KTH Computer Science
and Communication

Why so different?

Aspects of voice characteristics in operatic and musical theatre singing

EVA BJÖRKNER

Doctoral Thesis
Stockholm, Sweden 2006

TRITA-CSC-A 2006:23
ISSN 1653-5723
ISRN KTH/CSC/A--06/23—SE
ISBN 91-7178-518-3
ISBN 978-91-7178-518-3

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

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av filosofie doktorsexamen fredagen den 8 december kl. 14.00 i sal F3 Lindstedtsvägen 26, Kungliga Tekniska högskolan, Stockholm.

Tryck: Universitetsservice US-AB

Abstract

This thesis addresses aspects of voice characteristics in operatic and musical theatre singing. The common aim of the studies was to identify respiratory, phonatory and resonatory characteristics accounting for salient voice timbre differences between singing styles.

The velopharyngeal opening (VPO) was analyzed in professional operatic singers, using nasofiberscopy. Differing shapes of VPOs suggested that singers may use a VPO to fine-tune the vocal tract resonance characteristics and hence voice timbre. A listening test revealed no correlation between rated nasal quality and the presence of a VPO.

The voice quality referred to as "throaty", a term sometimes used for characterizing speech and "non-classical" vocalists, was examined with respect to subglottal pressure (P_{sub}) and formant frequencies. Vocal tract shapes were determined by magnetic resonance imaging. The throaty versions of four vowels showed a typical narrowing of the pharynx. Throatiness was characterized by increased first formant frequency and lowering of higher formants. Also, voice source parameter analyses suggested a hyper-functional voice production.

Female musical theatre singers typically use two vocal registers (chest and head). Voice source parameters, including closed-quotient, peak-to-peak pulse amplitude, maximum flow declination rate, and normalized amplitude quotient (NAQ), were analyzed at ten equally spaced subglottal pressures representing a wide range of vocal loudness. Chest register showed higher values in all glottal parameters except for NAQ.

Operatic baritone singer voices were analyzed in order to explore the informative power of the amplitude quotient (AQ), and its normalized version NAQ, suggested to reflect glottal adduction. Differences in NAQ were found between fundamental frequency values while AQ was basically unaffected.

Voice timbre differs between musical theatre and operatic singers. Measurements of voice source parameters as functions of subglottal pressure, covering a wide range of vocal loudness, showed that both groups varied P_{sub} systematically. The musical theatre singers used somewhat higher pressures, produced higher sound pressure levels, and did not show the opera singers' characteristic clustering of higher formants.

Musical theatre and operatic singers show highly controlled and consistent behaviors, characteristic for each style. A common feature is the precise control of subglottal pressure, while laryngeal and vocal tract conditions differ between singing styles. In addition, opera singers tend to sing with a stronger voice source fundamental than musical theatre singers.

Key words: operatic singing, musical theatre singing, voice source, subglottal pressure, flow glottogram, inverse filtering, formant frequencies, amplitude quotient (AQ), normalized amplitude quotient (NAQ), vocal registers, velum opening, throaty voice.

Contents

Abbrevations

AQ	amplitude quotient (U_{p-t-p}/ MFDR)
DEGG	differentiated EGG signal
d_{peak}	peak derivative of the glottal flow (=MFDR)
EGG	electroglottography
F0	fundamental frequency
F1	first formant frequency
Fn	n:th formant frequency
H1-H2	level difference between first the two partials in voice source spectrum
MFDR	maximum flow declination rate
MRI	magnetic resonance imaging
NAQ	normalized amplitude quotient (AQ/T0)
U_{p-t-p}	glottal peak-to-peak flow amplitude
P_{sub}	subglottal pressure
P_{sen}	normalized excess pressure
Q_{closed}	closed quotient (ratio glottal closed phase time to period time)
SPL	sound pressure level
T0	period time
Tcl	closed phase time
VPO	velopharyngeal opening

List of publications

This thesis is based on the following papers, referred to by letters A through E.

Paper A
Velum Behavior in Professional Classic Operatic Singing.
Peer Birch, Bodil Gümoes, Hanne Stavad, Svend Prytz, Eva Björkner & Johan Sundberg.
Journal of Voice, 2002; 16 (1): 61-71.

Paper B
Throaty Voice Quality: Subglottal Pressure, Voice Source, and Formant Characteristics.
Anne-Maria Laukkanen, Eva Björkner & Johan Sundberg.
Journal of Voice, 2005; 20 (1): 25-37.

Paper C
Voice Source Differences between Registers in Female Musical Theatre Singers.
Eva Björkner, Johan Sundberg, Tom Cleveland & Ed Stone.
Journal of Voice, 2006; 20 (2): 187-197.

Paper D
Subglottal Pressure and Normalized Amplitude Quotient Variation in Classically Trained Baritone Singers.
Eva Björkner, Johan Sundberg & Paavo Alku.
Logopedics Phoniatrics Vocology. In Press, Available online October 2006.

Paper E
Musical Theatre and Opera Singing – why so different? A study of Subglottal Pressure, Voice Source and Formant Frequency Characteristics.
Eva Björkner
Journal of Voice. Submitted 2006.

Other related papers by the author

Comparison of Two Inverse Filtering Methods in Parameterization of The Glottal Closing Phase Characteristics in Different Phonation Types.
Laura Lehto, Matti Airas, Eva Björkner, Johan Sundberg & Paavo Alku.
Journal of Voice. In Press, Available online 14 February 2006

An Amplitude Quotient Based Method to Analyze Changes in the Shape of the Glottal Pulse in the Regulation of Vocal Intensity.
Paavo Alku, Matti Airas, Eva Björkner & Johan Sundberg
J. Acoust. Soc. Am. 2006; 120(2); 1052–1062

Author's contribution to the papers

Paper A
Author EB performed all measurements, assisted during the recordings, and prepared the analysis. Coauthors JS, PB, HS, and BG planned the investigation. Co-author JS assumed the main responsibility for writing the report.

Paper B
Author EB carried out the major part of the MRI analysis, of the acoustic measurements, and of the analyses (area functions, flow glottogram characteristics, formant frequencies). The investigation was planned and the recordings were carried out by co-authors AML and JS. The manuscript was jointly authored by coauthors EB, JS and AML.

Paper C
The major part of the work (analysis and writing) was carried out by the author EB. Co-author JS assisted in the analysis and in writing the manuscript. The investigation was planned and the recordings were made by co-authors TC, ES and JS.

Paper D
The major part of the work (analysis and writing) was carried out by the author EB. Co-author JS assisted in editing the manuscript. The recordings were made for another study under the supervision of JS.

Paper E
This work was designed and carried out entirely by the author EB. JS assisted in editing the manuscript.

Introduction

The voice is the major tool in speech communication. Also, it is possibly the most flexible among musical instruments. The singing voice is unique in the sense that we can not only produce a wide range of pitches and voice qualities with it, but also add *words* to elucidate and complement our musical expression.

Phonation is produced when air expelled from the lungs causes the vocal folds to vibrate. These vibrations generate a pulsating airflow which constitutes an audible source of acoustic energy, i.e., sound. This source sound is controlled by the degree of constriction of the vocal folds, the subglottal pressure, the volume of the airflow, and is modified in the vocal tract. Typically, voiced sounds are all the vowels as well as many consonants. In spoken languages, for example in English, approximately 78% of the phonemes are voiced (Catford 1977). In singing this figure is considerably higher.

To produce voiced sounds, three basic systems are involved; the respiratory system, the voice source, and articulation. *The respiratory system* is a compressor-like system, controlling breathing and phonation. When speaking habitually, the elastic and muscular forces involved act at an unconscious level but as vocalizing becomes an art, like in stage speech or singing, the control of the respiratory muscles needs to be precise, and hence conscious and trained. The *voice source* is the pulsating transglottal airflow produced by the vibrating vocal folds. When the combination of *subglottal pressure* and glottal configuration are appropriate the vocal folds start to oscillate and sound is produced. The sound varies in terms of quality and frequency depending on the muscular and aerodynamic conditions in the larynx. The adjustment of the *articulators*, i.e., the pharynx, the tongue, the jaw opening, the soft palate, and the lips, changes the acoustic conditions in the vocal tract. These conditions in turn influence the spectral properties such that the sounds produced can be perceived and interpreted in terms of speech sounds and voice qualities.

This introduction will present descriptions of the anatomy and function of the voice organ, voice source analysis methods, differences between speech and singing, and will consider aspects of the key question of this thesis, the differences between singing styles.

Respiration

Respiration is the act of breathing. The respiratory apparatus consists of (a) an upper cavity, i.e., the thorax, formed by the rib cage and the pulmonary system, (b) a lower cavity, formed by the abdomen, and (c) the diaphragm, that separates these two cavities.

The lungs are located and suspended in the rib cage and respiratory events primarily result from modification of the rib cage dimensions. Chest wall movements are influenced by both active muscle forces and passive forces. The passive forces originate from (a) elastic recoil in the rib cage, (b) resistance to airflow by the airways, (c) gravity, and (d) from the inertial properties of the respiratory system (Rodarte & Rehder 1986). Air enters the body through the upper airways; the nose, the mouth, the pharynx (the throat), and travels down through the *larynx* (see section Larynx and phonation) and the *trachea* (the windpipe) into the lungs. During inspiration, an active contraction of the external intercostal muscles lifts the ribs and pulls them upward and outward and the *diaphragm* (the most important inhalatory muscle) lowers the "floor" in the thorax. These actions increase lung volume and create a pressure drop in the lungs, allowing air to rush in through the open airways.

In singing, but also in phonation in general, voluntary control of both inhalatory and exhalatory muscles is of paramount importance. In quiet expiration, by contrast, the inhalatory muscles automatically relax and the thorax-unit recoils back to its resting position. Vital capacity is crucial to a singer's maximum phrase duration; it is the amount of air in the lungs that can be expelled after maximum inhalation.

Non-singers and country singers have been found to use only slightly higher lung volumes than those used by speakers (Hixon *et al.* 1973; Hoit *et al.* 1996; Cleveland *et al*. 1997). Professional operatic singers, on the other hand, use notably higher portions of their vital capacity. Also, possible gender differences have been revealed; female operatic singers were found mostly to spend between 40-50 % of their vital capacity in a phrase while male singers spend only 20-30 % (Thomasson & Sundberg 1997). In addition, high lung volumes are associated with glottal abduction forces (Iwarsson *et al*. 1998).

Subglottal pressure

The *subglottal pressure* (P_{sub}), produced by the respiration system, is the pressure below the closed or the semi-closed glottis. P_{sub} is one of the main factors for vocal fold vibration and the primary factor contributing to vocal loudness (Gauffin & Sundberg 1989). Increasing P_{sub} in terms of increasing vocal loudness generally tends to raise fundamental frequency (F0) in speakers (Gramming 1988). In addition, the control of P_{sub} has been found to be less precise when male operatic singers applied a non-habitual inhalatory pattern (similar to that found in non-singers) rather than a habitual pattern (Thomasson 2003).

Direct determination of P_{sub} is a tricky and invasive procedure due to the need to reach a measurement position below the adducted vocal folds. An alternative technique to measuring P_{sub} in phonation (indirect and non-invasive) is to capture the intra-oral pressure during the occlusion for the consonant /p/. This can be done by inserting a tube connected to a pressure transducer, into the mouth. Data derived from such measurements are often also referred to as P_{sub}.

To initiate vocal fold vibration P_{sub} has to exceed a minimum pressure generally referred to as the phonation threshold pressure. This threshold pressure, as well as the P_{sub} range, varies substantially with pitch and, presumably, also with vocal fold thickness. Comparisons between subjects with different threshold pressures and P_{sub} ranges are facilitated by using the normalized excess pressure (P_{sen}) (Titze 1992), representing an attempt to compensate for phonation threshold differences.

For obtaining a detailed view of how P_{sub} influences the voice source a series of different P_{sub} values needs to be analyzed. In Papers C, D and E, the singers were asked to sing from loudest to softest degree of vocal loudness at different F0s. This yielded a set of P_{sub} values within each singer's total vocal loudness range. Then, ten equally spaced P_{sub} values, within the singer's total P_{sub} range, were selected. However, phonation threshold pressure is sometimes difficult to determine accurately, and errors easily result in quite misleading estimations of the P_{sen} range. In such cases a better and simpler alternative is to express P_{sub} as a percentage of the subject's total P_{sub} range. This however requires access to a fair number of pressure values. Thus, P_{sub} data can be expressed as (1) the actual pressure in cm H_20, (2) as P_{sen} , and (3) normalized with respect to the P_{sub} range.

Larynx and phonation

Cartilages

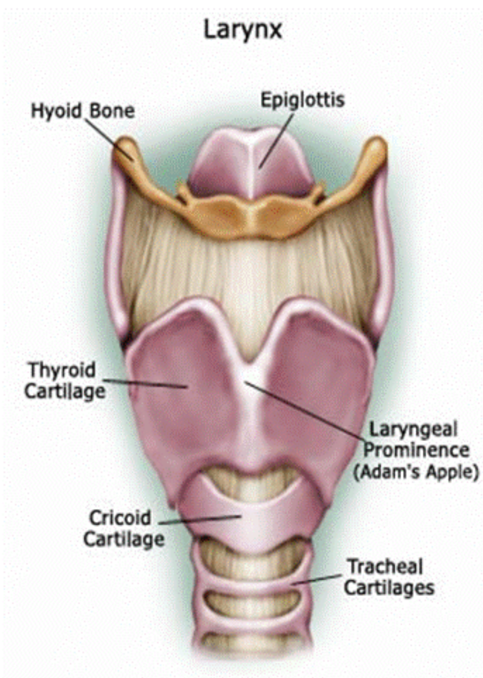


Figure 1. Front view of the larynx (from Netter F, Atlas of Human Anatomy 2nd ed. Novartis, East Hanover, New Jersey. 1997)

The larynx is a cartilaginous structure located between the top of the trachea, the tube leading from the lungs, and the hyoid bone (see Figure 1). It is composed of cartilages connected by ligaments and muscles. The *cricoid* is the uppermost cartilage of the trachea, immediately below the *thyroid*. The cricoid has the shape of a complete "signet" ring, as opposed to the other hoarseshoe-shaped tracheal cartilages, and its back is larger than its front. On top of the posterior signet part ride the much smaller *arytenoid* cartilages. They are shaped somewhat like triangles and can to a certain extent rotate vertically and horizontally, as well as slide posteriorly and anteriorly on the cricoid cartilage (Laver 1980). The *thyroid* is the big shield-like cartilage protecting the larynx which, in adult males, is protruding and is frequently referred to as the Adam's apple. The *hyoid bone* is the uppermost part of the laryngeal structure and is attached to the skull and the lower mandible. The *epiglottis* is the flap of cartilage lying behind the tongue and in front of the entrance to the larynx. At rest, the epiglottis is upright and allows air to pass through the larynx and into the rest of the respiratory system. During swallowing, it folds back to cover the entrance to the larynx, preventing food and drink from entering the trachea. The small tube inside the larynx is called the epilarynx tube.

Vocal folds

Inside the larynx, the *true vocal folds* and the *ventricular folds* are formed by the thyroarytenoid muscle. The true vocal folds are a complex structure containing muscles as well as layers of tissue. The *body* is the vocalis and thyroarytenoid muscles, anteriorly attached to the thyroid and posteriorly to the processes of the arytenoids (see Figure 2).

