acoustic and Aerodynamic Measurements. Folia Phoniatr Logop 1996;48:175-185.
25. Liu TA, Channell RW. Reliability of Pitch Discrimination of Vowel and Piano Tones. J Voice 1995;394-402.
26. Wapnick J, Ekholm E. Expert Consensus in Solo Voice Performance Evaluation. J Voice 1997;429-436.
27. De Bodt MS, Wuyts FL, Heyning Van de PH, Croux C. Test-Retest Study of the GRBAS Scale: Influence of Experience and Professional Background on Perceptual Rating of Voice Quality. J Voice 1997;74-80.
28. Oates J, Russell A. Learning Voice Analysis Using an Interactive Multi-media Package: Development and Preliminary Evaluaton. J Voice 1998;500-512.

Chapter 4

Discrimination of Three Basic Male Voice Types by Voice Range Profile Derived Parameters

H. Lycke, A. Ivanova, M.M. Van Hulle, W. Decoster, F.I.C.R.S. de Jong
Accepted for publication by Folia Phoniatrica et Logopaedica

Abstract

Aims: To verify the existence of individual- or combinations of *not* commonly used and *not* easily understandable clinical voice range profile (VRP) parameters (also called "features") with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic male voice categories.
Methods: The voice range profiles of 256 male conservatory singing students and professional singers were recorded, parameterized into more compact descriptions ('features'), and subjected to a cluster analysis.
Results: Based on all parameters the frequency dip of the register transition zone was shown to yield the best cluster separation for three basic male voice types.
Conclusions: This study demonstrates that parameter combinations of the VRP exist that generate a clear separation of voice clusters. This was also in the case with female voices as shown in a former study (Chapter 3). The clusters may be attributed to the three traditional basic male voice types, and in this way our results can provide a fresh angle on the issue on male voice classification.

Introduction

Traditionally, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor. A correct classification of a singer's voice is indispensable in order to achieve an optimum performance and to avoid functional and organic voice disorders.[1]
Despite of its importance, there is neither a generally accepted protocol for voice classification, nor a consensus about what parameters should be used. The question arises, therefore, if there exist three types of female and male voices by nature.
In a previous study a clear separation into three basic female voice types could be made, based on a statistical analysis of Voice Range Profile (VRP) parameters.[2] It is, therefore, plausible to investigate whether the male voice can also be discriminated into three basic voice types.
As an extension of the above mentioned paper on female voices, the aim of this study was to verify the existence of individual- or combinations of *not* commonly used and *not* easily understandable clinical VRP parameters (also called "features") with which the data can be partitioned into a number of clearly

separated clusters as a basis for discriminating between basic male voice categories.

Materials and Methods

Because this paper uses the same methodology as the one on female voices (Chapter 3), this section is kept short.

Participants and Data Acquisition

The data from 256 male subjects (18 – 52 years, mean age = 22, SD = 4.9), consisting of 9 young singing students, 17 professional singers, 61 professional choir singers and 169 with and without singing experience, was investigated.
The VRPs were performed according to the Union of European Phoniatricians' recommendations[3], using the lingWAVES Voice Diagnostic Center, version 2.5 with a Center 322 Data Logger Sound Level Meter[4], or the Voice Profiler 4.0. version 26.01.2005[5]. The data contains maximum and minimum intensity measurements of each subject's voice, taken at different frequency points. The frequency axis was converted in a linear scale so as not to favor any frequency range. The result is a VRP which has the characteristic shape of an ellipse (Fig. 1).

Parameter Construction

Two types of parameters, also called features in a statistical context, were developed. Firstly, the so-called clinical parameters as they are amenable to a clinically interpretation, e.g., the maximum and minimum voice frequency. We add also parameters that quantify the transition zone in the VRP such as its frequency and intensity. In total 10 such parameters were identified and we gave them mnemonic labels such as *Freq_Dip* for the frequency of the transition zone parameter. The second type of parameters characterize, in a compact manner, the *geometry of the entire VRP and the chest/head voice parts* of the VRP, the *register transition zone*, and the *linear characteristics of the minimum and maximum intensity curves*. This way 14 more parameters were defined.
Finally, 26 voice frequency and intensity ratios and differences based on some of the above features were defined such as the ratio of the surface area of the chest voice to the total surface area enclosed by the maximum and minimum intensity curves. Hence, in total, 50 features were defined. For a detailed description: see Addendum 5.

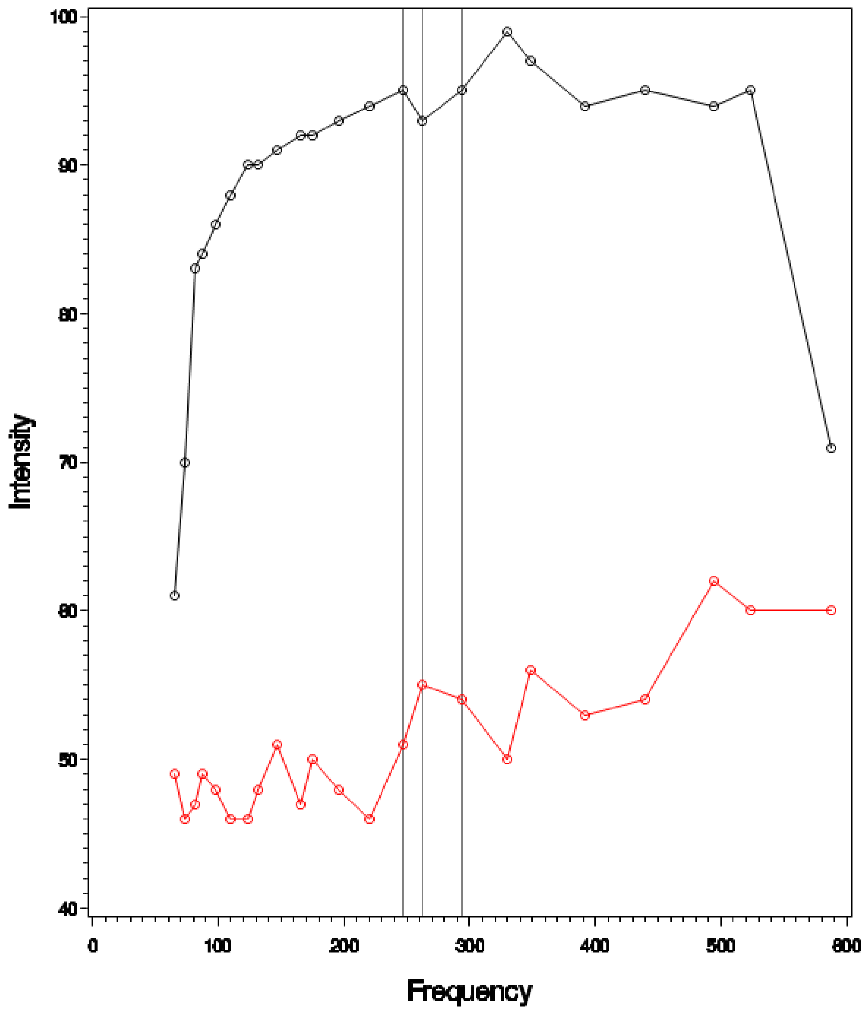


Figure 1. Voice Range Profile (VRP). The frequency axis was converted into a linear scale. The upper curve is for the maximum intensity points; the lower curve for the minimum intensity points. The transition zone is demarcated by the first and third vertical lines, the middle vertical line defines the dip of the transition zone.

Clustering and parameter selection

Prior to clustering, the redundant parameters (parameters for which the absolute value of the correlation is 0.95 or higher) were removed. This way 38 parameters remained. Then, the data set was standardized (zero mean and equal variance for each parameter, thus, converting data points into z-scores) to eliminate any bias from differences in scale (parameters with broadly distributed samples tend to dominate those with narrow distributions).

Next the following procedure was used: For the single parameter case, the Ward's minimum variance method was applied to define the optimal number of clustersK-means clustering was applied on all single, two and three parameter combinations to assign the voice data to clusters. The optimality of a given cluster solution was decided in statistical terms (R-squared (RS) and the semipartial R-squared (SPR).

RS quantifies the between cluster variability vs. the total variability. The between cluster variability assesses the squared distances, per feature dimension, taken between the cluster centers (averages, centroids) and the center (average, centroid) of the whole data. The total variability is the squared distance between each data point and the center (average, centroid) of the whole data. RS varies in the interval [0, 1]. “0” means that there is no difference, “1” means that there is a maximal difference between the clusters. SRS defines the homogeneity in case the cluster is merged) used for ranking purposes (search for the best cluster separation). Note that this procedure could lead to closely tied solutions. In that case, we further explore each one of them and decide based on the consistency of the cluster members across parameter combinations (cluster migration).
We then apply the K-means clustering technique in search of the optimal parameter combinations. First we attempted a forward selection approach: given the best single parameter the best combination with a second parameter is searched, and so on. However, as there were many single parameters which cluster solutions had similar R-squared and semipartial R-squared values, we were faced with a combinatorial problem in our search for optimal 2 and 3 parameter solutions as the bulk of them did not correspond to good cluster solutions. Therefore, it was decided to do the inverse and adopt a *backward selection* procedure: prune an individual feature away from the incumbent set so that the remaining feature combination has the highest R-squared. We starting from the optimal 3 parameter combination (i.e., with the highest R-squared cluster solution).
Finally, in order to decide on the *number* of clusters, in the case of the tie mentioned above, and verify the quality of the found solution, we determine the number of subjects that migrate (once, twice, or more) from their clusters when varying the number of selected parameters. We express this in terms of a cluster migration index. The cluster number that scores best in this sense is retained as the final solution.

The analyses were performed using Statistical Analysis SAS/STAT software (release 9.2).

Results

When applying the clustering procedure on single parameters and when plotting, on the same figure, the corresponding R-squared (RS) and semipartial R-squared

(SPR) values, it can be observed that, for most single features, the optimal number of clusters is 3 or 4. We further investigate both possibilities.

We start the backward selection procedure (Table 1) from the optimal 3 parameter combination. For the 3 cluster case we obtained the following combination: the frequency of the register dip (*Freq_Dip)*, the mean frequency of the register transition zone (*mean_freq_TrZone)*, and the length of the perimeter of the chest register part of the VRP (*Perim_Chest)*. For the 4 clusters case, we obtained the following parameter combination: the ratio of the surface area of the head voice part of the VRP divided by the total VRP surface area (*R2);* the ratio of the surface area of the head voice part of the VRP divided by the surface area of the chest voice part *(R10)*; and the ratio of the perimeter of the chest voice part of the VRP divided by the total perimeter of the VRP (*R4*).

Table 1. The best cluster solutions for the case of 3 and for 4 clusters by using a backward parameter selection procedure. *Freq_Dip* = frequency register dip; *mean_freq_TrZone* = mean frequency of the register transition zone; *Perim Chest* = length of the perimeter of the chest register; *R2* = ratio surface area head voice/total surface area; *R10* = ratio surface area head voice/ surface area chest voice*; R4* = ratio perimeter chest voice/perimeter total.

	3 clusters Set 1: *Freq_Dip, mean_freq_TrZone, Perim_Chest*	4 clusters Set 2: *R2, R10, R4*
1 (single) parameter	*Freq_Dip* RS (≡ SPR): 0.895	*R2* RS (≡ SPR): 0.901
Combination of 2 parameters	*Freq_Dip, mean_freq_TrZone* RS: 0.834 SPR: *Freq_Dip*: 0.858 *mean_freq_TrZone*: 0.811	*R2, R10* RS: 0.881 SPR: *R2:* 0.875 *R10:* 0.887
Combination of 3 parameters	*Freq_Dip, mean_freq_TrZone, Perim_Chest* RS: 0.802 SPR: *Freq_Dip*: 0.824 *mean_freq_TrZone*: 0.801 *Perim_Chest*: 0.780	*R2, R10, R4* RS: 0.858 SPR: *R2:* 0.863 *R10:* 0.878 *R4:* 0.833
Migration index	86 % (220 subjects) of the data in the same cluster The subjects that migrated did that only once	46 % (114 subjects) of the data in the same cluster The subjects that migrated did that only once

The outcome of the backward selection procedure can be followed in Table 1. In the case of 3 cluster solution, the obtained RS and SPR indicate a good to superior cluster separation: the RS of the optimal 3 parameter combination equals 80%, the RS and SPR of the optimal 2 parameter combination are larger than 80%, and finally, the RS of the optimal single parameter is equal to its SPR, and reaches almost 90%. Similar clustering results are obtained for the 4 cluster case so that the migration index is going to be decisive.

Note that 220 subjects out of 256 (about 86% of data) were, for the 3 cluster case, and each of the 3 parameter combinations, classified into the same cluster (migration index, see Table 1). Thirty-six subjects migrated from one cluster to another but only once. There was not a single subject that migrated twice. For the 4 cluster case the situation is completely different as only 114 subjects (46 %) remained in the same cluster. Based on this outcome it was decided that the 3 cluster case is optimal.

Figure 2 shows planar view of the common cluster members (220 subjects) of the 3 parameter solution for the optimal 3 cluster case.

Discussion

In a former paper, a discrimination of three basic female voice types by Voice Range Profile-derived parameters was found.[2] A three-cluster case provided the most consistent solution across all feature combinations, in both a forward and backward selection procedure. The feature that led to the best cluster separation was the ratio of the perimeter length of the chest voice part of the voice range profile versus the total perimeter length. The pertinent study, in which male voices were examined, is an extension of the above mentioned paper on female voices. Because of the numerous options to combine features, a backward selection procedure was applied. Also in the male voices a three-cluster case provided the most consistent solution across all feature combinations. The feature that led to the best cluster separation in the male voices was the frequency of the register dip. In the case of this parameter, the RS is equal to the SPR and it is almost 90% which is an indication of a degree of high cluster separation. These results demonstrate that certain parameter combinations of the voice range profile are indeed able to yield a clear 3 cluster separation.

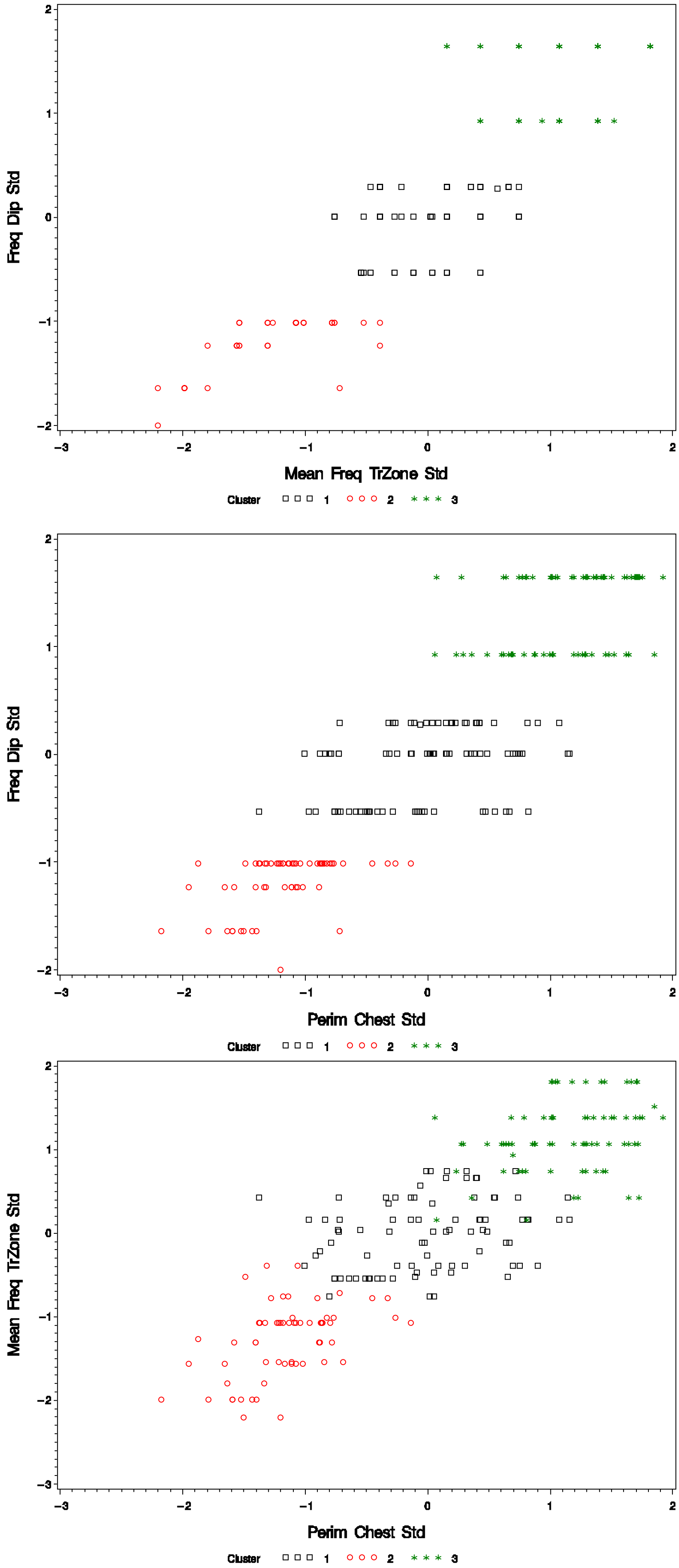


Figure 2. Planar views of the common cluster members (220 subjects) for the case of 3 clusters and 3 parameters.

The results of both studies show that different parameters (the ratio of the perimeter length of the chest voice versus the total perimeter length for female voices and the frequency dip of the register transition zone for male voices) yield a clear separation into three clusters of voice for each gender. Each of these formulas has to do with register transition. In voice classification, register transition has been connected to male and female voice categories.[6-13] However, the specific relationship with voice classification as a determinant parameter is not clear, nor the differences in gender. Although vocal registers are known to occupy separate portions of the total fundamental frequency range of the human voice[14], perceptual judgments alone invariably lead to many discussions and controversies[15] In this regard one should admit that a complete understanding of the physiology of register function is still lacking.[16] Miller and Schutte[17] recently addressed the question whether the female middle voice is a combination or balance of the primary registers: chest and falsetto. Keilmann and Michek[18] made an extensive study on the physiology of the female whistle register as a separate register on top of the falsetto register but others[19-21] question whether it can be considered as a true register. The recent study of Garnier et al.[25] supports the idea that the middle upper register (in the high soprano range) could be extended to higher pitches (up to C6) and that the upper register could be extended to lower pitches (down to C5).

According to Neumann et al.[12] “singers themselves are frequently unable to locate their point of transition and to distinguish clearly between registers”, and the judgment by experts of singing is also questioned. Hollien[21] stated that voice registers must be *operationally* defined: perceptually, acoustically, physiologically and aerodynamically. Titze[22] even expressed the need to describe registers also on the neuromuscular, biomechanical, and kinematic level. For obvious reasons some researchers prefer to limit their endeavour to only one aspect of the issue, e.g. the physiological description, admitting that “no single investigation can hope to address all the elements previously cited”.[19,21] According to Klingholz et al.[20] only acoustic methods provide a reliable and objective access to the registers. Frequency localization cannot be the sole characteristic of a vocal register.[14] That is why, in our opinion, vocal intensity characteristics, as objectively measured by phonetography, can open a new era of voice research. The eye casting dip(s) in the phonetogram at certain frequencies illustrate the different mechanisms based on the pre-phonatory larynx positions and specific muscle activities.[20,23,24] A recent study[25] showed that the region of voice instabilities, which can be spread over an octave (cfr. our specific methodology), is accompanied by SPL variations of up to 15 dB, even in classically trained singers, trying to avoid pitch breaks or jumps.

