Semioccluded Vocal Tract Exercises
A physiologic approach for voice training and therapy
MARCO GUZMAN

Semioccluded Vocal Tract Exercises
A physiologic approach for voice training and therapy

ACADEMIC DISSERTATION
To be presented, with the permission of the Board of the School of Education of the University of Tampere, for public discussion in the auditorium Virta 109, Åkerlundinkatu 5, Tampere, on 24 February 2017, at 12 o’clock.

UNIVERSITY OF TAMPERE
MARCO GUZMAN

Semioccluded Vocal Tract Exercises

A physiologic approach for voice training and therapy

Acta Universitatis Tamperensis 2265
Tampere University Press
Tampere 2017
The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Copyright ©2017 Tampere University Press and the author

Cover design by
Mikko Reinikka
CONTENTS

1. Introduction .............................................................................................................. 14
   1.1 Approaches of voice therapy ................................................................. 14
   1.2 Physiologic voice therapy programs ................................................. 16
   1.3 Semioccluded vocal tract exercises .................................................... 19
   1.4 Semiocclusions of vocal tract and energy transfer .......................... 23
   1.5 Radiological imaging methods ............................................................ 25
   1.6 Aerodynamic measurements of phonation ...................................... 27
   1.7 Electroglottography .............................................................................. 28
   1.8 High speed digital imaging ................................................................. 30
   1.9 Naso-laryngeal fibroscopy ................................................................. 32
   1.10 Aims of the study ................................................................................. 34

2. Material and methods .............................................................................................. 36
   2.1 Article I ..................................................................................................... 36
   2.2 Article II .................................................................................................. 44
   2.3 Article III ................................................................................................. 48
   2.4 Article IV ................................................................................................. 52

3. Results ....................................................................................................................... 56
   3.1 Article I ..................................................................................................... 56
   3.2 Article II .................................................................................................. 57
   3.3 Article III ................................................................................................. 59
   3.4 Article IV ................................................................................................. 60

4. Discussion ................................................................................................................... 62
   4.1 Influence of semioccluded exercises on aerodynamic variables ... 62
4.2 Influence of semioccluded exercises on glottal function .......... 66
4.3 Influence of semioccluded exercises on vocal tract function .... 71
4.4 Possible sources of error ............................................... 80
4.5 Future research .................................................................. 81

5. Conclusions ............................................................................. 83
6. Acknowledgements ................................................................. 85
7. Financial support ................................................................. 86
8. References .................................................................................. 87
LIST OF ORIGINAL ARTICLES


AUTHOR’S CONTRIBUTION

Article I
Vocal Tract and Glottal Function During and After Vocal Exercising with Resonance Tube and Straw.
All investigators participated in the research design of this study. The researcher in charge of CT samples was Dr. Petr Krupa. Acoustic, EGG and aerodynamic recordings were made by the author of this dissertation and Dr. Anne-Maria Laukkanen. Moreover, the author of this dissertation made all measurement from CT, acoustic, EGG, and aerodynamic samples. Auditory perceptual analysis procedures were also organized by Marco Guzman. He also wrote the manuscript and created the figures and tables. Dr. Anne-Maria Laukkanen, Dr. Sc. Jaromír Horáček, and Dr. Ahmed Genied provided their comments.

Article II
Original research design was made mainly by Marco Guzman and Dr. Anne-Maria Laukkanen. Thereafter, all researchers made contributions to obtain the final design. All investigators participated in different ways during experimental laryngoscopic procedures. HSDI samples were acquired by Dr. Matthias Echternach and Dr. Ahmed Geneid. High speed material was analyzed by Dr. Matthias Echternach. Dr. Daniel Muñoz performed statistical analysis of data. The author of the present dissertation entirely wrote the manuscript and was in charge
of figures and tables. Dr. Anne-Maria Laukkanen, Dr. Matthias Echternach, and Dr. Ahmed Geneid provided their comments.

Article III
Air pressure and contact quotient measures during different semi-occluded postures in subjects with different voice conditions.
The author of the present dissertation made the research design, participated in all acoustic, EGG and aerodynamic recordings, performed the analysis of all samples, and wrote the manuscript. He also was in charge of figures included in the original article. Dr. Christian Olavarria and Dr. Miguel Leiva performed laryngeal videostroboscopic procedures to corroborate laryngeal diagnosis of subjects. Christian Castro, Sofia Madrid, and Elizabeth Jaramillo collaborated in recruitment of subjects and experimental procedures. Statistical analysis was carried out by Dr. Daniel Muñoz. Comments and contributions on the manuscript during the writing process were made by Dr. Anne-Maria Laukkanen.