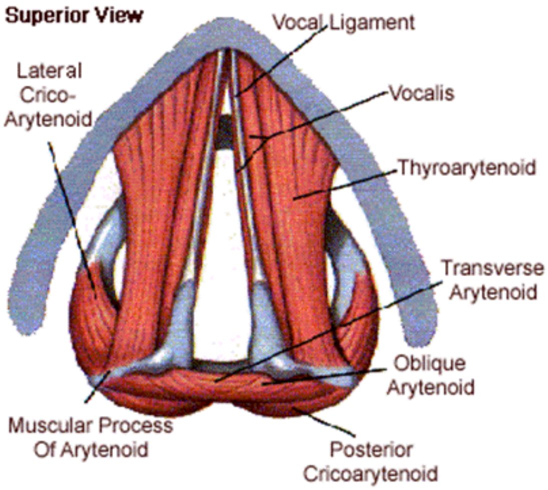


Figure 2. Laryngeal muscles. (from H M. Tucker, The Larynx, Thieme 1987)

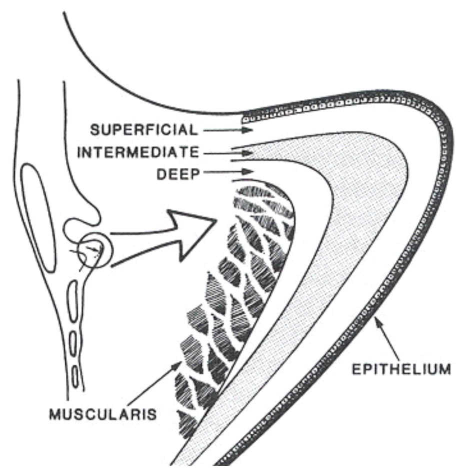


Figure 3. Vocal fold structure. (from Hirano, 1974)

The *cover*, as described by Hirano (Hirano 1974 ; Hirano 1977), is composed of a microscopic five-layered structure. The *deep lamina prop*ria is a fiber structure closest to the vocalis muscle. The *intermediate lamina propria* provides elasticity to the vocal fold. The *superficial lamina propria* also called "Reinke's space" consists of a gelatin-like substance, and the outermost layer is the *squamous epithelium* which serves to protect the underlying tissue and help regulate vocal fold hydration (see Figure 3). With respect to the different stiffness characteristics, the folds can also be divided into three subgroups, the *mucosa* (the cover) that includes the epithelium and the superficial lamina propria, the *vocalis ligament* (transition) including the intermediate lamina propria and the deep lamina propria, and the *body* of the vocal fold, i.e., the vocalis muscle. Due to the vocal fold structure the opening and closing of the *glottis*, the air space between the folds, is complex.

Glottal adduction, the action which closes the glottis, involves at least three muscular functions. The contraction of the *lateral cricoarytenoid* muscle swivels the arytenoid cartilages anteriorly and medially which adducts the vocal folds. The *interarytenoids* and the *transvserse arytenoids* help to close the posterior part of the glottis by a lateral gliding action (Laver 1980). Contraction of the lateral *thyroarytenoid muscles* produces medial compression of the glottis thus augmenting glottal adduction (van den Berg 1968). Vocal fold *abduction* is the muscular action that opens the glottis. It is normally performed by a single muscle pair, the *posterior cricoarytenoids*. Their contraction rotates the arytenoids outwards such that the vocal folds separate.

Two muscles regulate the stiffness in the vocal folds, the vocalis and the criocothyroid muscles (Hirano *et al.* 1970), which affects the fundamental frequency (F0). Vocal fold length is mainly regulated by the paired cricothyroid muscles which when contracted stretches the folds by tipping the thyroid downwards towards the criciod cartilage. Contraction of the vocalis thickens the vocal folds, particularly at low F0, which also thickens the loose cover tissue. If the vocal folds are stiffened F0 is increased. Hence, lower F0 is characterized by shorter, thicker and more flapping vocal folds and higher F0 by longer, thinner and stiffer folds. Thus P_{sub} needs to be adjusted to the current circumstances.

Vocal fold length differs between the genders and is typically between 15 – 20 mm in adult males and between 9 – 13 mm in adult females. This brings consequences for the pitch range; the longer the folds, the lower the pitch. In normal speech the typical F0 range for males is 80-200 Hz, and for females 150-350 Hz.

Voice source

The *voice source* is the pulsating transglottal airflow. A 'buzzing' source sound is generated by the periodic train of flow pulses, produced as the vibrating vocal folds chop the steady air stream from the lungs.

Vocal fold vibration starts when there is appropriate balance between the P_{sub} and the muscular tension in the vocal folds. The *myoelastic- aerodynamic theory* (van den Berg 1958), explains phonation as the result of three major factors, (1) the aerodynamic forces that affect the larynx, (2) the activation by nerve stimulation of laryngeal muscles regulating the elastic properties of laryngeal tissues, and (3) the acoustic coupling between the larynx and the sub- and supraglottal cavities, as well as the mechanical coupling between the folds.

The vibratory cycle is initiated when, by muscular action, the adducted vocal folds are forced apart by the P_{sub}. The folds separate with a vertical phase difference, such that the lower part opens before the upper part. This initiates a wave-like motion of tissue traveling from the inferior portion of the vocal fold cover to the superior, along the edges of the vocal fold (Hollien 1968). The wave-like motion is referred to as the *mucosal wave*.

As the folds open, the air flow accelerates through the glottal constriction. This causes a local drop in air pressure inside the glottis and the vocal folds are "sucked" medially (Broad 1977; Flanagan 1978). The combination of this effect and the tissue elasticity in the folds closes the glottis, where after a new glottal cycle begins. The "sucking" effect was earlier entirely ascribed to the Bernoulli effect but after extended research the current view emerged (Titze 1976; Stevens 1977; Titze 1988).

Mucosal waves are enhanced by vocalis contraction which increases the amount of loose cover tissue. Thus mucosal waves are more prominent in low pitched and loud phonation. Examination of the mucosal wave has become an important part of the assessment of vocal function (Titze *et al.* 1993).

Formants and articulation

Resonances in the vocal tract are referred to as *formants* and their frequencies and amplitudes shape the radiated spectrum. Hence, the location of the formant frequencies in the spectrum determines vowel quality and they also have an effect on voice quality (Laver 1980).

Articulation is the term used for all maneuvers that change the vocal tract shape. It is performed by the *articulators*, such as the pharynx, the tongue, the jaw opening, the soft palate (velum), and the lips. The frequencies of the first two formants, F1 and F2 determine the vowel quality, while the higher formants F3, F4 and F5 rather influence voice quality. F1 is particularly sensitive to jaw opening, F2 to the position of the body of the tongue, and F3 to the position of the tip of the tongue. Formant frequencies do not generally vary with F0. On the other hand, they are affected by vocal tract length. For any given vowel, adult women tend to have higher formant frequencies than adult males (Fant 1966).

Source - vocal tract interaction

In the 1960's, researchers started synthesizing speech in order to understand the phenomena behind voice production. A theory of fundamental importance was presented by Fant (Fant 1960). He introduced the *source-filter theory*, which assumes that the glottal source and the vocal tract filter are separate systems which do not interact. Fant used the theory for modelling synthesized speech and it is still used in speech modelling today. The source-filter theory implies that a time-varying vocal tract configuration has no effect on the shape of the glottal flow waveform. However, we now know that a *source-tract interaction* exist and affects the shape and the periodicity of the glottal waveform.

The interaction between the voice source and the vocal tract is not yet fully understood. For example, a nonlinear interaction between the glottal source and the impedance of the vocal tract during normal voicing can cause a *skewing* of the glottal flow pulse (Rothenberg 1973; Fant *et al.* 1985; Fant & Lin 1987). Further, *ripples* in the waveform during the glottal opening are due to absorption, or *glottal damping*, of the first formant energy (Fant 1993; Childers & Wong 1994). At the instant of glottal closure high frequency acoustic energy is generated. During the open phase, a considerable amount of the energy of the first formant is absorbed by the glottis which reduces the amplitude.

Titze studied the F0-F1 interaction in so-called *resonant voice* (Titze 2001). This type of phonation is perceptually defined as being produced with ease, adequate loudness, and vibrations in the facial tissues (Verdolini 1994). Titze found that the interaction provides maximum assistance to vocal fold vibration, which thereby increases the acoustic energy production of the voice source. This effect can increase the level of the radiated sound by as much as 10 dB (Titze 2004). On the other hand, as F0 approaches or passes F1 this positive source-tract interaction disappears (Fant 1993).

Vocal registers

The phenomenon and terminology of vocal registers is complex and somewhat confusing. In a summary of vocal registers in singing, Henrich (2006) reports that Garcia (1840) suggested the human voice is composed of three registers; *chest*, *falsetto-head*, and *counter bass*. Garcia defined the term register as follows: 'By the word register we mean a series of consecutive and homogeneous tones going from low to high, produced by the same mechanical principle, and whose nature differs essentially from another series of tones equally consecutive and homogeneous produced by another mechanical principle. All the tones belonging to the same register are consequently of the same nature, whatever may be the modifications of timbre or of the force to which one subjects them."(Henrich 2006). More than 160 years later, Garcia's suggestions are still highly valid though the terminology for the different registers have changed somewhat. For example, Hollien (1974) defined a vocal register to be characterized by a nearly identical vocal quality within a certain pitch range, and with little or no overlap in F0 between adjacent registers. His suggestion to define registers according to (1) perceptual, (2) acoustic, (3) physiologic, and (4) aerodynamic parameters reflects the complexity of the register phenomenon.

Today singing-voice registers are generally referred to as *chest register and head register.* Chest register is used in the lower part of the singing pitch range, up to 300-440 Hz, approximately, and is associated with a voice quality sometimes described as "thick" and "heavy." The head register is typically used above this pitch range (in classical singing also for lower F0) and is associated with a voice quality that can be described as "thinner" than that of the chest register.

The terms "thick" and "thin" can be motivated not only from descriptions of the vocal timbres but also from a muscular point of view. The chest and head register have been shown to be associated with different amount of vocalis contraction causing thickening or thinning of the vocal folds (Hirano *et al.* 1970).

According to Garcia the *falsetto* register is located between the chest and head registers. Today, however, the term falsetto is often used for a register appearing above, or even replacing, the head register particularly for the male voice (Henrich 2005).

The term *middle* register most often refers to the register used for the middle singing pitch range. Voice quality can be described as "mixture" of chest and head /falsetto registers. It is, however, unclear if it refers to perceptual or physiological parameters, or both. The terms *flute, whistle, flageolet* or *loft register* are mostly associated with the very highest singing pitch range. It is not accessible to all singers and is mostly used in improvised non-classical music. Henrich (2005) suggested that the term "registers" should be replaced with the term *laryngeal mechanism* and referred to in terms to numbers. Mechanism 1 is the register most commonly used in speech.

The terminology for the speaking-voice registers differs somewhat from the singing-voice registers. Phonation is referred to as (a) *vocal fry*/glottal fry, creak, or pulse register in the lowest vocal frequency range (b) *modal* or *chest* register in normal speaking or singing voice, and (c) *falsetto* or *loft register* at the highest vocal frequencies (Hollien 1974).

Continuous research over the last decades has managed to describe, partly or fully, several vocal terms with reference to their physical, acoustical and/or aerodynamical characteristics. Yet, the scientific community has failed to reach an agreement on the definition of "voice register". While some authors consider register as a purely laryngeal phenomenon, others define it in terms of voice quality similarity. Further, confusion has emerged as new knowledge has poured new meanings into existing terms. An updated voice-related vocabulary, common for all communities dealing with the voice, would undoubtedly be of great importance.

Spectral balance

The spectral balance and the location of the formant peaks reveal information about timbre and voice quality (Helmholtz 1877). The spectral slope of the voice source varies typically between 6 and 12 dB per octave depending on phonation type. High amplitude of the first partial (H1) contributes to a steeply sloping spectral envelope.

The level of the fundamental, as well as the *spectral balance* of the higher partials, is of great importance to the perception of voice quality. The balance between high and low frequency partials carries important information about vocal loudness (Sluijter 1997; Nordenberg & Sundberg 2004) and voice quality (Hammarberg *et al.* 1980). The level difference between the first and the second partial (H1-H2) has been extensively used in descriptions of voice source characteristics (Klatt & Klatt 1990; Hansson 1997). High values of H1-H2 reflect a dominant fundamental, sometimes resulting from breathy phonation (Klatt and Klatt 1990). Changes of the closed quotient of the glottal flow pulse have been found to affect the H1-H2 relation; higher closed quotients being associated with lower H1-H2 values (Holmberg *et al.* 1995, Sundberg *et al.* 1999). Sundberg & Högset (2001) found higher H1- H2 values in the falsetto register as compared to the modal register.

Our ears are particularly sensitive to frequencies in the region around 2500-3500 Hz. That is, a sound with an intensity level of 70 dB appears to be louder if occurring at 3000 Hz than at 1000 Hz, just because of the difference in frequency. With the exception of sopranos, this phenomenon is systematically used by operatic singers. The singers develop a technique by shaping their vocal tracts such that F3, F4 and F5 form a cluster in the 3 kHz region, while keeping a strong fundamental. The formant cluster, known as the "singer's formant" (Sundberg 1974), gives the partials in this region an extra boost that makes the voice easier to discern in the presence of a loud orchestral accompaniment. The epilarynx tube, including the laryngeal ventricle, has been suggested to be the primary contributor to this clustering effect (Sundberg 1974; Titze 2001).

As mentioned, variations of vocal loudness are associated with changes of the spectral balance. In soft voice the slope of the voice source spectrum is steeper than in loud voice. Therefore, a spectrum produced in soft voice and thus possessing weak high frequency partials cannot be converted into a loud voice simply by electronic amplification.

Variation of vocal loudness

Three basically distinct mechanisms control intensity regulation in the human voice (Titze 1994); (a) Below the larynx, the aerodynamic output of the lungs regulates intensity in terms of P_{sub} (Ladefoged and McKinney, 1963; Bouhuys *et al.*, 1968), (b) Within the larynx, intensity regulation can be performed by modifying the vocal fold vibration, which affects the conversion of aerodynamic flow into acoustic power. Increased vocal intensity corresponds to an increase of the flow amplitude and/or to a decrease in the length of the glottal closing phase, both typically caused by a raised subglottal pressure, (c) Above the larynx, vocal intensity can be modified by adjusting the resonances of the vocal cavity, especially when the first formant coincides with a harmonic of the glottal source (Titze 2004). This phenomenon, called formant tuning, is rarely used in speech but frequently in singing. When F0 approaches F1 as in high-pitched singing, singers, particularly sopranos, tend to tune their lower formants, increasing F1 such that it falls close to or slightly higher than F0. This technique substantially increases the loudness of their vocal output (Sundberg 1975). In addition, the frequency distance between F1 and the second and higher formants affects vocal intensity, such that vowels with a high F1 have higher intensities than vowels with a lower F1, all other things being equal.

Methods for voice analysis

Due to the awkward location of the larynx and the delicate function, based on a co-operation between muscle forces and aerodynamic effects, voice source analysis is far from straight-forward. Several methods have been developed over the years, both invasive and non-invasive. None of the methods can, however, cover all aspects of voice production alone, and a combination of two or more methods is beneficial.

Acoustical voice source measurements

Acoustical measurements are typically non-invasive and thus allow recording of almost habitual voice production. This is particularly valuable in analysis of singing, since voice use in singing would easily be disturbed by strange experimental conditions. To gain information about voice source characteristics in vowel production, the acoustic filtering effect of the vocal tract resonances must be eliminated. *Inverse filtering* (Miller 1959) is a method for retrieving the glottal flow from the speech pressure signal (or from the oral flow). This is done by eliminating the effects of the vocal tract filter, thus extracting the volume velocity waveform at the glottis. The idea behind the method is to first form a model for the vocal tract transfer function. By filtering the voice signal through the inverse of the model, the effects of vocal tract resonances are canceled. The result is an estimate of the glottal flow represented as a time-domain waveform, the *flow glottogram*, or volume velocity waveform (Lehto *et al.* 2006).