The findings of three natural clusters in female and male voice and the role that register transition related parameters play in clustering, is an indication that clustering may be a new angle on the issue of voice classification. However, further studies are necessary to link the results of the statistically obtained cluster separation to the three basic voice categories as commonly interpreted by most composers of vocal music and singing teachers.

Conclusions

This study demonstrates that certain parameter combinations of the male VRP are able to yield a clear cluster separation, as was also the case with the female VRP, which in turn can serve as a basis to discriminate between three basic male voice categories. This suggests that the complexity of voice classification may be reduced to a more compact, yet adequate formula, easily obtainable from the VRP. However, more studies are required to link the statistically obtained cluster separation to the three basic voice categories as they are traditionally used, even to date, and to determine the relevance of the found differences between the female and male parameter combinations.

Acknowledgments

Corneillie M, Gasparini L. Stemclassificatie bij mannen door middel van fonetografie Master Thesis Faculty of Medicine, Logopedische en Audiologische Wetenschappen. KU Leuven, 2012.

References

1. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209–216.
2. Lycke H, Decoster W, Ivanova A, Hulle Van MM, Jong de FICRS. Discrimination of three basic female voice types in female singing students by Voice Range Profile-derived parameters. Folia Phoniatr Logop 2012;64:80–86.
3. Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement/phonetography. Folia Phoniatr 1983; 35:286–288.
4. Franke I. Ling Waves Voice Diagnostic Center (VDC) + Voice Loading Test Handbook Version 2.5, Forchheim, 2005.
5. Pabon P. Manual Voice Profiler Version 4.0. Rotterdam: Alphatron Medical Systems, 2006.

6. Pahn J. Stimmübungen für Sprechen und Singen. VEB Verlag Volk und Gesundheit, Berlin, 1968.
7. Van Deinse JB, Goslings WRO. The technique of singing. Some remarks on the mechanism, the technique and the classification of the singing voice. Theodora Versteegh Foundation. Government Publishing Office, The Hague, 1982.
8. Janev R, Gjuleva L. Koorkunde, een handboek voor koordirigenten en koorzangers. G.F. Callenbach b.v., Nijkerk, 1984.
9. Butenschön S, Borchgrevink HM. Voice and Song. Cambridge University Press, Cambridge, 1982.
10. Wirth G. Stimmstörungen. Lehrbuch für Ärtzte, Logopäden, Sprachheilpädagogen und Sprecherzieher. Deutsche Ärtzte-Verlag, Köln, 1991.
11. Hertegard S, Gauffin J, Sundberg J. Open and covered singing as studied by means of fiberoptics, inverse filtering, and spectral analysis. J Voice 1990;4:220-230.
12. Neumann K, Schunda P, Hoth S, Euler HA. The interplay between glottis and vocal tract during the male passaggio. Pholiaphoniatr Logop 2005;57:308-327.
13. Bjôrkner E, Sundberg J, Cleveland T, Stone E. Voice source differences between registers in female musical theatre singers. J Voice 2006;20/2.
14. Colton RH. Vocal intensity in the modal and falsetto registers. Folia phoniat 1973; 25:62-70.
15. Röhrs M, Pascher W, Ocker C. Untersuchungen über das Schwingungsverhalten der Stimmlippen in verschiedenen Registerbereichen mit unterschiedlichen strobo-skopischen Techniken. Folia phoniat 1985;37:113-118.
16. Sundberg J. The Science of the Singing Voice. Northern Illinois University Press, Dekalb, Illinois, 1987.
17. Miller DG, Schutte HK. Mixing' the registers: Glottal source or vocal tract? Folia Phoniatr Logop 2005;57:278-291.
18. Keilmann A, Michek F. Physiologie undakustischeAnalysen der Pfeifstimme der Frau. Folia Phoniatr 1993;45:247-255.
19. Walker JS. An investigation of the whistle register in the female voice. J Voice 1988;2:140-150.
20. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimm-feld. Sprache - Stimme - Gehör 1985;9:109-111.
21. Hollien H. On vocal registers. Journal of Phonetics 1974;2:125-143.
22. Titze IR. A Framework for the Study of Vocal Registers. J Voice 1988;2:183-194.
23. Klingholz F, Martin F, Jolk A. Das dreidimensionale Stimmfeld. Laryng Rhinol Otol 1986;65.
24. Klingholz F, Jolk A, Martin F. Stimmfelduntersuchungen bei Knabenstimmen (Tölzer Knabenchor). Sprache - Stimme - Gehör 1989;13:107-111.
25. Garnier M, Henrich N, Crevier-Buchman L, Vincent C, Smith J, Wolfe J. Glottal behavior in the high soprano range and the transition to the whistle register. J Acoust Soc Am 2012;131/1;951–962.

Chapter 5

General discussion

General Discussion

The human voice is the instrument *par excellence* for oral communication. In day-to-day use the vocal load may not be heavy, in contrast to professional use. In singing, the vocal load can be extremely high, sometimes exceeding the physiological limits. To avoid damage and to optimize vocal performance it is important to know the possibilities and impossibilities of the voice. In other words, the vocal capacity must be estimated. Voice classification is a method to estimate the voice and composers of vocal music wrote and write repertoires that fit to the possibilities of the voice, in classical music indicated by voice classes. Traditionally, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor. Surprisingly, there exists no generally accepted method for voice classification, let alone an algorithm, as there is no consensus about what voice parameters to use. This calls for an objective method for voice classification. When this proves infeasible, one may rightfully question the existence of three basic female and male voice types by nature.

Voice research and singing

For many decades, studies on voice were limited to studies in laboratory settings. Brodnitz[1] cautioned: ”The bigger the machinery, the more the artificiality there is with a patient”, and von Leden[2] stated :” There must be careful study of the patient under conditions that are as close to his normal use of the voice as one can arrange. I could not stress this point too strongly”. Kitch et al. stated that the results obtained from research in naturalistic settings may be more valid in singing pedagogy than in laboratory settings.[3] They encourage investigators to explore the possibilities of taking the laboratory to performers and thus to consider larger groups of subjects for investigation. According to Hoffman-Ruddy et al. clinicians must visit the performer’s stage so that there is a better understanding of detrimental environmental situations they are working in.[4] However, in our experience, the access to stages and conservatories is not easy (Addendum 1). This limited access of scientists to the singing world could explain why so few studies are available in the literature, dealing with larger groups of singers and singing students outside the laboratory. Schutte and Seidner stated that “audiometer cabins or strongly dampened experimental rooms might influence the auditive self-control of the subject’s own voice”.[5] This is certainly the case in singers. Schultz-Coulon[6] and Wirth[7] too prefer an ‘acoustical living room atmosphere’. With the lingWAVES Voice Diagnostic Center we performed our

Voice Range Profile (VPR) measurements outside laboratory settings, in different rooms situated in conservatories and theatre companies, with a level lower than 45 dB(A) and according to the recommendations of the UEP.[5] Scientific studies on different categories of singing are scarce and were mainly focused on classical singing, for which voice classification has traditionally been essential. During the last two decennia, the scientific world witnessed profound changes in the world of singing education and practice but only a few studies appeared on different singing genres and techniques.[8-11]

Voice classification criteria

The results of our explorative studies (Addendum 1 and 2), can be seen in the light of the remarks of Radionoff et al. about the nomenclature disparity and the tremendous lack of consistency among curricula of commercial music degrees today.[12] The answers received to three different questionnaires (conform Addendum 1 and 2) showed a marked inter-individual difference in attitude towards voice classification by the various singing teachers and their students. Many different criteria for voice classification are applied today, even in classically oriented conservatories, while many singing students have their doubts on the correctness of the voice category as provided by their teachers.

Regarding contemporary music we found that many private singing teachers and their students do not feel the need of voice classification (Addendum 1). Scores of commercial music are usually adjusted (= transposed) to the capacities of the individual singer who has been chosen to sing a specific song. Usually, it is not necessary to know your voice type if you are singing for your own enjoyment. However, if you have the ambition to sing professionally or do some professional auditions, you must know your voice type. In most cases, at the audition the candidate will be asked about his voice type. Usually, the singing teacher is not present here. The jury will look for the most fit (voice) type for the role, without classification. Estimation of the voice type enables to find out what songs are most appropriate to your own voice.

Frequency range as a commonly used voice classification parameter

Frequency range is considered to be an important factor in voice classification, as indicated in many textbooks. Indeed, according to the outcome of our questionnaire (Addendum 2) among private singing teachers of different genres and among classical and musical theatre singing teachers at three European conservatories, frequency range proved to be the most commonly used voice classification parameter. However, when comparing a method based on the limits of the female frequency range, according to 38 authors with different backgrounds, to the results of voice classification by the singing teachers, the frequency range itself proved not to be panacea for voice classification (Addendum 3).

The riddle of voice classification

Until now, there is no generally accepted method for voice classification as singing teachers use various criteria (Addendum 1 and 2). In our study, frequency range proved not to be suitable as a voice classification parameter (Addendum 3). It is now questioned if three basic female and male voice types exist by nature. In an attempt to break out of the controversy, a new perspective is adopted in this study by letting the data speak for itself. Such an approach, called "data-driven", imposes minimal assumptions on the nature of the data, what elements to use for its analysis, and in our case even the existence of natural voice groups.

Phonetography as a basic tool for voice classification

According to Sundberg the voice source can be varied continuously along the phonatory dimensions of *vocal pitch* and *loudness*.[13] Plant and Younger found that "in general, intensity increases with higher subglottic air pressures, but there are also considerable variations both between individuals and within different portions of the frequency-intensity range for a given subject".[14] Wilson and Leeper[15] and Stathopoulos and Sapienza[16] too found large amounts of variability regarding changes in laryngeal airway resistance as a function of vocal sound pressure in both male and female subjects. Keilmann et al. came to the same conclusions for children.[17] In our opinion it seems obvious to assume that the interpretation of this *highly individual covariation of frequency and intensity*, as objectively measured in VRP analysis and providing *a new dimension in visualizing the voice*, could also give a clue for an objective voice classification. According to Frank and Donner voice classification by fundamental frequency - sound pressure levels (F°- SPL) measurement alone is not possible.[18] They feel there are no characteristics in the voice area in relationship with the different voice categories. Some authors,[19-22] however, assume that VRP provides much useful information for voice classification, but no specific interpretation of VRP results regarding voice classification can be found in the scientific literature. No specific curves are presented regarding specific voice categories. This has not only to do with difficulties in interpreting the VRP curves, but also with the general conviction that voice classification belongs to the domain of the professional singing voice.

The data driven method for discriminating between basic voice types based on VRPs

Initial method based on clinically frequently used and easily understandable parameters

In a pilot study, the VRP data from 206 female music conservatory singers was subjected to a clustering analysis to assess whether the data displays any natural clusters (groupings) without assuming their number and without using any prior voice classification by the singing teacher or by the students themselves (thus, adopting a “data driven approach”) (Chapter 2). We characterized hereto each VRP in terms of various clinically frequently used and easily understandable voice frequency/intensity parameters (10 parameters in total). All VRPs then become represented as points in this parameter space. We then applied Ward’s method, a tree-based clustering algorithm that allows us to objectively define the optimal number of clusters in this space by judging the heterogeneity of each found cluster and the homogeneity of the clusters when they are merged. Unfortunately, this cluster analysis revealed no significant clusters.

Refined method based on “intelligent” features

It was then hypothesized that perhaps there could exist other, more “intelligent” parameter (“feature”) combinations, even non-linear ones, of the VRP that could lead to clear clusters (Chapter 3 and 4). These combinations are less easily understandable in clinical terms, compared to the 10 parameters of the pilot study. They are used primarily for the analysis of determination of clusters. We defined parameters that characterize the geometry of the entire VRP and the chest/head voice parts of the VRP, the register transition zone, and the linear characteristics of the minimum and maximum intensity curves (14 parameters). Furthermore, we defined 26 voice frequency and intensity ratios and differences based on some of the foregoing features were defined such as the ratio of the surface area of the chest voice to the total surface area enclosed by the maximum and minimum intensity curves. In this way, 40 “intelligent” features were defined.

Ward’s method now revealed that there could be 3 or 4 clusters in the female data set. Both options turned out to be statistically plausible. In order to break the tie, we need an objective method. Furthermore, we wanted to identify which parameter (or a small set of them) is *crucial* for discriminating between the voice clusters which also adds to the usability of our method. We decided to develop a method that answers both questions: we identify the individual feature that best discriminates the clusters, the best feature pair (including the best individual feature), and so on, and observe if the individual singers migrate between the clusters (consistency of the cluster members across parameter combinations, cluster migration). We perform this analysis for the 3 and the 4 cluster cases and decide based on the cluster migration index. In principle, Ward's method could be used for this purpose but it is too computationally intensive to be practical. Also, when data of new singers would become available, the analysis would need to be done all over again to decide to which clusters the singers belong. Therefore, we used the K-means clustering technique. It is both computationally efficient and effective for assigning new data to already found clusters. Based on the migration index, we decided that the three cluster separation was optimal. The "intelligent" features turned out to be instrumental in reaching this conclusion. Finally, the analysis was repeated for the male voices and it also resulted in the three cluster solution being optimal.

In our study, trained and untrained, singing and non-singing voices are included. According to Sundberg et al., a major difference between speech and singing is that a precise and independent control of pitch and loudness is needed in singing but not in speech.[23] In acting and emotive speech, however, parallels between respiratory behaviour in speech and singing are obvious.[23-24] Principally, loudness demands in acting seem quite similar to those necessary in classical singing.[25-26] Furthermore, the age of our participants differs between the female and male groups. The results of this study demonstrate that even in different homogeneity of the groups different parameters of the VRP are able to yield a clear separation into three clusters for each gender.

Albeit that the used "intelligent" parameters turned out to be crucial for discriminating clusters, they are not readily amenable to a clinical interpretation. Therefore, the difference between the discriminating parameters in the female and male cases cannot be readily explained from a clinical viewpoint.

Nevertheless, the most discriminating features for both the female and male voices have to do with register transition. Additionally to the finding of three clusters in both the female and male groups, this is a second salient finding of our study. In light of this, a closer discussion about the register transition and voice classification is in order.

The frequency dip of the register transition zone as a separating feature in male voices is a well-known clinical parameter.[27-28] Why this separating feature is not the same in female and male voices in our study is not clear. Female voices seem to be more complex. The ratio of the perimeter length of the chest voice versus the total perimeter length as a distinctive feature in female voices could tentatively be interpreted in view of the results of studies on the high soprano range and the transition to the whistle register.[29-32] These studies suggest a possible individual varying range of up to three register transition zones in the female (soprano) voice, accompanied by a possible large vocal range, represented by perimeter length. Henrich concluded in her article on the history of vocal registers that further research is needed of the singers' resonance strategies in both classical and non-classical phonation types.[33] In our data driven study both classical and non-classical phonation types were included.

Voice research and the register enigma

Register transition has been an intensively debated item for years. Although there is a global idea of what is meant by "register", an exact definition seems to be almost impossible.[34] Some authors[35-36] still keep referring to the well-known register definition of Garcia (1840): "a series of succeeding sounds of equal quality on a scale from low to high, produced by the application of the same mechanical principle, the nature of which differs basically from another series of succeeding sounds of equal quality, produced by another mechanical principle", or to the more extended definition of Nadoleczny (1923), who assumed that the internally similar tones of one register were dependent upon a definite, invariable behaviour of the harmonics.

Authors have different opinions on how the modern register names should be understood. Titze discerns two types of transitions: a periodicity transition from a continuous tone (chest) to a series of pulses (vocal fry) and a timbre transition corresponding to the primary and secondary passages, based on changes in the closure conditions of the glottis.[37] Some scientists agree that voice registers must be operationally defined: perceptually, acoustically, physiologically and aero-dynamically.[38] Some authors make a distinction between singers and non-singers[38-39] and between speaking versus singing registers.[38] The singer seems to have the continued potential to use the voice registers if necessary or desirable.[38-40] So, trained singing voices could exhibit only a single register, while retaining their original set of registers.[38-39] Equalization of registers may be related to the laryngeal mechanism of medial compression.[36] A voice register is usually seen as exclusively a laryngeal event.[38-39,41-42] Spencer and Titze,[43] however, nuance this statement and define the register as the expression of a primarily laryngeal event. "Acoustic and myoelastic influences on register transition may exist along a continuum of blended interactions".[43] According to Klingholz et al. the position of the larynx, the form of the vocal fold mucosal wave, the aerodynamics of the vocal apparatus, the vocal quality and even the subjective feeling of the singer are possible indications.[27] The controversy is at a high level when the number of existing voice registers is discussed. The exact number of registers remains unknown and assertions go from one to as many as 107.[38,44] Schutte and Miller cite the historic Italian school, identifying several categories of registration timbres such as: (1) voce di petto (chest voice), (2) voce mista (mixed voice), (3) voce finta (feigned voice), (4) voce di testa (head voice) and (5) falsetto.[45] Hollien,[38] however, postulates that only three major registers exist: the pulse, modal and loft registersHowever, in the literature we still can find studies on the so-called whistle register (flute register, flageolet register, Pfeifregister)[46-47] and the Strohbasregister (pulse, vocal fry).[35,47] To many researchers[49-54] "the riddle of the middle register" still exists, while Miller et al.[55] proposed the name of mezza voce as a distinct register of the male singing voice. Miller and Schutte discuss the question whether the female middle voice is a combination or balance of the primary registers, chest and falsetto.[57]

Hirano et al. stressed the fact that register, pitch, and intensity are not independent parameters in the human voice.[57] According to Klingholz et al.,[27,58] the larynx is not able to produce all pitches of the voice range with the same mechanism and many adjustments of muscular and aerodynamic forces may occur in phonation.[20,21,30,36,37,57,59-62] Each voice register is characterized by specific properties and dependent on specific physiological mechanisms. Tarneaud emphasized the fact that untrained voices present a disproportion between the laryngeal sound and the resonance in the pharynx, which results in clearly perceptible registers and passagios.[63] With a well established muscular training, these passagios are camouflaged. In other words, the singer has learned to tune his pharynx to the sound produced by the larynx.

Resonance is relevant to the register phenomenon[37] and vocal tract resonances have a pronounced effect on the chest-falsetto transition, depending on the individual characteristics of the subject.[65] A strong correlation between the register transition and the source-filter interaction has been found.[33,64-66] Many voice breaks can be observed in untrained subjects while (classically) trained voices exhibit less pitch breaks or jumps.[32-33,64-66] Note that comparable frequency jumps are also frequent in a variety of animal vocalizations.[64]

The location of the register transition

Many textbooks on singing education stress the importance of the register transition in relation to the three basic female and male voice types. However, the results of our explorative study on voice classification show that only 9% of private singing teachers and only 54.5% (classical) and 60% (musical theatre) conservatory singing teachers use register transition as a voice classification criterion. Moreover, timbre is also mentioned as an important voice classification parameter by 56% private singing teachers and by all conservatory singing teachers (classical and musical theatre). As the timbre of the chest register is quite different from the timbre of the falsetto (head) register, it is at least remarkable to observe that no private singing teacher and only 36.4% (classical) and 20% (musical theatre) of the conservatory singing teachers in our study pay attention to the specific timbre of both registers as a voice classification criterion (Addendum 1).

Although the register transition zone has always been a major concern in singing education textbooks, its objective location on the voice range scale has been the subject of much debate. McGlone[67] prefers a method of self- determination of the shift between the registers, since that is "more reliable than having judges make the decision", while Neumann et al.[68] find that "singers themselves are frequently unable to locate their point of transition and to distinguish clearly between registers", and the judgment by musically trained persons is also questioned. Some explanation could be given for these apparently contradictory findings. One has to bear in mind the still existing practice of trying to locate registers by means of the subjective sensations felt by singers in some parts of their body.[20,27,35,41,44] In addition, the practice of register balance in educated singers interferes with locating the register transitions.[69] They sometimes are considered to lay outside the normal ranges on account of a given voice disorder.[28] Besides, persons with non-educated voices try to sing in the chest register as high as possible. As a result, a clear crack can be heard when entering the falsetto (head) register. In classical singing, much attention is paid to the register transition by trying to smooth out (camouflage) the register zone to make the voice sound as just one register.