Article IV
Laryngeal and Pharyngeal Activity During Semi-occluded Vocal Tract Postures in Subjects Diagnosed With Hyperfunctional Dysphonia.
Research design was made by Marco Guzman, Christian Castro, and Dr. Julia Gerhard. All laryngoscopic procedures were performed by Dr. Alba Testart. Marco Guzman wrote the entire manuscript, participated during all laryngoscopic procedures and was in charge of visual assessment of laryngoscopic samples by blinded judges. Christian Castro collaborated with recruitment of subjects and experimental procedures. Statistical analysis was carried out by Dr. Daniel Muñoz. English correction and comments on the manuscript were made by Dr. Julia Gerhard and Christian Castro.
LIST OF ABBREVIATIONS AND SYMBOLS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALR</td>
<td>Amplitude to length ratio</td>
</tr>
<tr>
<td>CIQ</td>
<td>Closing quotient</td>
</tr>
<tr>
<td>CQ</td>
<td>Contact quotient</td>
</tr>
<tr>
<td>CT</td>
<td>Computerized tomography</td>
</tr>
<tr>
<td>EGG</td>
<td>Electroglottography</td>
</tr>
<tr>
<td>F0</td>
<td>Fundamental frequency</td>
</tr>
<tr>
<td>F1</td>
<td>First formant</td>
</tr>
<tr>
<td>F2</td>
<td>Second formant</td>
</tr>
<tr>
<td>F3</td>
<td>Third formant</td>
</tr>
<tr>
<td>F4</td>
<td>Fourth formant</td>
</tr>
<tr>
<td>F5</td>
<td>Fifth formant</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast-Fourier Transform</td>
</tr>
<tr>
<td>HNR</td>
<td>Harmonic to noise ratio</td>
</tr>
<tr>
<td>HSDI</td>
<td>High speed digital imaging</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>LTAS</td>
<td>Long-term average spectrum</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>OQ</td>
<td>Open quotient</td>
</tr>
<tr>
<td>Poral</td>
<td>Oral pressure</td>
</tr>
<tr>
<td>Psub</td>
<td>Subglottic pressure</td>
</tr>
<tr>
<td>PTP</td>
<td>Phonation threshold pressure</td>
</tr>
<tr>
<td>Ptrans</td>
<td>Transglottal pressure</td>
</tr>
<tr>
<td>SOVTE</td>
<td>Semioccluded vocal tract exercises</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>SF</td>
<td>Spectral flatness</td>
</tr>
<tr>
<td>SPL</td>
<td>Sound pressure level</td>
</tr>
<tr>
<td>VLP</td>
<td>Vertical laryngeal position</td>
</tr>
</tbody>
</table>
Physiologic approach of voice therapy is commonly used by speech-language pathologists in treating patients with a wide variety of voice disorders. Voice rehabilitation programs based on physiologic approach are designed to modify the physiology of the vocal mechanism, including the three subsystems involved in voice production (breathing, phonation, and resonance) in an integrated or holistic way. Examples of physiologic voice therapy programs include: Resonant Voice Therapy, the Accent Method (of Voice Therapy), and Vocal Function Exercises. One common aspect to physiologic voice therapy programs is that they take advantage of semiocluded vocal tract exercises (SOVTE). This group of exercises includes phonation on voiced fricatives (consonants), nasals (consonants), lip and tongue trills, hand over mouth, and phonation into different tubes with the free end either freely in the air or submerged into a recipient filled with water. The general aim of the present dissertation was to determine how the three subsystems involved in voice production (breathing, phonation, and resonance) are affected by different SOVTE. If this group of exercises is considered as a physiologic approach for voice therapy, they should affect simultaneously glottal function, vocal tract function, and aerodynamic variables during exercising.

Specific aims of this dissertation were: 1) to investigate the vocal tract modifications and also the acoustic, aerodynamic and EGG characteristics of the voice when comparing vocal exercising with two different tubes (traditional Finnish glass tube and a thin stirring straw); 2) to observe the influence of tube
phonation into water on vocal fold vibration by using HSDI and phonovibrography; 3) to determine the effect of different artificial lengthening of the vocal tract (phonation through a tube with the free end in water and in air) on air pressure variables and vocal fold adduction in different types of subjects; and 4) to observe and compare the effects of different types of SOVTE on VLP, A-P laryngeal compression and pharyngeal width in a group of subjects diagnosed with hyperfunctional dysphonia.