The main criterion of a successful reproduction of the glottal flow is to achieve a maximally flat horizontal and ripple-free closed phase in the glottogram. In early manual inverse filtering, only the first two formants were adjusted. In later manual programs the user adjusts an appropriate number of formants and their bandwidths. This procedure is time-consuming and semiautomatic and automatic methods have therefore been developed. Both manual and semi-automatic methods require user interaction. In automatic inverse filtering methods, on the other hand, the user typically sets certain initial parameter values, after which the method estimates the voice source without any subjective user adjustments (Alku 1992).

The voiced signal can either be a flow signal, captured with a circumferentially vented flow mask (Rothenberg 1973), or an audio signal, recorded in an anechoic chamber or in free-field. Each recording technique has its benefits and limitations. Recordings with a flow mask can take place in any room but the mask distorts the auditory feedback and may affect extreme articulation. Audio recording, on the other hand, is vulnerable to room resonances and disturbing noise that can complicate inverse filtering.

By canceling the resonances in the signal an estimation of the pulsating transglottal airflow is obtained, represented by a flow glottogram showing the glottal volume velocity waveform. The flow glottogram reflects the glottal opening and closure in terms of time and amplitude. Resonances that the inverse filter failed to eliminate appear as ripples in the glottal waveform. This may in some cases complicate determination of the time instant of glottal closing and opening. Furthermore, a good separation between F0 and F1 facilitates inverse filtering which is why the vowel /ae/, produced with a high F1, is typically used in recording tasks. High pitches increase F0 interference with F1. Thus is it more difficult to estimate formant frequency location in high pitched female voices, children and tenors.

Parameterization of the voice source has been the target of intensive research during the past few decades. This has resulted in a large variety of methods to quantify the waveforms given by inverse filtering. One of the most commonly used approaches to parameterize the voice source is to divide the glottal flow waveform in time-based events.

Flow glottogram parameters

The *time-based* parameters of the flow glottogram yield information about fundamental frequency (F0), periodicity, and time patterns of the vibratory events (see Figure 4). They are typically referred to as *period time (T0), closed phase, open phase, closing phase,* and *opening phase*. The glottogram data are often normalized by dividing by T0, such that quotients are obtained. For example, the *open quotient*, equals the ratio between the duration of the open phase and T0, thus reflecting the portion during which the vocal folds are open. Similarly, the *closed quotient* reflects the portion of the period during which the folds are closed. These quotients are relevant to voice quality. For instance, a high open quotient typically refers to a breathy voice quality (Holmberg *et al.* 1988; Henrich *et al.* 2005), while a high closed quotient in speech typically refers to a pressed phonation type.

Time-based parameters are computed by measuring the *time lengths* between various events (i.e., glottal opening and closure as well as the instant of the maximal flow). The time-based parameters of the glottal closing phase can be combined with *amplitude-domain* values, extracted from the glottal flow and its first derivative. This approach is based on the voice source parameterization schemes developed by Fant (Fant & Lin 1988; Fant *et al.* 1994; Fant 1997). Based on Fant's findings, Alku and collaborators introduced two voice source measures; the *amplitude quotient* (AQ), defined as the ratio between the *peak-to-peak pulse amplitude (*$U_{p\text{-}t\text{-}p}$) and the *maximum flow declination rate (*MFDR*)*, also called d_{peak} (Alku & Vilkman (1996a, 1996b), and the *normalized amplitude quotient* (NAQ), defined as AQ/T0 (Alku *et al.* 2002). These quotients, extensively explored in the present thesis, have been found to be closely related to phonation mode (Alku et al. 1996; Alku et al. 2002).

$U_{p\text{-}t\text{-}p}$ correlates strongly with the amplitude of the fundamental (Gauffin & Sundberg 1989; Fant 1997). MFDR, reflecting the maximum speed of flow decrease during the glottal closure, has been shown to be closely related to voice characteristics such as vocal intensity (Fant *et al.* 1985), sound pressure level SPL (Gauffin & Sundberg 1989), and to the subglottal pressure P_{sub} (Sundberg *et al.* 1999).

Consequently, valuable information about voice production and voice quality can be obtained from flow glottogram data. It is noteworthy, however, that a relatively long closed phase can result either from (1) sufficient or firm glottal adduction, (2) thick vocal folds, (3) a contracted vocalis muscle, or (4) low F0. For this reason, it seems sensible to relate flow glottogram data to a control parameter. P_{sub} is a good candidate as it controls vocal loudness (Ladefoged 1961; Gauffin & Sundberg 1989).

Variation of P_{sub} is typically associated with contractions of various laryngeal muscles such as those controlling F0; speakers tend to raise their mean F0 when increasing vocal loudness (Gramming 1988). This type of automatic co-variation of phonatory characteristics is mostly unacceptable in singing.

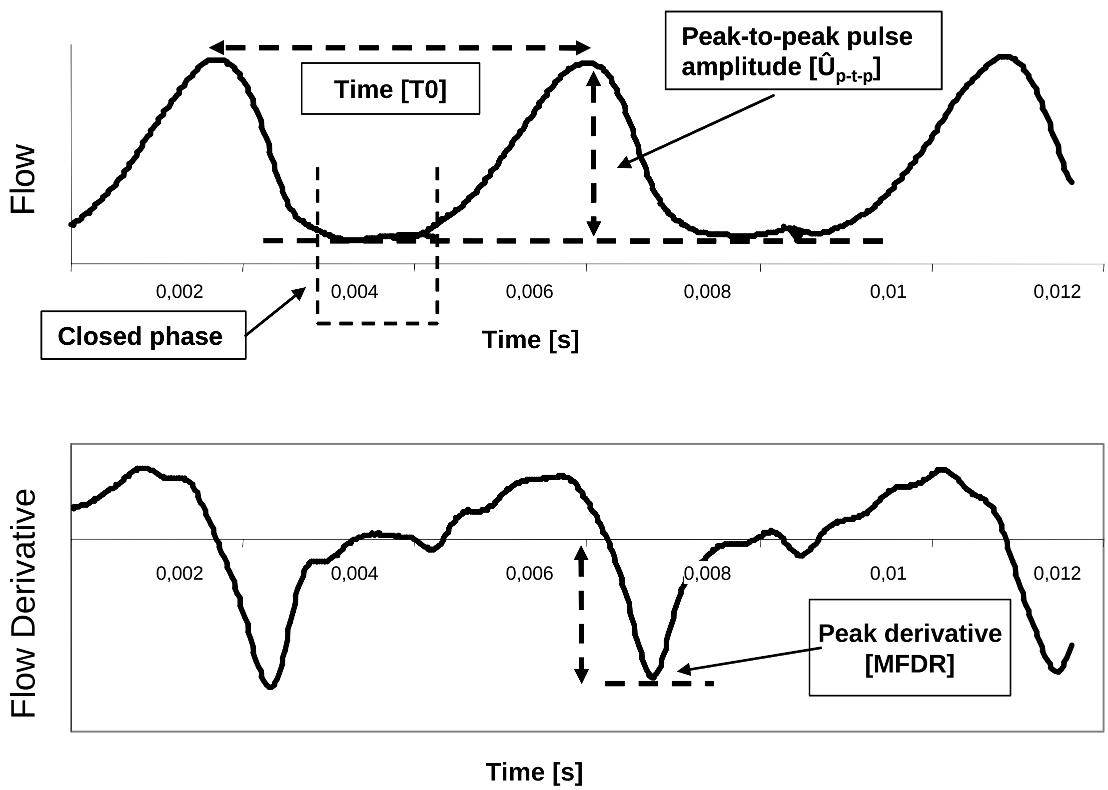


Figure 4. Flow glottogram and characteristic time and amplitude parameters.

Electroglottography

An *electroglottogram* (EGG) reflects vocal fold vibratory patterns in terms of variations of electrical impedance in the glottis (Fabre 1957; Fourcin & Abberton 1971). A high-frequency electrical circuit with a low voltage, i.e., physiologically safe, is used to send a small current between two electrodes placed on the skin of the neck at either side of the thyroid cartilage. As human tissue conducts electricity better than air, the amplitude of the electrical signal increases when the vocal folds are in contact, and hence decreases as the glottis opens and the impedance increases. Childers *et al.* (1986) showed that the EGG reflects the contact area of the vocal folds.

Vibratory analysis can be made both from the EGG signal and also from its derivate, the *DEGG* signal. The closing of the vocal folds is generally faster than the opening and closing therefore is reflected in the EGG signal by a steep slope that corresponds to a strong positive peak. The opening is reflected by a somewhat weaker maximum slope in the EGG signal which corresponds to a smaller negative peak in the DEGG signal. Henrich (2004) suggests that strong and weak peaks in the DEGG signal, in cases where they are single and precise, can be accurately related to instances of glottal closing and opening, respectively. However, occurrences of double peaks are quite common in DEGG signals and thus complicate the interpretation.

Since EGG is measured directly at the source (with no vocal tract influences) it is a straightforward and user-friendly method to gain information about vocal fold behavior. Nevertheless, it is not free from disadvantages. The placement of the electrodes is of great importance since a slight shift might introduce artifacts. Further, the vocal folds open with a vertical phase difference, often combined with a clear mucosal wave. Therefore, the instants of opening and closing shown by the EGG signal may not correspond fully to the onset and offset of the glottal flow. For these reasons, combining EGG with a different technique, acoustical or visual, is likely to yield more reliable data.

Magnetic resonance imaging

Magnetic resonance imaging (MRI) was introduced in 1971. An MRI scan is an imaging technique that produces high quality images of the inside of the human body. The method is based on nuclear magnetic resonance, a physical phenomenon in which magnetic fields and radio waves cause atoms to radiate weak radio signals. This implies that with MRI scanning the damaging radiation effects of the classical X-ray method are avoided. Today, MRI scanning is a commonly used method in medical settings. In voice analysis MRI is particularly used for describing vocal tract shapes, morphology, and dimensions (Baer *et al.* 1991; Fitch *et al*, JASA 1999).

Singing versus speech

Compared to spontaneous speech, singing is a much more accurately controlled phonation task. In spontaneous speech, interaction between laryngeal nerves, muscles, and aerodynamics cause effects which are intolerable in singing. As mentioned, speakers tend to raise their pitch when they increase vocal loudness while this linkage obviously is unacceptable in singing**.**

This means, that in order to turn the voice into a musical instrument, singers need to become aware of the various vocal parameters that are involved in voice production, and learn how to separately train and gain control over each of them. For example, singers need quick and accurate control of a number of different voice parameters.

(1) The respiratory system, which includes the inhalatory and exhalatory muscle activity and the passive recoil forces in the rib cage and the lungs, plus gravitation forces. These factors regulate lung volume which is of basic relevance to musical phrasing.

(2) Laryngeal muscle activity, which governs F0, glottal adduction, and phonation modes, and which must be tuned in accordance with P_{sub}.

(3) Articulation, which regulates the formant frequencies and hence affects the vocal output in terms of vowel and voice quality.

The P_{sub} range offers a striking example of the differences between speech and singing. For the loudest tones singers can use up to 60 cm H_2O or more, while speakers typically use no more than 20 cm H_2O. Other examples are lung volume range which is considerably wider in singing than in speech. The use of high lung volumes, particularly common in classical singing, entails the need to deal with substantial elasticity forces. In addition, singers obviously need accurate control of F0.

Summarizing, as compared to speakers singers develop a more independent, systematic, and accurate control and use extreme ranges of variation of various voice parameters. Therefore, relationships between different voice parameters are often more evident in singers' than in untrained speakers' voices. In other words, using professional singers as subjects would be quite rewarding in attempts to analyze the effects on the voice of a variation in a voice control parameter, e.g., P_{sub} or F0.

Singing versus singing – different singing styles

As mentioned, differences in voice qualities are reflections of variation in the muscular, aerodynamic, and acoustical conditions in the larynx and in the vocal tract. The subglottal pressure, the driving force in phonation, needs to be adapted in accordance with the laryngeal conditions.

Up to just a decade ago, most investigations of the singing voice were devoted to classical/operatic singing. However, the majority of young people (and not only them) generally do not listen to classical singing, but watch TV-shows like "Pop Idol" and want to be able to sing like their favourite singer. Thus, a growing interest for how different types of vocal styles are produced and how they should be taught to students has emerged. Noteworthy is also that, unlike classical singing which includes for example opera, early music, romance, and choir-singing, there is still no common definition of styles that are *not* classical singing. The term *non-classical* singing is typically used to describe singing in jazz, pop, blues, soul, country, folk, and rock styles. The search for finding a common term for these styles, that is more related to the music and the voice timbres rather than just being non-classical, has been going on for years. In the USA *Contemporary Commercial Music* (CCM) is being used by some vocal pedagogues (Lovetri & Weekly 2002).

During the last 10–15 years many research studies have been devoted to "non-classical" singing, and the number of investigations is constantly growing. In particular, *belting,* a timbral effect frequently used by female pop and musical theatre singers in high and loud notes, has been studied (Miles & Hollien 1990; Estill 1988; Sundberg *et al.* 1993; Bestebreurtje & Schutte 2000). Belting was found to be associated with high P_{sub}, a long closed phase, and generally high activity in the laryngeal and abdominal muscles. Further, in a female singer subjects, the NAQ parameter was found to reflect differences between singing styles and to correspond to perceived degree of phonatory pressedness (Sundberg *et al.* 2004).

The results gained in the present thesis (Papers C, D, and E) have revealed a number of typical voice differences between operatic and musical theatre singers. As compared with operatic singers musical theatre singers

a) use somewhat higher subglottal pressure,
b) produce higher MFDR,
c) produce higher sound pressure levels,
d) have higher closed quotient
e) have higher peak-to-peak flow glottogram pulse amplitude, and
f) have a less dominating voice source fundamental.

The voice research in this thesis attempts to reveal characteristics of singing styles in terms of acoustical and physiological facts. Such research should be beneficial for establishing a terminology applicable to a multitude of different singing styles as well as to speech. This would be of value to vocal pedagogy and for the mutual understanding between people active within the voice community.

Purpose of the studies

The purpose of the studies in this thesis was

- to explore voice function and voice source characteristics systematically with professional singers as subjects, who have acquired a high consistency in control of breathing, phonation, and articulation.
- to identify respiratory, phonatory and resonatory sources of the voice timbre differences between styles of singing.
- to examine the informative power of the amplitude quotient (AQ) & normalized AQ (NAQ) with respect to voice source characteristics.
- to examine the effect of subglottal pressure variation on AQ & NAQ.
- to expand knowledge about vocal registers in the female voice.
- to describe classical singers' control of articulation with respect to a velopharyngeal opening (VPO).
- to investigate whether a VPO generally is associated with a nasal vowel quality.
- to identify the phonatory, articulatory, and resonatory correlates to "throaty" voice quality.

Overview of the results

Paper A

Velum Behavior in Professional Classic Operatic Singing

Introduction

In vocal training and therapy, exercises involving velopharyngeal opening (VPO) have a long tradition. A classical exercise is to phonate on a nasal murmur or to initiate vowel phonation by such a murmur, e.g., [ma, mu, mi]. Resonance in the nasal and/or oral cavities may affect sound quality considerably. This seemingly suggests that a VPO may be beneficial in singing.

Determining whether or not there is a VPO present in singers is not a trivial task. Two commonly used methods of analyzing VPO are (1) visual evidence, e.g., nasofiberscope documentation, and (2) airflow measurements of nasal DC airflow. Both methods are invasive to some extent but do not prevent habitual singing. The methods are complementary but cannot be used simultaneously.