The analysis of register transitions and source-tract interaction is often studied using glissando singing.[32,65,70] This demonstrates that the supraglottis has an influence on the register transition.[64] Glissandos cause a larger "dip" and high frequencies can be obtained in the modal register by a singing technique, based on forcing.[71] This also means that the singer has a way, albeit limited, to camouflage or to accentuate the register transitions. Besides the choice between glissando singing or the production of sustained vowels, other factors like the adopted singing technique, the choice of the vowel and singing up or down can influence the location of the so-called register breaking point.[72] These findings also explain why authors accept overlaps in frequency between adjacent registers ("amphotere Klänge" in German publications).[38,73-75] Sundberg, for instance, situates the range of overlap between male modal and falsetto registers in the vicinity of 200 to 350 Hz, which is almost one octave.[76] It is also acknowledged that these ranges of register overlap and the register boundaries vary substantially among individuals. However, most authors omit to associate their findings with a specific voice classification. Some authors, however, provide more specific information on register transitions related to voice categories, but their findings are very divergent.[27,56,68,77-80] We may conclude that comparing findings on register transitions within and across studies proves to be difficult, primarily because of the discrepancies in the register systems used.[33,70,73]

The characteristics of the register transition zone, demonstrated on the VRP as distinctive features in the discrimination of basic vocal types

As frequency location could not be the sole characteristic of a vocal register, we hypothesized that more attention should be paid to the associated vocal intensity characteristics.[43] It is here that, in our opinion, VRP can open a new era of voice research. Of course, many scientists discovered the eye casting dip(s) in the VRP but interpretation remained obscure. The register transitions, however, clearly visible in the VRP at certain frequencies, illustrate the different mechanisms based on pre-phonatory larynx positions and specific muscle activities.[27,55-59,62]

Regarding the discussion about the number of registers, our methodology is based on Hollien's postulate that only three *major* registers exist: they are the pulse (= vocal fry), modal (= chest) and loft (=falsetto) registers.[38] Fundamental frequencies lower than 70 Hz are judged as vocal fry.[37-38,81] The register transition between chest and falsetto is commonly used in singing manuals and in specific studies on registers in singing.[38,55,69,74,77-78] This *main* register transition is also clearly visible in the VRP, which explains the choice of our methodology.

Instead of entering into discussions about the "exact" location of the register transition, we introduced a new methodology that covers the *main* register zone. The register *zone* (including the marked register dip) was measured from the last ascending maximum intensity point to the next maximum ascending intensity point. This methodology intercepts all possible influences on the register transition, mentioned above, which could bias a carefully balanced interpretation of the register transition phenomenon. Recent studies give indeed proof for one or more regions of voice instabilities, which can be spread over one octave.[32,68] This (large) zone of instability is caused by irregularities in vocal fold vibration, accompanied by SPL variations up to 15 dB, even in classically trained singers, trying to avoid pitch breaks or jumps. This methodology has been proven to be a valuable aid in discerning three distinct basic voice types per gender.

Clinical implications

This study demonstrates that parameters describing the register transition zone provide a clear cluster separation between three natural female and male voice categories. This suggests that the complexity of voice classification may be reduced to a more compact, yet adequate formula, easily obtainable from the VRP, and thus appropriate for clinical use.

Our study applies to male and female singing students and singers, ages 18 to 52, who are or have been using different types of singing techniques in their educational (conservatories) and professional (stages) environments. The results of our study concerning this mixed population of different gender, age, education and occupation could open a new area of research and clinical applications. We believe that our methodology can be applied to every kind of human vocalization, be it speaking or singing. After all, many authors consider that the acoustic principles of speaking and singing are basically the same[82-85] and do admit that the speaking and singing voice are strongly interdependent.[82] Many verbal productions are intermediary between singing and speaking (cfr. Sprechgesang, “Parlando” singing, recitatives).[63,84-86] Country singers use basically the same type of phonation when they sing and when they speak.[19]
Singing is supposed to involve different, and probably more careful patterns of control over both the vocal source and filter[11,68,86,88] as subglottal pressure not only increases intensity but also fundamental frequency.[88] Although the lung pressure attained in singing may be more than four times that used in speech, air flow must be kept at a level similar to that in speech.[87] Singers are more prone to “flow phonation”, thereby reducing excessive subglottic pressure.[89] Insufficient accuracy in subglottal pressure regulation will lead to singing out of tune.[86,88,90] Therefore, a perfect coordination between the laryngeal muscles and the respiratory muscles is essential.[90] This observation leaded to the conviction that professional singers are a “physiologically and neurologically unique group of individuals”.[91] However, the investigations of Thorpe et al. have shown a wide variability in the respiratory strategies employed by different singers.[92] The mean differences between the mean phonational range of a group of professional singers of all ages and the mean phonational range of an untrained group were statistically not significant.[93] There is little or no significant difference in maximum vocal intensity between a trained and untrained singer[20,94] and the spectrum of the voice source is about the same in ordinary speakers as it is in trained singers.[94] As Miller stated: ‘it is very confusing to singers and actors to think that they have two voices when they really

have two vocal folds that will speak or sing or yell or laugh'.[94]
Overlapping in categories also happens when one tries to divide individuals into singers, amateur-singers, singing students and professional singers. Schutte and Miller too, admit that the line between professional and dilettante is indistinctive.[96]
In discerning an area of the shouting voice next to the transition of chest and head registers, Hacki observed that the dynamics of this shouting voice area sometimes exceed the area of the singing voice.[79] Denk and Frank too came to the conclusion that the limits of the dynamics of the speaking- and singing voice area approach.[97]
The larger voice and intensity range in singing depends on the singing style, but even in an animated, expressive conversation, considerable extremes of frequency and intensity can be found.[97] In pathology, dissociation of one's speaking and singing voice is a frequently occurring cause of voice problems in singers.[98-99] These observations stress the importance of the register transition zone, in direct relationship with the three basic vocal types, as shown in our study.
Our study indicates that there exist three basic voice types, corresponding with classic voice types or not. This indicates the need of voice classification, also in modern music, where not a similar vocal repertoire has been written like in classic music.

We assume that our methodology can be useful, not only in determining a basic voice type for singers, but also in providing interesting cues for voice diagnosis and voice therapy in general, taking into account the relationship of all different human vocalizations.

Future research

More studies are required to link the statistically obtained cluster separation to the three basic female and male voice categories, as they are traditionally and to date used.

This study of clustering the voice is the consequence of the finding that there exists no generally accepted method for voice classification. Consequently, it is difficult to compare the results of "classic" voice classification with the clusters we found. The use of an "empirical" classification, i.e. the observation that singers sing for many years in a distinct voice type, may be an adequate approach.
Most parameters that have lead to the cluster separation, however, are not easily understandable in clinical terms. Therefore, it is not easy to link them to the clinical

situation, nor can the difference between these parameters that have lead to clustering of female and male voice readily be explained. This can be subject of future research. More explicitly, the location of important markers in the VPR, such as the optimal frequency level and the register transition zone, should, in our opinion, be considered in view of the basic voice type of each gender.

References

1. Brodnitz F. Panel Discussion: Medical Management of Voice Disorders. J Voice 1987;2:202-207.
2. von Leden H. Panel Discussion: Medical Management of Voice Disorders. J Voice 1987;2:202-207.
3. Kitch JA, Oates J, Greenwood K. Performance Effects on the Voices of 10 Choral Tenors: Acoustic and Perceptual Findings. J Voice 1996;217-227.
4. Hoffman-Ruddy B, Lehman J, Crandell C, Ingram D, Sapienza C. Laryngostroboscopic, Acoustic, and Environmental Characteristics of High-Risk Vocal Performers. J Voice 2001;543-552.
5. Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP) : standardizing voice area measurement / phonetography. Folia Phoniatr 1983; 35:286-288
6. Schultz-Coulon H-J. Zur routinemässigen Messung der stimmlichen Reaktion im Lärm. Sprache - Stimme - Gehör 1980;4:28-34.
7. Wirth G. Stimmstörungen. Lehrbuch für Ärtzte, Logopäden, Sprachheilpädagogen und Sprecherzieher. Deutscher Ärtzte-Verlag, Köln, 1991.
8. Bloothooft G, Plomp R. Spectral analysis of sung vowels. I. Variation due to differences between vowels, singers, and modes of singing. J Acoust Soc Am 1984; 75 (4):1259-1264.
9. Bjôrkner E, Sundberg J, Cleveland T, Stone E. Voice source differences between registers in female musical theatre singers. J Voice 2006;20:187-197.
10. Björkner E. Musical theatre and opera singing – why so different? A study of subglottal pressure, voice source, and formant frequency characteristics. J Voice 2008;22:533-540.
11. Hamdam AL, Deeb R, Tohme RA, Rifai H, Husseini S, Fuleihan N. Vocal technique in a group of Middle Eastern singers. Folia Phoniatr Logop 2008;60:217-221.
12. Radionoff SL, Satterfield TR, Lee E. Commercial Music : A survey of degree granting institutions in the United Kingdom. Choice for Voice 2008. Multidisciplinary Approaches to Performance, Health and Research in Voice, London:The British Voice Association 2008;1-22.
13. Sundberg J. Vocal Fold Vibration Patterns and Modes of Phonation. Folia Phoniatr Logop 1995;47:218-228.
14. Plant RL, Younger RM. The Interrelationship of Subglottic Air Pressure, Fundamental Frequency, and Vocal Intensity During Speech. J Voice 2000;170-177.
15. Wilson JV, Leeper HA. Changes in Laryngeal Airway Resistance in Young Adult Men and Women as a Function of Vocal Sound Pressure Level and Syllable Context. J Voice 1992;235-245.
16. Stathopoulos ET, Sapienza C. Respiratory and Laryngeal Function of Women and Men During Vocal Intensity Variation. J Speech and Hear Res 1993;64-75.
17. Keilmann A, Bader C-A, Bergler W. Aerodynamische Aspekte der kindlichen Phonation. Sprache-Stimme-Gehör 1994;125-129.

18. Frank F, Donner F. Die Bedeutung der Stimmfeldmessungen für den Gesangs-unterricht aus phoniatrischer und gesangspädagogischer Sicht. Sprache - Stimme - Gehör 1986;10:93-97.
19. Cleveland TF, Stone RE Jr, Sundberg J, Iwarsson J. Estimated Subglottal Pressure in Six Professional Country Singers. J Voice 1997;403-409.
20. Colton RH. Vocal Intensity in the Modal and Falsetto Registers. Folia phoniat 1973; 25:62-70.
21. Klingholz F, Martin F. Zum Entstehungsmechanismus des Vocal Fry (VF). Sprache - Stimme - Gehör 1983;7:16-21.
22. Schutte HK. Toward a Definition of Male 'Head' Register, Passagio, and 'Cover' in Western Operatic Singing. Folia Phoniatr Logop 1994;46:157-170.
23. Sundberg J, Elliot N, Gramming P, Nord L. Short-Term Variation of Subglottal Pressure for Expressive Purposes in Singing and Stage Speech: A Preliminary Investigation. J Voice 1993;227-234.
24. Leanderson R, Sundberg J, von Euler C. Breathing Muscle Activity and Subglottal Pressure Dynamics in Singing and Speech. J Voice 1987;258-261.
25. Watson PJ, Hixon TJ, Maher MZ. To Breathe or Not to Breathe -- That is the Question: An Investigation of Speech Breathing Kinematics in World Class Shakespearean Actors. J Voice 1987;269-272.
26. Williams RN. The Living, Breathing Actor. J Voice 1988;36-39.
27. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimm-feld. Sprache - Stimme - Gehör 1985;9:109-111.
28. Airainer R, Klingholz F. Quantitative Evaluation of Phonetograms in the Case of Functional Dysphonia. J Voice 1993;7:136-141.
29. Walker J.S. An Investigation of the Whistle Register in the Female Voice. J Voice 1988;2:140-150.
30. Keilmann A., Michek F. Physiologie und akustische Analysen der Pfeifstimme der Frau. Folia Phoniatr 1993;45:247-255.
31. Miller D.G., Schutte H.K. Physical Definition of the "Flageolet Register". J Voice 1993; 7:206-212.
32. Garnier M, Henrich N, Crevier-Buchman L, Smith CVJ, Wolfe J. Glottal behavior in the high soprano range and the transition to the whistle register. J Acoust Soc Am 2012;131/1:951–962.
33. Henrich N. Mirroring the voice from Garcia to the present day: Some insights into singing voice registers. Logopedics Phoniatrics Vocology 2006;31:3-14
34. Seidner W, Kröger H, Wernecke KD. Numerische Auswertung spektraler Stimmfelder. Sprache - Stimme - Gehör 1985;9:10-13.
35. Hollien H, Schoenhard C. The Riddle of the "Middle" Register in: Vocal Fold Physiology, Titze and Scherer Ed., Denver, 1983.
36. Large J. Acoustic Study of Register Equalization in Singing. Folia phoniat 1973;25: 39-61.
37. Titze IR. A Framework for the Study of Vocal Registers. J Voice 1988;2:183-194.
38. Hollien H. On vocal registers. Journal of Phonetics 1974;2:125-143
39. Röhrs M, Pascher W, Ocker C. Untersuchungen über das Schwingungsverhalten der Stimmlippen in verschiedenen Registerbereichen mit unterschiedlichen strobo-skopischen Techniken. Folia phoniat 1985;37:113-118.
40. Welch GF, Sergeant DC, MacCurtain F. Some Physical Characteristics of the Male Falsetto Voice. J Voice 1988;2:151-163.
41. Reinders A, Schutte HK. Registers in de zangstem. Logopedie en Foniatrie 1987;59: 358-360.
42. Walker JS. An Investigation of the Whistle Register in the Female Voice. J Voice 1988;2:140-150.
43. Spencer ML, Titze IR. An Investigation of a Modal-Falsetto Register Transition Hypothesis Using Helox Gas. J Voice 2001;15-24.

44. Colton RH, Hollien H. Perceptual Differentiation of the Modal and Falsetto Registers. Folia phoniat 1973;25:270-280.
45. Schutte HK, Miller R. Resonance Balance in Register Categories of the Singing Voice : A Spectral Analysis Study. Folia phoniat 1984;36:289-295.
46. Miller DG, Schutte HK. Formant Tuning in a Professional Baritone. J Voice 1990;4: 231-237.
47. Walker JS. An Investigation of the Whistle Register in the Female Voice. J Voice 1988;2:140-150.
48. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F, Wyke BD, Eds Disorders of Human Communication 6. Springer-Verlag Wien New York, 1982.
49. Hollien H, Schoenhard C. The Riddle of the “Middle” Register in: Vocal Fold Physiology, Titze and Scherer Ed, Denver, 1983.
50. Large J, Iwata S, von Leden H. The Primary Female Register Transition in Singing. Folia phoniat 1970;22:385-396.
51. Hollien H. That Golden Voice - Talent or Training ? J Voice 1993;7:195-205.
52. Röhrs M, Pascher W, Ocker C. Untersuchungen über das Schwingungsverhalten der Stimmlippen in verschiedenen Registerbereichen mit unterschiedlichen strobo-skopischen Techniken. Folia phoniat 1985;37:113-118.
53. Schutte HK, Miller DG. Belting and Pop, Nonclassical Approaches to the Female Middle Voice: Some Preliminary Considerations. J Voice 1993;7:142-150.
54. Large J, Murry T. Studies of the Marchesi Model for Female Registration. In: Large J, Ed. Contributions of Voice Research to Singing.College-Hill Press, Houston, Texas, 1980;183-199.
55. Miller DG, Schutte HK, Doing J. Soft Phonation in the Male Singing Voice: A Preliminary Study. J Voice 2001;483-491.
56. Miller DG, Schutte HK. 'Mixing' the Registers: Glottal Source or Vocal Tract? Folia Phoniatr Logop 2005;57:278-291.
57. Hirano M, Vennard W, Ohala J. Regulation of Register, Pitch and Intensity of Voice. An Electromyographic Investigation of Intrinsic Laryngeal Muscles. Folia phoniat 1970;22:1-20.
58. Klingholz F, Martin F, Jolk A. Das dreidimensionale Stimmfeld. Laryng Rhinol Otol 1986;65.
59. Klingholz F, Jolk A, Martin F. Stimmfelduntersuchungen bei Knabenstimmen (Tölzer Knabenchor). Sprache - Stimme - Gehör 1989;13:107-111.
60. Large J, Iwata S, von Leden H. The Primary Female Register Transition in Singing. Folia phoniat 1970;22:385-396.
61. Scherer RC. (Moderator and Editor). The Integration of Voice Science, Voice Pathology, Medicine,Public Speaking, Acting, and Singing. J Voice 1994;8:359-374.
62. Klingholz F. Die Akustik der gestörten Stimme. Georg Thieme Verlag, Stuttgart - New York, 1986.
63. Tarneaud J. Traité pratique de phonologie et de phoniatrie Librairie Maloine SA Paris 1961:2.
64. Isao T, Tokuda M, Kob ZM, Herzel H. Biomechanical modeling of register transitions and the role of vocal tract resonators. J Acoust Soc Am 2010;127/3:1528–1536.
65. Echternach M, Dippold S, Sundberg J, Arndt S, Zander MF, Richter B. High-Speed Imaging and Electroglottography Measurements of the Open Quotient in Untrained Male Voices' Register Transitions. Journal of Voice 2010;24/6:644–650.
66. Echternach M, Traser L, Richter B. Perturbation of Voice Signals in Register Transitions on Sustained Frequency in Professional Tenors. J of Voice 2012;26/5: 674.e9-674.e15.
67. McGlone RE. Air Flow in the Upper Register. Folia phoniat 1970;22:231-238.
68. Neumann K, Schunda P, Hoth S, Euler HA. The Interplay between Glottis and Vocal Tract during the Male Passaggio. Pholia phoniatr Logop 2005;57:308-327.
69. Vogelsanger GTh. Experimentelle Prüfung der Stimmleistung beim Singen. Folia phoniat 1954;6:193-227.

