Analysing and Understanding the Singing Voice: Recent Progress and Open Questions

Malte Kob*[,1], Nathalie Henrich[2], Hanspeter Herzel[3], David Howard[4], Isao Tokuda[5] and Joe Wolfe[6]

[1]Erich Thienhaus Institute, University of Music Detmold, Germany; [2]Department of Speech and Cognition, GIPSA-lab (UMR 5216: CNRS, INPG, University Stendhal, UJF), Grenoble, France; [3]Institute for Theoretical Biology, Humboldt-University, Berlin, Germany; [4]Department of Electronics, University of York, UK; [5]Department of Micro System Technology, Ritsumeikan University, Kusatsu, Japan; [6]School of Physics, University of New South Wales, Sydney, Australia

Abstract: The breadth of expression in singing depends on fine control of physiology and acoustics. In this review, the basic concepts from speech acoustics, including the source-filter model, models of the glottal source and source-filter interactions, are described. The precise control, the extended pitch range, the timbre control and, in some cases, the uses of alternate phonation modes all merit further attention and explanation. Here we review features of the singing voice and the understanding that has been delivered by new measurement techniques. We describe the glottal mechanisms and the control of vocal tract resonances used in singing. We review linear and nonlinear components of the voice and the way in which they are measured and modelled and discuss the aero-acoustic models. We conclude with a list of open questions and active fields of research.

Keywords: Singing voice, speech acoustics, vocal folds, vocal tract, source-filter interaction.

1. INTRODUCTION

The singing voice is probably the most versatile musical instrument of all. As a musical instrument, it covers a range of some 80 dB in intensity, two to four or five octaves in fundamental frequency and a broad range of timbre in numerous singing styles. In addition to the properties of a musical instrument it adds verbal components to the performance, which allows the expression of nearly all aspects of spoken language. The power source of the singing voice is the same as that used for breathing – one of our most vital and basic functions – which probably contributes to the close relation between the singer's mood and vocal expression.

The singing voice is probably our species' oldest musical instrument. Nevertheless, its production, performance and perception are fascinating fields of ongoing research in domains such as acoustics, voice, phoniatrics, musicology and education. Further, the requirement of precise control of pitch in much of singing requires a deeper understanding, and a consideration of behaviours, such as resonance tuning, that are observed in singing but that are probably rare in speech, with the possible exception of theatrical speech. Due to its interdisciplinary nature, the results of singing voice research often have impact in several domains. One example is the understanding of the production of the growl sound and the impact of its performance on voice health [1].

This article gives an overview of areas of singing voice research. It begins with an introduction to features of the singing voice. The second section introduces the source-filter theory of voice production. Sections 3 and 4 cover methods to study the voice source and the vocal tract filter and some results of these studies. In section 5 some aspects of nonlinearities in voice production are presented. Section 6 describes recent approaches to singing voice modelling. The review concludes with a summary and outlook.

2. FEATURES OF THE SINGING VOICE

2.1. Styles

The human singing voice is capable of a wide variety of sounds. This diversity is reflected in the numerous singing styles found around the world, including 'classical' singing styles (e.g. western operatic, lied, baroque, ...) to popular ones ('contemporary commercial music', e.g. rock, jazz, musicals, ...) and traditional ones (e.g. yodel, kulning, overtone singing, ...). In the past century, much research effort has aimed at understanding classical singing styles from both a physiological and a physical perspective [2]. This scientific knowledge has provided the basis for understanding popular and traditional singing styles. Pop, jazz, blues and musical-theatre voice source characteristics have been compared to those of classical singers [3-7], demonstrating consistent differences in sub-glottal-pressure management, glottal behaviour and sound-pressure level. Knowledge and modelling of nonlinear source-filter interactions have provided insights into belting style [8]. Voice-production characteristics of country singers have been compared with those of speech [9-12]. Some traditional singing styles, such as overtone singing, challenge the normal use of the voice apparatus [13,14].

---

*Address correspondence to this author at the University of Music Detmold, Erich Thienhaus Institute, Neustadt 22, 32756 Detmold, Germany;
Tel: +49-5231-975-646, Fax: +49-5231-975-689;
E-mail: kob@hfm-detmold.de

1574-8936/11 $58.00+.00

2.2. Registers

The variety of singing sounds is produced by a single instrument, qualitatively similar for everyone, which consists of three subsystems: a breathing apparatus (muscles, lungs, bronchi, trachea) delivers air flow, the larynx converts some of the aerodynamic energy in that flow into acoustic waves, and the vocal tract (the upper part of the larynx, pharynx, oral and nasal cavities) modifies these waves to produce a wide range of sounds. The perceived vocal qualities of the resulting sound can be categorised into singing-voice registers, each register corresponding to a pitch range with homogeneous timbre [15]. The singing-voice registers result from specific pneumo-phono-articulatory adjustments made by the singer to achieve a target vocal quality. Much remains to be understood regarding these adjustments [16].

Recent studies have explored articulatory behaviors in the main singing-voice registers for both male and female operatic singers by means of a dynamic Magnetic Resonance Imaging technique [17-19]. For the studied singers, the transition from modal to falsetto registers resulted in only minor modifications of vocal-tract shape, such as an elevation and tilting of the larynx and a lifted tongue dorsum. In comparison, the transition from modal to voix-mixte registers resulted in major modifications, such as a pharynx widening, lip and jaw openings, and increased jaw protrusion.

The laryngeal nature of singing-voice registers, to which the term register may sometimes solely refer [20], can be described in terms of vocal-fold biomechanics, glottal-flow properties and non-linear dynamics (see section 5). From a physiological point of view, many researchers, including the present authors, described human voice production in terms of four laryngeal mechanisms (M0, M1, M2, M3), each associated with a different biomechanical configuration of the laryngeal vibrator over the voice frequency range [21]. The main mechanisms (M1 and M2) are used in classical and non-classical singing styles. Modal, chest and male head voice are produced in M1, falsetto and female head voice are produced in M2. The use of M0 (synonymous with vocal fry and strohbass) can be found in Rhythm and Blues singing (e.g. Clarence ‘Frogman’ Henry). The laryngeal mechanism M3 (synonymous with whistle, or flute register) might be used in classical singing to reach the highest notes in the top range of light sopranos. It is commonly used in jazz (e.g. Mimi Perrin, ‘Les double six’).

2.3. Ranges

The tonal pitch is defined by the vibratory frequency of the vocal folds (f_o). In singing, the different laryngeal mechanisms can produce frequencies from a few Hz to more than 2000 Hz, depending on the laryngeal mechanism in use [20,22]. According to Roubeau [22], who measured it on a group of 42 subjects, men in M1 sing from about 78 Hz (D#2) to 370 Hz (F#4) and women from about 147 Hz (D3) to 392 Hz (G4). In M2, men sing from about 165 Hz (E3) to 660 Hz (E5) and women from about 196 Hz (G3) to 1046 Hz (C6). The voice of children has not been intensively studied yet. However, the importance of voice training [23] and the particularities in physiology and voice properties such as voice range profile [24] of children are subjects of current research.

2.4. Intonation

In many styles of singing, precise pitch control is a desirable skill. The perceived pitch of a sung sound is largely determined by the fundamental frequency of vibration of the vocal folds and less directly by the intensity, overall timbre and duration of the sound being produced. Auditory feedback allows the singer to monitor and thus to adjust the pitch appropriately while singing. Quantitative monitoring of fundamental period to an accuracy of 1μs of vocal fold vibration is now reported [25]. Subjects were found to be more accurate when they tune a note against a recording of themselves singing a reference note as opposed to tuning to a reference note from a female singer or a non-vocal complex tone [26]. In-tune singing and accurate ensemble intonation is essential for a cappella (unaccompanied) choral singing and singers tend towards just intonation, i.e. to a non-equal temperament with commensurate ratios among notes in a chord [27]. A consequence is that pitch may drift when music changes key in some ways [28]. Performers can either stay in-tune or in-pitch; the latter requiring a pitch shift which has been commented on since the sixteenth century [29].

2.5. Vibrato

In singing, vibrato consists of a modulation of fundamental frequency. It is a very common feature in singing that is especially associated with Western opera [30]. Sundberg [30] gives four parameters that quantify vibrato: rate (typically 5-7.5 Hz), extent (typically 1 semitone for opera singers), regularity (generally consistent but tends to vary most during the negative phase) and waveform (approximately sinusoidal). Vibrato in singing is usually accompanied by amplitude variation, which may in part be due to source-filter interactions. Only a few cycles of vibrato may appear in short sung notes. The perceived pitch of such notes appears to be a function of the vibrato phase at the end of the note [31], but when sinusoids rather than synthesized /a/ vowels are used, there are variations in the pitch changes perceived [32].