In this study three methods were used for detection of a VPO: (a) nasofiberscopy, (b) simultaneous measurements of nasal and oral airflow by means of a divided flow mask, and (c) comparison of the level of the fundamental in the nasal and oral airflow signals.

Aim

The purpose of the study was to investigate to what extent professional opera singers use a VPO during singing, and if such an opening is typically associated with a nasal quality of the voice timbre.

Method

Seventeen professional operatic singers of different classifications (soprano, tenor, baritone, bass) all premiere opera soloists, volunteered as subjects. Their task was to repeatedly sing the words [panta, puntu, pinti] in mezzo forte on each tone in an ascending A-major triad, extended over their entire pitch range. The singers sang this entire material twice, first for recording oral and nasal airflow and then for recording the VPO by means of a nasofiberscope. In total, 714 vowel samples were analyzed.

An overall assessment of the degree of “nasal quality” was collected from a listening test of the audio signal that was recorded during the naso-fiberscope session. The task of the panel, consisting of six conservatory singing teachers, was to rate to what extent they found that resonance in the nasopharynx contributed to the timbre.

In addition, quantitative visual estimates of VPO were obtained from four phoniatricians. Their task was to rate the degree of VPO from the video-recorded nasofiberscope images.

Results and discussion

For the vowels [a] and [u] nasal flow was observed to a lesser or greater degree in all singers but varied between pitches (see Figure 1). For the vowel [i] nasal flow was only observed in one tenor (tenor 2). The results indicate that in these cases the singers sang with a VPO. In addition, all three tenors interestingly showed signs of a VPO for the passagio pitches C#4 and E4. This may indicate that a VPO facilitates a seamless timbral transition in this pitch range.

The video recordings of nasofiberscopy revealed several shapes of VPO. The openings could be grouped into three types, (1) one extending along the *coronal* direction with retracted sidewalls and with the distance between the velum and pharyngeal wall being small or nil; (2) one extending along the *sagittal* direction with advanced sidewalls and greater distance between the velum and pharyngeal wall; and (3) a *constricted type*, with advanced sidewalls and a narrow distance between the velum and posterior pharyngeal wall showing the Passavant's ridge.

The listening test revealed a lack of correlation between the airflow data and the ratings of perceived nasal quality. The fact that the airflow data and the audio material used in the test were not recorded in the same session may, however, have influenced the result.

Further, the perceived nasal quality did not show any relationship with the phoniatricians' visual ratings of VPO. Only in tenor 2, was a clear nasal quality perceived (see Figure A1); however, only a minor VPO was observed in his video recording. Also, a rather wide VPO was observed in some singers without causing a particularly high perceived degree of nasal quality. This supports the assumption that the degree of nasal quality is not related to the VPO size. In other words, singers seem capable of using even a wide VPO without adding a nasal quality to their vowel timbre.

The airflow data show that many professional operatic singers undoubtedly sing with a VPO on the vowels [a] and [u]. A later investigation showed certain acoustical consequences of a VPO, which seem to explain the varied shape and size of the VPO (Sundberg *et al*).

Conclusions

Given the difficulties in determining the presence of a VPO, our conclusions need to be conservative. Yet, clear evidence of a VPO was found in the vowels [a] and [u] for all singer classifications, at least under some conditions. Three main shapes of VPOs were observed by nasofiberscopy; a constricted opening, or an opening extending in the coronal or sagittal directions. This suggests that singers may use a VPO to fine-tune the vocal tract resonance characteristics and hence voice timbre, without contributing to a perceived nasal quality.

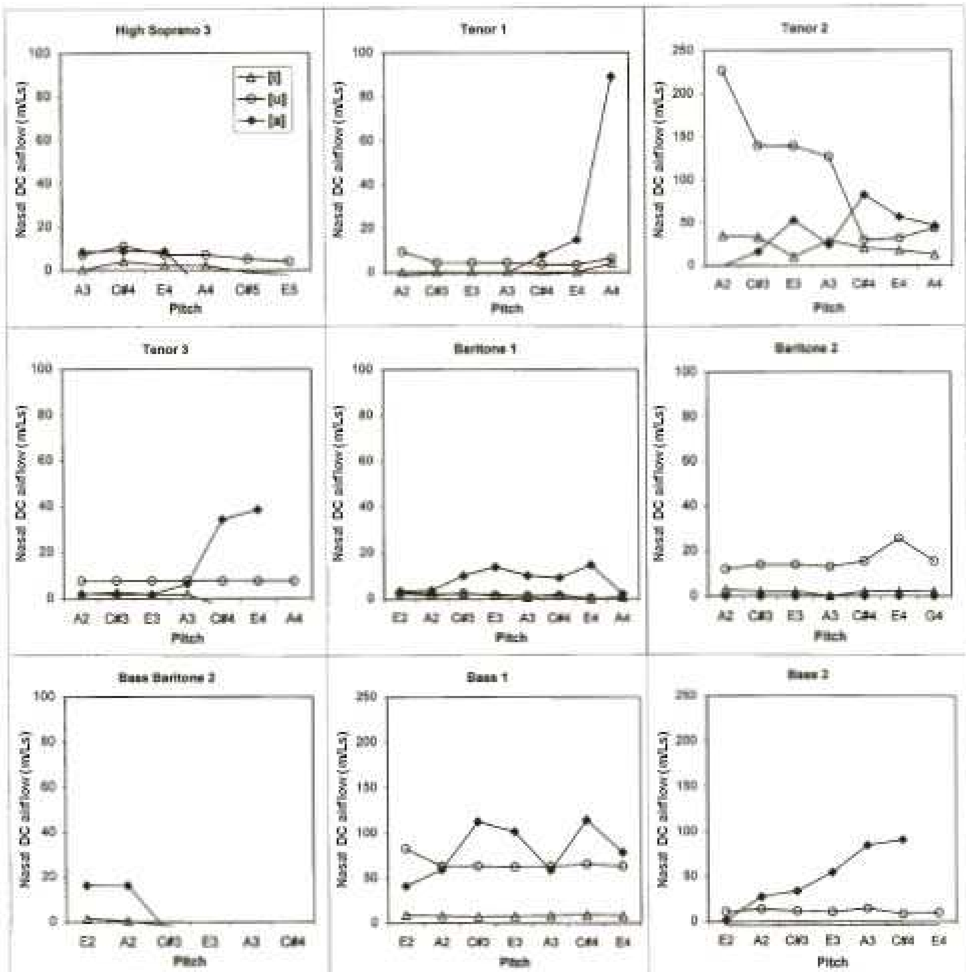


Figure A1 (Fig. 2 in paper A) Nasal DC airflow recorded during the second vowel in the test words [panta], [puntu], and [pinti] sung at middle degree of vocal loudness.

Reference:

Johan Sundberg, P. Birch, B. Gümoes, H. Stavad, S. Prytz and A. Karle Experimental Findings on the Nasal Tract Resonator in Singing. In Press *J. Voice* Available online 28 February 2006

Paper B

Throaty Voice Quality: Subglottal Pressure, Voice Source, and Formant Characteristics

Introduction
Voice quality is determined by formant frequencies and voice source characteristics. The perceptual characteristics associated with "throaty" voice quality are sometimes described as "strained, strangled, hypertense, swallowed, dark, tight, guttural" and even dysphonic. In addition, many voice pedagogues and therapists consider throaty quality as undesirable or even harmful to the voice. However, by contrast, throaty voice is often mentioned in music reviews of some non-classical vocalists as a positive characteristic of the singer's voice quality, thus suggesting that it is used as a timbral effect.

In this study data on formant frequencies, subglottal pressure, flow glottogram characteristics, and spectral data were collected, as well as area functions based on Magnetic Resonance imaging (MRI) in order to arrive at a multi-faceted description of throaty voice quality.

Aim
The purpose of the study was to identify the main acoustic characteristics of throaty voice quality and to elucidate its phonatory, articulatory, and resonatory correlates.

Method
One male and one female subject read a standard Swedish text twice; the first time with their habitual voice, and the second time with what they considered to be a throaty voice quality. During a second recording of each quality, in which the initial consonants of certain syllables were replaced by the consonant [p], the subjects held a plastic tube in the corner of the mouth to capture an estimation of P_{sub} during the p-occlusion.

A first listening test was run with a panel of five voice and speech specialists who evaluated sixteen syllables with respect to throatiness. Vowels rated as clearly throaty were selected for analysis. Voice source characteristics and formant frequencies were analyzed by means of inverse filtering. Long-term average spectrum was used to analyze the average spectrum characteristics.

A second listening test was carried out to test the relevance of formant frequencies to the perception of throatiness. In this test, listeners rated the throatiness of synthetic vowel stimuli produced by means of the KTH MUSSE synthesizer. The synthesis applied the formant frequencies measured in those vowels that had shown the greatest difference in perceived throatiness between the habitual and the throaty versions in the first listening test.

In a separate experiment using MRI, the male subject sustained the vowels [a:, i:, u:, ae:] in habitual and throaty quality. From this material, area functions were calculated and formant frequencies determined.

Results

The throaty samples of the male subject were perceived as somewhat throatier than those of the female subject. In contrast, the habitual voice samples produced by the male subject were perceived as less throaty than those of the female subject.

The voice source analysis revealed that for the male subject all parameters except MFDR showed significant differences between habitual and throaty voice quality (see Figure B1). In throaty voice as compared with habitual voice, P_{sub} was higher, pulse peak-to-peak amplitude lower, Q_{closed} higher, NAQ lower, and H_1-H_2 was lower. These differences suggest a more hyperfunctional/ pressed type of phonation in the male subject's throaty samples. In the female subject's throaty samples, MFDR was higher than in her habitually spoken samples. However, there were no other significant differences, neither in the flow glottogram parameters, nor in P_{sub}.

LTAS showed a higher level between 1 and 3 kHz in the male's throaty samples and between 1 and 4 kHz in the female voice. The spectral level in the region of F_0 was lower in the throaty samples, which implies a relatively weaker fundamental whereas the level near 0.7 kHz was higher. In habitual, but not in throaty voice, the male voice showed a sharp peak near 3.3 kHz, possibly an example of a "speaker's formant."

The formant frequencies in the throaty samples tended to show higher F_1 values than in habitual voice. For F_2 and F_3, the differences were small, even though lower values were observed for F_2 in front vowels. F_4 tended to be clearly lower in the throaty versions.

The MRI data revealed that the area functions for the male subject's various vowels showed expected characteristics (see Figure B2). Vowel [*a*:] had a narrow pharynx and a wide mouth cavity, [i:] had a wide pharynx and a narrow mouth cavity, and [u:] showed a large mouth cavity and constrictions near the velum and at the lips. For all throaty vowels, the lower part of the pharynx was consistently narrower than for habitual articulated vowels. This difference was particularly evident for [u:] and [*a*:].

Conclusions

Our throaty voice samples resulted from a narrowing of the pharynx, which may be combined with a somewhat hyperfunctional type of phonation. Acoustically, the narrow pharynx was characterized by an increase of F_1, a decrease of F_2 in front vowels, and a decrease of F_4. The synthesis experiment suggested that these formant frequency characteristics are reasonably typical for a throaty voice timbre. The voice source characteristics lead to an attenuation of the fundamental and an increase of the spectrum level between 1 and 3 kHz. The combination of formant characteristics and hyperfunctional phonation appears to be perceptually important for producing a throaty voice quality.

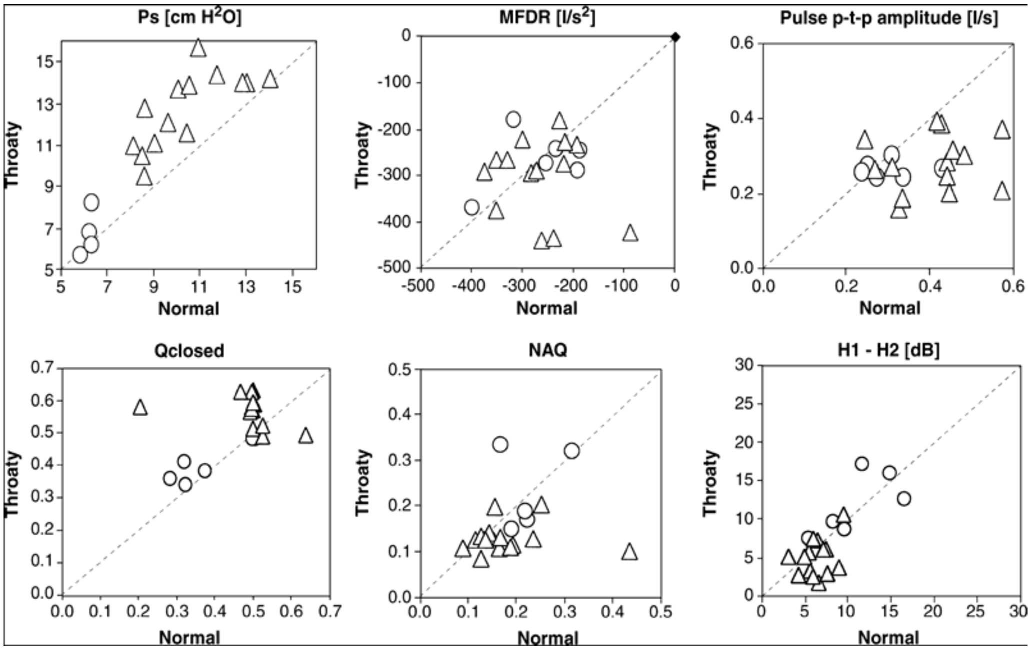


Figure B1 (Fig.2 in paper B). Scatter plots comparing P_{sub} and the indicated flow glottogram parameter values observed in the reading with habitual and throaty voices. For each vowel, the values for habitual voice are plotted along the x-axis and those for throaty voice along the y-axis. Circles and triangles refer to the values of the female and the male subjects, respectively. The graphs include data only for those samples that differed by more than 600 units (out of 1000) in mean rating of throatiness.

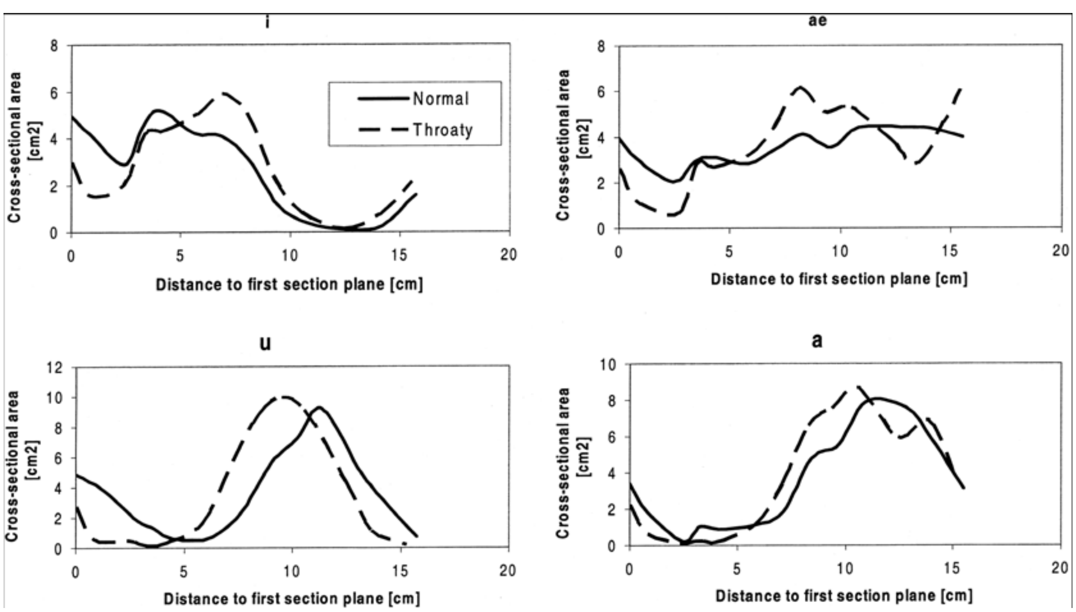


Figure B2 (Fig. 1 in paper B). Vocal tract cross-sectional areas as a function of distance to the center of gravity of section 1 for the productions of the indicated vowels in habitual and in throaty voice for the male subject.