Study I assessed a male subject with classical singing training. CT was carried out when the subject produced the vowel [a:] at comfortable speaking pitch, phonating into the resonance tube (the distal end in the air) and when repeating [a:] after the exercise. A similar procedure was performed with a narrow straw after fifteen-minute silence. Anatomic distances and areas were measured from CT midsagittal and transversal images. Acoustic, perceptual, EGG and Poral parameters were also investigated. Study II evaluated eight subjects via HSDI while phonating into a silicon tube with the free end submerged into water. Two test sequences were studied: 1) phonation pre, during and post tube submerged 5 cm into water, with tube phonation lasting for five minutes, and 2) phonation into the tube submerged 5 cm, 10 cm, and 18 cm into water. Several glottal area parameters were calculated using phonovibrograms. Study III assessed forty-five participants representing four vocal conditions: subjects diagnosed with 1) normal voice and without voice training, 2) normal voice with voice training, 3) muscle tension dysphonia, and 4) unilateral vocal fold paralysis. Participants phonated into different kinds of tubes (drinking straw, 5 mm in inner diameter; stirring straw, 2.7 mm in inner diameter; silicon tube, 10 mm in inner diameter) with the free end in air and in water. Aerodynamic, acoustic and EGG signals were captured simultaneously. Mean values of the following variables were considered: EGG CQ, F0, Psub, Poral, and Ptrans. Study IV evaluated twenty subjects with hyperfunctional dysphonia. All participants were required to produce eight
different SOVTE during observation with flexible endoscope: lip trills, hand-over-mouth, phonation into four different tubes, and tube phonation into water using two different immersion depths. Each exercise was produced at three loudness levels: habitual, soft, and loud. A evaluation test with three blinded laryngologists was performed to determine the VLP, anterior-posterior laryngeal (A-P) compression, and pharyngeal width. Judges rated the three endoscopic variables using a 5-point scale. To assess intra and inter-rater agreement, an intraclass correlation coefficient was used. A multivariate linear regression model considering VLP, pharyngeal width, and A-P laryngeal compression was conducted. Correlation analysis between variables was also conducted.

Study I demonstrated that during and after phonation into the tube or straw the velum closed the nasal passage, the vertical laryngeal position was lower and the hypopharyngeal area became wider compared to condition pre tube/straw. Furthermore, the ratio between the inlet of the lower pharynx and the outlet of the epilaryngeal tube became larger during and after tube/straw phonation. A stronger spectral prominence in the singer/speaker’s formant cluster region after tube and straw phonation was showed by acoustic results. Listening test revealed a better voice quality after both straw and tube compared to condition before exercises. CQ$_{EEG}$ decreased both during the tube and the straw and remained lower after exercising. Psub showed an increase during straw and remained higher after it. Larger changes were observed during straw. Study II demonstrated certain trends. Amplitude to length ratio, harmonic to noise ratio and spectral flatness (derived from glottal area) decreased for all tube immersion depths, and glottal closing quotient increased for 10 cm immersion. Contact quotient in turn increased for 18 cm immersion. CQ$_{EEG}$ decreased during phonation into the tube in 5 cm depth, and jitter decreased during and after it. Study III showed that all SOVTE had a significant effect on Psub, Poral, Ptrans, and CQ$_{EEG}$. Phonation into a 55 cm silicon tube submerged 10 cm in water and phonation into a stirring straw (the outer end of
it in the air) resulted in the highest values for CQ, Psub, and Poral compared to baseline for all groups of subjects, regardless of the vocal status. Poral and Psub correlated positively. Study IV indicated that VLP, A-P laryngeal compression and pharyngeal width changed differently throughout the eight semi-occluded postures. A narrower aryepiglottic opening, a wider pharynx, and a lower larynx than resting position were produced by all SOVTE. Tube into the water and narrow tube into the air demonstrated more prominent changes. Correlation analysis showed that VLP significantly correlated with pharyngeal width and A-P laryngeal compression. Additionally, pharyngeal width significantly correlated with A-P laryngeal compression.