2.6. Choir Sound/Choral Blend

Choir singing requires singers to blend with each other such that no one voice dominates and ‘individual voices are imperceptible’ [33]. This is achieved by reducing the levels of the second and third formants [34] and using a narrower vibrato than in solo singing [34,35]. Relative positioning of choir singers can alter blend and singers often have a preference for being mixed with other parts. Tuning and blend can be directly affected by the performance space itself and by the positions of the singers relative to each other. Daugherty [36] found that the spacing of singers relative to each other had a greater impact on the overall ensemble sound than whether or not the singers were in sections or in mixed formation. In terms of the acoustics of the performing space, Aspaas *et al.* [37] found no difference in the overall diffuse-field auditorium spectrum when choirs sang in different formations, but singers were able to hear each other better in the mixed formation. Perceptually, musicians can consistently label good and bad blend but they are more consistent for altos and basses than for sopranos and tenors [38].

Another element found to influence the ability of a choir singer to sing in-tune is the extent to which singers can hear their own output against those of the singers around them, and Ternström [39] introduced the term Self-to-other ratio to quantify this effect. Singers sang vowels accompanied by a synthesised version whose output level was controlled by the singer's level picked up by a microphone. The singer controlled the level of the synthesised vowel by moving towards or away from the microphone. Preferences were reported to be 2 dB with considerable individual variation. Measurements have also been made for members of an opera chorus singing in the Sydney Opera House [40]. Here the values were higher at +10 dB to +15 dB; each singer was easily able to hear him/herself but less able to hear other singers and the orchestra.

3. VOICE SOURCE ASSESSMENT

Three basic musical parameters of the singing voice – pitch, loudness and duration – are determined by the laryngeal sound production process. This process also contributes to other singing voice properties, such as timbre, projection and text intelligibility. Since the glottis cannot be easily investigated during singing without severe disturbance of the singer, a number of indirect methods have been developed. A comprehensive overview of voice analysis methods is given by [41].

3.1. The Source-Filter Model

The source-filter model treats the vocal tract as an acoustical duct leading from the larynx to the sound radiation field outside, passing via the mouth, nose or both. Fig. (**1**) represents the model. The glottis is the source and the envelope of its spectrum is usually monotonically decreasing with increasing frequency. The acoustic current output at the lips encounters the radiation impedance, which increases monotonically with frequency. Between the two, resonances of the vocal tract introduce stronger frequency dependence [42] (also referred to as 'formants' in voice sciences. We return to discuss terminology, resonances and formants in section 4.1).

The frequency (R1) of the first resonance lies somewhere between 0 and 1 kHz, that of the second (R2) lies between about 1 and 2 kHz. The values of R1 and R2 depend strongly on the articulation: opening the mouth increases R1 (and to a lesser extent R2) [43]). R2 is a strong function of the position (forward or back) of the point in the mouth at which the tongue constricts the tract and on the rounding of the lips.

Each vowel sound is associated with a particular range of (R1,R2) and therefore can be produced by a particular articulation. The higher resonances are associated with the timbre of the voice, rather than with phonemes.

3.2. Inverse Filtering

From a signal point of view, the glottal source is the acoustic flow at the glottis, which can be represented by an anti-causal two-pole low-pass filter with a -12 dB/oct. spectral slope [44]. The vocal-tract transfer function is approximated by an all-pole filter, and the lip radiation by a high-pass filter with a +6 dB/oct spectral slope. An assumption of linearity is commonly made, which means

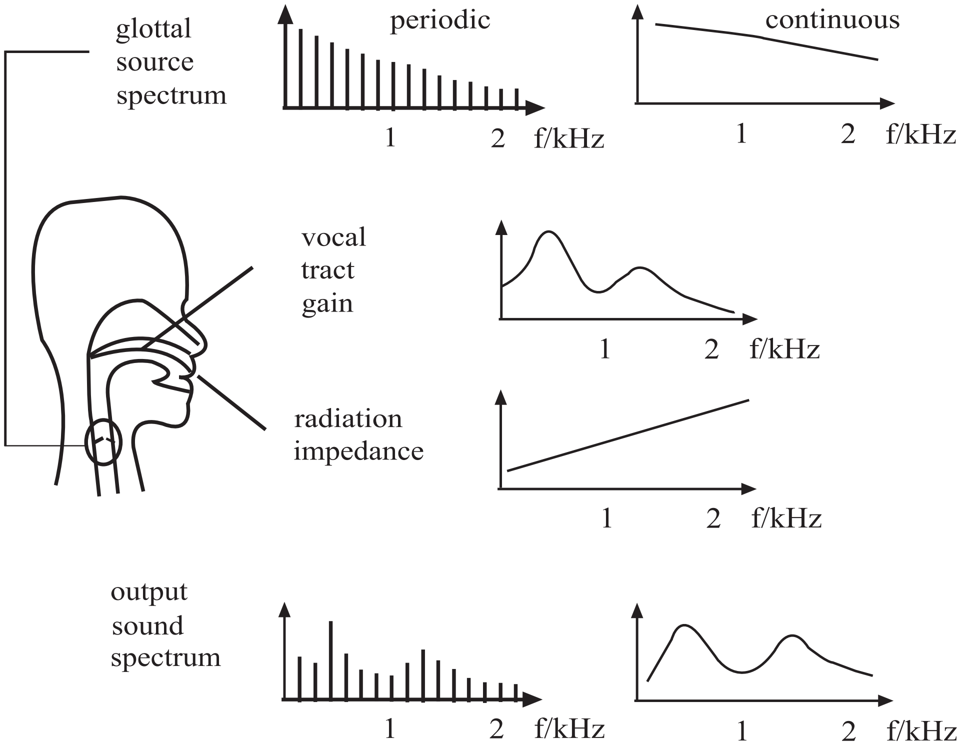

Fig. (1). A cartoon representation of the source filter model (from [138]). Normal phonation produces a periodic vibration and thus a harmonic spectrum at the glottis. A whisper produces a quasi-continuous spectrum. These are the glottal source input to the vocal tract. Resonances in the vocal tract give it a gain that varies strongly with frequency. The radiation impedance at the lips increases at about 6 dB/octave. (All the spectra are sketched in the logarithmic dB scale, which is why the effects add rather than multiply). The spectrum of the output sound is determined by the frequency dependence of source, gain and radiation impedance.

that glottal source, vocal-tract filter and lip radiation are assumed to be independent from each others. The vocal-tract filtering effect can thus be estimated and removed from the output voiced sound, resulting in an estimate of the glottal flow derivative, or the glottal flow itself by additional integration. This technique is referred to as inverse-filtering [45-47]. The main limitation of this technique in singing-voice assessment is that the spectral spacing between voice harmonics becomes greater when pitch increases [48].

3.3. Stroboscopy/High-Speed – Endoscopic Visualisation

The most direct methods to assess the physiological properties of the voice source are endoscopic visualisation through the oral or the nasal cavities (using rigid or flexible endoscopes respectively). While inverse-filtering techniques estimate transglottal airflow, videoendoscopic techniques estimate glottal area [49]. For a few decades, colour videoendoscopy and videostroboscopy have been common tools in both voice science and medicine [50]. These visualisation techniques have been refined by the introduction of videokymography [51], which enhances time resolution in display by using a single line of the image, chosen on the glottal plane. Spatial information can be added by strobophotoglottographic transillumination (SPGG) [52]. Recent advances in endoscopy use imaging systems capable of acquiring several thousands of image per second [53]. Such systems allow to examine the basic physiological mechanisms of different singing styles [54]. Originally in monochrome, the high-speed images can be obtained via CCD colour camera. Image resolution at high speed requires bright illumination. High-speed sequences produce large numbers of data, and data reduction and display are currently the main challenges. Advanced image-processing algorithms have been proposed for glottal-edge detection, and kymograms or phonovibrograms are used to display temporal data within one single image [55].

3.4. Electroglottography – Vocal Fold Contact Measurement

Among the indirect methods to measure glottal properties in phonation, electroglottography (EGG) is the most widely used. Two electrodes placed on the singer’s neck at the glottal level measure the electrical admittance, whose varying component is largely due to the variation in vocal-fold contact. EGG allows a direct and non-invasive measurement of vocal-fold contact area [56]. It is thus complementary to imaging techniques. EGG has been used to identify laryngeal mechanisms [21]. Glottal closing and opening instants can be determined from the differentiated EGG signal, and hence glottal parameters such as fundamental frequency, open, closed or contact quotients can be derived [57]. The technique can be refined by the use of several electrodes [58,59] which can give access to larynx position and supra glottal contact such as ventricular-fold vibration. Developing new methods for analysing and displaying electroglottographic signals and their first derivative is the subject of on-going research [60].

3.5. Voice Range Profile

The Voice Range Profile (VRP), also called phonetogram, plots the frequency range and the corresponding SPL of a singer’s voice [22,61-67]. It can be supplemented with acoustical voice-quality parameters [62], or with perceptual data [68]. The pitch range of a VRP is strongly dependent on the laryngeal mechanism [21]. The VRP upper and lower limits will depend on the sung vowel [69], and on the singing task [70].

4. THE ROLE OF THE TRACT-FILTER

The source-filter model introduced above treats the vocal tract as an acoustical duct leading from the larynx to the sound radiation field outside, passing via the mouth, nose or both. The tract geometry is not only complicated, but varies rapidly in time. Further, evolution has given us not only exquisite muscular control of that geometry, but also audition sensitive enough to distinguish the subtle changes in the voice that are produced by even small variations in articulation. Because of these subtleties and because of the enormous importance of the voice to human communication and culture, we are interested in quantifying its acoustic properties in a large number of different configurations, and in quantifying them precisely.