Paper C

Voice Source Differences between Registers in Female Musical Theatre Singers

Introduction

Vocal register is a phenomenon of great relevance in vocal art. In the female singing voice a register transition generally occurs somewhere in the middle of the pitch range D4 – C5 (294–523 Hz). Most voices can produce several pitches in both registers near this transition. This implies that singers, depending on genre, vocal conditions, and musical expression, can choose to sing the same pitches in different registers.

Musical theatre singing typically requires women to use two vocal registers. Within the genre of musical theatre singing, the styles of singing can differ considerably depending on the epoch of time during which the piece was written. In the songs of early musical theatre productions (~1930-1960), the head register is commonly used like in classical/operatic singing. With the introduction of the rock musical in the 1960's singers were required to sing with more heavy/chest register. Today the whole spectrum of musical theatre repertoire from 1930 and onwards are played in musical theatre houses. This puts high demands on singers' ability to exert excellent control of both registers.

It is generally agreed that vocal registers are reflected in voice source characteristics, such as the relative duration of the closed phase, Q_{closed}, the peak-to-peak pulse amplitude (U_{p-t-p}), and the maximum flow declination rate (MFDR). These parameters are in turn heavily influenced by two physiological voice control parameters: subglottal pressure (P_{sub}) and glottal adduction. For example, P_{sub} regulates vocal loudness, and with increased glottal adduction Q_{closed} tends to increase, and U_{p-t-p} tends to decrease. Hence, it seems important to take these parameters into account in a study of vocal registers. In addition, the normalized amplitude quotient (NAQ) has been found to reflect glottal adduction. In a single- subject investigation of a female singer, NAQ was found to correlate with the degree of perceived phonatory pressedness and showed differences between styles of singing.

This investigation studies the register function in female musical theatre singers by analyzing their voice source characteristics and by paying special attention to the influence of P_{sub} on these characteristics, including NAQ.

Aim

The purpose of the study was to investigate the relation between subglottal pressure and voice source characteristics in chest and head registers in the female voice.

Method

Seven professional female musical theatre singers, aged between 17 and 43 years, all classically trained, volunteered as subjects. Their task was to sing a sequence of the syllable /pae:/ on a pitch where they could use both the chest and the head registers. This pitch varied between C4 (262 Hz) and G4 (392 Hz) for different subjects. The subjects initiated the sequence at high lung volume and at maximum degree of vocal loudness, and continued while gradually decreasing vocal loudness until softest possible. The oral pressure during the p-occlusion was used as an estimation of P_{sub}.

Recordings were made with a flow mask for capturing oral flow. For recording oral pressure (estimation of P_{sub}) the subject held a plastic tube in the corner of her mouth. The audio signal was recorded at a distance of 30 cm from the lips.

Ten equally spaced ideal values were selected over each singer's total P_{sub} range. The measured P_{sub} values closest to these ideal values were then identified from the three takes and the subsequent vowel was selected for analysis.

Informal listening to the samples revealed that some subjects produced very small differences in timbre between the registers. Hence, a computerized listening test was run with a panel of three voice experts. Their task was to rate how representative the 280 sung samples were of chest and head register. The results revealed that a considerably greater number of the samples were perceived as being sung in head than in chest register. Therefore, 34 phonations from five singers, rated as the most typical samples were analyzed (17 chest register and 17 head register samples).

Inverse filtering was performed, and the parameters Q_{closed}, $U_{p\text{-}t\text{-}p}$, MFDR, and NAQ were derived from the resulting flow glottograms.

Results

Figure C1 illustrates the differences between the registers for the means across the 34 clear cases. The mean and SD of P_{sub} were higher for the chest register samples, which also had higher Q_{closed} and lower NAQ means, and somewhat higher $U_{p\text{-}t\text{-}p}$ values.

The relationship between the singers' ten P_{sub} values and Q_{closed} for the two registers showed obvious register differences. For a given pressure, Q_{closed} tended to be higher in chest register, although the differences were smaller at lower P_{sub} values. The relationship between the ten P_{sub} values and the other glottal parameters for each singer showed that for increasing P_{sub}, (a) MFDR became more strongly negative in both registers, (b) $U_{p\text{-}t\text{-}p}$ tended to increase with increasing P_{sub} (the intersubject scatter could reflect inter-individual differences, eg, with respect to vocal fold length), and (c) NAQ decreased with increasing P_{sub}, and chest register values were lower than head register values.

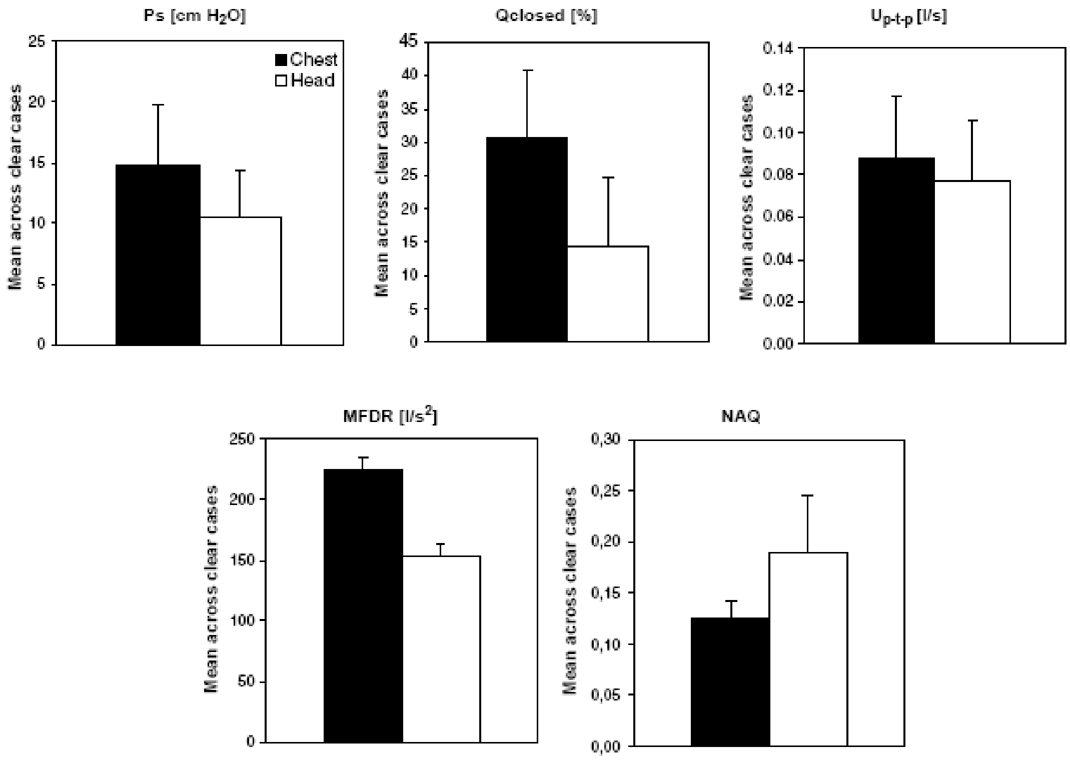


Figure C1 (Fig. 4 in paper C)**.** Means across the 34 clear cases of the indicated parameters for the two registers. The bars represent one standard deviation.

Given the relevance of P_{sub} to glottal flow waveform properties, comparisons of glottal parameters at identical P_{sub} values are informative. The subjects' P_{sub} ranges differed considerably, but four of the singers used a P_{sub} value of about 11 cm H_2O somewhere in their recordings. Figure C2 shows the four subjects' 11-cm H_2O samples. The sound level was higher in chest register (Fig. C2 a), which seems to be in agreement with the typical observation that head register at low fundamental frequency is difficult to combine with loud phonation. The higher sound level in chest register corresponded to a more negative MFDR (Fig. C2 b).The closed phase was clearly longer in chest register (Fig. C2 c), which should produce stronger overtones including the second partial. U_{p-t-p} did not differ consistently between the registers (Fig. C2 d), and varied among singers, and NAQ was consistently lower in chest than in head register (Fig. C2 e).

Conclusions

The results showed that for typical tokens performed by our female musical theatre singers, MFDR, U_{p-t-p} and Q_{closed} were significantly greater, whereas NAQ were significantly lower in chest register than in head register. At the pitches used, the second partial is close to F1, enhancing F1, which could have contributed to the overall higher sound level in chest than in head register.

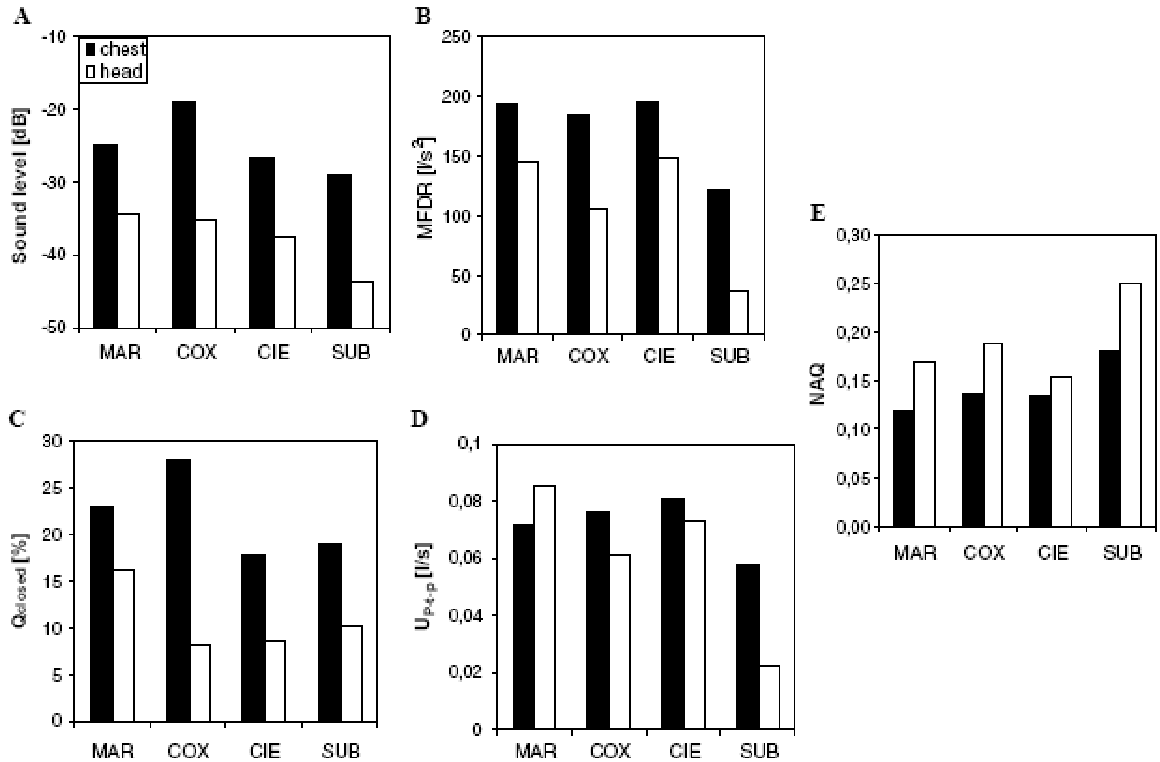


Figure C2 (Fig. 8 in paper C). (a) Sound level (b) MFDR, (c) Q_{closed}, (d) Up-t-p, and (e) NAQ for the phonations of the singers at a P_{sub} of about 11 cm H_2O in chest and head register.

Paper D

Subglottal Pressure and Normalized Amplitude Quotient Variation in Classically Trained Baritone Singers

Introduction

Since its introduction, voice source parameters related to glottal flow like amplitude quotient (AQ) and its normalized version (NAQ) have been used in a number of studies of speech. Separate studies have addressed voice quality, vocal dynamic extremes, and emotional colouring. Both AQ and NAQ have been found to differ between phonation modes. Compared to spontaneous speech, singing is a much more accurately controlled phonation task. For instance, speakers tend to raise their pitch with increasing vocal loudness, a habit obviously unacceptable in singers. In this study voice source characteristics of five professional baritone singers were analyzed.

AQ and NAQ are closely associated with glottal conditions since they are computed from the basic flow waveform parameters $U_{p\text{-}t\text{-}p}$ and MFDR. NAQ has only been used in two earlier studies on singing, both concerning its relationship with P_{sub} and NAQ in female singers. In a single-subject study, NAQ was found to correlate with perceived degree of phonatory pressedness, and to differ between both singing styles and phonation modes (Sundberg *et al.* 2004). Another study, based on register differences in five musical theatre singers, showed that for a given P_{sub} value NAQ differed between registers (Paper E, 2006).

A fuller understanding of the information that AQ & NAQ offer about the voice source can be ideally be promoted by using professional singers as subjects, who have learnt a consistent phonatory behaviour. In the present study a group of male operatic singers were used to examine the effect of P_{sub} variation on AQ & NAQ. In a companion paper (paper E) an extended set of parameters were used for comparison with a group of male musical theatre singers.

Aim

The purpose of the study was to investigate the relationships between the subglottal pressure and the amplitude quotient (AQ) and the normalised amplitude quotient (NAQ), respectively, in professional singers for wide ranges of pitch and loudness.

Method

The singers' task was to sing the syllable [pae:] during a diminuendo, starting at maximum vocal loudness and gradually decreasing to softest possible. The singers were asked to sing the task at a high, medium, and low pitch (278, 196, and 139 Hz) and repeat each F0 three times. The recording setup included a flow-mask, microphone, and a pressure transducer, connected to a plastic tube in the singers' mouth. Ten equally spaced P_{sub} values were selected from the three repeats, and the corresponding vowels were inverse filtered. Data on $U_{p\text{-}t\text{-}p}$, MFDR, AQ and NAQ were calculated.

The P_{sub} values were converted to the normalized excess pressure (P_{sen}). It represents a relative measure of the P_{sub}, basically compensating for the F0 dependence of P_{sub}. This normalization is preferable since a doubling of F0 is typically associated with a doubling of P_{sub}.

Results and discussion

The five singer's P_{sub} data were highly structured and correlated strongly with F0; P_{sub} was consistently higher the higher the pitch (see Figure D1). The P_{sub} ranges clearly differed between subjects. However, they all showed a similar relationship between P_{sub} and F0; all singers approximately doubled their P_{sub} for the higher F0 as compared to the lower F0, i.e., for an octave interval.

Both AQ and NAQ decreased with increasing P_{sub}, particularly at low P_{sub} values, and reached an asymptote-like value at high P_{sub} (see Figure D2). AQ and NAQ also differed between the F0 values. Surprisingly, the non-normalised AQ parameter, showed much less variation with F0 than the normalised NAQ parameter.

The interpretation of these results is not clear-cut. If we assume that NAQ reflects phonation mode, the results suggest that the singers kept phonation mode constant when they varied vocal loudness, but that they varied phonation mode depending on F0. This interpretation seems unlikely as subjects were professional singers. A more plausible interpretation would be that AQ reflects changes between phonation mode within a singer. Under that assumption, the results suggest that the singers kept phonation mode independent of both loudness and pitch. However, the relationships between both AQ and NAQ, on the one hand, and perceived degree of pressedness, on the other, remain open questions for future investigation.