70. Roubeau B, Henrich N, Castellengo M. Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. J ournal of Voice 2008;23/4:425-438.
71. Pabon JPH. Kwaliteitsaspecten in het fonetogram. Logopedie 1997;10:20-33.
72. Gramming P. The Phonetogram. An Experimental and Clinical Study. Thesis. Department of Otolaryngology, University of Lund, Malmö General Hospital, Malmö, 1988.
73. Keidar A, Hurtig RR, Titze IR. The Perceptual Nature of Vocal Register Change. J Voice 1987;1/ 3:223-233.
74. Bloothooft G. De professionele zangstem. Log en Fon 1985;57:237-242.
75. Lerman JW, Duffy RJ. Recognition of Falsetto Voice Quality. Folia phoniat 1970;22: 21-27.
76. Sundberg J. The Science of the Singing Voice. Northern Illinois University Press, Dekalb, Illinois, 1987.
77. Pahn J. Stimmübungen für Sprechen und Singen. VEB Verlag Volk und Gesundheit, Berlin, 1968.
78. van Deinse JB, Goslings WRO. The technique of singing. Some remarks on the mechanism, the technique and the classification of the singing voice. Theodora Versteegh Foundation. Government Publishing Office, The Hague, 1982.
79. Hacki T. Die Beurteilung der quantitativen Sprechstimmleistungen. Das Sprechstimm-feld im Singstimmfeld. Folia Phoniatr 1988;40:190-196.
80. Wirth G. Stimmstörungen. Lehrbuch für Ärtzte, Logopäden, Sprachheilpädagogen und Sprecherzieher. Deutscher Ärtzte-Verlag, Köln, 1991.
81. Hollien H, Michel JF. Vocal Fry as a Phonational Register. JSHR 1968;11:3.
82. Baer T, Bell-Berti F, Rubin P. Articulation and voice quality. In: Lawrence VL, ed. Transcripts of the sixt symposium “Care of the Professional Voice”. New York: The Voice Foundation, 1978;48-53.
83. Frank F. Sprech- und Singstimmbeurteilung über stimmliche Parameter. Sprache - Stimme - Gehör 1993;17:43-46.
84. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. Raven Press, New York, 1991;49-68.
85. Habermann G. Stimme und Sprache. Eine Einführung in ihre Physiologie und Hygiene. Für Ärtzte, Sänger, Pädagogen und alle Sprechberufe. Georg Thieme Verlag, Stuttgart - New York, 1986.
86. Leanderson R, Sundberg J, von Euler C. Breathing Muscle Activity and Subglottal Pressure Dynamics in Singing and Speech. J Voice 1987;258-261.
87. Rothenberg M, Miller D, Molitor R, Leffingwell D. The Control of Air Flow During Loud Soprano Singing. J Voice 1987;262-268.
88. Sundberg J, Elliot N, Gramming P, Nord L. Short-Term Variation of Subglottal Pressure for Expressive Purposes in Singing and Stage Speech: A Preliminary Investigation. J Voice 1993;227-234.
89. Carroll LM, Sataloff RT, Heuer RJ, Spiegel JR, Radionoff SL, Cohn JR. Respiratory and Glottal Efficiency Measures in Normal Clasically Trained Singers. J Voice 1996; 139-145.
90. Leanderson R, Sundberg J. Breathing for Singing. J Voice 1988;2-12.
91. Kmucha ST, Yanagisawa E, Estill J. Endolaryngeal Changes During High-Intensity Phonation Videolaryngoscopic Observations. J Voice 1990;346-354.
92. Thorpe CW, Cala SJ, Chapman J, Davis PJ. Patterns of Breath Support in Projection of the Singing Voice. J Voice 2001;86-104.
93. Brown WS Jr, Morris RJ, Hicks DM, Howel E. Phonational Profiles of Female Professional Singers and Nonsingers. J Voice 1993;7:219-226.
94. Sataloff RT. Rational Thought: The Impact of Voice Science upon Voice Care. J Voice 1995;9:215-234.
95. Miller R. The Perspective of the Singing Teacher and Voice Pedagogue. In: Scherer, RC, Ed. The Integration of Voice Science, Voice Pathology, Medecine, Public Speaking, Acting, and Singing. J Voice 1994;359-374.

96. Schutte HK, Miller DG. Belting and Pop, Nonclassical Approaches to the Female Middle Voice: Some Preliminary Considerations. J Voice 1993;7:142-150.
97. Denk DM, Frank F. Die Wertigkeit verschiedener Stimmfeldmessungen an jeweils demselben Probanden. Sprache - Stimme - Gehör 1993;17:169-172.
98. Teachey JC, Kahane JC, Beckford NS. Vocal Mechanics in Untrained Professional Singers. J Voice 1991;5:51-56.
99. Sundberg J, Titze I, Scherer R. Phonatory Control in Male Singing: A Study of the Effects of Subglottal Pressure, Fundamental Frequency, and Mode of Phonation on the Voice Source. J Voice 1993;7:15-29.

Addendum 1

Voice classification in practice 1: criteria in contemporary singing education

Parts of this study are published in:
Stemclassificatie in de praktijk. Een exploratieve studie
H. Lycke , W. Decoster and F.I.C.R.S. de Jong
Nederlands Tijdschrift voor Stem- Spraak- en Taalpathologie
Accepted for publication

Abstract

Aims: To explore how contemporary singing teachers deal with voice classification and which criteria they use.
Study Design: explorative study using questionnaires.
Methods: One questionnaire was sent to 200 singing teachers via internet and a second questionnaire to 22 singing teachers of one Classical Belgian conservatory and two Musical Theatre conservatories.
Results: Of the 200 singing teachers, 72 responded (36%). In 61.1% voice classification was important for at least one reason, while 38.9% did not find voice classification an important issue. Most used acoustical parameters for voice classification were frequency range/tessitura (56.0%), voice quality/timbre (56.0%), volume (12.1%) and register transition (9.0%). The conservatory singing teachers classified their students (n = 165). In the conservatory singing teachers, voice classification was an important issue in singing education. Frequency range/tessitura, voice quality, register transition, and volume were the most frequently used criteria. However, each singing teacher reported a varying individual set of voice classification criteria, depending on the singing student and on the specialty of the department.
Conclusions: The results of this study indicate that there is a marked difference of attitude towards voice classification in singing teachers and that different criteria for voice classification are used. Apparently, there is no consensus about the advisability and criteria of voice classification among the various singing teachers.

Introduction

At present, music education and performance can be divided into two broad categories: Classical Music and Commercial Music. Classical Music represents various genres, such as Opera, Lied, and Oratorio. Commercial Music represents genres such as Pop, Rock, Jazz, Country, Rhythm and Blues, Hip-Hop, Rap, Gospel, and Musical Theatre. According to Gilman et al. contemporary Commercial Music has become the largest and conceivably the most popular genre of music in the United States.[1] It may be assumed that the same trend occurs in European and other countries, probably due to the emergence of commercial industry reality TV shows such as “Star Academy” and “X-Factor”.
Many singing students are taking private singing lessons, which are not curriculum-bound. According to Radionoff et al., contemporary Commercial Music

singers often complain that their singing teachers do not understand the vocal styles and demands of a contemporary Commercial Music singer[2]. In their study on Commercial Music, they concluded that "along with nomenclature disparity a tremendous lack of consistency exists among curriculums of Commercial Music degrees."

Traditionally, in Classical Music, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor. There are, however, many subtypes, according to different roles and based on the characteristics of the voice, such as intensity, timbre, mobility, vibrato, temperament, expression and personality. In the first place, classifying a voice means to determine the frequency and intensity voice range in which a subject can work without harming or fatiguing his voice and to which repertoire he should be assigned by the singing teacher.[3-9] Correct classification of the singer's voice is indispensable in order to achieve optimum performance. Coleman already stated the general consensus that singing and speaking outside a given physiological pitch or intensity range is a potential hazard[4] .

Voice specialists stressed the importance of an correct voice classification before voice education starts.[7-14] Incorrect voice classification can enhance functional and organic voice disorders.[7-8,15-19] The biographies of famous and less famous singers very frequently mention examples of the pernicious outcomes for their voice and for their career caused by incorrect voice classification, and recent studies show the great prevalence of voice disorders by incorrect voice classification among singing students, singers and singing teachers.[1,20-22]

In classical singing education great emphasis is put upon voice classification, but little is known how the relatively new music institutions and individual singing teachers deal with voice classification.

The purpose of this study was to explore how contemporary singing teachers of Classical and Commercial Music deal with voice classification today and which criteria they use to classify their singing students.

Methods

This study is part of a larger explorative study on voice classification by the Centre of Excellence for Voice of the Dep. Exp. ORL of the KU Leuven, Belgium.

A questionnaire (questionnaire 1) was sent to all 200 private singing teachers, who were registered at the commercial Dutch Internet site www.vocalisten.nl, and who mentioned their e-mail address. The singing teachers who cooperated in this study returned their answers via e-mail. These singing teachers recommended themselves for a total of 134 specialties and styles of singing. Each of them proclaimed to master a great variety of genres and singing styles, such as Belting, Blues, Classic, Close Harmony, Country, Disco, Easy Listening, Evergreens, Funk, Fusion, Gypsy, Hard Rock, Jazz, Latin, Opera, Musical Theatre, Pop, Rhythm and Blues, Salsa, Soft Rock, and World Music. The singing teachers were asked if voice classification was important to them and why. They were also asked what criteria for voice classification they applied.

At three conservatories (one Belgian Classical conservatory specialized in Opera, Lied, and Oratorio, and one Dutch and one British conservatory specialized in Musical Theatre) an other questionnaire (questionnaire 2) was distributed among singing teachers via the Head of the Department. The answers to the questionnaires were collected by one of the singing teachers. The 22 singing teachers from three conservatories who cooperated in this study classified a total of 165 singing students: 81 singing students (58 females and 23 males) at the Belgian classic conservatory, 63 singing students (55 females and 8 males) at the Dutch conservatory (Musical Theatre), and 21 singing students (9 females and 12 males) at the British conservatory (Musical Theatre). The singing students were aged between 18 and 28 years, mean age 21 years. The singing teachers were asked to classify their singing students and to indicate on what criteria their voice classification was based.
Descriptive statistics were performed by SPSS 16.00.

Results

Questionnaire 1

From the 200 singing teachers 72 answers were received from 58 female (80.6%) and 14 male (19.4%) singing teachers. Consequently, the response rate is 36.0%. Although most of these singing teachers had a classical singing training, each of them recommended him/herself on the website for many specialties and musical styles (up to 24 in one teacher). Most popular styles were Musical Theatre (66.7%), Classic (50.0%), Pop and Jazz (each 43.8%), Dutch Pop (19.0%), a Capella (17.0%), Close Harmony (16.0%), Lied (15.0%) and Opera (13.0%).
Fourty-four singing teachers (61.1%) found that voice classification is important for at least one reason, while 28 singing teachers (38.9%) stated that voice

classification was not an important issue for their teaching. The reported arguments pro and contra voice classification are listed in figure 1 and 2.

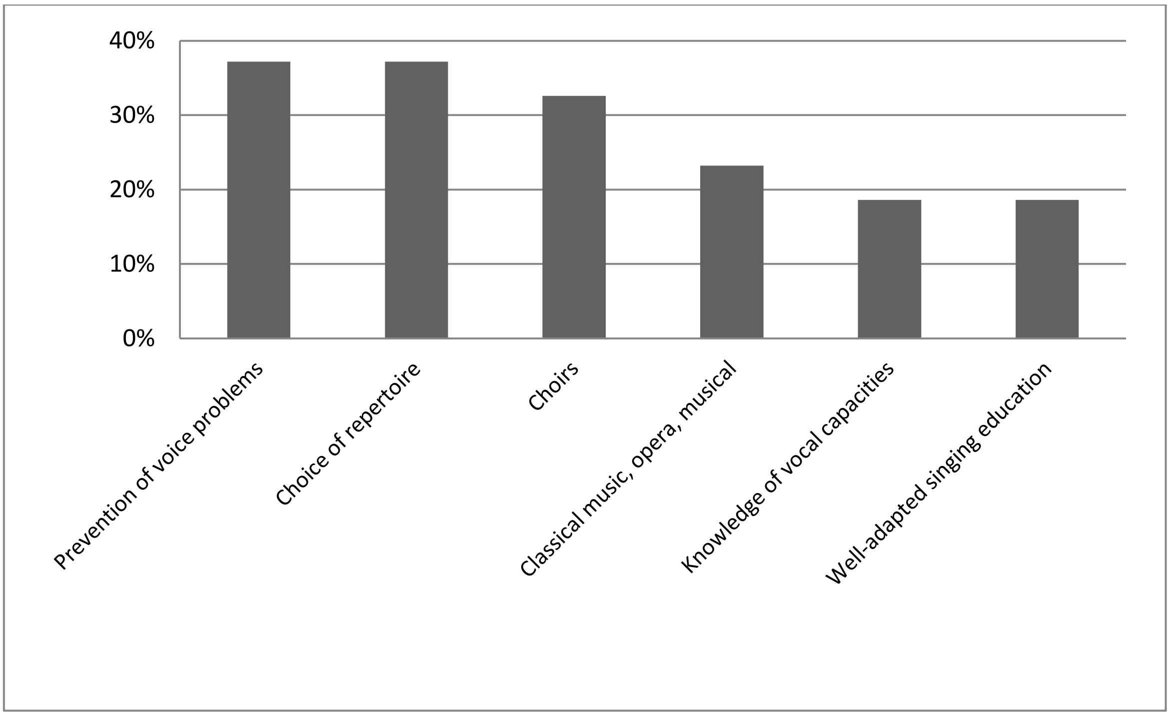


Figure 1. Arguments used pro voice classification reported by 72 singing teachers (percent).

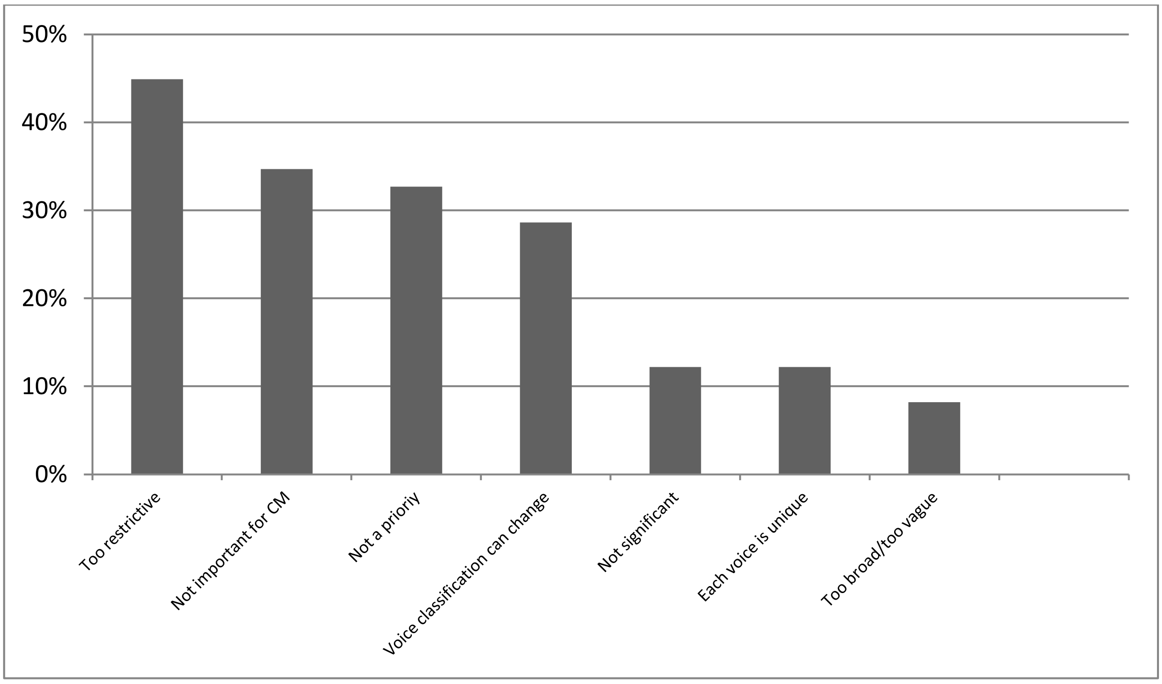


Figure 2. Arguments used contra voice classification reported by 72 singing teachers (percent).

Most singing teachers (n=66, 91.7%) provided information about their criteria for voice classification. These criteria for voice classification can be sorted into: physical features, acoustical features, specific methods, miscellaneous, and "other factors". Six respondents (8.3%) did not mention any criteria for voice classification.

The criteria used for voice classification by the singing teachers are shown in figure 3. Most frequently used acoustical parameters for voice classification were frequency range/tessitura (56.0%), voice quality/timbre (56.0%). One singing teacher used a purely commercial approach: voice classification on demand.

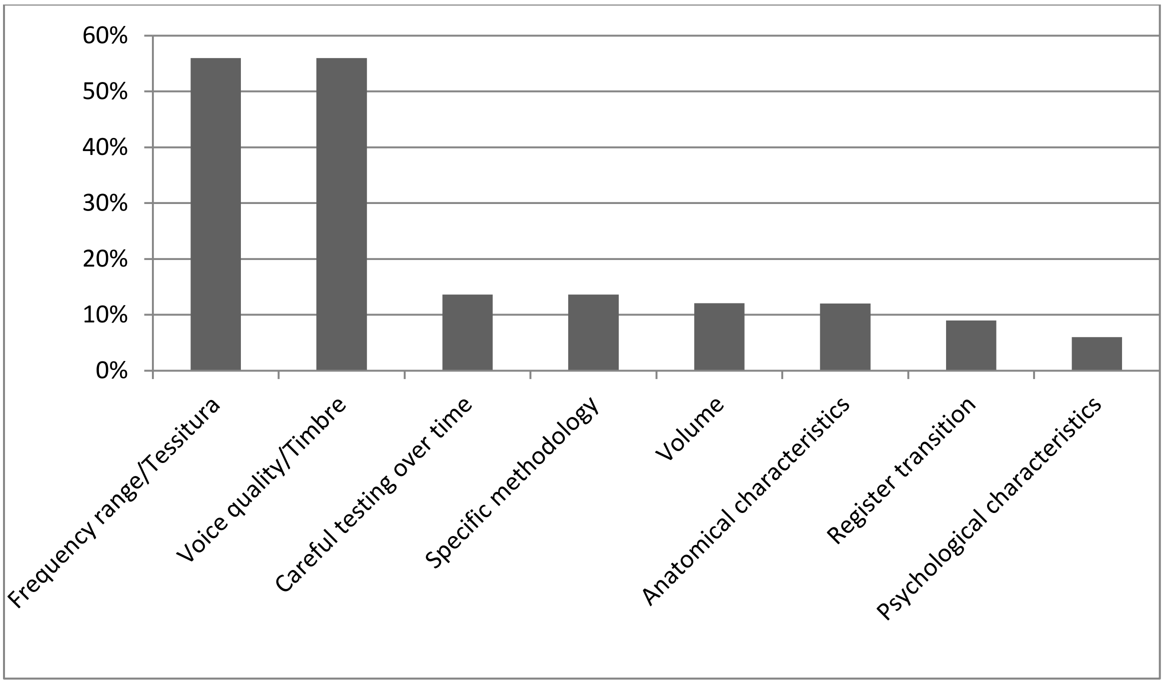


Figure 3. Criteria for voice classification used by 72 private singing teachers (percent).

Questionnaire 2

Table 1 represents the voice classification criteria used by singing teachers of classic and musical conservatory. In Table 2 and 3 the voice classification criteria used by the individual singing teachers of the two types of conservatory are listed (Classical conservatory in Table 1 and Musical Theatre conservatory in Table 2). Frequency range/tessitura and voice quality/timbre were used by all conservatory singing teachers as voice classification criteria, except for one Classical singing teacher who used no voice classification criteria at all. Singing teachers of the three conservatories used a different set of voice classification criteria per singing student. Moreover, voice classification criteria also appeared to be dependent on the type of conservatory.

Table 1. Voice classification criteria used by singing teachers of classic and musical conservatory (percent).

Voice classification criteria	Classical conservatory (N = 12)	Musical Theatre conservatory (N=10)
Frequency range/tessitura	100	100
Timbre , quality, color, character	100	100
Register transition	54.5	60.0
Frequency range of belt	0.0	50.0
Volume	54.5	50.0
Quality in specific register	36.4	20.0
Vocal development	0.0	60.0
Speaking voice	27.3	10.0
Ease	36.4	10.0
Physics	9.1	30.0
Feeling of the singer	18.2	10.0

Table 2. Individual voice classification criteria used by 11 classical singing teachers.