The voice itself is a very useful probe of the frequency response of the vocal tract. In the section on the source-filter model, we saw that the spectral envelope of the sound radiated from the mouth depends on the spectrum of the voice source at the larynx, the gain function of the tract, the spectrum of radiation impedance at the mouth and interactions among these.

The radiation impedance is relatively simple: it is dominated by the inertia of the air just outside the mouth and so, to a good approximation, it increases by 6 dB per octave [71]. The voice source spectrum and the gain spectrum are more complicated and of particular interest to us. They are both unknown and somewhat difficult to measure in a speaking or singing voice.

Fig. (**2**) shows four spectra associated with tokens the vowels /3/ and /o/, produced at the same pitch by the same subject. The top line shows the spectra of electroglottograph (EGG) measurements. Of course, an EGG spectrum can not be compared to the spectrum of the acoustic source at the glottis: it is included here to show the similarities in glottal contact between different phonemes.

The second pair of spectra shows the spectrum of the voice radiated from the mouth. This signal has passed through the filter of the vocal tract and the spectral envelope shows several broad maxima. Usually, each of these maxima is associated with a resonance in the vocal tract, and *vice versa*.

4.1. Formants and Terminology

The word ‘formant’ may be used with three different meanings. In general acoustics, it usually means a broad maximum in the spectral envelope of a sound. In phonetic sciences however, it usually means either the acoustic resonance itself or the resonance of a filter that can be used to model the vocal tract (with F1, F2, F3 etc. usually labeling the first, second, third filter in parallel). In most cases, the resonant frequencies and bandwidths of the physical resonance and the filter model are similar. They give rise to a marked peak in the spectral envelope.

There are cases where a resonance does not give rise to a spectral maximum in the voice output signal, such as when a voice, with say f_0 at 500 Hz (B4 a note near the top of the tenor range) sings a vowel such as /u/ with the first tract resonance at 300 Hz. For these reasons, the word ‘formant’ should therefore be defined when used.

4.2. Periodic & Non-Periodic Phonation

The second pair of spectra in Fig. (**2**) shows the voice sound generated by normal phonation. Here, the series of maxima in the spectral envelope allow us to estimate the resonance frequencies of a filter model that would produce them. We should remember, however, that the source spectrum also has a dependence on frequency, as does the radiation impedance. The spectral slope of the source is negative, which effect would tend to make the spectral maximum in the speech sound occur at a frequency lower than that of the resonance. The spectral slope of the radiation impedance, however, is positive, which tends to offset that relatively small effect. The precision available, however, is limited by the spacing between the harmonics, which equals f_o. In this case, f_o is relatively low (about 140 Hz) so one might expect to estimate the resonances with a precision of about 70 Hz or so. Speaking voices, particularly those of children, often use much higher values of f_o and soprano singers may use f_o as high as 1 kHz or more.

The problem of limited frequency sampling may be overcome by using a quasi-continuous source spectrum. The whisper (third row in Fig. **2**) is created by turbulent airflow through the glottis, so we have a quasi-continuous spectrum produced very near the glottis. For a few reasons, however, it is not the ideal probe of the vocal tract that it might seem at first glance. The first problem is that its source spectrum is unknown. The second is that that spectrum varies unpredictably with time: it is, after all, noise. The third is that changing the source – here changing the mode of phonation from normal to whisper – changes the properties of the filter [72]. The glottis is on average larger for whispering than for normal phonation, and this raises the resonances of the tract. Similar comments could be made about the use of the creak [73], synonymous with vocal fry [74,75] or mechanism 0, as a source to probe the tract gain function, although here the frequency shift is smaller than for whisper.

4.3. Broad Band Excitation

An alternative method of determining the tract resonances uses a synthetic source signal placed at the lips. This has the disadvantage that the source is ‘at the wrong end’ of the tract, but the advantages that the source is known and that its frequency resolution may be good. To make such measurements, a source of acoustic current is synthesised and input at the lips, next to a microphone [76-78]. Typically, a calibration measurement is made with the mouth closed, and the microphone measurement made during speech or singing is divided by that made during calibration. This ratio is approximately equal to the parallel combination of the acoustic impedances of the tract and the radiation field at lips, divided by the acoustic impedance of the radiation field.

Such measurements are shown in the last row of Fig. (**2**). The quasi-continuous curve is the impedance ratio. Superposed on this are the harmonics of the voice. This technique has been used to study how singers use tract resonances, a topic to which we return below.

4.4. Tomography and Modelling

An indirect method for determining resonances and other transfer functions of the tract is to acquire a three

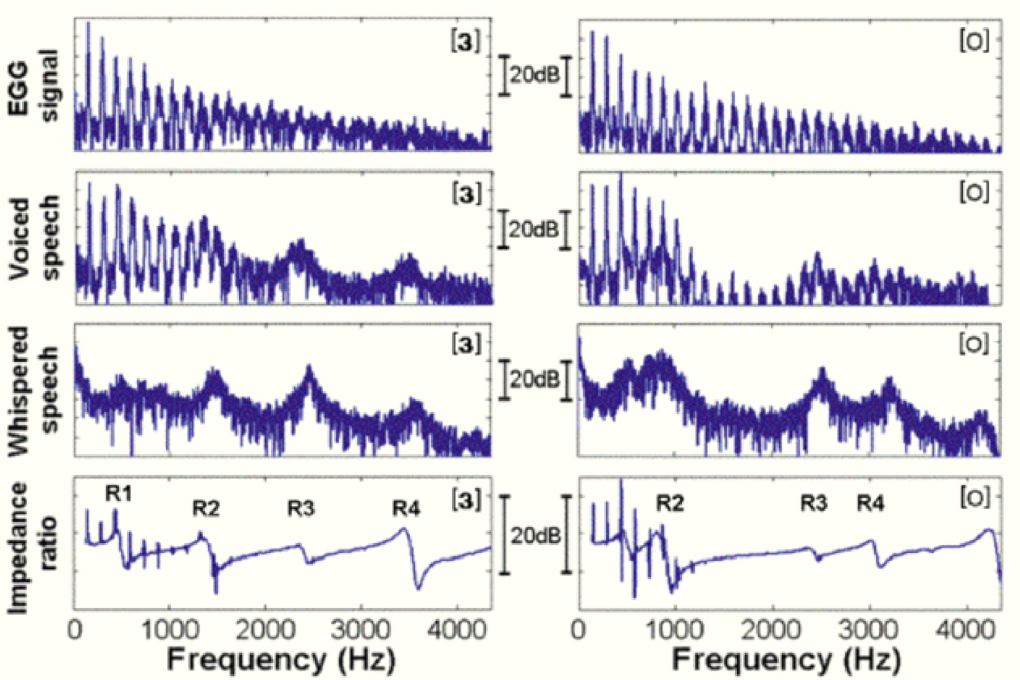

Fig. (2). The spectra at left were all recorded for the vowel /3/ and those at right for /o/. They show respectively the electroglottograph signal from near the larynx, the sound radiated from the mouth during normal and whisper phonation and the impedance ratio measured using a broad band source. After [138].

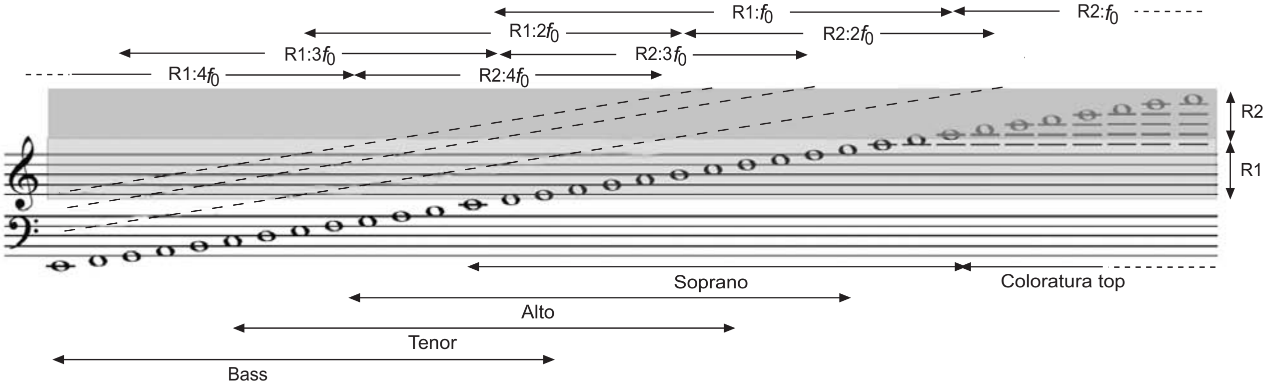

Fig. (3). The fundamental frequencies of four vocal ranges are indicated by notes, their next three harmonics by dashed lines. The ranges of the first two vocal tract resonances R1 and R2 are shaded. The arrows above the figure show the ranges of possible resonance tuning strategies. For bases, R1 and f_o hardly overlap. For all other ranges, R1:f_o tuning may be used for some notes. For the lower voices, R1:2f_o, R1:3f_o etc are possible. For the highest soprano range (called here the coloratura top, above C6), the only resonance strategy possible is R2:f_o.

dimensional image of the tract by magnetic resonance tomography [79-81]. From this image, the shape of the vocal tract may be estimated and, using a model, the acoustic properties of the tract may be calculated e.g. [82,83]. For analysing the singing voice, this has the complications that MRI images are usually acquired with the subject lying down, which may change the vocal geometry, and in the presence of high levels of noise.