Conclusion

Our baritone singers showed a linear increase of P_{sub} with F0. AQ and NAQ were found to be principally unaffected by increases of P_{sub}. However, NAQ differences were found between the high and the low F0 value. The AQ values, by contrast, remained basically unaffected by F0. It seems likely that classically trained singers keep phonation mode independent of loudness and pitch. If so, AQ is likely to reflect phonation mode. Nevertheless, more studies are needed to fully understand the information about voice production that is offered by AQ and NAQ.

Reference:

Sundberg J, Thalén M, Alku P, Vilkman E. Estimating perceived phonatory pressedness in singing from flow glottograms, *J Voice*. 2004:18; 56-62

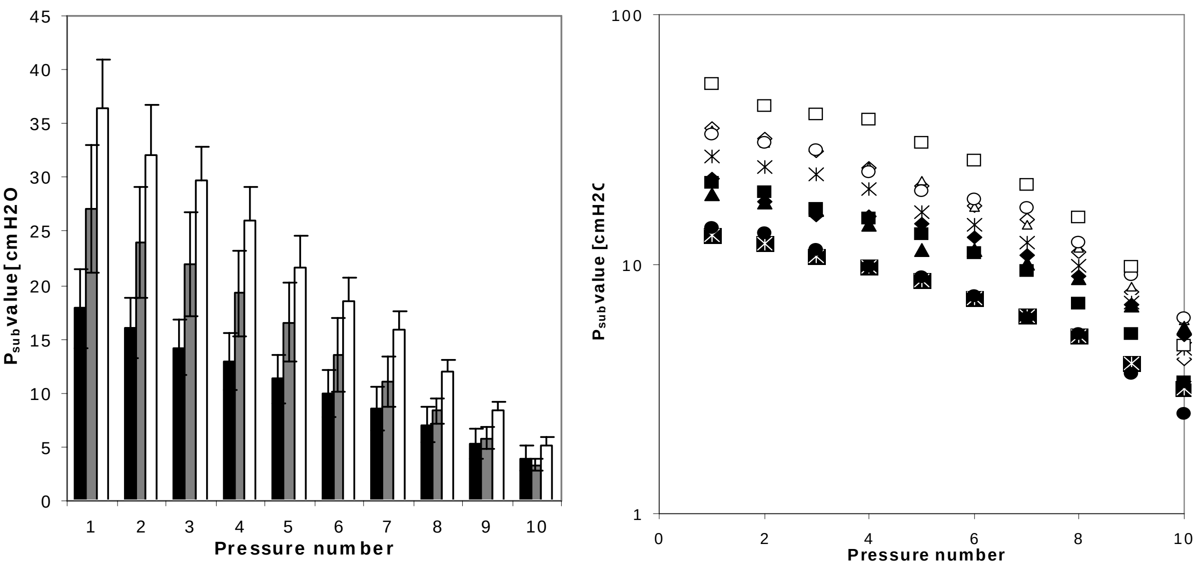


Figure D1 (Fig. 3 in paper D). (a.) The singers' mean subglottal pressures P_{sub} for each of the ten selected pressures. Black, grey and white columns represent the fundamental frequencies of 139 Hz, 196 Hz and 278 Hz, respectively. (b) Each singer's P_{sub} values, plotted along a logarithmic scale, for each of the ten selected pressures. Filled and open symbols refer to the fundamental frequencies of 139 Hz and 278 Hz, respectively.

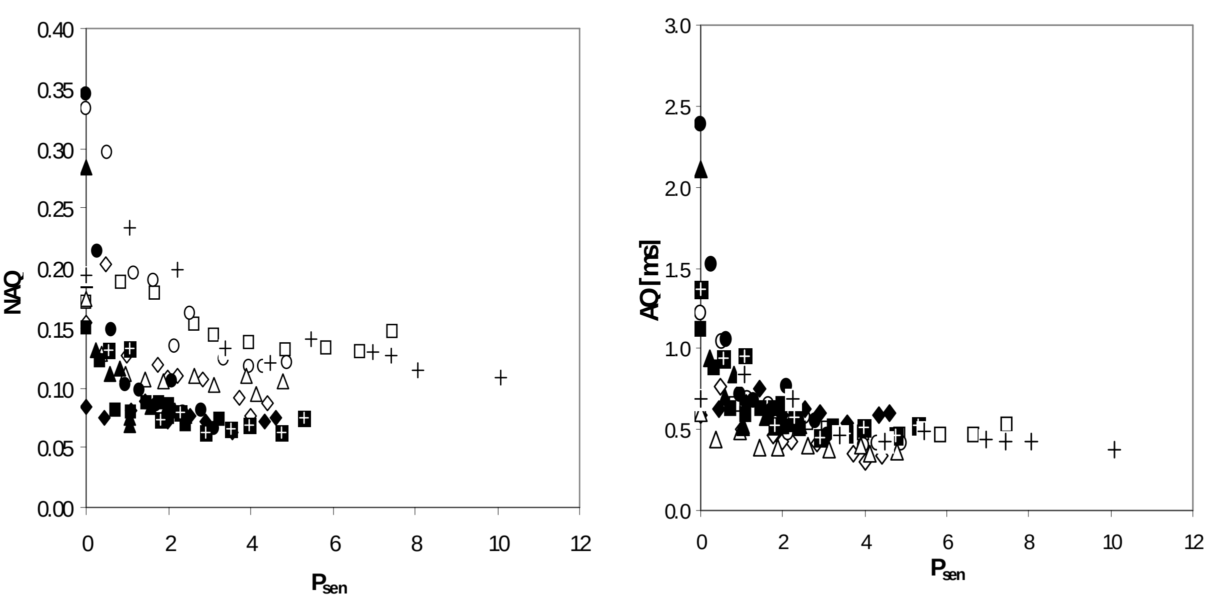


Figure D2 (Fig. 6 in paper D) NAQ and AQ values (left and right graph) for each of the five singers´ ten selected samples as function of normalized excess pressure P_{sen}. Each symbol represents one subject. Filled and open symbols refer to low and high F0 (139 and 278 Hz).

Paper E

Musical Theatre and Opera singing – Why so different? A study of subglottal pressure, voice source and formant frequency characteristics

Introduction

Vocal music is a large artistic field where pop, rock, soul, world music, jazz occupy an important portion. The voice timbre varies widely depending on style of singing. For instance, the timbre of musical theatre (MT) and of opera singers differs considerably. The origin of this difference is however unclear. Furthermore, there are not only vocal differences associated with the styles. Different working demands are also put on opera singers and on MT singers: MT singers use microphones and might be required to sing 7-8 performances a week, while opera singers sing without electronic amplification and rarely have more than two or three performances per week. Consequently, depending on style singers use their vocal instrument differently.

The majority of scientific studies of the singing voice have been devoted to the classically trained male voice. Only in the last two decades MT singing has been studied, in particular female belting. The term belting most commonly refers to a powerful voice quality used at higher pitches. It typically occurs as a timbral effect in female MT singing.

In view of the limited knowledge about voice production in MT singing, this study was designed to investigate (a) male singers, (b) a wide range of vocal dynamics, (c) a wide pitch range, and (d) differences between MT and operatic singing.

The investigation uses material recorded for Paper D for the operatic singers while for the MT singers new recordings were made. Identical experimental protocol was applied in both sets of recordings.

Aim

The purpose of this study was to gain a better understanding of the phonatory and resonatory origins of the salient timbral differences between musical theatre (MT) singers and of western operatic singers. The investigation examines voice production characteristics over a wide range of vocal intensities and fundamental frequencies (F0).

Method

The voice source parameters analyzed in Paper D (P_{sub}, AQ, and NAQ) were complemented by peak-to-peak pulse amplitude ($U_{p\text{-}t\text{-}p}$), maximum flow declination rate (MFDR), and closed quotient (Q_{closed}). The analysis also included some acoustic characteristics, such as the relation between the first and the second harmonic (H1-H2), the formant frequencies, as well as the sound pressure level (SPL). A combination of electroglottography (EGG) and inverse filtering was used to gain information about voice source characteristics, specifically the closed phase. This extending of the parameter set was motivated by the need for covering the entire process of voice production, including air flow control, voice source, spectral balance, and vocal output.

The experimental protocol, recording methods and analysis procedures were identical with those used for paper D. Audio, P_{sub}, and EGG signals were recorded while the MT subjects sang the same tasks. Thus, the singers' task was to repeatedly sing the syllable [pae:] during a diminuendo, starting at maximum vocal loudness and gradually decreasing to softest possible. The performed this task at a high, medium, and low pitch (278, 196, and 139 Hz) and repeated each F0 three times. The results were compared with those obtained from the professional opera singers in paper D. Flow glottogram characteristics were analyzed for the high and the low F0 only.

Results and discussion

Both groups varied P_{sub} quite systematically, and used higher P_{sub} not only for increased vocal loudness but also for increased F0. However, the MT singers tended to use higher P_{sub} than the opera singers for both F0 values studied (see Figure E1). MFDR increased with increasing P_{sub}, as expected.

Clear differences could be observed between the groups. The MT singers' higher P_{sub}, values were associated with clearly higher MFDR values, especially for 278 Hz. Further, the MT singers' peak flow amplitudes $U_{p\text{-}t\text{-}p}$ were surprisingly high as compared to the opera singers.

The closed quotient Q_{closed} increased with increasing P_{sub} and seemed to saturate for the loudest phonations, i.e. at high P_{sub} (see Figure E2). The MT singers showed higher Q_{closed} values than the opera singers, who showed notably lower values at high F0. As expected, the closed phases were longer (higher Q_{closed} values) for the low F0 in both groups.

In speech, a long closed phase is often associated with pressed phonation. However, the MT singers' consistently higher Q_{closed} values combined with the high values of $U_{p\text{-}t\text{-}p}$ do not necessarily indicate a pressed type of phonation. This assumption is supported by the MT singers' relatively strong fundamental as reflected in their high H1-H2 values (see Figure E3). Yet, the opera singers showed consistently higher H1-H2 values than the MT singers, although they used lower P_{sub} values. This means that, for a given relative P_{sub} value the opera singers produced a stronger fundamental than the MT singers. The difference was rather small, particularly in loud voice, but may nevertheless contribute quite importantly to the typical timbral differences between operatic and MT singers.

A physiological difference between the two singer groups underlying this difference may be vocal fold thickness, achieved by a stronger contraction of the vocalis muscle. A firmer contraction of vocalis would lead to thicker vocal folds.

The MT singers showed higher formant frequencies than the opera singers, presumably because of a somewhat higher larynx position. Unlike the opera singers the MT singers tended to raise their formant frequencies slightly with rising F0. Further, they did not, as the opera singers, tune F3, F4, and F5 closely together into a singer's formant cluster.

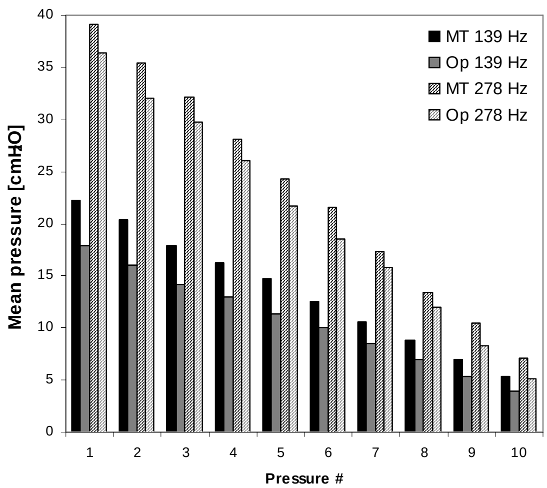
For a given P_{sub} the MT singers produced higher SPL values, a likely consequence of their higher MFDR values (Fig E4). These in turn would depend on the fact that for a given P_{sub} the MT singers reached a higher $U_{p\text{-}t\text{-}p}$ and a higher Q_{closed}. Both these differences might be related to a combination of thicker vocal folds and looser mucosa, possibly resulting from a firmer contraction of the vocalis muscle.

Both groups of singers changed MFDR and $U_{p\text{-}t\text{-}p}$ proportionally, producing similar NAQ and AQ values for a given P_{sub}. Interestingly, the NAQ parameter, normalizing the $U_{p\text{-}t\text{-}p}$ / MFDR ratio with regard to F0 period, differed substantially between the high and the low F0 while the un-normalized AQ value was similar for both F0 values. Another noteworthy and unexpected result was that neither AQ, nor NAQ differed between the singers groups.

In earlier investigations these parameters have been shown to reflect phonation mode, such that hyperfunctional voices produce low NAQ values. Most voice experts would agree that opera singers tend to avoid habitual use of hyperfunctional voice. The similarity in NAQ and AQ seems to suggest that neither the opera singers, nor the MT singers used hyperfunctional voice. The full understanding of the information offered by the NAQ and AQ parameters appear to require some further investigation. Thereby, it may be worthwhile to take into consideration also the relationship between $U_{p\text{-}t\text{-}p}$ and Q_{closed}.

Conclusions

Both the MT and the opera singers varied their P_{sub} systematically, tending to double their P_{sub} for a doubling of F0. For a given value of P_{sub} the MT singers produced higher values of MFDR, of Q_{closed}, and of $U_{p\text{-}t\text{-}p}$. In addition, for a given relative P_{sub}, they produced a smaller value of H-H2. All these phonatory differences between the singer groups can be expected if the MT singers sang with a somewhat firmer contraction of the vocal muscle causing thicker vocal folds and a looser mucosa. With respect to the resonatory characteristics, the MT singers had higher formant frequencies and did not show the opera singers' characteristic clustering of F3, F4, and F5.



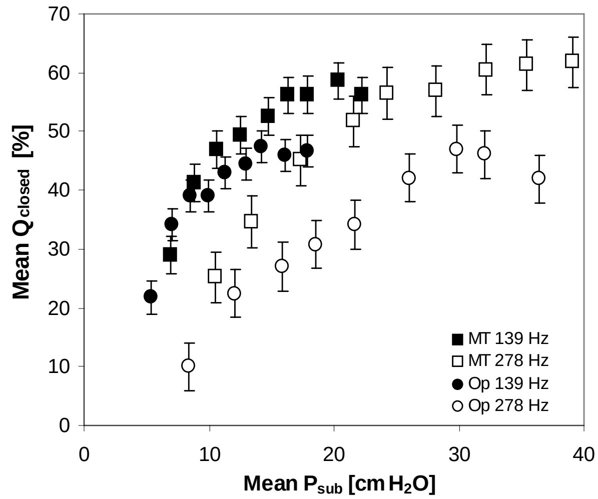
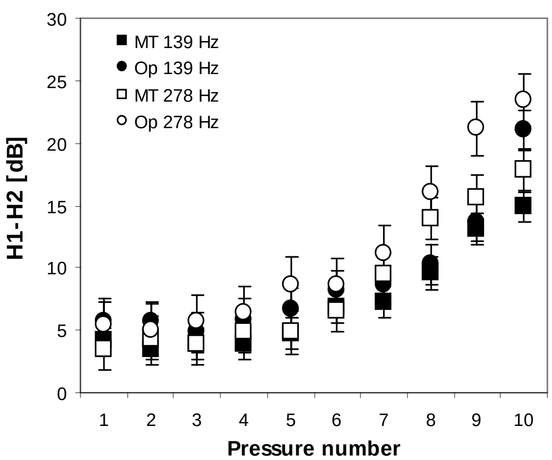


Figure E1 (Fig. 3 in Paper E*) Mean pressures for the ten selected P_{sub} values observed at the indicated F0 values as produced by the musical theatre (MT) and opera (Op) singers.*

Figure E2 (Fig. 6 in Paper E). Mean Q_{closed} as a function of mean P_{sub} for the musical theatre (MT) and the opera (Op) singers. The bars represent ± one standard error.