Singing teachers	1	2	3	4	5	6	7	8	9	10	11
Frequency range/tessitura	x	x	x	x	x	x	x	x	x	x	x
Timbre , quality, color	x	x	x	x	x	x	x	x	x	x	
Register transition	x	x		x	x	x			x		
Frequency range of belt											
Volume		x	x	x		x				x	x
Quality in specific register			x	x		x	x				
Vocal development											
Speaking voice		x	x				x				
Ease		x		x		x	x				
Physics								x			
Feeling of the singer						x				x	

Singing teachers of the Belgian Classical conservatory used two to seven different criteria (Table 2). Except for three singing teachers, using only frequency range/tessitura, timbre, and register transition as voice classification criteria, all

other singing teachers used quite different sets of criteria, while one classical singing teacher used no classification criteria at all. In Classical singing teachers vocal development and frequency range of belt were not used as voice classification parameters.
Singing teachers of the Dutch and the British Musical Theatre conservatories used two to 11 different criteria (Table 3). They all used frequency range /tessitura and voice quality/timbre as voice classification criteria. The frequency range of belt was specifically mentioned by five of the 10 singing teachers of Musical Theatre.

Table 3. Individual voice classification criteria used by 10 Musical Theatre singing teachers.

Singing teachers	1	2	3	4	5	6	7	8	9	10
Frequency range/tessitura	x	x	x	x	x	x	x	x	x	x
Timbre , quality, color	x	x	x	x	x	x	x	x	x	x
Register transition	x	x	x	x				x	x	
Frequency range of belt	x	x	x	x	x					
Volume	x	x	x	x		x				
Quality in specific register				x						x
Vocal development	x	x	x	x	x					x
Speaking voice				x						
Ease				x						
Physics				x					x	x
Feeling of the singer				x						

Discussion

Because of the lack of studies on this subject, it is difficult to compare the results of this study to those of other authors. Studies on different categories of singing are scarce and were mainly focused on classical singing, which emphasizes the importance of voice classification in Classical singing. During the last two decennia the scientific world noticed the changing world of singing education and practice and few studies appeared on different singing genres and techniques.[15,23-26]
While many singing teachers in the internet enquiry (questionnaire 1) had a Classical singing education themselves, they paid few or no attention to voice classification, which was formerly very important in Classic voice education. Almost 40% of the internet singing teachers stated that voice classification was no important issue for their teaching. Additionally, not all of the reported arguments

contra voice classification seem to be valid. After all, if voice category could change anyway, classification is still important in order to watch the instant physiological limits of the voice, just if the voice is unique. Therefore, voice classification is important in all types of singing. These findings indicate a worrisome trend of altered attitude towards voice classification by many private singing teachers today. It could be argued that many singing teachers try to attract as many singing students as possible by an extensive offer of singing genres for which they thought voice classification is not absolutely necessary. Furthermore, in this internet enquiry 91.7% of the respondents provided information about their criteria for voice classification, while 38.6% of them did not find voice classification an important issue for their teaching. Most used acoustical parameters for voice classification were frequency range/tessitura (56.0%), voice quality/timbre (56.0%), volume (12.1%) and register transition (9.0%). Obviously, in this internet study there is no consensus about the criteria for voice classification.

Radionoff et al.[2] pointed at the tremendous lack of consistency among curriculums of commercial music degrees. It has been assumed that professional singers are at risk for voice problems, which may result in vocal pathologies. Recent studies showed the great prevalence of voice disorders among singing students, singers and singing teachers by incorrect voice classification.[1,20-22] Gilman et. al.[1] investigated the perceptions and barriers to seek voice care among contemporary Commercial Music performers. Although most subjects reported that their voice was a critical part of their profession, the lack of available and affordable voice care and education about the importance of voice care proved to be an important barrier to seeking appropriate help for voice problems. The apparent trend of altered attitude towards voice classification as a basic principle of voice education may be seen as a consequence of the changed music scene today. A consequent decrease of care of vocal hygiene and ergonomics may be expected. These findings must have consequences for the prevention and care of voice disorders.

The results of the questionnaires among conservatory singing teachers indicated that, at least in some Classical and Musical Theatre conservatories, voice classification is still an important issue in singing education. However, each singing teacher of each type of conservatory used a varying individual set of voice classification criteria, depending on the singing student and on the specialty of the department. One singing teacher wrote: “I trust my own judgment”, without mentioning which criteria she used and another one claimed: “The voice type is what the singer tells you what he/she is”. Singing students were classified by a combination of two to 11 criteria. The criteria enumerated by the singing teachers were not always the same for each of their students. The criteria could be sorted into 11 major groups. Frequency range/tessitura, voice timbre, quality, color and character were most frequently used as voice classification criteria, followed by register transition, volume. These criteria are also well known as basic voice parameters. Only few singing teachers mentioned the exact dimensions of the frequency range/ tessitura or situated the register transition on the frequency scale. The use of belting techniques in Musical Theatre makes the frequency range of belt an important issue, which is absent in Classical singing education. In some cases, descriptions were given about the nature of the student's voice, its actual development (which could make a classification premature or uncertain), and the ease and comfort of tone production.

While in Classical and in Musical Theatre conservatories voices are still classified according to well known traditional criteria, this is much less the case in private singing education: frequency range/tessitura (100.0% versus 56.0%), quality/timbre (100.0% versus 56.0%), register transition (57.1% versus 9.0%), and volume (52.4% versus 12.1%). Private singing teachers also prefer a more careful testing over time (13.6%) and specific methodologies (12.0%). It is also quite possible that private singing teachers do not feel the need to classify, nor do their singing students feel the need to be classified.

The results of questionnaire 2 must be cautiously interpreted as they only provide an insight in the responses 22 singing teachers at three conservatories in three European countries. Access to conservatories for scientific research is not easy. One classical conservatory and two conservatories specialized in Musical Theatre agreed to cooperate in this study. However, various other departments of these conservatories, which are also involved in singing education, were not willing to cooperate. In the three conservatories, singing voices were still systematically classified. However, during this investigation, spread over one year, only 75.3% of the classical singing students and 88.3% of the Musical Theatre students had been classified.

Conclusions

Next to the traditionally classically oriented music conservatoires many new music education institutions emerged these last years. In the ever changing world of music genres and styles music institutions and singing teachers need to adapt to the new demands of their students. Many formerly classically trained singing teachers are confronted with the demands for new or mixed singing techniques and singing styles. Formally, voice specialists stressed the importance of an exact voice classification before voice education starts. Incorrect voice classification could induce functional and eventually, organic voice disorders. However, the results of this study showed a marked difference in attitude towards voice classification and different criteria for voice classification in various kinds of singing teachers today. A highly subjective approach to voice classification by many singing teachers apparently demonstrated common feelings of uncertainty. Further research is needed to understand the actual trends in the singing world which will have, no doubt, important implications for clinical practice.

References

1. Gilman M, Merati AI, Klein AM, Hapner ER, Johns MM. Performer's attitude toward seeking health care for voice issues: understanding the barriers. J Voice 2009; 23:225-228.
2. Radionoff SL, Satterfield TR, Lee E. Commercial Music: A survey of degree granting institutions in the United Kingdom. Choice for Voice 2008. Multidisciplinary Approaches to Performance, Health and Research in Voice, London: The British Voice Association 2008:1-22.
3. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F, Wyke BD, Eds. Disorders of Human Communication 6. Wien New York:Springer-Verlag; 1982.
4. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209-216.
5. Garde E. La Voix. Paris: Presses Universitaires de France;1954.
6. Lycke H. Stemclassificatie en Fonetogram. Dynamiek tussen morfologie en functie. In: de Jong F, Ed. VOX 2007: Spreken en zingen. Twee werelden apart? Nijmegen: Drukkerij Graficolor, 2007: 209-221.
7. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
8. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Paris:Librairie Maloine SA,1961[2].
9. Thurmer S. The tessiturogram. J Voice 1988;4:327-329.
10. Dejonckere Ph. Le phonétogramme, son intérêt clinique. Les cahiers d'ORL 1977;12: 865-872.
11. Gross M. Stimmfeldmessung, eine objektivere Methode der Stimmdiagnostik. Sprache - Stimme – Gehör 1980;4:100-101.
12. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimmfeld. Sprache - Stimme – Gehör 1985;9:109-111.

13. Pabon JPH, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. J Speech Hear Res 1988;31:710-722.
14. Seidner W, Kröger H, Wernecke KD. Numerische Auswertung spektraler Stimmfelder. Sprache - Stimme – Gehör 1985;9:10-13.
15. Burns P. Clinical management of Country and Western singers. J Voice 1991;5:349-353.
16. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.
17. Sundberg J. Vocal Tract Resonance. In: Sataloff RT, Ed. Professional Voice. The Science and Art of Clinical Care. New York: Raven Press; 1991:49-68.
18. Abitbol J. L'odyssée de la voix. Paris: Editions Robert Laffont SA; 2005.
19. Barthélémy Y. La voix libérée, Une nouvelle technique pour l'art lyrique et la rééducation vocale. Paris: Editions Robert Laffont, SA;1984.
20. Blanchard R. et de Candé R. Dieux et divas de l'opéra. Paris: Librairie Plon;1987.
21. Matheopoulos H. Les Ténors du siècle: de Caruso à Pavarotti. Paris: Editions de la Martinière; 1999.
22. Segond A. Divines divas. Paris: Gallimard; 2002.
23. Bjôrkner E, Sundberg J, Cleveland T, Stone E. Voice source differences between registers in female musical theatre singers. J Voice 2006;20:187-197.
24. Björkner E. Musical theatre and opera singing – why so different? A study of sub-glottal pressure, voice source, and formant frequency characteristics. J Voice 2008; 22:533-540.
25. Hamdam AL, Deeb R, Tohme RA, Rifai H, Husseini S, Fuleihan N. Vocal technique in a group of Middle Eastern singers. Folia Phoniatr Logop 2008;60:217-221.
26. Hanayama EM, Camargo ZA, Tsuji DH, Pinho SMR. Metallic Voice: physiological and acoustic features. J Voice 2009;23:62-70.

Questionnaire 1

Questionnaire

1. Do you think voice classification is important? Yes / No / No opinion

2. Why?
   ……………………………………………………………………….

3. On which criteria is your voice classification based?
   ……………………………………….....................................…

Questionnaire 2

Identification of the subject:

Identification of the singing teacher:

Voice classification:

BASS	.		ALTO	.
BARITONE	.		MEZZO-SOPRANO	.
TENOR	.		SOPRANO	.
?	.		?	.

On what criteria is this voice classification based?

....

....

....

....

....

Addendum 2

Voice classification in practice 2: The classification by conservatory singing teacher and the opinion of singing students

Parts of this study is published in:
Stemclassificatie in de praktijk. Een exploratieve studie
H. Lycke , W. Decoster and F.I.C.R.S. de Jong
Nederlands Tijdschrift voor Stem- Spraak- en Taalpathologie
Accepted for publication

Abstract

Aims: Exploration of voice classification by conservatory singing teachers and how the singing students think about their voice classification by their teachers.
Study design: Explorative study using questionnaires.
Methods: In one questionnaire, 22 singing teachers at one Belgian classical conservatory and two Musical Theatre conservatories, one in the Netherlands and one in the United Kingdom were asked to classify their students (N = 165). In an other questionnaire the classified singing students responded about their voice classification. First year students (n = 73) in the Master's degree programs in speech-language pathology at a Belgian university were used as controls.
Results: Some singing students, 23.4% in classical singing training and 11.9% in Musical Theatre training, reported not to know their voice category. In the control group, this figure was 83.6%. The methodology and the results of voice classification were different according to the type of conservatory. Musical Theatre students were most frequently classified in the middle voice categories, while higher voices were more present in the classical conservatory group. The results of the two questionnaires showed apparently feelings of uncertainty about voice classification among both singing teachers and their students.
Conclusions: The results of this study indicate the hazard of neglecting careful watch on the physiological limits of the voice. Further research is needed to understand the actual trends in the singing world, which can have important implications for future clinical practice.

Introduction

Recent studies pointed at the many differences in singing techniques between Western Classical Opera and Musical Theatre.[1-2] As voice classification has always been a basic principle of voice education[3-9], the question arises if there is a trend of changing attitude towards voice classification.
Lycke et al. explored voice classification in private singing teachers who presented themselves on the commercial Dutch website www.vocalisten.nl (see: Addendum 1 of this thesis). The answers of 72 singing teachers who filled in the questionnaire presented a good example of today's wide variety of available singing specialities, singing techniques, and voice classification. The question: "Is voice classification important to you?" was answered in many different ways. About 61.0% found voice classification an important issue for singing education, while 39.0 % did not.

The singing teachers reported various arguments pro and contra voice classification and various criteria for classification. Obviously, in this internet study there was no consensus about the criteria for voice classification, and there was apparently a trend of altered attitude towards voice classification as a basic principle of voice education.
Furthermore, Lycke et al. assessed voice classification criteria in 165 students by 22 singing teachers of one Classical conservatory and two Musical Theatre conservatories (see: Addendum 1 of this thesis). Voice classification appeared to be an important issue in singing education at these conservatories. However, each singing teacher reported a varying individual set of voice classification criteria, depending on the singing student and on the specialty of the department. Obviously, a marked difference in attitude towards voice classification and different criteria for voice classification between singing teachers of these conservatories was observed.
A correct voice classification has been advocated by many authors. In order to be able to carry through a careful watch on the physiological limits of the voice, both singing teacher and student should be aware of the voice category and should both agree on that. According to Radionoff et al., contemporary commercial music singers often complain that their singing teachers do not understand the vocal styles and demands of a contemporary Commercial Music singer.[10]
The purpose of this study was to investigate the results of voice classification by singing teachers of two different types of conservatories (Classical singing and Musical Theatre) and how the singing students think about their voice classification by their singing teachers.

Methods

This study is part of a larger explorative study on voice classification by the Centre of Excellence for Voice of the Dep. Exp. ORL of the KU Leuven, Belgium.
At three conservatories: one Belgian classical conservatory specialized in Opera, Lied, and Oratorio, and one Dutch and one British conservatory specialized in Musical Theatre a questionnaire (questionnaire A) was distributed among singing teachers via the Head of the Department. The answers to the questionnaires were collected by one of the singing teachers. The 22 singing teachers from three conservatories who cooperated in this study classified a total of 165 singing students: 81 singing students (58 females and 23 males) at the Belgian classic conservatory, 63 singing students (55 females and 8 males) at the Dutch

conservatory (Musical Theatre), and 21 singing students (9 females and 12 males) at the British conservatory (Musical Theatre). The singing students were aged between 18 and 28 years, mean age 21 years. The singing teachers were asked to classify their own singing students.

A second questionnaire (questionnaire B) was distributed among the singing students of all levels of the above indicated three conservatories. A total of 165 singing students, 122 female students and 43 male students, aged between 18 and 28 years, mean age 21 years, filled in the questionnaire about their voice classification. During this investigation, spread over one year, 75.3% of the classical singing students and 88.3% of the Musical Theatre students were classified by their singing teacher. The students who cooperated in this study got enough time to read and to sign an informed consent. The students filled in the questionnaire and handed it over to the investigator. There was also an opportunity to ask more information about the aim of the study. Seventy-three female first year students in the Master's degree programs in speech-language pathology at a Belgian university (aged between 18 and 20 years, mean age 18 years) were used as a control group. Descriptive statistics were performed by the SPSS 16.00.

Results

Questionnaire A

In the classical music conservatory 51.7% of female singing students were classified as soprano, 36.2% as mezzo. There were no singing students classified as alto. In various cases more than one category was filled in, that indicates an apparent doubt. In 5.2% of the students the singing teachers doubted between soprano and mezzo. There were also doubts concerning the differentiation between mezzo and alto (1.7%) and between soprano, mezzo and alto (1.7%), while 3.4% of the female singing students were classified as 'voice category unknown'.

Male singing students were classified as 39.1% tenors and 34.8% as baritones. There were no singing students classified as bass, but in 21.7% of the singing students there was doubt concerning the differentiation between baritone and bass and in 4.3% there was a doubt concerning the differentiation between tenor and baritone (Table 1).

Table 1. Voice classification of students by 12 Classical singing teachers and 10 Musical Theatre singing teachers.

	Classical singing teachers		Musical Theatre singing teachers	
	n	Percent	n	Percent
Soprano	30	51.7	7	10.9
Mezzo	21	36.2	33	51.6
Alto	0	0.0	7	10.9
Soprano/Mezzo	3	5.2	7	10.9
Mezzo/Alto	1	1.7	1	1.6
Soprano/Mezzo/Alto	1	1.7	4	6.3
Voice category unknown	2	3.4	5	7.8
Total	58	100	64	100
Tenor	9	39.1	4	20.0
Bariton	8	34.8	10	50.0
Bass	0	0.0	0	0.0
Tenor/Bariton	1	4.3	5	25.0
Bariton/Bass	5	21.7	1	5.0
Voice category unknown	0	0.0	0	0.0
Total	23	100	20	100

The results of voice classification by the 10 singing teachers of the Musical Theatre conservatories are shown in Table 1. In the Musical Theatre conservatory 10.9% of the singing students were classified as sopranos, 51.6% as mezzos and 10.9% as altos. Singing teachers doubted concerning the differentiation between soprano and mezzo (10.9%), between mezzo and alto (1.6%), and between soprano, mezzo and alto (6.3%), while 7.8% of the female singing students were classified as ‘voice category unknown’. Male singing students were classified as 20.0% tenors and 50.0 as baritones. There were no basses. There was doubt concerning the differentiation between tenor and baritone (25.0%) and between baritone and bass (5.0%).

Questionnaire B

In the group of singing students at the Classical conservatory 19 (23.4%) indicated not to know their voice category. This number was 10 (11.9%) in the Musical Theatre singing students, 61 (83.6%) of the control group reported not to know their voice category. Of the classical singing students 29 (50.0%) reported to be classified as sopranos, 9 (15.6%) as mezzo-sopranos, 6 (10.3%) as altos, 8 (34.8%) as tenors and 10 (43.5%) as baritones. In the Musical Theatre group 11 (17.2%) of the singing students reported to be classified as sopranos, 25 (39.1%)

as mezzo-sopranos, 18 (28.1%) as altos, 11 (55.0%) as tenors and 9 (45.0%) as baritones.

Table 2. Voice classification mentioned by Classical singing students, Musical Theatre singing students and controls.

	Classical singing students		Musical Theatre singing students		Controls	
	n	Percent	n	Percent	n	Percent
Soprano	29	50.0	11	17.2	4	5.5
Mezzo	9	15.6	25	39.1	2	2.7
Alto	6	10.3	18	28.1	6	8.2
Voice category unknown	14	24.1	10	15.6	61	83.6
Total	58	100	64	100	73	100
Tenor	8	34.8	11	55.0		
Baritone	10	43.5	9	45.0		
Bass	0	0.0	0	0.0		
Voice category unknown	5	21.7	0	0.0		
Total	23	100	20	100		

Of the 73 female university speech-pathology students in the control group 12 (16.4%) students who were singing in an university choir were classified by the conductor of the choir. In this control group 4 (5.5%) of the female students were classified as sopranos, 2 (2.7%) as mezzos and 6 (8.2%) as altos (Table 2).

A clinician was consulted for voice classification in 8 (3.4 %) of all 238 subjects (singing students and controls). A total of 28 (34.6%) of classical singing students and 5 (5.9%) of Musical Theatre students classified their own voice. One control, singing in a choir, classified her own voice.
All singing students answered to be subjectively classified by their singing teacher or choir conductor by doing some exercises (e.g. vocalises) during lessons or rehearsals.
Concerning the question: ‘Do you think your voice classification is correct?’, 19 (23.5%) of the Classical singing students, 5 (6.0%) of the Musical Theatre singing students and 6 (8.2%) of the controls were of the opinion that their voice classification was not correct.
Doubts about a correct voice classification was reported by 9 (11.1%) of the Classical singing students, 7 (8.3%) of the Musical Theatre singing students and 6 (8.2%) of the controls.