4.5. Source-Filter Interaction

We have mentioned above how different types of voice source (normal phonation, whisper or creak) change the properties of the filter. It is also possible that the voice source may be changed by changing the filter.

In mathematical models of autonomous aero-acoustic oscillators, the acoustic load of the upstream or downstream ducts can influence the vibration of the oscillator [84-86]. It is obvious, for example, that the motion of the lips of a trombonist is strongly influenced by the resonances in the downstream duct: the trombone in this case. The vibration of the vocal folds is somewhat similar to that of the trombonist's lips: both are valves that tend to open under increasing upstream pressure and which tend to close when air flows between them. However, the frequency of vibration of the vocal folds is driven by the biomechanical adjustments within the folds, while the trombonist's lips usually vibrate at a frequency determined by the downstream acoustical load. The extent to which the vocal folds are affected by acoustic loads has not been determined experimentally, but both mathematical models and indirect experiments suggest an effect [87,88].

4.6. Resonance Tuning

Fig. (**2**) shows that tract resonances can augment the power in some harmonics of the voice and thus increase the overall output power. For the voice shown, with low f_0, the close spacing between harmonics ensures that at least one harmonic receives a useful boost. The situation is more difficult for higher voices.

The standard soprano range, from C4 to C6, roughly coincides with the normal range of the first resonance R1 – see Fig. (**3**). Thus, especially for close vowels with low R1 and high notes, the normal value of R1 often lies below f_0. A soprano who sang with f_0 above R1 would lose the power boost available. Singing with R1 lying midway between f_o and 2f_o would also lose this helpful power boost. From observations of how sopranos increased the area of their mouths as they ascended in pitch, Sundberg and colleagues [89] concluded that they were altering vowels so as to tune R1 near f_0.

Measurements using the broadband technique showed that this indeed occurred and that it is practised by sopranos with a wide range of backgrounds [90-92]. Tomography studies have shown the articulation involved [18, 93]. If the fundamental frequency of the voice passes from below to above the tract resonance frequency, the acoustic impedance varies from being largely inertive (pressure leads flow) to largely compliant (flow leads pressure) and it is argued that this can produce instabilities in the voice [88].

As Fig. (**3**) indicates, R1:f_o is not the only possible resonance tuning strategy: it is possible for most voice ranges to tune R1 to higher harmonics (R1:2f_o, R1:3f_o etc.). Systematic R1:2f_o tuning has been observed for altos, and some men singers tune R1:nf_o over parts of their range [91]. The very high range of the soprano voice (above C6) is used by coloratura sopranos and some pop and jazz singers. In this range, no mouth geometry can raise R1 sufficiently to match f_o. Some of the singers who perform in this range tune the second resonance to the fundamental – R2:f_o tuning [92].

4.7. Harmonic Singing

With a low f_o, it is possible to tune R2 to one of the higher harmonics. This is practised in harmonic singing, in which a relatively low note is sustained as a drone, which makes it less noticeable, while different harmonics of the drone are selectively enhanced by R2:nf0 tuning. This makes available notes in the harmonic series, which in turn allows playing the notes available to the bugle or natural horn [94].

A number of different singing styles are gathered under the term harmonic singing. Among these two styles are found two different sound production methods: Kargyraa, a rather rough and low sound that potentially originates from joint vocal fold and ventricular fold action, and Sygyt, a whistle-like sound that features a distinct melody pitch and a rather suppressed drone [95]. Research addresses the acoustic relation between vocal tract shape and melody pitch selection [14,96].

5. NONLINEAR PHENOMENA

The human voice production system can be considered as a system of coupled nonlinear oscillators (e.g. left and right fold, different modes of vibrations, ventricular folds, aerodynamic resonators). Vocal fold vibrations are driven by the airflow provided by lung pressure. The interaction of air flow and tissue movement is highly nonlinear. One reason is that, when kinetic energy is dissipated in turbulence, the pressure drop is proportional to the square of the flow. Moreover, vocal fold collisions and their visco-elastic properties are governed by nonlinear relations [97]. In this section we emphasise that nonlinear dynamics provides an appropriate framework to describe many features of the human voice.

5.1. Normal Phonation as Limit Cycle Oscillations

Even regular steady phonation can be discussed from the point of view of nonlinear dynamics: an energy source, the airflow, leads to self-sustained periodic vibrations of the vocal folds with a well-defined period and amplitude. Such an autonomous oscillator is also termed a limit cycle. Limit cycles require energy supply and the steady amplitude is due to losses in nonlinearities such as collisions. The transfer of energy from the air flow to the vocal folds is achieved via wave-like vibration patterns of the folds [15]. The transition from a pre-phonatory standstill to self-sustained oscillations is a manifestation of a ‘Hopf bifurcation’ [98]. The description of regular phonation as a limit cycle implies that all vibratory modes are synchronised leading to a ‘clean’ or periodic voice with only tiny perturbations (jitter and shimmer below 1%). In the following we discuss more complex phenomena found in the voice production system.

5.2. Subharmonics, Biphonation, and Deterministic Chaos

The theory of coupled oscillators predicts that regular oscillations (limit cycles) can bifurcate to rather complex signals. For example, period-doubling bifurcations may lead to subharmonic oscillations and a secondary Hopf bifurcation might induce another independent sound. Many terms have been introduced to describe these phenomena including octave jump, double harmonic break, diplophonia, biphonation, and noise concentrations (see [99] for references). Here we will use the term subharmonics for frequency components with a fixed ratio such as 1:2 or 1:3 and biphonation for two simultaneously produced independent sounds. Perceptually, these nonlinear phenomena are typically associated with ‘rough’ sounding voices [100].

Interestingly, the whole plethora of nonlinear phenomena have been observed in a wide variety of vocalisations: newborn cries [101], pathological voices [99], normal speech [102], Russian lament [103], animal communication [104,105], and in contemporary vocal music [106].

In some cases the nonlinear phenomena can be traced back to specific physiological mechanisms such as left-right asymmetry [107], co-vibrations of ventricular folds [108], anterior-posterior modes within the vocal folds [109], and source-tract interactions [110]. In jazz, rock music and contemporary vocal music, nonlinear phenomena are exploited to extend the range of traditional principles of phonation. Numerous examples of nonlinear phenomena in vocal improvisations are discussed in [111].

Interesting applications of nonlinear dynamics theory are register transitions. Regular phonation in different registers can be considered as limit cycles with certain characteristics. Variations of parameters such as muscle tension might induce sudden register-transitions (`bifurcations’). There is experimental and theoretical evidence [112,113] that these transitions can be accompanied by nonlinear phenomena. Below we discuss an example in some detail.

6. SINGING VOICE MODELLING

The motivation of modelling musical instruments is two-fold. Either the essential physical process of generation of the musical sound is implemented and compared to the original principle for deeper understanding of the sound generation process, or the musical sound that is perceived by a listener is modelled for sound synthesis and music production purposes. The first method can be called `physical modelling’, and requires deep insight into the physics of musical instruments, and the realism of the resulting sound depends on the complexity and accuracy of the modelled principles. The second method can be called `sound modelling or synthesis’, and is optimised through most accurate shaping of the musical sound signal for each desired style, timbre and pitch. In the ideal case, both concepts converge, i.e. the modelling of the physics of a particular instrument configuration will produce a wave form that corresponds exactly to the original sound.

Current synthesisers for music production employ a number of concepts, ranging from FM synthesis via sample-based sound generation to physical models of various instruments. Among these instruments solo voices are not available yet, but choir sounds as back voices are found that give the impression of a blended choir sound without differentiation of single voices. These sounds are produced using pre-recorded and shaped voice samples, and are mostly limited to the production of one phoneme, e.g. /a/. Obviously, physical voice synthesis can not yet produce realistic voice sounds. However, some aspects of voice production can be modelled accurately. A review of singing voice simulation techniques is given by [96].

6.1. Concepts and Challenges

Nowadays, hardware models of voice generation are mostly used for education purposes. Simple models can be used to demonstrate the source-filter concept, whereas complex models can produce coherent phonemes and simple words since several centuries [114]. The naturalness of the produced sounds is, however, very far from the original. For a number of open research questions hardware models are used and serve as experimental reference for numerical approaches, mostly based upon the famous scaled model of the glottis ‘M5’ [115-117], or 1: 1 vocal fold models [118].

Numerical models for voice production have evolved significantly in the last years. In this issue some recent approaches are reviewed. Due to the high number of degrees of freedom, numerical models using finite element or similar methods have not yet been used for singing voice synthesis. Numerical models with a smaller number of elements, multiple- or discrete-mass models, have been used successfully for the modelling of various voice phenomena. In [119] a multiple-mass model of the vocal folds and wave propagation through the vocal tract using a wave guide-approach is described. It has recently been applied to the modelling of vocal fold nodules, a frequent voice disease, mainly in female singers [120]. The following paragraph describes the application of a similar model to the simulation of register transitions.