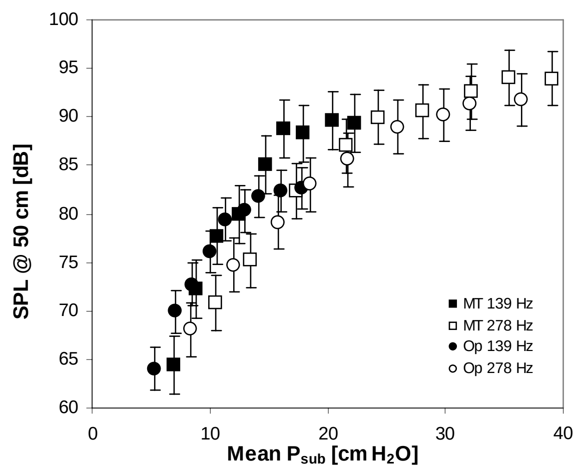


Figure E3 (Fig. 7 in Paper E). Mean H_1-H_2 as a function of ten selected pressure numbers, i.e., in terms of percentage of the singers' total pressure ranges. Low numbers refer to high pressures. The bars represent ± one standard error.

Figure E4 (Fig. 9 in Paper E). Mean SPL as a function of mean P_{sub} for the musical theatre (MT) and the opera (Op) singers. The bars represent ± one standard error.

General discussion

Terminology is a notorious problem in the voice area. The terms used do not necessarily have the same meaning to different people, and indeed not even to all voice experts. A promising solution is to use terms that allude to physiological landmarks. An example is "throaty" voice, a term sometimes used to describe a special timbral effect occurring in speech but also in some female pop singers. However, the relationship between the timbral characteristics of this type of voice and the associated physiological properties was far from well established. An attempt to an unprejudiced approach was taken in Paper B. The results revealed that the term "throaty" is quite well-chosen in the sense that it alludes to an articulatory setting actually associated with a pharyngeal constriction. As expected it was found to affect the lowest formant frequencies. On the other hand the results also suggested that throaty voice was produced with a slightly raised P_{sub}, a higher Q_{closed} and a lower $U_{p\text{-}t\text{-}p}$, all phonatory characteristics typical of hyperfunctional voice.

Another example of a term alluding to a physiological landmark is "nasal quality", the topic examined in Paper A. In this case, the results provided little evidence in support for the term. No relationship whatsoever was found between an expert panel's perception of "nasal quality" and the occurrence of a velopharyngeal opening to the nasal cavity in professional operatic singers. The results also indicated that such an opening was quite common in the basses, baritones, tenors and mezzo-sopranos, but not in the sopranos.

A later investigation revealed that a velopharyngeal opening may entail resonatory effects enhancing the singer's formant, a spectrum envelope peak occurring in the neighborhood of 3 kHz in classically trained male singers. It is interesting that few of the eight female singers included in the investigation showed signs of a velopharyngeal opening. The results thus show that the term "nasal quality" has little relationship with the coupling between the oral and nasal cavities.

Summarizing, Paper B identified a relationship between a perceptual term and the associated physiological characteristics; throaty was associated with a narrow throat. Paper A, however, found little relationship between the term "nasal quality" and an opening connecting the nasal cavity to the vocal tract. In other words, it seems risky to trust the allusions to physiology, suggested by some terms used to describe voice qualities. A better alternative would be to describe the phonatory and resonatory characteristics. On this ground a more useful and reliable terminology can be built.

Given the problematic state of the terminology used for voice timbres it is sometimes important to have a listening panel assess the recorded stimuli. Thus, in Paper A an expert panel was used to evaluate the degree of nasal quality, and the outcome of this test was then compared with the physiological characteristics, as mentioned. In Paper B, an expert panel was used to select typical examples of throaty voice, and in Paper C an expert panel was asked to rate how representative sound examples were of chest and head register. In Papers D and E, however, no listening test was needed, since the subjects used in these investigations were earning their livelihood from performing in opera and in musical theatre; it seemed meaningless to question whether they were typical representatives for opera and MT singing, respectively.

The subglottal pressure P_{sub} has a major influence on the voice source. Consequently it was important to take P_{sub} into account in most papers. In some papers it was expressed in real values of cm H_2O. This yielded valuable information in Paper C and E where it revealed that for a given pressure, differences existed between registers and singer categories. In some papers P_{sub} was given also in terms of the normalized excess pressure P_{sen} proposed by Titze (Titze 1992). P_{sen} basically compensates for the fact that the threshold pressure increases with increasing F0, relating P_{sub} to the phonation threshold pressure, which, however, is sometimes difficult to determine accurately. The opera singers' lower threshold pressures in Paper E resulted in a considerably wider P_{sen} range than that of the MT singers. This appeared to be somewhat misleading, as the MT singers' total pressure ranges exceeded that of the opera singers only marginally. A third alternative for specifying P_{sub} is to express subglottal pressures in terms of numbers corresponding to percentages of the individual singer's total pressure range. This turned out to be particularly rewarding in Paper E, where the different singer groups used different ranges of P_{sen}. Indeed, if only P_{sub} or P_{sen} had been considered in this investigation, the fact that the opera singers' voice source fundamental was stronger than that of the MT singers' would have been obliterated.

Pressed phonation is a voice quality often assumed to characterize singing in the non-classical popular repertoire. In vocology the term is used for phonatory hyperfunction, an abnormality, frequently leading to vocal fold pathology. In singing, on the other hand, a pressed voice quality is often used as timbral ornament, particularly in rock music and in rock musicals. This poses the question whether singing in the non-classical popular repertoire is a threat to vocal health.

A high Q_{closed} is sometimes regarded as an indicator of a pressed voice. This assumption was refuted by the results in Paper E, showing that an increase of Q_{closed} may result from a variety of phonatory conditions. For example, for low and moderate P_{sub} values an increase of P_{sub} increases Q_{closed}. Further, a thickening of the vocal folds resulting from an increased contraction of the vocalis muscle was assumed to underlie the male MT singers' high Q_{closed} values in Paper E as well as the female MT singers' chest register in Paper C. In any event, a high Q_{closed} value alone is not a reliable indicator of a pressed voice.

Another sign of a pressed phonation mode is a low pulse amplitude of the flow glottogram. Operatic singers are generally assumed to avoid pressed phonation. In Paper E their $U_{p\text{-}t\text{-}p}$ values were actually found to be slightly lower than those of the MT singers. This supports the idea that pressed phonation is not a characteristic of MT singing.

AQ and NAQ have been proposed as criteria of phonation modes, a lower value indicating a more pressed phonation. Almost identical values of these parameters were observed in the operatic and in the musical theatre singers. This is a further indication that both singer groups avoided pressed phonation.

Conclusions

The question posed in the title of this thesis was why voice timbre differs between musical theatre and opera singers. The included studies aimed at identifying underlying phonatory and resonatory characteristics. A number of common or diverging characteristics were found:

- Both operatic and musical theatre singers tend to double their subglottal pressure for a doubling of fundamental frequency.

- Compared with operatic singers musical theatre singers
  a) use somewhat higher subglottal pressure,
  b) produce higher maximum flow declination rate,
  c) produce higher sound pressure levels,
  d) have higher closed quotient
  e) have higher peak-to-peak flow glottogram pulse amplitude, and
  f) have a less dominating voice source fundamental.

- Operatic singers tend to use a carefully varied velopharyngeal opening which, however, shows little relationship with the perceptual quality of "nasal resonance".

- Female subjects' shifting from modal/chest to middle/head register is associated with decrease of subglottal pressure, of MFDR, of closed quotient, of peak-to-peak flow glottogram pulse amplitude, and with an increase of NAQ.

- When singers increase subglottal pressure, NAQ and AQ decrease and soon reach an asymptotic value.

- When singers double F0, NAQ increases while AQ remains basically constant.

- Throaty voice is associated with a narrow pharynx and typically seems to be combined with a hyperfunctional voice.

Acknowledgements

This work would never have been possible without help from several generous people. I particularly want to thank:

Johan Sundberg, my supervisor and dear friend, for your deep trust in me and your great supervision, generosity, inspiration, and humour.

Paavo Alku, my co-supervisor, for accepting me as a doctoral student in the HOARSE-network that funded this thesis and made it a reality, for your supportive supervision, and your great sense of humour.

Anders Askenfelt, my supervisor, for your steady clog in "Formantorkestern", and your devoted support in completing this thesis.

Svante Granqvist, for your wonderfully clever and user-friendly programs making my research possible, for your never ending patience with my never ending questions, and for being such a great friend.

I would also like to thank *Mattias Heldner*, for kind help with the statistical analysis, and *Peta Sjölander*, for proof reading this thesis and for your friendship, and *Sten Ternström*, not for only interesting conversations about acoustics, but also for your inspiring musicality and joy for vocal music.

This work was carried out at three research laboratories:

In Finland at the Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing (TKK), funded by the European Community's Human Potential Programme under contract HPRN-CT-2002-00276 [HOARSE-network],

in Sweden at the Royal Institute of Technology, Department of Speech, Music and Hearing (TMH),

in England at the University of York, Department of Electronics, with an Institutional Grant IG 2002-2049 from STINT, the Swedish Foundation for International Cooperation in Research and Higher Education.

I would also like to acknowledge *Patricia Grammings minnesfond* that contributed to some of the traveling and conference expenses.

During my time at three departments in three different countries I have come to meet and to get to know many wonderful and colorful people. To acknowledge and put into words what each one of you has meant to me would possibly make another thesis so I simply wish that you know in your hearts what I would have written. Thank you for all the unforgettable moments of laughter, music, and leisure:

all members at TMH, especially my colleagues and friends in the music acoustics group during the years, *Monica Thomasson, Peta Sjölander, Svante Granqvist, Roberto Bresin, Anders Friberg, Erwin Schoonderwaldt, Anick Lamarche, Sofia Dahl, and Kjetil Falkenberg Hansen.*

all members at TKK, especially, *Matti Airas, Henri Penttinen* and *Laura Lehto*.

and all colleagues in York, especially *Jonathan Thorpe* and *Sandra Pauletto*.

and outside work, my wonderful friends whom I could not do without.

Finally, I would like to thank my family for unconditional love and support; my parents *Curt* and *Berit*, my sister *Christine,* and my brother *Tom*.

This thesis is dedicated to my mother, who inspired me to devote my life to singing.
The choir of angles has got itself a top soprano.

Berit Björkner

1940 - 2006

References

Addington D W. (1968) The relationship of selected vocal characteristics to personality perception. *Speech Monogr*. 35: 492–503.

Adler D R. (2204) A Love Affair. Tribute to Ivan Lins, 2004. Available at http:// www.allaboutjazz.com/php/article.php?id=1691

Airas M, Alku P. (2004) Emotions in short vowel segments: effects of the glottal flow as reflected by the normalized amplitude quotient. Affective Dialogue Systems Workshop 04, Kloster Irsee, Germany, June 2004**.**

Alku P. (1992) Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. *Speech Comm*. 11: 109–118.

Alku P, Vilkman E. (1996a) A comparison of glottal voice source quantification parameters in breathy, normal, and pressed phonation of female and male speakers. *Folia Phoniat*. 48: 240–254.

Alku P, Vilkman E. (1996b) Amplitude domain quotient for characterization of the glottal volume velocity waveform estimated by inverse filtering. *Speech Comm*.18: 131–138.

Alku P, Bäckström T, Vilkman E. (2002) Normalized amplitude quotient for parameterization of the glottal flow. *J Acoust Soc Am*. 112: 701–710.

Austin S. (1997) Movement of the velum during speech and singing in classically trained singers. *J Voice.* 11: 212–221.

Austin S. (2000) Nasal resonance—fact or fiction? *NATS J Sing.* Nov/Dec: 33–34.

Baer T, Gore J C, Gracco L C, Nye PW. (1991) Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *J Acoust Soc Am*. 90: 779–828.

Bele I. (2002) Professional Speaking Voice. A Perceptual and Acoustic Study of Actors' and Teachers' Voices. (PhD thesis) Department of Special Needs Education, University of Oslo, Norway.

Berg J van den. (1958) Myoelastic-aerodynamic theory of voice. *J Speech Hear Res.* 1(3): 227-244.

Berg J van den. (1968) Mechanisms of the larynx and the laryngeal vibrations. In Malmberg B. Manual of Phonetics. Amsterdam: *North Holland Publishing Company* 278- 308.

Bestebreurtje M E, Shutte H K (2000) Resonant strategies for the belting style: Results of a single female subject. *J Voice*. 14:1 94-204

Boone DR. (1991) Is Your Voice Telling on You? How to Find and Use Your Natural Voice. *San Diego, CA: Singular Publishing Group*.

Björkner E, Sundberg J, Cleveland T, Stone E. (2006) Voice source characteristics in different registers in classically trained female musical theatre singers. *J Voice*. 20(2): 187-197.

Björkner E, Sundberg J, Alku P. Subglottal pressure and Normalized Amplitude Quotient variation in classically trained baritone singers. In Press *Logoped Phoniatr Vocol*. 2005.

Broad D J (1979) The new theories of vocal fold vibration. *Speech and Language: Advance in Basic Research and Practice* 2: 203-256.

Brodnitz F. (1965) Vocal Rehabilitation. *Rochester. Minn: Custok Printing Inc.*

Bouhuys A, Mead J, Proctor D F, Stevens K N. (1968) Pressure-Flow Events during Singing. In *Annals of the New York Academy of Sciences* (M. Krauss, M. Hammer, & A. Bouhuys, editors) 155: 165-176.

Campbell N, Mokhtari P. (2003) Voice Quality: the 4th prosodic dimention. In Proc. of the 15th ICPhS Barcelona 2003: 2417-2420.

Catford J C. (1977) Fundamental Problems in Phonetics. Edinburgh: *Edinburgh University Press.*

Childers D G, Hicks D M, Moore G P, Alsaka A Y. (1986) A model for vocal fold vibratory motion, contact area, and the electroglottogram. *J Acoust Soc Am*. 80 (5):1309-1320.

Childers D G, Wong C-F. (1994) Vocal source-tract interaction. *IEEE Transactions on Biomedical Engineering*. 41(7): 663 -671.

Cleveland T, Sundberg J. (1985) Acoustic analyses of three male voices of different quality, in A Askenfelt, S Felicetti, E Jansson, & J Sundberg, (ed.): *SMAC 83. Proceedings of the Stockholm International Music Acoustics Conference, Stockholm: Royal Swedish Academy of Music.* 46 (1): 143-156.

Cleveland T F, Stone R E, Sundberg J, Iwarsson J. (1997) Estimated subglottal pressure in six professional country singers. *J Voice.* 11: 403-409.

Coward H. *Choral Technique and Interpretation*. Chapter 3: Voice. Available at: http://www.hartenshield.com/choral_technique_03.html

Engwall O. (2002) Tongue talking studies in intraoral speech synthesis. (PhD thesis) Stockholm, Sweden: Department of Speech, Music and Hearing, KTH.

Estill J. (1988) Belting and classic voice quality: some physiological differences. *Med Probl Perform Artists.* 3: 37-43.

Evetts E, Worthington R. (1928) The Mechanics of Singing. *London: JM Dent & Sons Ltd*.

Fant G. (1960) Acoustic Theory of Speech Production. *The Hague, Mouton.*

Fant G. (1966) A note on vocal tract size factors and nonuniform *F*-pattern scalings. *STL-QPSR*, 1966/1: 22–30.*

Fant G. (1982) Premiminaries to analysis of the human voice source. *STL-QPSR* 1982/4: 1-27.*

Fant G. (1993) Some problems in voice source analysis. *Speech Comm*. 13: 7-22.