From the present dissertation it is possible to conclude that: 1) vocal exercises including semiocclusions and artificial lengthening of the vocal tract have a simultaneous influence on variables related to phonation, resonance, and aerodynamic aspects. There is a clear interaction between the three subsystems, each one influencing the two others during voicing; 2) Most changes produced during SOVTE constitute relevant therapeutic and voice training tools for subjects with voice disorders and for professional voice users, respectively. These changes seem to lead to a more economic voice production; 3) Subjects with organic or behavioral voice disorders, as well as healthy vocally untrained, and healthy trained subjects behave similarly under similar exercise conditions. 4) Aerodynamic aspects are mainly affected during SOVTE by increasing air pressure measures such as Poral and Psub, the latter being a compensation of the former. An increase in Poral, in turn, is expected since semiocclusions decrease the airflow rate. 5) Influence of SOVTE in phonation is mainly related to the degree of glottal adduction during exercising. This is mostly determined by the degree of resistance to the airflow that each exercise offers and compensations caused by resistance. High resistance seems to promote a greater vocal fold adduction compared to a lower resistance exercise. Therefore, higher resistance exercises could be more
beneficial for patients with diagnosis such as vocal fold paralysis or presbyphonia. On the other hand, glottal function of patients with hyperadduction or vocal fatigue could be improved using exercises that involve lower airflow resistance; 6) Important therapeutic goals are reached during SOVTE regarding vocal tract configuration. A lower larynx, a wider pharynx and a higher velum is produced during this type of exercises compared to vowel phonation. The changes are more prominent when the airflow resistance is higher. These modifications are linked to positive changes in acoustic and perceptual output of voice.
1 INTRODUCTION

1.1 Approaches of voice therapy

Over the years a variety of approaches (orientations) have emerged for the treatment of voice disorders. Stemple (2000) classified the voice therapy approaches into four major categories: hygienic, psychogenic, symptomatic, and physiologic.

Hygienic orientation of voice therapy is based on two main aspects: 1) many functional voice disorders are initiated and maintained by behaviors or vocal habits that damage the laryngeal structures; 2) elimination of harmful and traumatic behaviors will improve vocal performance. Therefore, hygienic approach of voice therapy focuses on the identification and elimination of harmful vocal behaviors followed by the development of proper vocal behaviors (Thomas and Stemple 2007). Common components of hygienic approach of voice therapy include: hydration of the larynx, voice rest, silent cough, avoidance of screaming and yelling, control of vocal loading, etc. (Stemple 2000).

Psychogenic approach of voice therapy supposes that underlying emotional or psychosocial problems may cause voice disorders (Stemple 2000). This approach of voice therapy focuses on identification and modification of the emotional and psychosocial problems related to the onset and maintenance of the voice disorder. According to this approach, when the psychogenic causes are resolved, the voice disorder is eliminated (Aronson 1990).
Symptomatic voice therapy is based on the modification of vocal symptoms. Voice therapy under this approach focuses on the modification of aberrant vocal symptoms related to pitch, loudness and voice quality. Symptomatic voice therapy is based on the belief that modification and correction of phonation, respiratory, and resonance characteristics produce an improvement in the voice condition (Thomas and Stemple 2007). Symptomatic approach involves several vocal exercises to modify aberrant vocal symptoms (Boone 1977). In the recent years additional facilitating voice exercises have been added to the original list of symptomatic exercises (Boone and McFarlane 1988; Boone, McFarlane, and Von Berg 2005). Some of the traditional and commonly used facilitating exercises in voice therapy are: pushing exercises, humming, chewing exercise, yawn-sigh, change of loudness, inhalation phonation, digital manipulation/digital pressure, relaxation, establishing a new pitch, and “placement of the voice” techniques.

On the other hand, physiologic voice therapy is based on the belief that voice disorders are best treated by modifying the underlying physiology of voice production (Stemple 2000; Stemple, Graze, and Klaben 2000). Stemple et al. (2000) suggest that the physiologic approach involves three key components: 1) improving the balance among the main subsystems involved in voice production: respiration, phonation, and resonance (vocal tract configuration and sensations related to “voice placement”); 2) improving the strength, balance, tone, and stamina of laryngeal muscles; and 3) developing a healthy mucosa covering of the true vocal folds. Evidence suggests that physiologic methods of voice therapy have greater scientific support (larger number of studies and higher level of evidence) than other approaches of voice therapy (Thomas and Stemple 2007). Examples of physiologic voice therapy programs include: Vocal Function Exercises (Stemple 2000), the Accent Method of Voice Therapy (Kotby 1995), and Resonant Voice Therapy (Verdolini 1998). Each program approaches the voice condition in a
holistic or integrated manner with the aim of altering the overall physiology of voice production.

1.2 Physiologic voice therapy programs

One common aspect to physiologic voice therapy programs mentioned above is that all of them take advantage of SOVTE. This group of exercises includes phonation on voiced fricatives, nasals, close vowels, lip and tongue trills, hand over mouth, and phonation into different tubes with the free end either freely in the air or submerged into a recipient filled with water.