6.2. Modelling Register Transitions

In the study of voice production, numerous models have been developed to simulate the vocal fold vibrations ranging from simplified low-dimensional models [97,107,121-124] to complex high-dimensional models [125].

Up to now, registers have been modelled using distinct sets of biomechanical parameters corresponding to the phonation condition of each register [119,121]. Based on such models, register transitions can be realised by simulating a muscle activity that corresponds to the transfer between two distinct sets of parameters [126]. f_o gliding has been simulated by several models [88,113,127,128]. Moreover, hysteresis of transitions between chest and falsetto registers and voice instabilities observed during the register transitions in excised larynx experiments [129-131] have been also studied [113,128]. The hysteresis implies coexistence of two registers within the same physiological condition of the vocal folds, which is essential for realising an abrupt transition between the registers. As one of few models that simulate the register transitions, a four-mass body-cover polygon model is considered here [128]. This model was developed to replicate as closely as possible the sudden chest-falsetto transitions and the accompanying phenomena observed in a singing Voice

The four-mass model is based on the body-cover differentiation proposed by Story and Titze [121] with one-mass representation of the body and three-mass representation of the cover [113] (see Fig. **4**).

A smooth vocal fold geometry as in Lous *et al.* [123] is also utilised. The three-mass structure in the cover layer was designed by adding one more mass on top of the two-mass model [97,123]. Addition of the third mass divides the upper part of the cover layer into two portions, which can vibrate out of phase. These phase differences can simulate the mucosal waves, which are observed in the videokymograms of both chest and falsetto registers [129,131,132]. In particular, during the falsetto register, the waves are visible only on the thin upper medial portion of the vocal folds and on the upper vocal fold surface. With the three cover masses of the present model, such oscillatory mode can be simulated by the anti-phase oscillations between the upper and the middle masses [113]. Compared with the standard two-mass model, which was not designed to model such small oscillations of the upper vocal fold, the three-mass structure in the cover layer has the advantage of modelling the falsetto register based on the upper vocal fold oscillations, which can easily coexist with the chest register.

The main modelling assumptions are

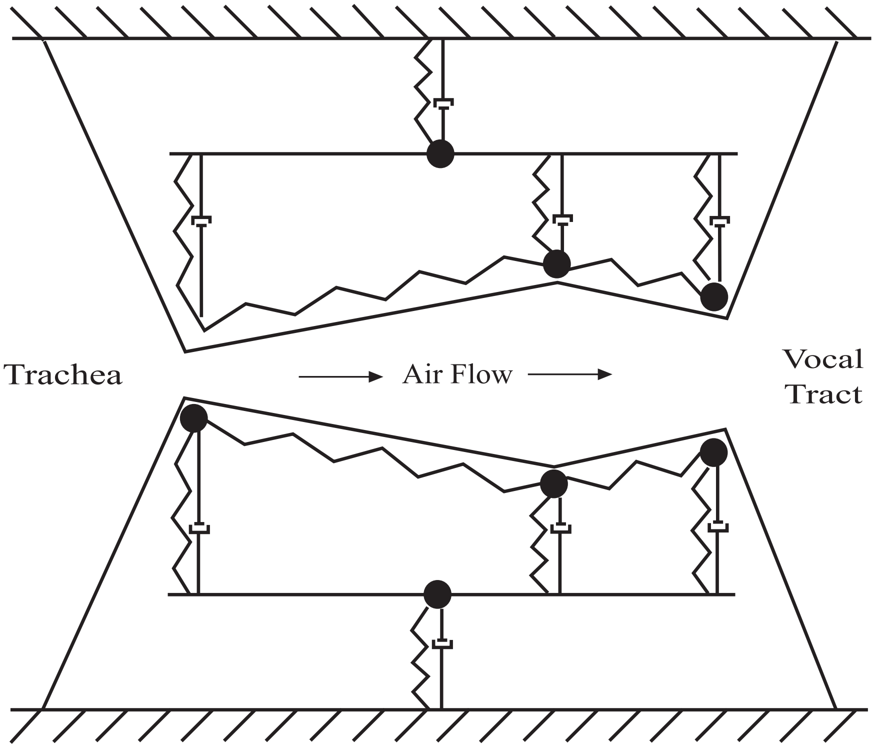

Fig. (4). Schematic illustration of the four-mass polygon model of the vocal folds. The left and right vocal folds have a symmetric configuration. Each vocal fold is composed of four masses (one mass in body layer and three masses in cover layer) coupled by linear springs. The air flow coming from the lungs is described by the ‘Bernoulli term’ $\frac{\rho v^2}{2}$ below the narrowest part of the glottis.

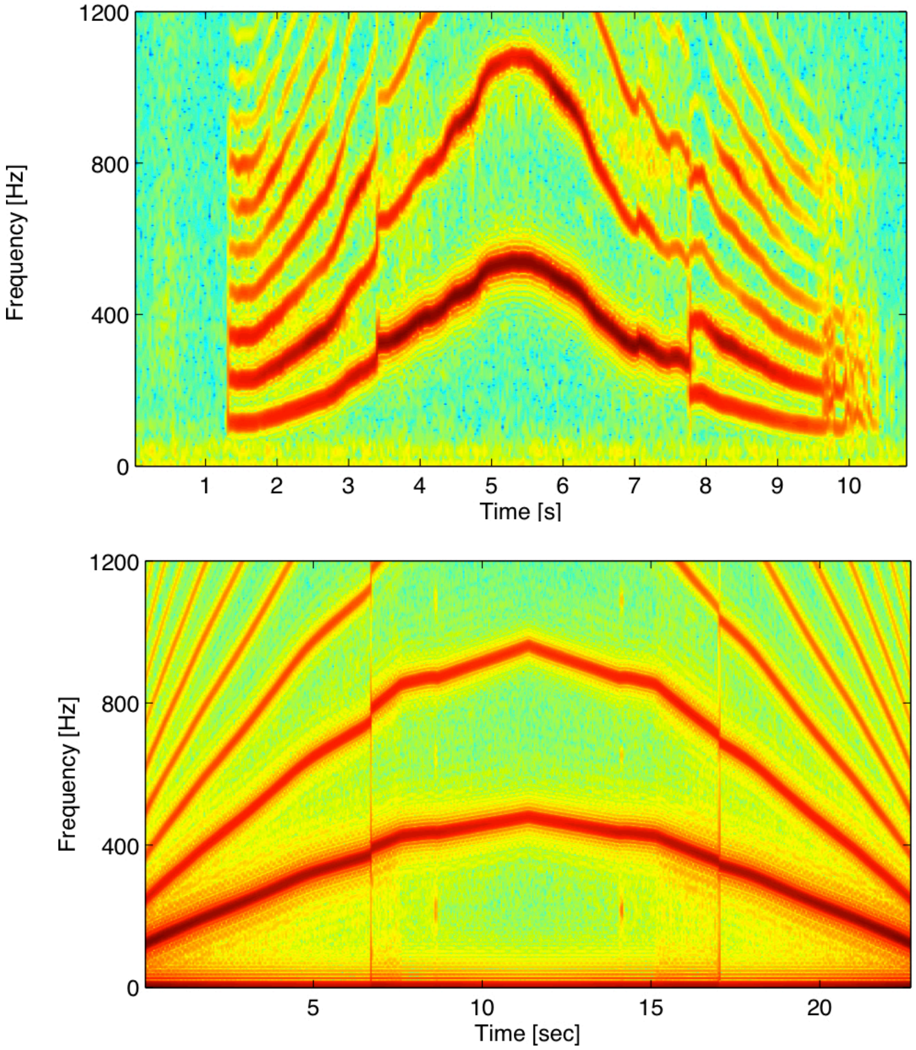

Fig. (5). (**a**) A spectrogram of a human voice with a gliding fundamental frequency. (**b**) A model simulation of the gliding fundamental frequency. The tension parameter is increased from T=1 over $t < 11.5$ s and then decreased back to T = I over 11.5 s < t < 22.5 s.

1. the four masses are coupled by linear springs,
2. the air flow inside the glottis is described by the so-called ‘Bernoulli term’ $\frac{\rho v^2}{2}$ below the narrowest part of the glottis [107],
3. the left and the right vocal folds are antisymmetric images of each other.

To take into account the influence of the vocal tract, sub-glottal and supra-glottal resonances are modelled by the wave-reflection analog systems [133-136], which were coupled to the vocal folds model in an interactive fashion [88,136]. To simulate singing a glissando, a tension parameter T is controlled, where T determines the size and the stiffness of the four masses in a way that linearly controls the natural frequency of the four masses. The other parameter values are adopted from the standard values established in the two-mass models [97,107,123].

Fig. (**5**) compares a glissando of an untrained singer with simulations of a corresponding f_o glide of the four-mass model. The singer’s glissando in Fig. (**5a**) exhibits register transitions with frequency jumps at around 3.3 s and 7.8 s with slightly different pitches. There is an abrupt phonation onset at 1.2 s and a more smooth offset with some irregularities at 10 s.