Fant G, Lin Q, Gobl G. (1985) Notes on glottal flow interaction. *STL-QPSR*, 1985/2-3: 21-45.*

Fant G, Lin Q. (1987) Glottal source - Vocal tract acoustic interaction, *STL-QPSR,* 1987/1: 133- 127.*

Fant G, Lin Q. (1988) Frequency domain interpretation and derivation of glottal flow parameters. *STL-QPSR* 1987/2–3: 1–21.*

Fant G, Kruckenberg A, Liljencrants J, Båvegård M. (1994) Voice source parameters in continuous speech. Transformation of LF parameters. In Proceedings of the

International Conference on Spoken Language Processing 1994 (Yokohama) : 451–1454.
Fant G. (1997) The voice source in connected speech. *Speech Comm.* 22: 125–139.
Fant G, Liljencrants J, Lin Q. (1985) A four-parameter model of glottal flow. *STL-QPSR***,** 1985/4: 1–13.*
Fabre P. (1957) Un proce´de´ e´lectrique percutane´ d'inscription de l'accolement glottique au cours de la phonation: Glottographie de haute fre´quence,'' *Bulletin de l'Acade´mie Nationale de Me´decine*, : 66–69.
Fisher H B. (1975) Improving Voice and Articulation. *2nd ed. Boston MA*: Houghton Mifflin.
Fitch W T, Giedd J. (1999) Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J Acoust Soc Am.* 106: 1511-1522.
Fourcin A J, Abberton E R M. (1971) First applications of a new laryngograph. *Medical and Biological Illustrated*. 21: 172-182.
Flanagan J L, Ishizaka K. (1978) Computer model to characterize the air flow displacement by the vibrating vocal cords. *J Acoust Soc Am*. 63: 1559-1565.
Fröschels E. (1931) Lehrbuch der Sprachheilkunde. 3. Aufl. *Leipzig and Wien: F Deulicke.*
Gauffin J, Sundberg J. (1989) Spectral correlates of glottal voice source waveform characteristics. *J Speech Hear Res*. 32: 556-565.
Gauffin J, Sundberg J. (1978) Clinical applications of acoustic voice analysis - acoustical analysis, results and discussion. In: Buch N H, editor. *Proc of the Int Assoc Log Phon Congr 15–18th August 1977, Copenhagen, Denmark*. Herning, Denmark: 489–502.
Gobl C, Ní Chasaide A. (2003) Amplitude-based source parameters for measuring voice quality. *Voqual´03* Genevea, 151-156.
Gramming P, Sundberg J, Elliot N, Nord L. (1993) Does the nose resonate in singing? In: Friberg A, Iwarsson J, Jansson E, Sundberg J, (1994) eds. *SMAC 93. Proceedings of the Stockholm Music Acoustics Conference.* Stockholm, Sweden: Royal Swedish Academy of Music; Publication No. 79. 166–171.
Gramming P. (1988) The phonetogram: An experimental and clinical study. (PhD thesis) Malmö Sweden, Dept. Otolaryngology, University of Lund.
Hammarberg B, Fritzell B, Gauffin J, Sundberg J, Wedin L. (1980) Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngol*. 90(5-6): 441-51.
Hammarberg B. (1986) Perceptual and Acoustic Analysis of Dysphonia. (PhD thesis). *Stockholm, Sweden, Karolinska Institutet*.
Hansson H. (1997) Glottal characteristics of female speakers: Acoustic correlates. *J Acoust Soc Am.* 101: 466-481.
Helmholtz H. (1877) On the Sensations of Tone. 2nd edition. Reprint by *Dover Publications*, Inc., New York, 1954.
Henrich N. (2006) Mirroring the voice from Garcia to the present day: Some insights into singing voice registers. *Log Phon Vocol*. 31: 3-14

Henrich N, d'Alessandro C, Doval B, Castellengo M. (2004) On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. *J Acoust Soc Am.* 115: 1321-1332.

Henrich N, d'Alessandro C, Doval B, Castellengo M. (2005) Glottal open quotient in singing: Measurements and correlation with laryngeal mechanism, vocal intensity, and fundamental frequency. *J Acoust Soc Am*. 117: 1417-1430.

Hertegård S, Gauffin J, Lindestad P-Å . (1995) A comparison of subglottal and intraoral pressure measurements during phonation. *J Voice*. 9: 149–155.

Hertegård S, Gauffin J. (1992) Acoustic Properties of the Rothenberg mask. *STL-QPSR* 1992/2-3: 9–18.*

Hoit J D, Jenks C L, Watson P J, Cleveland T. (1996) Respiratory function during speaking and singing in professional country singers. *J Voice*. 10 (1): 39-49.

Hirano M. (1974) Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniat.* 26: 89–94.

Hirano M. (1977) Structure and vibratory behavior of the vocal folds. In: Dynamic Aspects of Speech Production: 13-27. Tokyo: *University of Tokyo Press*.

Hirano M, Vennard W, Ohala J. (1970) Regulation of register, pitch and intensity of voice. A electromyographic investigation of intrinsic laryngeal muscles. *Folia Phoniat.* 22: 1-20.

Hixon T, Goldman M D, Mead J. (1973) Kinematics of the chest wall during speech production: Volume displacements of the ribcage, diaphragm, and abdomen. *J Speech Hear Res.* 13: 78-115.

Hollien H, Moore G P. (1968) Stroboscopic laminography of the larynx during phonation. *Acta Otolaryngol*. 65: 209–215.

Hollien H. (1974) On vocal registers. *J Phonetics*. 2: 125-143.

Holmberg E B, Hillman R, Perkell J, Guiod P, Goldman S. (1995) Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *J Speech Hear Res*. 38: 1212 - 1223.

Holmberg E B, Hillman, R E, Perkell J S. (1988) Glottal air flow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am.* 84: 511–529.

Iwarsson J, Thomasson M, Sundberg J. (1998) Effects of lung volume on the glottal voice source. *J Voice.* 12: 424 -433.

Jakobson R, Fant G, Halle M. (1952) Preliminaries to Speech Analysis. *Cambridge, MA: MIT Press.*

Jiang J J, Titze I R. (1994) Measurement of vocal fold intraglottal pressure and impact stress. *J Voice.* 2: 132–144.

Klatt D, Klatt L. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am.* 87: 820–857.

Ladefoged P, McKinney N P. (1963) Loudness, sound pressure, and subglottal pressure in speech. *J Acoust Soc Am.* 35: 454-60.

Ladefoged P (1961) Subglottal activity during speech. In *Proceedings of the 4th Internat. Congr. of Phonetic Sciences*, Helsingfors: 73-91.

Laver J. (1975) *Individual Features in Voice Quality*. (PhD thesis) Edinburgh, UK: University of Edinburgh.

Laver (1980) The phonetic description of voice quality. *Cambridge University Press*.

Large J. (1968) An acoustical study of isoparametric tones in the female chest and middle registers in singing. *NATS Bull* :12–15

Laukkanen A-M, Vintturi J, Vilkman E, Sala E, Siikki I, Lukkarila P, Sihvo M. (2001) Perceptual, acoustic and self-reported correlates of vocal loading. *Proceedings of the 25th World Congress of the International Association of Logopedics and Phoniatrics*. Montreal, Canada.

Lehto L, Airas M, Björkner E, Sundberg J, Alku P. (2006) Comparison of two inverse filtering methods in parameterization of the glottal closing phase characteristics in different phonation types. *J Voice*. in press, available online.

Leino T. (1994) Long-term average spectrum study on speaking voice quality in male actors. In: Friberg A, Iwarsson J, Jansson E, Sundberg J, eds. *SMAC93*, *Proceedings of the Stockholm Music Acoustics Conference 1993*. Stockholm, Sweden: Royal Swedish Academy of Music 79: 206–210.

Lohmann P.(1938) *Stimmfehler–Stimmberatung.* Mainz: Schott Verlag

LoVetri J, Weekly EM. (2002) Contemporary Commercial Music (CCM) survey: Who's teaching what in nonclassical music? *J Voice.* 17: 207-215

Löfqvist A, Carlborg B, Kitzing P. (1982) Initial validation of an indirect measure of subglottal pressure during vowels. *J Acoust Soc Am*. 72: 633-635.

Maffei G C. (1956) Delle lettere del Signor G. C. M. da Solofra libri due: dove tra gli altri bellissimi pensieri di Filosofia e di Medicina v'è un discorso della voce e del modo d'apparar di garganta senza maestro, Naples, Italy: Amato 1562. New edition In: Bridgman N. G.C. Maffei et sa lettre sur le chant. *Revue de musicologie* 38:10–34.

McIver W, Miller R. (1996) A brief study of nasality in singing. *NATS J Sing* March/April: 21–26.

Miles B, Hollien H. (1990) Whither belting? *J Voice.* 4: 64-70.

Miller R L. (1959) Nature of the vocal cord wave. *J Acoust Soc Am.* 3l: 667-679.

Miller R. (1996) The velopharyngeal (palato-pharyngeal) port during singing. *NATS J Sing.* Sept/Oct: 27–29.

Millet B, Dejonckere P. (1995) Singing Voice and Position of the Velum. [Videofilm]. Presentation at *the International Association of Logopedics and Phoniatric*s. Amsterdam.

Nawka T, Anders L C, Cebulla M, Zurakowski D. (1997) The speaker's formant in male voices. *J Voice*. 11: 422–428.

Nolan F. (1983) The phonetic bases of speaker recognition. *Cambridge, UK: Cambridge University Press*.

Nordenberg M, Sundberg J. Effect on LTAS of vocal loudness variation. *Log Phon Vocol*. 29(4): 183-191.

OSIRIS. Available at:
http://www.expasy.org/www/UIN/html1/projects/osiris/osiris.html

Pasculli K. Review of Brian Vander Ark: Resurrection, 10.8.03.Available at: http://www.discoveringartists.com/html/reviews/cds/index8.asp
Reid C. (1933) A Dictionary of Vocal Terminology. An Analysis. *New York: Music House.*
Rodarte J R, Rehder K. (1986) Dynamics of respiration. In: A Fishman, P Macklem and J Meads (Eds), *Handbook of Physiologoy. Section 3: The respiratory system. Vol III, ,part 1: Mechanics of Breathing.* (pp 131-144). Bethesda, MA: American Physiology Society.
Rothenberg M. (1973) A new inverse filtering technique for deriving the glottal airflow waveform during voicing. *J Acoust Soc Am.* 53: 1632-1645.
Sarraf C. Review of Holly Cole, Musicfolio 09/99. Available at: http://www.musicfolio.com/jazz/holly_cole.html
Schade G. (2005) Systematische Messung der Geschwindigkeiten der horizontalen Stimmlippenkonturen bei Veränderung des Schalldruckpegels und der stimmlichen Grundfrequenz in der Schließungsphase des phonatorischen Schwingungszyklus, Habilitation, University Hospital Hamburg-Eppendorf.
Sluijter, A M C., van Heuven, V J, Pacilly, J.A. (1997) Spectral balance as a cue in the perception of linguistic stress. *J Acoust Soc Am.* 101: 503-513.
Stevens K N. (1977) Physics of laryngeal behavior and larynx modes. *Phoneti*ca. 34: 264–279.
Stone R E Jr, Cleveland T, Sundberg J, Prokop J. (2003) Aerodynamic and acoustical measures of speech, operatic, and broadway styles in a professional female singer. *J Voice.* 17: 283-297.
Sundberg J, Andersson M, Hultqvist C. (1999) Effects of subglottal pressure variation on professional baritone singer' voice sources, *J Acoust Soc Am.* 105: 1965-1971.
Sundberg J, Högset C. (2001) Voice source differences between falsetto and modal registers in counter tenors, tenors and baritones. *Log Phon Vocol*. 26(1): 26 – 36.
Sundberg J. (1974) Articulatory interpretation of the 'singing formant'. *J Acoust Soc Am.* 55: 838–844.
Sundberg J, Kullberg Å. (1999) Voice source studies of register differences in untrained female singing. *Logoped Phoniatr Vocol*. 24: 76–83.
Sundberg J, Thalén M, Alku P, Vilkman E. (2004). Estimating perceived phonatory pressedness in singing from flow glottograms. *J Voice.* 1(18): 56-62.
Sundberg J. (1975) Formant technique in a professional female singer. *Acustica.* 32: 89–96.
Sundberg J, Gramming P, Lovetri J. (1993) Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. *J Voice.* 7: 301-310.
Sundberg J, Fahlstedt E, Morell A. (2005) Effects on the glottal voice source of vocal loudness variation in untrained female and male subjects. *J Acoust Soc Am.* 117: 879-885
Sundberg J, Titze I, Scherer R. (1993) 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.* 7 (1): 15-29.

Sundberg J. (1989) Synthesis of singing by rule. In: Mathews M, Pierce J, editors. *Current Directions in Computer Music Research*. System Development Foundation Benchmark series, Cambridge, MA: The MIT Press: 45–55: 401–403.

Thomasson M, Sundberg J. (1997) Lung volume levels in professional classical singing. *Log Phon Vocol*. 22: 61-70.

Thomasson M, Sundberg J. (1999) Consistency of phonatory breathing pattern in professional operatic singers. *J Voice.* 13(4): 529-541.

Thomasson M. (2003) Belly-in or belly-out? Effects of inhalatory behaviour and lung volume on voice function in male opera singers. *TMH – QPSR* Vol 45: 1-9.*

Thomasson M. (2003) From Air to Aria (PhD thesis) Stockholm, Sweden: KTH.

Ternström S. (1992) *Soundswell-Signal Workstation Software Manual Version 3.06*. Stockholm, Sweden: Soundswell Acoustics.

Titze I R. (1976) On the mechanics of vocal fold vibration. *J Acoust Soc Am* 60: 1366–1380.

Titze I R. (1988) The physics of small amplitude oscillation of the vocal folds. *J Acoust Soc Am.* 83: 1536–1552.

Titze I R. (1989) On the relation between subglottal pressure and fundamental frequency in phonation. *J Acoust Soc Am*. 85: 901-906.

Titze I R. (1992) Phonation threshold pressure: A missing link in glottal aerodynamics. *J Acoust Soc Am.* 91: 2926-2935.

Titze I R. (2001) Acoustic interpretation of resonant voice. *J Voice.* 15: 519–528.

Titze I R. (2004) A theoretical study of f0-f1 interaction with application to resonant speaking and singing voice. *J Voice.* 18: 292–298.

Titze I. (1994) Principles of Voice Production. *Englewood Cliffs*, NJ: Prentice-Hall.

Titze I R. (2004) A theoretical study of f0-f1 interaction with application to resonant speaking and singing voice. *J Voice.* 18(3): 292–298.

Titze I R. Jiang J J, Hsiao T Y. (1993) Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration. *Ann Otol Rhinol Laryngol*. 102(1): 58-63

Titze I. A few thoughts about longevity in singing. Available at http:// www. nvcs.org./singers/longevi.pfd.

Vennard W. (1967) *Singing—The Mechanism and the Technique. N*ew York: Carl Fischer.

Verdolini K, Druker D, Palmer P, Samawi H. (1998) Laryngeal adduction in resonant voice. *J Voice.* 12(3): 315–327.

Vilkman E, Alku P, Vintturi J. (2002) Dynamic extremes of voice in the light of time domain parameters extracted from the amplitude features of glottal flow and its derivative. *Folia Phoniat.* 54: 144-157.

Vilkman E, Alku P, Laukkanen A-M. (1995) Vocal fold collision mass as a differentiator between registers. *J Voice.* 9: 66-73.

* STL-QPSR - Speech Transmission Laboratory Quarterly Status and Progress Report, Dept. of Speech, Music and Hearing, Royal Institute of Technology, Stockholm, Sweden, available at http://www.speech.kth.se/qpsr/qpsr1960-1996.html