Vocal function exercises program is based on series of four steps, all of them using a semiclosed configuration of the vocal tract. In the first step (warm-up), the patient produces a sustained vowel [i:] at a predetermined pitch. In the second and third steps, the patient is required to produce glides up and glides down respectively using a sustained bilabial consonant [β:] with an exaggerated inverted megaphone shape in the vocal tract. In the final step, the patient has to sustain the consonant [β:] with different predetermined pitches using the same inverted megaphone shape that step two and three. Patients are required to produce all exercises as softly as possible with a forward “placement of the voice”. Stemple et al. (2000) stated that by improving the strength, endurance and coordination of the systems involved in voice production (breathing, phonation, and resonance), the exercises help to rehabilitate the voice.

The Accent Method focuses on the improvement of the respiratory, phonatory, and articulatory aspects of speech using a holistic manner. This method is also considered holistic or physiologic due to it simultaneously involves various voice features such as pitch, loudness and timbre (Bassiouny 1998). Better control of voice production is considered as the main goal of the Accent Method. Breath
support is expected to be affected by respiratory accents during exercises in this method. The improvement of breath support is believed to reduce excessive laryngeal muscular effort (Kotby et al. 1991). The Accent Method is based on three main principles: 1) optimal breath support; 2) rhythmic practice of accentuated vowels with progressive carryover to connected speech; and 3) rhythmic body movements (Bassiouny 1998). In practical terms, Accent Method protocol is based on the training the abdominodiaphragmatic breath while producing rhythmic voiced fricative consonants and close vowels (all considered semioccluded postures of the vocal tract). Once the respiratory-phonation connection is well established, the enhanced respiratory-phonation pattern is generalized to connected speech.

Resonant voice therapy, which is another physiologic program of voice rehabilitation, teaches how to modify the voice production so that the result is a more resonant voice quality. This program defines resonant voice as voice production involving vibratory sensations, usually on the anterior region of face and mouth in the context of easy phonation (Stemple et al. 2000). The goal of this therapy program is to achieve the strongest possible voice with the least effort and vocal folds impact to minimize the probability of vocal fold injury (Verdolini et al. 1998). Verdolini (1998) suggested that resonant voice is associated with vocal folds that are barely adducted or barely abducted. This glottic configuration is expected to produce a relatively strong voice while providing some protection against vocal fold tissue injury. In the resonant voice therapy program, the patient moves through several stages within a hierarchy of increasing complexity. Advancement through the hierarchy demands evidence of skill acquisition at earlier stages (Roy et al. 2003). The methodology of this approach focuses on the processing of sensory information. The therapy procedures begin with stretching and breathing maneuvers followed by basic training gestures (voice exercises) (Chen et al. 2007). In practical terms, phonatory exercises in the resonant voice therapy program includes nasal and voiced fricative consonants (considered as
semioccluded postures of the vocal tract) while the patient phonates different phonatory tasks from a sustained tone to connected speech (words, sentences, etc). All steps involve sensory motor learning principles.

Even though Vocal Function Exercises, Resonant Voice Therapy and Accent Method are the most known and studied physiologic voice therapy programs, there are a number of programs or voice exercise sequences, based on the same physiologic principles, which speech therapists use in different countries around the globe. Many of those programs also take advantage of SOVTE. One of the most commonly used semioccluded exercise is the tube phonation, either with the free end in the air or submerged into the water. The latter is called water resistance therapy.

Several types of tubes are currently used for voice therapy and training, some of them are: 1) the traditional Finnish resonance glass tube (26-28 cm in length and 9 mm in inner diameter for adults, 24-25 cm, 8 mm for children); 2) commercial plastic drinking straws; and 3) plastic stirring straws (narrow straws to stir coffee). Furthermore, two main versions of water resistance therapy (tube phonation into water) have been used: 1) Phonation into traditional Finnish resonance glass tube submerged in water in a bowl (Laukkanen 1992); and 2) Lax Vox technique, which consists of phonation in a flexible silicone tube (35 cm in length and 9-12 mm in inner diameter) submerged into a bottle filled with water (Sihvo 2006). In the glass tube method, tubes are kept 1-2 cm or 5-15 cm below the surface of the water depending on the patient’s needs. With the Lax Vox method, the participant is instructed to keep the tube 1-7 cm in water for voice therapy (Sihvo 2006). Additionally, some voice therapists or voice trainers use commercial plastic drinking straws with the free end in a glass or bottle filled with water as a simple version of water resistance therapy.

A considerable number of studies have been carried out to disclose the underlying physics and physiology of tube phonation and other SOVTE. Some of
them have explored changes in the vocal fold vibration, others in vocal tract configuration, and there are also studies focusing on the aerodynamic variables.