Glissando is simulated in Fig. (**5b**) by varying the tension parameter T from 1 to 5.5 and then back. In low-frequency regimes (0 s - 6.7 s; 17 s - 22.5 s), the three cover masses exhibits chest-like vibrations with a complete closure of the glottis (no figure shown). In high-frequency regime (6.7 s - 17 s), on the other hand, falsetto-like vibrations with diminished closure of the glottis are discernible (no figure shown). Note that a frequency jump is observed at 6.7 s ($f_0 = 390$ Hz), whereas a backward transition takes place at 17 s ($f_0 = 350$ Hz). These differences between chest-falsetto and falsetto-chest transitions are an indication of hysteresis (bifurcations leading to the hysteresis are discussed in detail in [113]). Hysteresis indicates that there are coexisting vibratory regimes (`limit cycles’) for a range of parameters. Moreover, hysteresis implies that there are voice breaks instead of *passagi* of trained singers.

In addition to the register transitions, subharmonics are occasionally observed, e.g., at 8.7 s and 14.1 s. It has been discussed earlier [113,129] that register transitions are often accompanied by nonlinear phenomena such as subharmonics and chaos. Such phenomena are reproduced in the model simulations. It is also well known that register transitions in untrained singers are accompanied by vocal breaks (see, e.g., [137]). In the present simulations, sudden jumps of pitch and amplitudes were indeed found while varying the tension parameter *T* smoothly. In this sense, the four-mass model shows a nice agreement with the glissando of untrained singers. It should be noted that these findings are not sensitive to the modelling details, since the four-mass model represents just the core mechanisms of the vocal folds oscillations. Gross features of the register transitions simulated by the present model are expected to be found commonly in other vocal fold models.

7. SUMMARY AND OUTLOOK

Singing voice research has advanced significantly in the last few decades, in part because of new tools for voice assessment. Sophisticated models of the glottal-source reproduce many of the features of the singing voice. The role of the 'tract-filter' in resonance tuning and harmonic singing is better understood, but the interaction between the source and the filter is the subject of ongoing research.

Some other singing styles, such as undertone singing, growl and the whistle voice are much less understood. Tools including EGG and high-speed endoscopy have been used to study these styles. However, a number of questions remain unanswered, such as: How are the ventricular folds brought to oscillation in undertone singing? What are the control parameters of register changes? To what extent do non-linear effects occur in singing voice? These questions remain to be answered by new interdisciplinary research. Another important area needing work is making the new understanding available to singers and teachers, in a way that facilitates not only understanding but also learning and training.

REFERENCES

[1] Eckers C, Hütz D, Kob M, Murphy P, Houben D, Lehnert B. Investigation of voice production in death metal singers. In: NAG/DAGA International Conference on Acoustics Rotterdam 2009; p. 1747-50.

[2] Sundberg J. The science of the singing Voice Dekalb, Illinois: Northern Illinois University Press; 1987.

[3] Sundberg J, Gramming P, Lovetri J. Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. J Voice 1993; 7(4): 301-10.

[4] Thalén M, Sundberg J. Describing different styles of singing: a comparison of a female singer's voice source in "Classical", "Pop", "Jazz" and "Blues". Logoped Phoniatr Vocol 2001; 26(2): 82-93.

[5] Stone RE, Cleveland TF, Sundberg PJ, Prokop J. Aerodynamic and acoustical measures of speech, operatic, and Broadway vocal styles in a professional female singer. J Voice 2003; 17(3): 283-97.

[6] Björkner E, Sundberg J, Cleveland T, Stone E. Voice source differences between registers in female musical theater singers. J Voice 2006; 20(2): 187-97.

[7] Björkner E. Musical theater and opera singing-why so different? A study of subglottal pressure, voice source, and formant frequency characteristics. J Voice 2008; 22(5): 533-40.

[8] Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing. J Acoust Soc Am 2009; 126(3): 1530-40.

[9] Hoit JD, Jenks CL, Watson PJ, Cleveland TF. Respiratory function during speaking and singing in professional country singers. J Voice 1996; 10(1): 39-49.

[10] Stone RE, Cleveland TF, Sundberg J. Formant frequencies in country singers. J Voice 1999; 13(2): 161-7.

[11] Sundberg J, Cleveland TF, Stone RE, Iwarsson J. Voice source characteristics in six premier country singers. J Voice 1999; 13(2): 168-83.

[12] Cleveland TF, Sundberg J, Stone RE. Long-term-average spectrum characteristics of country singers during speaking and singing. J Voice 2001; 15(1): 54-60.

[13] Klingholz F. Overtone singing: productive mechanisms and acoustic data. J Voice 1993; 7(2): 118-122.

[14] Adachi S, Yamada M. An acoustical study of sound production in biphonic singing, Xöömij. J Acoust Soc Am 1999; 105(5): 2920-32.

[15] Titze IR. Principles of voice production. New York: Prentice-Hall; 1994.

[16] Miller DG. Registers in singing. Rijksuniversiteit Groningen, the Netherlands; 2000.

[17] Echternach M, Sundberg J, Arndt S, *et al.* Vocal tract and register changes analysed by real-time MRI in male professional singers - a pilot study. Logoped Phoniatr Vocol 2008; 33: 67-73.

[18] Echternach M, Sundberg J, Arndt S, Markl M, Schumacher M, Richter B. Vocal tract in female registers - a dynamic real-time MRI study. J Voice 2010; 24: 133-9.

[19] Echternach M, Sundberg J, Markl M, Richter B. Professional opera tenors' vocal tract configurations in registers. Folia Phoniatr Logop 2010; 62(6): 278-87.

[20] Henrich N. Mirroring the voice from Garcia to the present day: some insights into singing voice registers. Logoped Phoniatr Vocol 2006; 31(1): 3-14. Available from: http://hal.archives-ouvertes.fr/-hal-00344177/fr.

[21] Roubeau B, Henrich N, Castellengo M. Laryngeal vibratory mechanisms: The notion of vocal register revisited. J Voice 2009; 23(4): 425-38. Available from: http://hal.archives-ouvertes.fr/hal-00319915/fr.

[22] Roubeau B, Castellengo M, Bodin P, Ragot M. Laryngéal registers as shown in the voice range profile. Folia Phoniatr Logop 2004; 56(5): 321-33.

[23] Fuchs M, Meuret S, Thiel S, Täschner R, Dietz A, Gelbrich G. Inuence of singing activity, age, and sex on voice performance parameters, on subjects' perception and use of their voice in childhood and adolescence. J Voice 2009; 23(2): 182-9.

[24] Schneider B, Zumtobel M, Prettenhofer W, Aichstill B, Jocher W. Normative voice range pro_les in vocally trained and untrained children aged between 7 and 10 years. J Voice 2010; 24(2): 153-60.

[25] Howard DM. Electrolaryngographically revealed aspects of the voice source in singing. Logoped Phoniatr Vocol 2010; 35(2): 81-9.

[26] Moore RE, Estis J, Gordon-Hickey S, Watts C. Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli. J Voice 2008; 22(4): 399-407.

[27] Helmholtz H. On the Sensations of Tone. 2nd ed. New York: Dover Publications; 1877. Translation by A.J. Ellis.

[28] Howard DM. Intonation Drift in a Capella Soprano, Alto, Tenor, Bass Quartet Singing With Key Modulation. J Voice 2005; 21(3): 300-15.

[29] Lindley M. Just Intonation. In: New Grove Dictionary of Music and Musicians. 2nd ed. Oxford: Oxford University Press; 2001.

[30] Dejonckere PH, Hirano M, Sundberg J, eds. Vibrato, San Diego: Singular Publishing Group; 1995.

[31] d'Alessandro C, Castellengo M. The pitch of short duration vibrato tones. J Acoust Soc Am 1994; 95(3): 1617-30.

[32] van Besouw RM, Brereton JS, Howard DM. Range of tuning for tones with and without vibrato. Music Perception 2008; 26(2): 145-55.

[33] Ekholm E. The effect of singing mode and seating arrangement on choral blend and overall choral sound. J Res Music Edu 2000; 48(2): 123-35.

[34] Goodwin AW. An acoustical study of individual voices in choral blend. J Res Music Edu 1980; 28(2): 119-128.

[35] Rossing TD, Sundberg J, Ternström S. Voice timbre in solo and choir singing: Is there a difference? J Res Singing Appl Vocal Pedagogy 1985; 8(2): 1-8.

[36] Daugherty JF. Choir spacing and formation: Choral sound preferences in random, synergistic and gender-specific chamber choir placements. Intl J Res Choral Singing 2003; 1(1): 48-59.

[37] Aspaas C, McCrae CR, Morris RJ, Fowler L. Select acoustic and perceptual measures of choral formation. Int J Res Choral Singing 2004; 2(1): 11-21.

[38] Killian JN, Basinger L. Perception of choral blend among choral, instrumental and non-music majors using continuous response digital interface. J Res Music Edu 2007; 55(4): 313-25.

[39] Ternström S. Preferred Self-to-other in choir singing. J Acoust Soc Am 1999; 105(6): 3562-74.

[40] Ternström S, Cabrera D, Davis P. Self-to-other ratios measured in an opera chorus in performance. J Acoust SocAm 2005; 118(6): 3903-11.