Discussion

The results of this study cannot be compared to those of other authors because of lack of studies on this subject. The results of both questionnaires must be cautiously interpreted because only 22 singing teachers and 165 singing students at three conservatories (one Classical conservatory and two conservatories specialized in Musical Theatre) in three European countries were enquired.

A former questionnaire sent to private singing teachers revealed that only 61.0% stated that voice classification was an important issue (see: Addendum 1 of this thesis). In the three conservatories which cooperated in this study, however, singing voices were systematically classified. Musical Theatre students were most frequently classified in the middle and lower voice categories. Quite the opposite was seen in the Classical singing conservatory, where the highest categories dominated. An explanation may be that middle voices are more preferred in Musical Theatre, while higher voices are most favoured in Classical conservatories, each attracting in a way its own clientele of singing students. There is also the possibility that each type of conservatory tries to train its own preferred vocal types by specifically adapted singing techniques and gives less thought to a correct voice classification.

The percentage of singing students who could not be classified according to the six basic voice categories (alto-mezzo-soprano for female voices and bass-baritone-tenor for male voices) is remarkably high: 12.0% female and 26.0% male students in the Classical conservatory and 26.6% female and 30.0% male students in the Musical Theatre conservatories. The singing students answered to be subjectively classified by their singing teacher or choir conductor by doing some exercises (e.g. vocalises) during lessons or rehearsals. This indicates that there is no generally accepted protocol for voice classification. Furthermore, many singing students expressed their doubts about voice classification: 17.6% of the singing students did not know their voice category, 9.7% had doubts and 14.5% found that their voice classification was wrong. These high percentages demand serious consideration because voice specialists have always stressed the importance of an exact voice classification before voice education starts. Incorrect voice classification may induce functional and eventually, organic voice disorders. Klingholz states that male singers have less problems with their singing voice than female singers because female vocal folds vibrate two times more than male vocal folds[11]. Moreover, he stated that female voices very often are classified as a voice type which is too high. In a study on “vocal attrition” (vocal pathology and reduced vocal functions associated with behavioral, biogenic, and psychological factors), 62 of the 74 of the university female voice students (84%) said to be sopranos, 11 (15.0%) were mezzo-sopranos and only one student declared to be a contralto. Only 10 (13%) of these singing students proved to be free of symptoms, 19 (25%) had few and 45 (61%) had multiple symptoms.[12] Miller testified: “young singers press for louder and louder and higher and higher sounds, no matter what their bodies can do comfortably and efficiently”.[13] Sataloff stated: “singers are habitually unhappy with the limitations of their voices. In many situations, voice teachers are to blame. Both singer and teacher must resist the impulse to show off the voice in works that are either too difficult for the singer’s level of training or simply not suited to the singer’s voice”.[14] According to McKinney, misclassification can be a major cause of dysfunction in the young adult voice. As “every aspiring young singer knows that the larger incomes are in the high notes, so regardless of statistical evidence that most of them are baritones and mezzo-sopranos, they push for the higher voice classifications quite early.”[15] These observations are corroborated by a study on mechanical stress in phonation by Titze, who found that the largest mechanical stresses in vocal fold vibration are the tensile stresses required for pitch increase.[16]

The majority of the first year students in speech-language pathology did not know their voice category. This could indicate that young people in general are not aware of their own voice classification. Voices seem to be classified only when there is a need to (for instance, when presenting as a member of a choir or in preparation for a (Classical) singing career).

Conclusions

The results of this study give an impression how singing teachers and their students deal with voice classification today. Musical Theatre students were most frequently classified in the middle and lower voice categories, while in Classical singing students the highest categories dominated. However, there is no generally accepted protocol for voice classification and many singing students expressed their doubts about voice classification. Furthermore, a large number of singing students reported not to know their voice category. The results of this study indicate the hazard of neglecting careful watch on the physiological limits of the voice. Further research is needed to understand the actual trends in the singing world, which can have important implications for future clinical practice.

References

1. Bjôrkner E, Sundberg J, Cleveland T, Stone E. Voice source differences between registers in female musical theatre singers. J Voice 2006;20:187-197.
2. Björkner E. Musical theatre and opera singing – why so different? A study of sub-glottal pressure, voice source, and formant frequency characteristics. J Voice 2008; 22:533-540.
3. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F, Wyke B, Eds. Disorders of Human Communication 6. Wien New York:Springer-Verlag;1982.
4. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209-216.
5. Garde E. La Voix. Paris: Presses Universitaires de France;1954.
6. Lycke H. Stemclassificatie en Fonetogram. Dynamiek tussen morfologie en functie. In: de Jong F. Ed. VOX 2007: Spreken en zingen. Twee werelden apart? Nijmegen: Drukkerij Graficolor, 2007:209-221.
7. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
8. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Paris: Librairie Maloine SA,19612.
9. Thurmer S. The tessiturogram. J Voice 1988;4:327-329.
10. Radionoff SL, Satterfield TR, Lee E. Commercial Music: A survey of degree granting institutions in the United Kingdom. Choice for Voice 2008. Multidisciplinary Approaches to Performance, Health and Research in Voice, London:The British Voice Association 2008:1-22.
11. Klingholz F, Martin F, Jolk A. Die Bestimmung der Registerbrüche aus dem Stimm-feld. Sprache - Stimme – Gehör 1985;9:109-111.
12. Sapir S. Vocal attrition in voice students: survey findings. J Voice 1993;7:69-74.
13. Miller R. The Perspective of the Singing Teacher and Voice Pedagogue. In: Scherer,R.C. Ed. The Integration of Voice Science, Voice Pathology, Medicine, Public Speaking, Acting, and Singing. J Voice 1994;8:359-374.
14. Sataloff RT. The Professional Voice: Part III. J Voice 1987;1:283-292.
15. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.
16. Titze IR. Mechanical stress in phonation. J Voice 1994;8:99-105.

Questionnaire A

Identification of the subject:

Identification of the singing teacher

Voice classification:

BASS		ALTO	
BARITONE		MEZZO-SOPRANO	
TENOR		SOPRANO	
Voice category unknown		Voice category unknown	

Questionnaire B

Q1. Do you know your voice classification?

- ❑ No
- ❑ Yes : what is your voice classification?

  - o bass
  - o baritone
  - o tenor
  - o alto
  - o mezzo
  - o soprano

Q2. Who has determined your voice classification?

- o myself
- o singing teacher
- o conductor
- o other :

Q3. How was your voice classification determined?

……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………

Q4. Do you think your voice classification is correct?

- ❑ Yes
- ❑ No

Explain :

……………………………………………………………………………………
……………………………………………………………………………………
……………………………………………………………………………………
…………………………………………..

Addendum 3

Frequency range as a parameter for female voice classification. An explorative study

Is Frequency Range/Tessitura a Reliable Parameter for Voice Classification?

H. Lycke, W. Decoster and F.I.C.R.S. de Jong

Unpublished results

Abstract

Aims: To compare the results of voice classification by singing teachers with the results of voice classification based on an algorithm deduced from the limits of frequency range/ tessitura as found in the literature.
Methods: Sixteen singing teachers of three European conservatories classified 99 of their students. An algorithm, that is based on the limits of the female frequency range, according to 38 authors with different backgrounds, was elaborated.
Results: The voice classification by the singing teachers did not fit in the algorithm.
Conclusions: The results of this study demonstrated that frequency range alone proved to be not suitable as the parameter for voice classification.

Introduction

Traditionally, in Classical Music, female voices are classified into three principal categories: alto, mezzo-soprano, and soprano. Classifying a voice means to determine the frequency and intensity range in which a person can work without harm or fatigue of his voice and to which repertoire he should be assigned.[1-9] Incorrect voice classification may lead to singing beyond the physiological limits of the voice and can consequently induce functional and organic lesions.[7-11] Sapir[12] reports on the great prevalence of voice disorders by incorrect voice classification among singing students, singers and singing teachers. Therefore, voice specialists stress the importance of correct voice classification before the start of voice education.[7-9]
Frequency range has always been a major determining parameter in voice classification, which is described in many specialized writings.[1-9] The results of two recent studies demonstrated that frequency range was used as an important voice classification parameter by 56.% of private singing teachers of different genres and by every Classical and Musical Theatre singing teachers at three European conservatories (Addendum 1 and 2 of this thesis). The frequency range that is most convenient to a voice is called tessitura. Tessitura is not only different for each voice category but also for each individual[8]. In voice classification, frequency range was regularly used in combination with other parameters, such as voice quality/timbre, register transition, intensity. The combination of these parameters is often used together with other characteristics, such as vocal development, speaking voice, ease of voice production, and feeling of the subject. This makes discrimination which parameter is decisive in voice classification difficult.

Moreover, the results of previous studies showed that every single singing teacher used a varying individual set of parameters, dependent not only on the singing student but also on the type of conservatory and education. (Addendum 1 and 2 of this thesis). Information about voice range / tessitura is scarce and diverse. Lycke[6] synthesized the (extreme) limits of voice ranges / tessitura of the three basic female voice types, as mentioned by many authors with different backgrounds (Table 1).

Table 1. Female frequency range / tessitura, according to 38 authors.

Classification	Lowest frequency	Lowest frequency (Hz)	Highest frequency	Highest frequency (Hz)
Soprano	c – c1	131– 262	g2 – c4	784 – 2093
Mezzo-soprano	B - a	123 – 220	f2 – e3	698 – 1319
Alto	A - g	110 – 196	d2 – e3	587 – 1319

According to this synthesis there is a difference of more than one octave between the limits of frequency range. This means that there are many overlaps between the different voice categories, not only between two adjacent voice types, but even between all three basic voice categories. If one tries to classify a voice by ear, based on the limits of frequency range alone, some problems will occur around specific frequencies. Between 123 Hz and 131 Hz there is an overlap between mezzo-soprano and soprano; between 196 Hz and 220 Hz there is an overlap between mezzo-soprano and alto, and between 131 Hz and 196 Hz there is even an overlap between the three basic female voice types. For the highest frequencies, there is an overlap between the three basic female voice types between 698 Hz and 784 Hz and between 784 Hz and 1319 Hz. In order to classify subjects unambiguously, an algorithm has to be created without the many zones of overlap between the limits of frequency range / tessitura.
The purpose of this study was to compare the results of voice classification by singing teachers with the results of voice classification based on an algorithm deduced from the limits of frequency range/ tessitura as found in the literature.

Materials and methods

Sixteen singing teachers (9 female and 3 male singing teachers at one Classical conservatory and 3 female and 1 male teachers at one Musical Theatre conservatory) were asked to classify their female singing students (aged between

18 and 28 years, mean age 20.6 years). The singing teachers filled in a questionnaire about the voice classification of their students (see: questionnaire).
The frequency range of the singing students was determined by the Voice Range Profile (LingWaves Voice Diagnostic Center, version 2.5;2007, with a Center 322 Data Logger Sound Level Meter) according to the methodology proposed by the UEP.[13]
The results of voice classification by the singing teachers were compared with the results of voice classification based on the limits of female frequency range / tessitura, according to 38 authors of different backgrounds (Table 1) and using the following algorithm[14] (Table 2):

Table 2. Algorithm for voice classification based on frequency range / tessitura (cf. Table 1).

Lowest frequency	alto	<	123 Hz	≥	mezzo-soprano	<	131 Hz	≥	soprano
Highest frequency	alto or mezzo-soprano	≤	1319 Hz	>	soprano				

For the data description SPSS 16.00 was used.

Results

Of the 99 female singing students, 86 (86.9%) were classified by their singing teacher. The singing teachers classified 29.0% of their students as sopranos, 46.5% as mezzo-sopranos and 10.5% as altos. Doubts about the voice type were expressed for 14.0% of their singing students.
Table 3 shows the results of voice classification by the singing teachers compared to those of the algorithm, based on the limits of frequency range, as found in the literature.

Table 3. Voice classification of 86 singing students by their singing teacher.

	Singing teacher	Algorithm
Sopranos	25	11
Mezzo-sopranos	40	58
Altos	9	13
Doubts	12	4

The 11 sopranos assigned by the algorithm were classified by the singing teachers as six mezzo-sopranos, four altos and one doubt.
The 58 mezzo-sopranos according to the algorithm were classified by the singing teachers as 23 sopranos, 23 mezzo-sopranos, two altos and 10 doubts.
The 13 altos assigned by the algorithm were classified by the singing teachers as two sopranos, seven mezzo-sopranos, three altos and one doubt.
In the literature 110Hz is mentioned as the lowest female frequency. However, three subjects produced 98 Hz and one subject produced even 87 Hz. In the literature 2093 Hz is mentioned as the highest female frequency. In this study, however, seven singing students produced 1568 Hz as highest frequency.

Discussion

Information about frequency range / tessitura is diverse. According to Tarneaud the normal singing range is along 13 to 14 tones, but sometimes comprises 2 ½ to 3 octaves.[8] Brown et al.[15] found no statistically significant differences in mean phonational range between a group of professional singers of all age groups and an untrained group. Other studies demonstrated that singing training result in a significant increase of frequency range.[16-21]
The results of two previous studies (Addendum 1 and 2) showed that of all voice classification criteria, which were mentioned by singing teachers, frequency range constituted the only parameter that could be measured objectively. Although frequency range was considered as the most important parameter for voice classification, only few singing teachers mentioned the exact dimensions of the frequency range / tessitura of their students.
Singing teachers tended to classify their students in the highest (sopranos) and middle (mezzo-sopranos) voice categories (75.6%) while, according to the algorithm, 82,6% of the same students were classified in the middle and lower voice categories (mezzo-sopranos and altos).
Singing teachers had 14.0% doubts about voice classification of their students versus 4.7% according to the algorithm. Singing teachers classified 29.0% of their singing students as sopranos versus 12.8% according to the algorithm. They classified 46.5% as mezzo-sopranos versus 67,4% according to the algorithm, and 10.5% altos versus 15.1% according to the algorithm. None of the sopranos assigned by the algorithm were classified as sopranos by the singing teachers. Only 27 of the 58 mezzo-sopranos and only three of the 13 altos according to the algorithm were classified as such by the singing teachers. There was even no

consensus about extreme voice types: sopranos according to the algorithm were classified as altos by the singing teachers and vice versa. This huge discrepancy between classification by singing teachers and classification according to the algorithm could be explained by the fact that in this study every singing teacher used his/her own criteria for frequency range, even depending on each individual singing student. Between the two methods of voice classification there was most disagreement concerning the labels sopranos and mezzo-sopranos. This could be explained by the tendency of young singers to produce increasingly higher sounds[22], taking into account that singers are habitually unhappy with the limitations of their voices. According to Sataloff[23], in many situations voice teachers are to blame. There is also the statement of McKinney[24] that every aspiring young singer push for the higher voice classifications quite early, knowing that the larger incomes are in the high notes. Therefore, female voices are frequently classified as a voice category which is too high. The results of this study confirm this statement.

Conclusion

Although voice classification based on frequency range is still very popular among contemporary singing teachers, the results of this study demonstrate that frequency range proved not to be suitable as a voice classification parameter, if used as a single parameter. While frequency range is an objective parameter the combination with other, but subjective criteria makes voice classification an even more subjective issue.

Acknowledgments

Malfroid A., Riedmüller S. Stemclassificatie in de praktijk. Een exploratieve studie. Master Thesis Faculty of Medicine, Logopedische en Audiologische Wetenschappen. KU Leuven, 2009.

References

1. Scotto di Carlo N. Le classement vocal des chanteurs débutants. TIPA 1980;7.
2. Scotto di Carlo N. Le bilan vocal des chanteurs. Folia Phoniatr 1982;34:300-304.
3. Bunch M. Dynamics of the Singing Voice. In: Arnold GE, Winckel F, Wyke B, Eds. Disorders of Human Communication 6. Springer-Verlag, Wien New York: 1982.

4. Coleman RF. Performance demands and the performer's vocal capabilities. J Voice 1987;1:209-216.
5. Garde E. La Voix. Presses Universitaires de France, Paris:1954.
6. Lycke H. Stemclassificatie en Fonetogram. Dynamiek tussen morfologie en functie. In: de Jong F. Ed. VOX 2007: Spreken en zingen. Twee werelden apart? Drukkerij Graficolor, Nijmegen: 2007:209-221.
7. Stalmans R, De Bodt M. Classificatie van stemmen. Logopedie 2003;31-36.
8. Tarneaud J. Traité pratique de phonologie et de phoniatrie. Librairie Maloine SA, Paris: 1961[2].
9. Thurmer S. The tessiturogram. J Voice 1988;4:327-329.
10. Burns P. Clinical management of Country and Western singers. J Voice 1991;5:349-353.
11. Barthélémy Y. La voix libérée, Une nouvelle technique pour l'art lyrique et la rééducation vocale. Editions Robert Laffont, SA, Paris:1984.
12. Sapir S. Vocal attrition in voice students: survey findings. J Voice 1993;7:69-74.
13. Schutte H.K., Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement / phonetography. Folia Phoniatr 1983;35:286-288.
14. Malfroid A., Riedmüller S. Stemclassificatie in de praktijk. Een exploratieve studie. Master Thesis Faculty of Medicine, Logopedische en Audiologische Wetenschappen. K.U.Leuven, 2009.
15. Brown WS Jr, Morris RJ, Hicks DM, Howel E. Phonational Profiles of Female Professional Singers and Nonsingers. J Voice 1993;7:219-226.
16. Akerlund L, Gramming P. Average Loudness Level, Mean Fundamental Frequency, and Subglottal Pressure: Comparison Between Female Singers and Nonsingers. J Voice 1994;8:263-270.
17. Mendes AP, Brown WS Jr, Rothman HB, Sapienza C. Effects of Singing Training on the Speaking Voice of Voice Majors. J Voice 2004;18:83-89.
18. Sulter AM, Schutte HK, Miller DM. Differences in Phonetogram Features Between Male and Female Subjects With and Without Vocal Training. J Voice 1995;9:363-377.
19. Frank F, Donner F. Die Bedeutung der Stimmfeldmessungen für den Gesangs-unterricht aus phoniatrischer und gesangspädagogischer Sicht. Sprache - Stimme - Gehör 1986;10:93-97.
20. Siupsinskiene N. Quantitative analysis of professonally trained versus untrained voices. Medicina 2003;39:1.
21. Siupsinskiene N, Lycke H. Effects of Vocal Training on Singing and Speaking Voice Characteristics in Vocally Healthy Adults and Children Based on Choral and Nonchoral Data. J Voice (in press).
22. Miller R. The Perspective of the Singing Teacher and Voice Pedagogue. In: Scherer RC. Ed. The Integration of Voice Science, Voice Pathology, Medicine, Public Speaking, Acting, and Singing. J Voice 1994;8:359-374.
23. Sataloff RT. The Professional Voice: Part III. J Voice 1987;1:283-292.
24. McKinney J. The singing/acting young adult from a singing instruction perspective. J Voice 1997;11:153-155.

Questionnaire

Identification of the subject:

Identification of the singing teacher:

Voice classification:

BASS	
BARITONE	
TENOR	
Voice category unknown	

ALTO	
MEZZO-SOPRANO	
SOPRANO	
Voice category unknown	

On which criteria is this voice classification based?