1.3 Semi-occluded vocal tract exercises

SOVTE have been widely used in voice therapy, vocal warm up and voice training. It has been stated that the vocal exercises, such as artificial lengthening of the vocal tract or narrow anterior constriction of the vocal tract, increase the vocal tract impedance, specifically the inertive reactance which may favorably influence the vocal fold vibration (Titze 1998; Story et al. 2000). Increased inertive reactance changes the glottal flow amplitude and pulse shape (Rothenberg 1986; Story et al. 2000; Titze 2006a; Titze 2006b). Furthermore, the PTP (the Psub required to barely initiate and sustain phonation) is reduced by increased vocal tract inertance (Titze 1998).

Vocal tract impedance can affect the voice in two ways: 1) through an acoustic-aerodynamic interaction, and 2) through a mechano-acoustic interaction (Rothenberg 1986; Story et al. 2000).

Regarding acoustic-aerodynamic interaction, the shape of the glottal flow pulse is affected by the acoustic pressures in the vocal tract (Rothenberg 1981, Rothenberg 1986; Titze and Story 1997; Titze and Laukkanen 2007). When fundamental frequencies are below the first formant, the acoustic pressures in the supraglottis cause a skewed flow pulse compared to variation of the glottal area. Thus, the glottal airflow is suppressed at glottal opening and maintained during the glottal closing phase. This increased skewing of the glottal flow waveform leads to strengthening of the higher harmonics (less spectral tilt), increase in SPL (Bickley and Stevens 1987; Laukkanen et al 2008; Titze 2008), and to a more resonant voice quality, i.e. a brighter and louder sound, which is characterized by enhanced
vibratory sensations in the frontal part of face and mouth and easy voice production (Titze 2004; Titze 2006a; Titze 2006b; 2008) pointed out that, since the skewing of the glottal airflow signal is one of the determinants of vocal intensity, the source-filter interaction can be used to increase intensity rather than vibrational amplitude of the vocal folds, thus avoiding an increase in the vocal fold impact stress.

The second way that voice production can be affected by vocal tract impedance is the mechano-acoustic interaction of the vocal tract pressures and the vocal folds (Rothenberg 1986; Titze and Story 1997). When the vocal tract inertance is increased, it may affects the vocal fold vibration itself. Specifically, the inertive reactance lowers the PTP (Titze and Story 1997). A low PTP would promote ease of phonation (decrease in perceived phonatory effort). According to Story et al. (2000) this occurs when F0 approaches F1. Hence, when phonation is produced at a frequency or close to F1 may allow for an efficient voice production that could be associated with lower phonatory effort.

Related to this, Titze (1988; 2008) states that if the optimal condition for phonation is that of maximum inertive reactance, F0 should coincide with the peak of the reactance curve which is not located at F1 but occurs at a frequency somewhat below it. To maintain favorable impedance conditions across pitches and loudness, a vocalist may want to decrease the epilaryngeal area. When the narrowed epilarynx is combined with a wide pharynx, the reactance never goes negative below about 3000 Hz, which means that the acoustic load is inertive for all possible values of F0 regardless of vowel or consonant (Titze and Story 1997).

In a recent modeling study with excised larynges, Conroy et al. (2014) evaluated PTP during nine conditions: control, two tube diameters, three tube lengths, and three levels of flow input. A significantly decreased in PTP was detected for the longest tube and narrower tubes. These findings corroborate the earlier statements related to the mechano-acoustic interaction of the vocal tract pressures and the vocal folds (e.g. Titze and Story 1997).
Several studies with subjects phonating into resonance tubes of varying lengths and diameters have been performed by Laukkanen to explore glottal function (Laukkanen 1992; Laukkanen et al. 1995a; Laukkanen et al. 1995b; Laukkanen et al. 1996; Laukkanen et al. 2007; Laukkanen et al. 2008). Changes in several different variables both during and after tube phonation were reported across these studies, including glottal waveform shape, glottal efficiency, laryngeal resistance, and perceived vocal effort among others. According to Laukkanen et al. (1995a), electromyographic evaluation showed that the average laryngeal muscle activity during phonation into a tube was the equal or lower compared to phonation during sustained vowel. These results seem to imply that during increased impedance of the vocal tract there is a decreased laryngeal muscle activity and it tends to stay lower in phonation immediately after exercises.

Related to glottal function, Titze et al. (2002) reported lower amplitude and lower glottal closed time (obtained from EGG signal) during phonation into straws compared to sustained vowel phonation. Authors suggested that the use of high subglottal pressures needed in singing are possible with narrow straws, having minimal vocal folds collision. Bickley and Stevens (1987) reported similar findings in their study of EGG waveforms produced by a group of participants phonating on voiced fricatives. Authors reported that for production of voiced fricative consonants, the open phase of the glottal cycle increased by over 20% relative to open postures of the vocal tract. Gaskill et al. (2008) studied the effect of a voiced lip trill on \(CQ_{\text{EGG}}\) in classically trained singers and vocally untrained participants. Most participants showed a tendency for a reduction in \(CQ_{\text{EGG}}\) during the lip trill, with a more pronounced change in the untrained participants.