[41] Baken RJ, Orlikoff RF. Clinical measurement of speech and voice San Diego: Singular Publishing Group; 2000.

[42] Fant G. Acoustic theory of speech production. The Hague, The Netherlands: Mouton & Co.; 1960.

[43] Fant G. Speech Sounds and Features. MIT Press; 1973.

[44] Doval B, d'Alessandro C, Henrich N. The spectrum of glottal flow models. Acta Acustica united with Acustica 2006; 92: 1026-46. Available from: http: //hal.archives-ouvertes.fr/hal-00368131/fr.

[45] Rothenberg M. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. J Acoust Soc Am 1973; 53(6): 1632-45.

[46] Airas M. TKK Aparat: An environment for voice inverse-filtering and parameterization. Logoped Phoniatr Vocol 2008; 33(1): 49-64. Available from: http://dx.doi.org/10.1080/14015430701855333.

[47] Alku P, Magi C, Bäckström T. Glottal inverse filtering with the closed-phase covariance analysis utilizing mathematical constraints in modelling of the vocal tract. Logoped Phoniatr Vocol 2009; 34(4): 200-209. Available from: http://dx.doi.org/10.3109/14015430902913519.

[48] Arroabarren I, Carlosena A. Inverse filtering in singing voice: A critical analysis. IEEE Trans Audio Speech Language Proc 2006; 14: 1422-31.

[49] Granqvist S, Hertegård S, Larsson H, Sundberg J. Simultaneous analysis of vocal fold vibration and transglottal airflow: exploring a new experimental setup. J Voice 2003; 17(3): 319-30.

[50] Dejonckere P, Bradley P, Clemente P, *et al.* A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. guideline elaborated by the committee on phoniatrics of the European Laryngological Society (ELS) Eur Arch Otorhinolaryngol 2001; 258: 77-82.

[51] Svec JG, Schutte HK. Videokymography: High-speed line scanning of vocal fold vibration. J Voice 1996; 10: 201-5.

[52] Hess MM, Ludwigs M. Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns. J Voice 2000; 14(2): 255271.

[53] Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and Evolution. Folia Phoniatr Logop 2008; 60: 33-44.

[54] Hertegård S. What have we learned about laryngeal physiology from high-speed digital videoendoscopy? Curr Opin Otolaryngol Head Neck Surg 2005; 13(3): 152-156.

[55] Lohscheller J, Eysholdt U, Toy H, Döllinger M. Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-D diagrams for visualizing and analyzing the underlying laryngeal dynamics. IEEE Trans Med Imaging 2008; 27: 300-9.

[56] Childers DG, Hicks DM, Moore GP, Eskenazi L, Lalwani AL. Electroglottography and vocal fold physiology. J Speech Hear Res 1990; 33(2): 245-54. Available from: http://jslhr.asha.org/cgi/content/abstract/33/2/245.

[57] Henrich N, d'Alessandro C, Doval B, Castellengo M. On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation. J Acoust Soc Am 2004; 115(3): 1321-32. Available from: http://hal.archives-ouvertes.fr/hal-00344175/fr.

[58] Rothenberg M. A multichannel electroglottograph. J Voice 1992; 6(1): 36-43. Available from: http://www.rothenberg.org/Multichannel/Multichannel.htm.

[59] Kob M, Frauenrath T. A system for parallel measurement of glottis opening and larynx position. Biomed Signal Proc Control 2009; 4(3): 221-8. New trends in voice pathology detection and classification - M & A of vocal emissions. Available from: http://www.sciencedirect.com/science/article/B7XMN-4WF2NR2-1/2/ 4761d1fb12c7aa684cf4d1f3f6732703.

[60] Herbst CT, Fitch WTS, Švec JG. Electroglottographic wavegrams: A technique for visualizing vocal fold dynamics noninvasively. J Acoust Soc Am 2010; 128(5): 3070-8. Available from: http://link.aip.org/link/?JAS/128/3070/1.

[61] Coleman RF, Mabis JH, Hinson JK. Fundamental frequency-sound pressure level profiles of adult male and female voices. J Speech Hear Res. 1977; 20(2): 197-204.

[62] Pabon JP, Plomp R. Automatic phonetogram recording supplemented with acoustical voice-quality parameters. J Speech Hear Res 1988; 31(4): 710-22.

[63] Titze IR. Acoustic interpretation of the voice range profile (phonetogram). J Speech Hear Res 1992; 35(1): 21-34.

[64] Sulter AM, Wit HP, Schutte HK, Miller DG. A structured approach to voice range profile (phonetogram) analysis. J Speech Hear Res 1994; 37(5): 1076-85.

[65] Holmberg EB, Ihre E, Södersten M. Phonetograms as a tool in the voice clinic: changes across voice therapy for patients with vocal fatigue. Logoped Phoniatr Vocol 2007; 32(3): 113-127. Available from: http://dx.doi.org/10.1080/14015430701661632.

[66] Lamarche A, Ternström S, Pabon P. The singer's voice range profile: female professional opera soloists. J Voice 2010; 24(4): 410-26. Available from: http://dx.doi.org/10.1016/j.jVoice2008.12.008.

[67] Pabon P, Ternström S, Lamarche A. Fourier descriptor analysis and unification of voice range profile contours: method and applications. J Speech Lang Hear Res 2010; Available from: http://dx.doi.org/10.1044/1092-4388(2010/08-0222).

[68] Lamarche A, Ternström S, Hertegård S. Not just sound: supplementing the voice range profile with the singer's own perceptions of vocal challenges. Logoped Phoniatr Vocol 2009; 34(1): 3-10. Available from: http://dx.doi.org/10.1080/14015430802239759.

[69] Lamesch S. Laryngeal mechanisms and vowels in singing, viewed with respect to voice range profiles, glottal and spectral parameters, and laryngeal mechanism transition. Université Pierre et Marie Curie, Paris 6; 2010. Available from: http://tel.archives-ouvertes.fr/tel-00488701/fr.

[70] Lamarche AMJ. Putting the singing voice on the map - towards improving the quantitative evaluation of voice status in professional female singers. KTH, Stockholm, Sweden; 2009.

[71] Rayleigh JWS. The theory of sound. No. 2 in the theory of sound. Macmillan; 1896. Available from: http://books.google.de/books?id=pkIIAAAAIAAJ.

[72] Swerdlin Y, Smith J, Wolfe J. The effect of whisper and creak vocal mechanisms on vocal tract resonances. J Acoust Soc Am 2010; 127: 2590-8.

[73] Hollien H, Michel JF. Vocal fry as a phonational register. J Speech Hear Res 1968; 11: 600-4.

[74] Blomgren M, Chen Y, Ng ML, Gilbert HR. Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. J Acoust Soc Am 1998; 103(5): 2649-58. Available from: http://link.aip.org/link/?JAS/103/2649/1.

[75] Miller DG, Sulter AM, Schutte HK, Wolf RF. Comparison of vocal tract formants in singing and nonperiodic phonation. J Voice 1997; 11(1): 1-11.

[76] Epps J, Smith JR, Wolfe J. A novel instrument to measure acoustic resonances of the vocal tract during phonation. Meas Sci Technol 1997; 8: 1112-21.

[77] Dowd A, Smith J, Wolfe J. Learning to pronounce vowel sounds in a foreign language using acoustic measurements of the vocal tract as feedback in real time. Language Speech 1997; 41(1): 1-20.

[78] Kob M, Neuschaefer-Rube C. A method for measurement of the vocal tract impedance at the mouth. Med Eng Phys 2002; 24(7-8): 467-71. Available from: http://www.sciencedirect.com/science/article/B6T9K-46RDKN1-5/2/c429a6ac3991ae2d1753731d78719988.

[79] Story BH, Titze IR, Homan EA. Vocal tract area functions for an adult female speaker based on volumetric imaging. J Acoust Soc Am 1998; 104(1): 471-87.

[80] Story BH, Titze IR, Homan EA. The relationship of vocal tract shape to three voice qualities. J Acoust Soc Am 2001; 109(4): 1651-67.

[81] Tom K, Titze IR, Homan EA, Story BH. Threedimensional vocal tract imaging and formant structure: varying vocal register, pitch, and loudness. J Acoust Soc Am 2001; 109(2): 742-7.

[82] Sulter AM, Miller DG, Wolf RF, Schutte HK, Wit HP, Mooyaart EL. On the relation between the dimensions and resonance characteristics of the vocal tract: a study with MRI. Magn Reson Imaging 1992; 10: 365-73.

[83] Takemoto H, Honda K, Masaki S, Shimada Y, Fujimoto I. Measurement of temporal changes in vocal tract area function from 3D cine-MRI dat. J Acoust Soc Am 2006; 119: 1037-49.

[84] Rothenberg M. An interactive model for the voice source. Dept Speech, Music and Hearing. KTH, Stockholm, 1981; 22. Quarterly Progress Status Report.

[85] Fletcher NH. Autonomous vibration of simple pressurecontrolled valves in gas ows. J Acoust Soc Am 1993; 93: 2172-80.

[86] Titze IR. Acoustic interpretation of resonant voice. J Voice 2001; 15: 519-28.