………………………………………………………………………………………

………………………………………………………………………………………

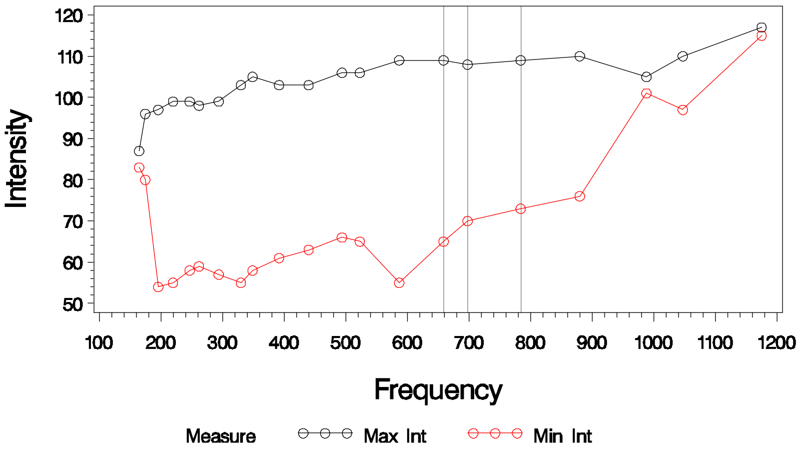
………………………………………………………………………………………

………………………………………………………………………………………

Addendum 4

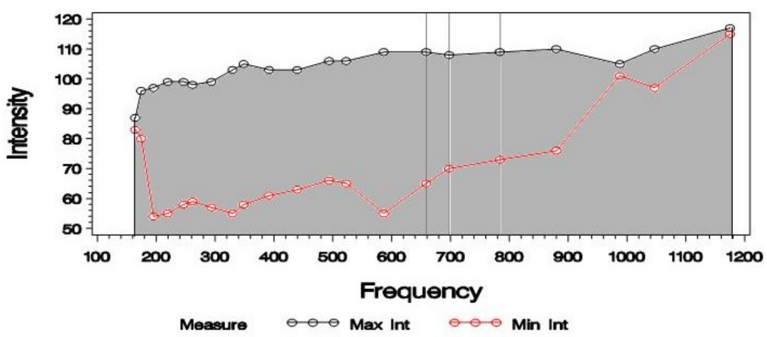
Features females (Chapter 3)

Appendix 1

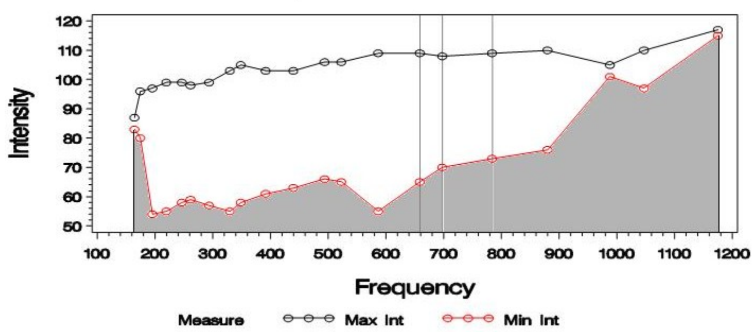


Total Ellipse:

1. Square_upper – surface area of enclosed region, upper part, for max intensity;



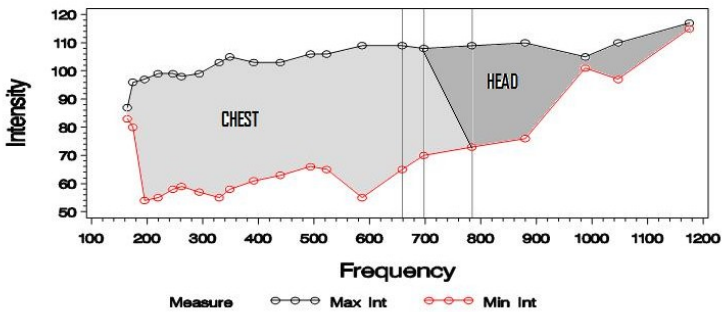
2. Square_lower – surface area of enclosed region, lower part, for min intensity;



3. Square_total – surface area of ellipse;
4. Perim_upper – perimeter of upper part, for max intensity;

5. Perim_lower – perimeter of lower part, for min intensity;
6. Perim_total – perimeter of elipse: Perim_upper + Perim_lower;
7. Eigenvalue1 – the eigenvalue corresponding to the frequency axis by using Principal Component Analysis (PCA);
8. Eigenvalue2 – the eigenvalue corresponding to the intensity axis by using Principal Component Analysis (PCA);
9. Freq_Mean, Intensity_Mean – the centre of ellipse;
10. Max_freq, min_freq – maximum and minimum frequency of the total ellipse;
11. Max_int, min_int – maximum and minimum intensity of the total ellipse;
12. Freq_range – max_freq – min_freq;
13. Intensity_range – max_int – min_int.

Register Transition Zone:
Remark : to divide the chest voice- from the head voice measurements, we used the dip of the transition zone for the max intensity and the maximum value of the min intensity of the transition zone (we call it later: break point). If there was more than 1 maximum value then we took the value with the smallest frequency.



1. Dip_TrZone – intensity of dip of transition zone;
2. Freq_Dip – frequency of transition zone;
3. Break_min_int – intensity of break point;
4. Freq_Break – frequency of break point;
5. Square_TrZone – surface area of ellipse of Transition Zone;
6. Perim_TrZone – perimeter of Transition Zone.

Chest voice, head voice:
1. Square_Chest – surface area of enclosed region, corresponding to the chest voice;
2. Square_Head - surface area of enclosed region, corresponding to the head voice;

3. Perim_Chest – perimeter of upper and lower curvature of ellipse corresponding chest voice;
4. Perim_Head – perimeter of upper and lower curvature of elipse corresponding head voice.

Slopes:

1. Slope_MaxInt - Slope of the regression line for maximum intensity;
2. Slope_MinInt - of the regression line for minimum intensity;
3. Tang_f0_f05 – the tangent of the line going through the points f_0 and f_0^{5th};

   Remark: For example: If the measurements of max intensity of the subjects begin with A (110 Hz), then the frequency that corresponds to the point f_0^{5th} is e (165 Hz).

Ratios o f traditional voice frequency and intensity parameters:

1. R1 - ratio surface area chest voice/total surface area;
2. R2 - ratio surface area head voice/total surface area;
3. R3 - ratio surface area transition zone voice/total surface area;
4. R4 – ratio perimeter chest voice/perimeter total;
5. R5 – ratio perimeter head voice/perimeter total;
6. R6 – ratio perimeter transition zone/perimeter total;
7. R7 – ratio perimeter head voice/perimeter chest voice;
8. R8 – ratio perimeter transition zone/perimeter chest voice;
9. R9 – ratio perimeter transition zone/perimeter head voice;

   If the value of perimeter of head voice is 0, we assume the value of ratio to be missing;
10. R10 – ratio surface area head voice/ surface area chest voice;
11. R11 – ratio surface area transition zone/ surface area chest voice;
12. R12 – ratio surface area transition zone/ surface area head voice;

    Remark: If the value of the surface area of the head voice is 0, then we assume that the value of ratio R12 is missing;
13. R13 – ratio maximum frequency/minimum frequency total;
14. R14 – ratio maximum intensity/minimum intensity total;
15. R15 – ratio maximum frequency/minimum frequency of chest voice;
16. R16 - ratio maximum intensity/minimum intensity of chest voice;
17. R17 – ratio maximum frequency/minimum frequency of head voice;
18. R18 - ratio maximum intensity/minimum intensity of head voice;

Remark: If the value of the surface area and perimeter of the head voice for a certain subject equals 0, then we assume that the values of the ratios R17 and R18 are missing;

19. R19 - ratio maximum frequency/minimum frequency of transition zone;
20. R20 - ratio maximum intensity/minimum intensity of transition zone;

    Remark: We assumed that the transition zone exists for every subject, even if it is a single point;
21. R21 – ratio Eigenvalue2/Eigenvalue1 of total ellipse.

Differences of traditional voice frequency and intensity parameters:

1. Diff1 – difference in frequency between dip of transition zone and break point;
2. Diff2- difference in intensity between dip of transition zone and break point;

Remark: Since we assumed that transition zone exist for every subject, even if it is a single point, we can calculate always the value of Diff1 and Diff2;

Appendix 2. Redundant features (features with correlation of 95% or higher).

Square_upper	Square_lower	Perim_upper	Perim_lower	Perim_total	Eigenvalue1	Perim_head	Perim_trzone	R2	R7
Square_lower Perim_upper Perim_lower Perim_total Eigenvalue_1 Perim_head	Square_upper Perim_upper Perim_lower Perim_total Eigenvalue1 Perim_head	Square_upper Square_lower Perim_lower Perim_total Eigenvalue1 Perim_head	Square_upper Square_lower Perim_upper Perim_total Eigenvalue1 Perim_head	Square_upper Square_lower Perim_upper Perim_lower Eigenvalue1 Perim_head	Square_upper Square_lower Perim_upper Perim_lower Perim_total Perim_head	Square_upper Square_lower Perim_upper Perim_lower Perim_total Eigenvalue1 R7 R17	R19	R10	Perim_head R17

R8	R10	R17	R19
R19	R2	Perim_head R7	Perim_trzone R8

Features females

Appendix 3. 3 clusters solution

- Single feature
- Cluster solution for feature R4.

Variable	TotStd	WithinStd	RSquare	RSqRatio
r4	1.00000	0.32120	0.897836	8.788185
OVER-ALL	1.00000	0.32120	0.897836	8.788185

- Combination of two features
- Cluster solution for features R4, perim_head.

Variable	TotStd	WithinStd	RSquare	RSqRatio
r4	1.00000	0.36827	0.865700	6.445993
perim_head	1.00000	0.42040	0.824989	4.713910
OVER-ALL	1.00000	0.39520	0.845344	5.465965

- Combination of three features
- Cluster solution for features R4, perim_head, R9.

Variable	TotStd	WithinStd	RSquare	RSqRatio
r4	1.00000	0.35212	0.877220	7.144679
perim_head	1.00000	0.48434	0.767699	3.304754
r9	1.00000	0.46166	0.789221	3.744295
OVER-ALL	1.00000	0.43548	0.812280	4.327076

Appendix 4. : 3 clusters solution

- Single feature

Feature	Over-All RSquare
r4	**0.898**
r21	0.896
r2	0.890
Perim_Head	0.854
Square_Head	0.854
r9	0.842
Square_TrZone	0.841
r16	0.840
diff2	0.834
r12	0.825
r14	0.823
Break_min_int	0.823
freq_mean	0.823
Dip_TrZone	0.823
eigenvalue2	0.823
r20	0.821
r18	0.820
square_total	0.812
Square_Chest	0.807
Perim_Chest	0.805
tang_f0_f05	0.804
Freq_break	0.803
Freq_Dip	0.796
intensity_mean	0.792
r15	0.788
r11	0.786
r3	0.778
diff1	0.757

Features females

Feature	Over-All RSquare
r8	0.757
r6	0.754
r13	0.751
Perim_TrZone	0.747
slope_maxint	0.724
slope_minint	0.706

- Combination of two features

Combination of features	Over-All Rsquare
r4,Perim_Head	**0.845**
r2,Square_Head	0.821
r4,r12	0.820
r4,r2	0.804
r16,r14	0.804

Details:

Combination of features	**Variable**	**RSquare**
r4,Perim_Head	**r4**	**0.866**
r4,Perim_Head	**perim_he**	**0.825**
r4,Perim_Head	**OVER-ALL**	**0.845**
r2,Square_Head	r2	0.809
r2,Square_Head	square_h	0.833
r2,Square_Head	OVER-ALL	0.821
r4,r12	r4	0.802
r4,r12	r12	0.841
r4,r12	OVER-ALL	0.820
r4,r2	r4	0.781
r4,r2	r2	0.827
r4,r2	OVER-ALL	0.804
r16,r14	r16	0.815
r16,r14	r14	0.793
r16,r14	OVER-ALL	0.804

- Combination of three features

Combination of features	Over-All RSquare
r4,r9,Perim_Head	**0.812**
r4,Perim_Head,freq_mean	0.793
r4,freq_mean,r9	0.782
Perim_Head,r12,r4	0.781

Details:

Combination of features	Variable	RSquare
r4,r9,Perim_Head	**r4**	**0.877**
r4,r9,Perim_Head	**r9**	**0.789**
r4,r9,Perim_Head	**perim_head**	**0.768**
r4,r9,Perim_Head	**OVER-ALL**	**0.812**
r4,Perim_Head,freq_mean	r4	0.844
r4,Perim_Head,freq_mean	perim_head	0.833
r4,Perim_Head,freq_mean	Freq_Mean	0.702
r4,Perim_Head,freq_mean	OVER-ALL	0.793
r4,freq_mean,r9	r4	0.840
r4,freq_mean,r9	Freq_Mean	0.713
r4,freq_mean,r9	r9	0.796
r4,freq_mean,r9	OVER-ALL	0.782
Perim_Head,r12,r4	perim_head	0.735
Perim_Head,r12,r4	r12	0.746
Perim_Head,r12,r4	r4	0.857
Perim_Head,r12,r4	OVER-ALL	0.781

Addendum 5

Features males (Chapter 4)

Data

The final data contains 256 subjects that we use for the investigation.
Parameters:
The parameters for the analysis were divided into 2 groups:

- clinical parameters: the ones that can be easily clinically interpreted;
- intelligent parameters (intelligent features).

Appendix 1.

Clinical parameters:

1. ***Freq_Mean, Intensity_Mean –*** the average frequency and intensity of the subject.
2. ***Freq_max, freq_min*** – maximum and minimum frequency of the subject;
3. ***Intensity_max, Intensity_min*** – maximum and minimum intensity of the subject;
4. ***Freq_range*** = ***max_freq – min_freq;***
5. ***Intensity_range*** =***Intensity_max – Intensity_min;***
6. ***Dip_TrZone*** – intensity of the dip of the transition zone;
7. ***Freq_Dip*** – frequency of the transition zone;
8. ***Break_min_int*** – intensity of the break point;
9. ***Freq_Break*** – frequency of the break point;
10. ***Mean_freq_TrZone*** – mean frequency of the transition zone.

Intelligent parameters:
Total Ellipse:

1. ***Square_upper*** – surface area of the enclosed region, upper part, for max intensity;
2. ***Square_lower*** – surface area of the enclosed region, lower part, for min intensity;
3. ***Square_total*** – surface area of the ellipse
4. ***Perim_upper*** – perimeter of the upper part, for max intensity;
5. ***Perim_lower*** – perimeter of the lower part, for min intensity;
6. ***Perim_total*** – perimeter of the ellipse: ***Perim_upper + Perim_lower***
7. ***Eigenvalue1*** – the eigenvalue corresponding to the frequency axis obtained by using a Principal Component Analysis (PCA);
8. ***Eigenvalue2*** – the eigenvalue corresponding to the intensity axis obtained by using a Principal Component Analysis (PCA);

Transition Zone:

Remark 2: In order to divide the chest voice- from the head voice measurements, we used the dip of the transition zone for the max intensity and the maximum value of the min intensity of the transition zone (we call it later: ***break point***). If there was more than 1 maximum value then we took the value with the smallest frequency.

1. ***Square_TrZone*** – surface area of the ellipse of the ***Transition Zone***;
2. ***Perim_TrZone*** – perimeter of the ***Transition Zone.***

Chest voice, head voice:

1. ***Square_Chest*** – surface area of the enclosed region, corresponding to the chest voice;
2. ***Square_Head*** - surface area of the enclosed region, corresponding to the head voice;
3. ***Perim_Chest*** – perimeter of the upper and lower curvature of the ellipse corresponding to the chest voice;
4. ***Perim_Head*** – perimeter of the upper and lower curvature of the ellipse corresponding to the head voice

Ratios of existing features:

1. ***R1 -*** ratio of the surface area of the chest voice/total surface area;
2. ***R2 -*** ratio of the surface area of the head voice/total surface area;
3. ***R3 -*** ratio of the surface area of the transition zone voice/total surface area;
4. ***R4*** – ratio of the perimeter of the chest voice/perimeter total;
5. ***R5*** – ratio of the perimeter of the head voice/perimeter total;
6. ***R6*** – ratio of the perimeter of the transition zone/perimeter total;
7. ***R7*** – ratio of the perimeter of the head voice/perimeter chest voice;
8. ***R8*** – ratio of the perimeter of the transition zone/perimeter chest voice;
9. ***R9*** – ratio of the perimeter of the transition zone/perimeter head voice;
   If the value of the perimeter of the head voice is 0, then we assume that the value of the ratio is missing.
10. ***R10*** – ratio of the surface area of the head voice/surface area chest voice;
11. ***R11*** – ratio of the surface area of the transition zone/surface area chest voice;
12. ***R12*** – ratio of the surface area of the transition zone/surface area head voice;
    Remark: If the value of the surface area of the head voice is 0, then we assume that the value of ratio ***R12*** is missing.

The ratios R13 and R14 are based on the clinical parameters:

13. ***R13*** – ratio of the maximum frequency/minimum frequency total;
14. ***R14*** – ratio of the maximum intensity/minimum intensity total;
15. ***R15*** – ratio of the maximum frequency/minimum frequency of the chest voice;
16. ***R16*** - ratio of the maximum intensity/minimum intensity of the chest voice;
17. ***R17*** – ratio of the maximum frequency/minimum frequency of the head voice;
18. ***R18*** - ratio of the maximum intensity/minimum intensity of the head voice;
    Remark: If the value of the surface area and the perimeter of the head voice for a certain subject equals 0, then we assume that the values of the ratios ***R17*** and ***R18*** are missing.
19. ***R19*** - ratio of the maximum frequency/minimum frequency of the transition zone;
20. ***R20*** - ratio of the maximum intensity/minimum intensity of the transition zone;
    Remark: We assumed that the transition zone exists for every subject, even if it is a single point.
21. ***R21*** – ratio ***Eigenvalue2***/***Eigenvalue1*** of the total ellipse.

Differences of existing features:

1. ***Diff1*** – difference in frequency between the dip of the transition zone and the break point.
2. ***Diff2***- difference in intensity between the dip of the transition zone and the break point.
   Remark: Since we assumed that the transition zone exists for every subject, even if it is a single point, we can always calculate the values of ***Diff1*** and ***Diff2***.

Slopes:

4. ***Slope_MaxInt*** - Slope of the regression line for maximum intensity.
   1. ***Slope_MinInt*** - of the regression line for minimum intensity.
   2. ***Tang_f0_f05*** – the tangent of the line going through the points f_0 and f_0^{5th}.
      Remark : We have only two values to estimate the regression line → it is better to use the angle of the tangent.

Appendix 2. Redundant features (features with correlation of 95% or higher).

Freq_max	**Freq_range**	**Freq_mean**	**Intensity_range**	**Square_upper**	**Square_lower**	**Perim_upper**
Freq_range	Freq_max	Freq_max	R14	Freq_range	Freq_range	Freq_range
Freq_mean	Square_upper			Freq_max	Freq_max	Freq_max
Square_upper	Square_lower			Perim_upper	Perim_upper	Square_upper
Square_lower	Perim_upper			Perim_lower	Perim_lower	Square_lower
Perim_upper	Perim_lower			Perim_total	Perim_total	Perim_lower
Perim_lower	Perim_total			Eigenvalue1	Eigenvalue1	Perim_total
Perim_total	Eigenvalue1					Eigenvalue1
Eigenvalue1	Perim_head					Perim_head

Perim_lower	Perim_total	Eigenvalue1	Perim_trzone	Perim_head	R1	R2
Freq_range	Freq_range	Freq_range	R19	Freq_range	R2	R1
Freq_max	Freq_max	Freq_max		Perim_upper		
Square_upper	Square_upper	Square_upper		Perim_lower		
Square_lower	Square_lower	Square_lower		Perim_total		
Perim_upper	Perim_upper	Perim_upper				
Perim_total	Perim_lower	Perim_lower				
Eigenvalue1	Eigenvalue1	Perim_total				
Perim_head	Perim_head					

R4	R5	R7	R8	R11	R14	R16
R5	R4	R17	R11	R8	Intensity_range	R14
					R16	

R17	R19
R7	Perim_trzone

Summary

The vocal capacity must be gauged for correctly assessing the possibilities and impossibilities of the voice to optimize its performance and to avoid damage. Voice classification characterizes the vocal capacity of a singer and composers of vocal music write repertoires in accordance with the possibilities of the voice of the targeted singer. Traditionally, in classical music, voices are classified into three principal categories: for the female voice alto, mezzo-soprano, and soprano, and for the male voice bass, baritone, and tenor.