Some earlier studies have also assessed changes in the vocal tract shape during SOVTE. Vampola et al. (2011) using CT observed the vocal tract shape in a female subject before, during, and after phonation into a tube. Finite-element models were made on the basis of CT registrations to study changes in vocal tract input.
impedance. CT results indicated that the most dominant change during phonation into the tube was the expansion of the cross-sectional area of the oropharynx and the oral cavity due to a different tongue position. CT images also revealed that the velum rose to seal the nasopharyngeal port during tube phonation, and this change remained after the exercises. Moreover, the total volume of the vocal tract was considerably larger after phonation into the tube. Calculations made with finite-element model showed an increased vocal tract input inertance and an increased radiated acoustic energy after phonation into tube. Laukkanen et al. (2010) observed similar vocal tract shape modifications during glass tube phonation in a female subject using magnetic resonance imaging. The velar closure improved and the midsagittal area of the vocal tract became larger during and after phonation into a straw. The ratio between the transversal area of the lower pharynx and the epilarynx area increased both during and after the straw. Moreover, acoustic results showed an increase in the energy of the speaker’s formant cluster and in the overall sound pressure level. The frequencies of F2, F4 and F5 decreased, while F3 increased. There was a decrease on distances between F4 and F3 and F5 and F4. Authors suggested that straw phonation helps establishing a speaker’s formant cluster, which increases loudness and thus improves vocal economy.

Another effect of SOVTE that has been explored is the modifications produced in air pressures during and after exercises. Titze et al. (2002) stated that when a semi-occlusion, such as tube phonation is produced, there is a change in flow resistance depending on the size of the tube. Thus, the Poral is positive, and this in turn, would reduce the Ptrans (difference between subglottic and oral pressure), unless Psub is raised. Findings from an experiment using different types of straws (different diameters and lengths) evidenced that stirring straws offer more flow resistance than drinking straws. Data from Psub and Poral demonstrated that the higher is the flow resistance of the tube, the greater are the subglottic pressures that subjects would need to generate. The increase in Psub was, in the same study,
corroborated in humans. Authors stated that Psub was inversely related to the straw
diameter (Titze et al. 2002). Furthermore, in a study designed to investigate
phonation during varying artificial extension of the vocal tract, Laukkanen et al.
(2007) demonstrated that Psub was higher during phonation into longer tubes.
They suggested that increased Psub could reflect higher vocal effort used to
compensate for increased vocal tract loading. Radolf et al. (2014) found that the
mean Poral increased about 4 times in habitual comfortable phonation and about 9
times in soft phonation, during voice production into a resonance tube submerged
10 cm below the water surface. The Psub also increased, likely due to a
compensation for the increase in supraglottal resistance. In a physical model of
voice production study by Horáček et al. (2014), similar results were reported. Both
Poral and Psub increased with a resonance tube 10 cm in water. Greater changes
were observed for soft phonation.

1.4 Semiocclusions of vocal tract and energy transfer

Voice rehabilitation programs based on SOVTE propose that a resonant manner of
voice production promote a maximal vocal economy (maximum acoustic output
with minimum degree of vocal fold impact stress). Furthermore, According to Titze
(1998) the origin of vibratory sensations in the front part of face and mouth in
resonant voice could be due to the efficiency of the energy conversion process at
the glottis (from aerodynamic to acoustic energy). When this process is efficient,
the vibrations are distributed all over the face and head areas. On the other hand,
when the energy conversion process is poor, the vibrations will remain mostly in
the laryngeal area. Titze (2003) stated that: “resonant voice engage the vocal tract
for maximum transfer of power from the glottis to lips, and ultimately all the way
to the listeners”. In a computer simulation study, Titze and Story (1997) suggested
that phonation can be made more efficient through impedance matching between
the voice source and the vocal tract. Impedance matching is related to the
maximum power transfer theorem and it can be accomplished by using
semiocclusions at the lips or a combination of adjustments in vocal fold adduction
and epilaryngeal constriction.