[87] Childers DG, Wong CF. Measuring and modeling vocal source-tract interaction. IEEE Trans Biomed Eng 1994; 41(7): 663-71.

[88] Titze IR. Nonlinear source-filter coupling in phonation: Theory. J Acoust Soc Am 2008; 123: 2733-49.

[89] Sundberg J, Skoog J. Dependence of jaw opening on pitch and vowel in singers. J Voice 1997; 11: 301-6.

[90] Joliveau E, Smith J, Wolfe J. The tuning of vocal tract resonances by sopranos. Nature 2004; 427: 116.

[91] Henrich N, Smith J, Wolfe J. Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. J Acoust Soc Am 2011; 129: 1024-35.

[92] Garnier M, Henrich N, Smith J, Wolfe J. Vocal tract adjustments in the high soprano range. J Acoust Soc Am 2010; 127: 3771-80. Available from: http://hal.archives-ouvertes.fr/hal-00480078/fr.

[93] Sundberg J. Articulatory configuration and pitch in a classically trained soprano singer. J Voice 2009; 23(5): 546-51.

[94] Bloothooft G, Bringmann E, Cappellen MV, Luipen JBV, Thomassen KP. Acoustics and perception of overtone singing. J Acoust Soc Am 1992; 92(4): 1827-35.

[95] Levin TC, Edgerton ME. The Throat Singers of Tuva. 1999; URL: http://www.sciam.com.

[96] Kob M. Analysis and modelling of overtone singing in the sygyt style. Appl Acoust 2004; 65(12): 1249-59. Available from: http://www.sciencedirect.com/science/article/B6V1S-4D5JVMR-4/2/0c9aea4543f5c6d8ffde99e9dc141cc8.

[97] Ishizaka K, Flanagan JL. Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell Sys Tech J 1972; 51: 1233-68.

[98] Mergell P, Herzel H, Wittenberg T, Tigges M, Eysholdt U. Phonation Onset: Modelling and High Speed Glottography. J Acoust Soc Am 1998; 104: 464-70.

[99] Herzel H, Berry D, Titze IR, Saleh M. Analysis of vocal disorders with methods from nonlinear dynamics. J Speech Hearing Res 1994; 37: 1008-19.

[100] Herzel H, Reuter R. Quantifying correlations in pitch and amplitude contours of sustained phonation. Acta Acoust United With Acoust 2000; 86: 129-35.

[101] Mende W, Herzel H, Wermke K. Bifurcations and chaos in newborn infant cries. Phys Lett A 1990; 145: 418-24.

[102] Dolansky L, Tjernlund P. On certain irregularities of voiced speech forms. IEEE Trans AU. 1968; 16: 51-6.

[103] Mazo M, Ericson D, Harvey T. Emotion and expression: temporal data on voice quality in russian lament. In: Fujimura O, Hirano M, eds. Vocal Fold Physiology: Voice quality control. San Diego: Singular Publishing Group; 1995. p. 173-8.

[104] Wilden I, Herzel H, Peters G, Tembrock G. Subharmonics, biphonation, and deterministic chaos in mammal vocalization. Bioacoustics 1998; 9: 171-96.

[105] Fitch WT, Neubauer J, Herzel H. Calls out of Chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. Animal Behav 2002; 63: 407-18.

[106] Edgerton M, Neubauer J, Herzel H. Nonlinear phenomena in contemporary vocal musical composition and performance. Perspectives New Music 2003; 41: 30-65.

[107] Steinecke I, Herzel H. Bifurcations in an asymmetric vocal fold model. J Acoust Soc Am 1995; 97: 1571-8.

[108] Fuks L. From Air to Music - Acoustical, physiological and perceptual aspects of reed wind instrument playing and vocal-ventricular fold phonation. Royal Institute of Technology, Stockholm, Sweden; 1999.

[109] Neubauer J, Herzel H, Mergell P, Eysholdt U. Spatiotemporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. J Acoust Soc Am 2001; 110: 3179-92.

[110] Hatzikirou H, Fitch WT, Herzel H. Voice instabilities due to source-tract interactions. Acta Acoust United With Acoust 2006; 92: 468-75.

[111] Neubauer J, Edgerton M, Herzel H. Nonlinear phenomena in contemporary vocal music. J of Voice 2004; 18: 1-12.

[112] Miller DG, Švec JG, Schutte HK. Measurment of characteristic leap interval between chest and falsetto registers. J Voice 2002; 16: 8-19.

[113] Tokuda IT, Horáček J, Svec JG, Herzel H. Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. J Acoust Soc Am 2007; 122: 519-31.

[114] von Kempelen W. Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine. Wien: Degen; 1791.

[115] Scherer RC, Shinwari D, Witt KJD, Zhang C, Kucinschi BR, Afjeh AA. Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. J Acoust Soc Am 2001; 109(4): 1616-30. Available from: http://link.aip.org/link/-?JAS/109/1616/1.

[116] Li S, Scherer RC, Wan M, Wang S, Wu H. Numerical study of the effects of inferior and superior vocal fold surface angles on vocal fold pressure distributions. J Acoust Soc Am 2006; 119(5 Pt 1): 3003-10.

[117] Kucinschi BR, Scherer RC, Dewitt KJ, Ng TTM. An experimental analysis of the ressures and flows within a driven mechanical model of phonation. J Acoust Soc Am 2006; 119(5 Pt 1): 3011-21.

[118] Thomson SL, Tack JW, Verkerke GJ. A numerical study of the flow-induced vibration characteristics of a voice-producing element for laryngectomized patients. J Biomech 2007; 40(16): 3598-3606. Available from: http://dx.doi.org/10.1016/j.jbiomech.2007.06.007.

[119] Kob M. Physical modeling of the singing voice [Dissertation]. RWTH Aachen University. Logos-Verlag, Berlin; 2002.

[120] Dejonckere PH, Kob M. Pathogenesis of vocal fold nodules: new insights from a modelling approach. Folia Phoniatr Logop 2009; 61(3): 171-9. Available from: http://dx.doi.org/10.1159/0002-19952.

[121] Story BH, Titze IR. Voice simulation with a body-cover model of the vocal folds. J Acoust Soc Am 1995; 97: 1249-60.

[122] Pelorson X, Hirschberg A, van Hassel RR, Wijnands APJ, Auregan Y. Theoretical and experimental study of quasisteady flow separation within the glottis during phonation. J Acoust Soc Am 1994; 96: 3416-31.

[123] Lous NJ, Hofmans GC, Veldhuis RNJ, Hirschberg A. A symmetrical two mass vocal fold model coupled to vocal tract and trachea, with application to prothesis design. Acta Acoust United With Acoust 1998; 84: 1135-50.

[124] Sciamarella D, d'Alessandro C. On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters. Acta Acoust United With Acoust 2004; 90: 746-61.

[125] Alipour F, Berry DA, Titze IR. A finite-element model of vocal-fold vibration. J Acoust Soc Am 2000; 108: 3003-12.

[126] Titze IR, Story B. Rules for controlling low-dimensional vocal fold models with muscle activation. J Acoust Soc Am 2002; 112: 1064-76.

[127] Adachi S, Yu J. Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing. J Acoust Soc Am 2005; 117: 3213-24.

[128] Tokuda IT, Zemke M, Kob M, Herzel H. Biomechanical modeling of register transitions and the role of vocal tract resonators. J Acoust Soc Am 2010; 127(3): 1528-36. Available from: http://dx.-doi.org/10.1121/1.3299201.

[129] Berry DA, Herzel H, Titze IR, Story BH. Bifurcations in excised larynx experiments. J Voice 1996; 10: 129-38.

[130] Švec JG, Schutte HK, Miller DG. On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges. J Acoust Soc Am 1999; 106: 1523-31.

[131] Horáček J, Švec JG, Veselý J, Vilkman E, Klepáček I, Vetešník A. Measurement of the vocal-fold vibration behaviour in excised human larynges. In: Proceedings of 2nd International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications 2001; p. 1-6.

[132] Švec JG, Šram F, Schutte HK. Videokymography in voice disorders: What to look for? Ann Otol Rhinol Laryngol 2007; 116: 172-80.

[133] Kelly JL, Lochbaum C. Speech synthesis. In: Proceedings of the Fourth International Congress on Acoustics, 1962; p. 1-4. Paper G42.

[134] Liljencrants J. Speech synthesis with a reflection-type line analog. Department of Speech Communication and Music Acoustics, Royal Institute of Technology; 1985.
[135] Story BH. Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract. Department of Speech Pathology and Audiology, University of Iowa; 1995.
[136] Titze IR. Myoelastic Aerodynamic Theory of Phonation. Iowa City, IA: National Center for Voice and Speech; 2006.
[137] Švec JG, Pešak J. Vocal breaks from the modal to falsetto register. Folia Phoniatr 1994; 46: 97-103.
[138] Wolfe J, Garnier M, Smith J. Voice acoustics: an introduction; 2009. Downloaded 27/10/2010. Online publication. Available from: www.phys.unsw.edu.au/jw/voice.html.

Received: December 02, 2010 Revised: December 20, 2010 Accepted: December 23, 2010