From clinical experience and our exploratory studies (**Addendum 1 and 2**) it became clear that there exists no generally accepted method for voice classification. There is a marked difference in attitude towards voice classification by various singing teachers and there is no consensus about what voice parameters to use. This calls for an objective method for voice classification.

Frequency range is considered to be an important factor in voice classification, as shown in many textbooks. In another exploratory study (**Addendum 3**), however, the frequency range by itself proved not to be panacea for voice classification.

At this point, one may rightfully question the existence of three basic female and male voice types by nature. In an attempt to break out of the controversy, a new perspective is adopted in this study by letting the data speak for itself. Such an approach, called data-driven, imposes minimal assumptions on the nature of the data, what elements to use for its analysis, and in our case even the existence of natural voice groups.

In a pilot study (**Chapter 2**), Voice Range Profile (VRP) - derived parameters that are commonly applied and easily understandable in clinical practice are used (i.e. highest frequency, lowest frequency, maximum intensity, minimum intensity, and frequency when entering and exiting the register transition zone). In order not to favor any frequency range in the analysis, the frequency axis was transformed from the nonlinear note scale into a linear frequency scale (Hz). The data from 327 female students from different conservatories between 18 and 25 years were analyzed. The aim of this study was to verify the existence of individual- or combinations of commonly used and easily understandable VRP parameters with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female voice categories.

We used two complementary clustering procedures: Ward's minimum variance method, to determine the number of clusters, and K-means clustering, to assign

the subjects to the clusters. After applying Ward's procedure, we concluded that there are *possibly* two or three clusters in the data as the measure of cluster separation -- overall R-squared -- was very low in both cases.

However, it could very well be that the necessary information is not present in the considered parameters to discriminate natural voice groups. There could exist other, more "intelligent" parameter ("feature"*)* combinations, or even non-linear ones, of the voice range profile, that lead to clear clusters.

These results have lead to further studies, which form the core of this thesis (**Chapter 3 and 4**). The aim of these studies was to verify the existence of individual- or combinations of not commonly used and not easily understandable clinical VRP parameters (also called "features") with which the data can be partitioned into a number of clearly separated clusters as a basis for discriminating between basic female and male voice categories. In these studies less easily clinically understandable parameters were used, derived from

- the *geometry of the VRP* such as the surface area enclosed between the maximum and minimum intensity curves, their frequency ranges, and their perimeter lengths;
- the *register transition zone* such as the intensity of the dip in the maximum/minimum intensity curves between the chest and head voice parts of the VRP and the frequency at which it occurs;
- the *geometry of the chest/head voice parts* of the VRP such as their surface areas and their perimeter lengths;
- the *linear characteristics of the minimum and maximum intensity curves* such as the slopes of the regression lines through the maximum and minimum intensity curves.

Additionally, a number of voice frequency and intensity ratios and differences were defined based on some of the above parameters such as the ratio of the surface area of the chest voice to the total surface area enclosed by the maximum and minimum intensity curves. Another example is the ratio of the perimeter length of the chest voice part of the VRP to the total perimeter length.

The data from 206 female conservatory singing students (18 - 25 years) and 256 male subjects (18 – 52 years), consisting of 9 young singing students, 17 professional singers, 61 professional choir singers and 169 with and without singing experience, was investigated.

Statistical Analysis and Methods

After preprocessing the data we applied Ward's minimum method to assess whether the data displays any natural clusters (groupings), this means without assuming their number and without using any prior voice classification (cf. data-driven approach).

However, Ward's method could not be decisive as it could return more than one statistically plausible cluster solution. In order to break the tie, we need an additional method. Furthermore, we want to identify which parameter (or a small set of them) is **crucial** for discriminating between the voice clusters. We decided to use K-means clustering in combination with a selection procedure (i.e., forward or backward feature selection) to define the discriminative parameters and the cluster migration index to decide which cluster solution is more consistent across discriminative parameter combinations identified, and adopt that as the final cluster solution.

The analyses were performed using Statistical Analysis SAS/STAT ® software (release 9.2).

Results

At first, the female voices were examined (**Chapter 3**). Ward's procedure indicated that there could be three or four clusters in the data. In using K-means clustering, both a forward and backward feature selection procedure was applied to both clustering options. Based on the migration index, the three-cluster solution turned out to be the most consistent one. The parameter that led to the best three-cluster separation was the ratio of the perimeter length of the chest voice part of the voice range profile versus the total perimeter length. In the case of this single variable, the overall R-squared is equal to the partial R-squared, and it is larger than 80% which is an indication of a high cluster separation degree.

Secondly, male voices were examined (**Chapter 4**). Again, Ward's procedure indicated that there could be three or four clusters in the data. Because of the numerous options to combine parameters, only a backward selection procedure was applied in combination with K-means clustering. Based on the migration index, also for the male voices, the three-cluster solution turned out to be the most consistent one across all parameter combinations. The parameter that led to the best three-cluster separation in the male voices was the frequency of the register dip. In the case of this 1 (single) parameter, the overall R squared is equal to the partial R squared, and it is almost 90% which is an indication of a degree of high cluster separation.

Conclusions

The results of this study demonstrate that different parameters of the VRP are able to yield a clear separation into three clusters for each gender. Such a result is remarkable, since this may not be expected from biological variables. One can wonder if ancient composers of vocal music had an innate feeling about the existence of three natural basic human voice categories. Most “intelligent” parameters that have lead to the cluster separation, however, are not easily understandable in clinical terms. Therefore, it is not easy to link them to the clinical situation, nor can the difference between these parameters that have lead to the clustering of female and male voices readily be explained. A second salient result of this study is the finding that each of these features has to do with register transition.

Clinical applications

The results of the pertinent study may provide a basis for settling the issue of voice classification.

Future research. Further studies are necessary to link the three statistically obtained clusters to the three basic female and male voice categories as commonly interpreted by most composers of vocal music and singing teachers.

Investigation of the difference between the parameters that have lead to clustering of female and male voice.

Samenvatting

De vocale capaciteit moet worden ingeschat om de mogelijkheden en beperkingen van de stem correct vast te stellen, teneinde de performance te optimaliseren en stemschade te voorkomen. Stemclassificatie typeert de stemcapaciteit van de zanger en componisten van vocale muziek schrijven repertoires in overeenstemming met de mogelijkheden van een welbepaalde zanger. Traditioneel worden in de klassieke muziek stemmen geclassificeerd in drie hoofdcategorieën: voor de vrouwelijke stem alto, mezzo-sopraan en sopraan, en voor de mannelijke stem bas, bariton en tenor.

Vanuit klinische ervaring en uit onze exploratieve studies (**Addendum 1 & 2)** werd duidelijk dat er geen algemeen aanvaarde methode voor stemclassificatie bestaat. Er is een duidelijk onderscheid in attitude ten opzichte van stemclassificatie bij verschillende zangpedagogen en er is geen consensus omtrent de keuze van de te gebruiken parameters. Deze vaststelling vraagt om een objectieve methode voor stemclassificatie.

De frequentiespanne van de stem wordt beschouwd als een belangrijke factor bij stemclassificatie, zoals weergegeven in vele handboeken. Een andere exploratieve studie **(Addendum 3)** wijst er op dat de frequentiespanne van de stem op zichzelf geen universele oplossing biedt voor stemclassificatie.

Aldus kan men zich terecht vragen stellen omtrent het bestaan in de natuur van drie basis stemsoorten bij man en vrouw. In een poging om uit de controverse te geraken werd in deze studie uitgegaan van een nieuw perspectief door de data voor zichzelf te laten spreken. Een dergelijke benadering, data driven genoemd, stelt minimale veronderstellingen voorop betreffende de aard van de data, welke de te gebruiken elementen zijn voor de analyse ervan en, in ons geval, zelfs het bestaan van natuurlijke stemgroepen.

In een pilootstudie **(Hoofdstuk 2)** werden uit het fonetogram afgeleide parameters gebruikt die algemeen toegepast worden en die in de klinische praktijk eenvoudig te begrijpen zijn (namelijk hoogste frequentie, laagste frequentie, maximum intensiteit, minimum intensiteit en de frequentie bij het binnenkomen en het verlaten van de registerovergangszone). Teneinde in de analyse geen enkele frequentie te bevoordelen werd de frequentie-as omgezet vanuit de niet-lineaire toonhoogteschaal in een lineaire frequentieschaal (Hz). De data van 327 vrouwelijke zangstudenten, tussen 18 en 25 jaar, uit verschillende conservatoria werden geanalyseerd. Het doel van de studie was na te gaan of de representatie van iemands stem in de ruimte bepaald door deze algemeen toegepaste klinische frequentie-intensiteit gerelateerde parameters van de stem toelaten een onderscheid te maken tussen drie basis vrouwelijke stemcategorieën. We gebruikten twee complementaire clustering procedures: Ward’s minimum variance method teneinde het aantal clusters te bepalen en K-means clustering teneinde de personen aan de clusters te koppelen. Na het toepassen van de Ward's procedure concludeerden we dat er zich mogelijks twee of drie clusters in de data bevonden, aangezien de maat van clusterscheiding– overall R-kwadraat- in beide gevallen zeer laag was.

Het kon echter zeer goed mogelijk zijn dat de nodige informatie om een onderscheid te maken tussen natuurlijke stemgroepen niet aanwezig is in de gebruikte parameters. Er konden misschien andere, meer "intelligente" parameter ("feature") combinaties, of zelfs niet-lineaire, van het fonetogram bestaan die konden leiden tot drie clusters.

Deze resultaten gaven aanleiding to verdere studies die de kern van deze thesis vormen **(Hoofdstuk 3 en 4)**. De doelstelling van deze studies was vast te stellen of individuele of combinaties van parameters uit het fonetogram in staat zijn om tot een duidelijke clusterafscheiding te komen waarmee drie basis vrouwelijke of mannelijke stemgroepen kunnen onderscheiden worden.

In deze studies werden minder eenvoudig te begrijpen klinische parameters gebruikt, afgeleid uit de *geometrie van het fonetogram* zoals de oppervlakte omsloten door de maximum en minimum intensiteitcurven, hun frequentie-omvang en hun lengte-perimeter; de *registerovergangszone* zoals de intensiteit van de dip in de maximum/minimum intensiteitcurven tussen het borst-/falsetstem gedeelte van het fonetogram en de frequentie waarop dit plaats vindt; de *geometrie van het borst-/falsetstemgedeelte* van het fonetogram zoals hun oppervlakten en hun lengteperimeters; de *lineaire karakteristieken van de minimum en maximum intensiteitcurven* zoals de hellingen van de regressielijnen door de maximum en minimum intensiteitcurven. Aanvullend werden een aantal frequentie- en intensiteits ratio's en verschillen van de stem gedefinieerd, gebaseerd op enkele van de bovenvermelde parameters zoals de ratio van de oppervlakte van de borststem t.o.v. de totale oppervlakte omsloten door de maximum en minimum intensiteitcurven. Een ander voorbeeld is de ratio van de perimeter lengte van het borststemgedeelte in het fonetogram ten opzichte van de totale perimeter lengte.

De data van 206 vrouwelijke conservatorium zangstudenten (18-25 jaar) en 256 mannelijke subjecten (18-52 jaar), bestaande uit 9 jonge zangstudenten, 17 professionele zangers, 61 professionele koorzangers en 169 personen met en zonder zangervaring, werden geanalyseerd.

Statische analyse en methodiek

Na de preprocessing van de data werd Ward's minimum methode toegepast teneinde vast te stellen of de data enige natuurlijke clusters (groeperingen) vertoonden, dus, zonder hun aantal te veronderstellen en zonder gebruik te maken van enige voorafgaande stemclassificatie (conform data-driven benadering).

De methode van Ward bleek echter niet afdoende doordat zij meer dan een statistisch plausibele clusteroplossing opleverde. Om de knoop door te hakken is er een aanvullende methode vereist. Bovendien wensen wij te identificeren welke parameter (of een kleine set parameters) **cruciaal** is voor de discriminatie tussen de stemclusters. We beslisten om K-means clustering te gebruiken in combinatie met een selectieprocedure (namelijk voorwaartse of achterwaartse feature selectie) teneinde de discriminerende parameters en de cluster migratie-index te laten beslissen welke clusteroplossing meer consistent is tussen de geidentificeerde discriminerende parameter combinaties, en deze als finale clusteroplossing te aanvaarden.

Resultaten

Vooreerst werden vrouwenstemmen onderzocht (**Hoofdstuk 3)**. Ward's procedure toonde aan dat er zich drie of vier clusters in de data bevonden. Door middel van K-means clustering werden zowel een voor-, alsook een achterwaartse feature selectieprocedure toegepast voor beide clusteropties. Gebaseerd op de migratie-index, bleek de drie clusteroplossing de meest consistente. De parameter die leidde tot de beste drie clusterscheiding was de ratio van de perimeter lengte van het borststem gedeelte van het fonetogram en de totale perimeter lengte. In het geval van deze enkelvoudige variabele is het totale R-kwadraat gelijk aan het partiële R-kwadraat en is groter dan 80%, hetgeen een indicatie is voor een hoge cluster separatiegraad.

Nadien werden de mannenstemmen onderzocht **(Hoofdstuk 4).** Ward's procedure toonde opnieuw aan dat er zich drie of vier clusters konden bevinden in de data. Omwille van de talrijke opties om parameters te combineren werd enkel een achterwaartse selectieprocedure toegepast in combinatie met K-means clustering. Gebaseerd op de migratie-index, bleek, ook voor de mannenstemmen, de drie clusteroplossing de meest consistente te zijn onder alle parameter combinaties. De parameter die tot de beste drie clusterscheiding leidde bij de mannenstemmen was de frequentie van de registerdip. In het geval van deze enkelvoudige parameter is R-kwadraat gelijk aan het partiële R-kwadraat en is bijna 90%, hetgeen een indicatie is voor een hoge cluster separatiegraad.

Conclusies

De resultaten van deze studie tonen aan dat verschillende parameters van het fonetogram in staat zijn een duidelijke scheiding in drie clusters te bekomen voor beide geslachten. Een dergelijk (discontinu) resultaat is opmerkelijk aangezien dit niet verwacht wordt bij biologische variabelen. Men kan zich afvragen of de oude componisten van vocale muziek een ingeboren aanvoelen hadden over het bestaan van drie natuurlijke basiscategorieën van de menselijke stem. De meeste "intelligente" parameters die geleid hebben tot de clusterscheiding zijn echter niet eenvoudig te begrijpen in klinische termen. Om die redenen is het niet eenvoudig hen te koppelen aan de klinische situatie, noch kan het verschil tussen deze parameters die geleid hebben tot de clustering van vrouwen- en mannenstemmen eenvoudig verklaard worden. Een tweede opvallend resultaat van deze studie is de vaststelling dat elk van deze features te maken heeft met de registerovergang.

Klinische toepassingen

De resultaten van deze studie kunnen mogelijk een basis vormen voor het oplossen van het probleem van de stemclassificatie.

Toekomstig onderzoek

Verdere studies zijn noodzakelijk teneinde de drie statistisch bekomen clusters te koppelen aan de drie basis stemcategorieën van mannen en vrouwen, zoals die gewoonlijk geïnterpreteerd worden door de meeste componisten van vocale muziek en zangpedagogen.

Onderzoek naar het verschil tussen de parameters die geleid hebben tot de clustering van vrouwen- en mannenstemmen.

Curriculum vitae

Hugo Lycke was born in Ghent in 1938.
He wrote his master thesis as a speech and voice therapist on the ‘Dysodia - functional disorders of the singing voice’ in 1963. He strongly emphasized the need for further study on phonetography as an adequate tool for voice classification.
He obtained master degrees in Leisure Studies, Theatre Sciences, Dance Research and Biomedical Sciences (Gerontology) at the Free University of Brussels and gave conferences and workshops on the singing voice in 12 European countries and in the USA.
He was founder/director of three Rehabilitation Centers, a private Center for Creative Communication (including a record company) and a service-residence for elderly people.
As a vocal coach he worked in conservatoria and in opera and musical theatre companies, evaluating thousands of singers, actors and dancers during auditions in Belgium and abroad. He performed more than thousand phonetograms from all kinds of performing artists and specialized in voice therapy for singers and actors.
Since July 2009 he happily lives in France with his wife, owning a “maison d’hôtes” (“Maison Conchette” in Thiers), gives workshops on voice and runs a practice at home for vocal coaching.
In 2006 he started his PhD study on voice classification (Promoter: Prof. Dr. Felix de Jong, Co-promoteres: Prof. dr. Wivine Decoster en Prof. dr. Marc Van Hulle).

List of publications

Lycke H, Decoster W, Ivanova A, Van Hulle MM, de Jong FICRS. Discrimination of Three Basic Female Voice Types in Female Singing Students by Voice Range Profile-Derived Parameters. Folia Phoniatr Logop 2012;64:80–86.
DOI: 10.1159/000337042.

Lycke H, Ivanova A, Van Hulle MM, Decoster W, de Jong FICRS. Discrimination of Three Basic Male Voice Types by Voice Range Profile-Derived Parameters. Folia Phoniatr Logop (Accepted for publication by Folia Phoniatr Logop).

Lycke H, Decoster W, de Jong FICRS. Stemclassificatie in de praktijk. Een explorative studie. Nederlands Tijdschrift voor Stem-, Spraak- en Taalpathologie. (Accepted for publication).

Šiupšinskienė N, Lycke H. Effects of Vocal training on Singing and Speaking Voice Characteristics in Vocally Healthy Adults and Children Based on Choral and Nonchoral Data. Journal of Voice 2011;25/4:177-189.

Šiupšinskienė N.; Lycke, H. Singing voice quality measured by spectral and voice range profiles in vocally healthy trained and untrained voice adults and children. Sveikatos mokslai=Health sciences. ISSN 1392-6373. 2010, t. 20, Nr. 5(71), p. 3458-3464. [Index Copernicus].

Presentations

<u>Parts of this study have been presented at the following conferences and symposia:</u>

- VOX , Leuven, 2007
- Pan European Voice Conferences, Groningen, 2008
- Choice for Voice, London, 2008
- Choice for Voice, London 2010
- Collegium Medicorum Theatri, Frankfurt, 2011
- Pan European Voice Conferences, Marseille 2011
- European Laryngological Society, Helsinki 2012

Acknowledgments

I am most indebted to special persons who were very helpful with the realization of this project:

First of all, my promoter, Prof. dr. Felix de Jong, whose enthusiastic support in every aspect of the realization of this study I could fully appreciate during the years.

Prof.dr. Wivine Decoster, co-promoter, offered useful technical comments and logistic support.

Prof. dr. Marc Van Hulle, co-promoter, provided interesting new insights in the methodology of our research, leading to the present (unexpected) amazing discovery of three natural basic voice types of the human voice, which hopefully can give an answer to the age-old intriguing problems of voice classification.

Anna Ivanova M.Sc. was very helpful with the complex statistical analysis of the phonetographic results of our many subjects.

Diny Helsper took great care of the lay out of this manuscript.

This study is dedicated to Magda, my loving wife, who provided me for more than fifty years, with the warm and caring nest in which this study could grow and ripen.

Hugo Lycke
Thiers and Louvain, January 2013.