The maximum power theorem in electronic circuits and transmission systems
states that the internal impedance of the source should match the impedance of the
load for maximum power transfer and/or minimize reflections from the load (Titze
2002). The concept of impedance matching was originally developed for electrical
power, but can be applied to any other field where a form of energy (not
necessarily electrical) is transferred between a source and a load. An alternative
application of the maximum power theorem is on the “voice circuit” where the
glottis is the source impedance and vocal tract the load impedance. Titze (2002)
pointed out that: “in order to apply the maximum power transfer principle to voice
production, we must realize that pressure and flow are not as simply related to each
other as voltage and current are in a resistive circuit”. Since the acoustic load is a
complex impedance, the quantitative nature of the maximum power theorem is not
the same, but qualitatively the principle is expected to be preserved: for maximum
acoustic power to be delivered to the vocal tract, the internal glottal impedance
should be of the same order of magnitude as the acoustic load impedance of the
vocal tract. On the other hand for maximum efficiency (acoustic
power/aerodynamic power), the load impedance should be considerably higher
than the glottal source impedance.

In practical terms, impedance of the vocal tract can be calculated when the size
and shape of the vocal tract is known. Moreover, the relation between subglottic air
pressure and flow during phonation or the amount of contact between the vocal
folds may reflect glottal impedance. Titze (2002) hypothesizes that tight adduction
of the vocal folds requires a narrower supraglottal airway, whereas looser
adduction requires a wider airway to maximize the output power. Using simulation models to calculate mean flows, peak flows, and oral radiated pressure (sound pressure that comes from the mouth to the atmosphere) for an impedance ratio between the vocal tract and the glottis, the author reported that when the impedance ratio between source and load impedance approaches 1.0, maximum power is transferred and radiated from the mouth.

1.5 Radiological imaging methods

Two main techniques have been used to obtain images of the vocal tract configuration in speech and voice research. One technique is the magnetic resonance imaging and the other is computerized tomography. These methods allow to acquire a series of image slices, in one or more anatomical planes (coronal, transversal, and sagittal), through a desired volume of the human body. In addition, three-dimensional shape information about the vocal tract volume and associated airways can obtained from these image slices.

The use of both, MRI and CT has several advantages and disadvantages. In the present dissertation CT was used, therefore, the focus will be on that method. Some characteristics of CT imaging are: the air-tissue interface is defined more accurately than MRI, the scanning time required for acquisition of a full image set is much lower than MRI, and both teeth and bones are precisely captured (Inohara et al. 2010). The main disadvantage of CT main is the exposure to ionizing radiation (Story et al. 1998), and for this reason; researchers avoid this technique for vocal tract analysis.

Imaging techniques have been applied in voice and speech science to explore normal morphology and development of the vocal tract (Moore 1992; Story et al. 1998; Fitch et al. 1999; Vorperian et al. 2005; Story 2008), to reconstruct three-
dimensional vocal tract models (Maeda 1982; Story et al. 1996; Whalen et al. 1999; Clement et al. 2007; Howard et al. 2009; Inohara et al. 2010), to create laryngeal computer models (Chen 2012), to explore articulatory adjustments of the vocal tract in speech and voice production (Crary et al. 1996; Story et al. 1998; Demolin et al. 2003; Narayanan et al. 2004; Takemoto et al. 2006; Vasconcelos et al. 2011; Miller et al. 2012a; Miller et al. 2012b), for laryngeal imaging (Silverman et al. 1995; Ahmad et al. 2009), comparing vocal tract morphometry in normal and pathological voices (Yamasaki et al. 2011), voice improvement following surgical treatment (Galgano et al. 2009), studying vocal tract configuration during singing (Story et al. 2001; Sundberg et al. 2007; Echternach et al. 2008; Echternach et al. 2010; Echternach et al. 2011), effect of voice therapy and training (Vampola et al. 2011), and in surgical planning (Raappana et al. 2008).

In the present dissertation, CT was used to explore changes of the vocal tract configuration before and during phonation into different tubes and after the tubes have been removed. The fact that the air-tissue interface is defined more accurately in CT than MRI, is the main reason that support the use of CT in instead of MRI in our study. The exposure to ionizing radiation was carefully controlled by experimenters not to produce any harmful effect on the participant, who was informed about potential risks related to CT examination.

CT examination is based in X-ray physical principles. During CT procedure, the interaction of X-ray photons (transmitted to the patient) produced multiple images. Two types of photons can be transmitted: 1) primary photons, passing through the tissue without interacting, or 2) secondary photons, produced from an interaction with tissues. The secondary photons will in general be deflected from their original direction and result in scattered radiation (Brenner et al. 2007).
1.6 Aerodynamic measurements of phonation

Aerodynamic measurements are a method to perform objective and noninvasive exploration of vocal function. While there are a number of methods that can be used to measure these aerodynamic parameters, current clinical and research instrumentation is typically designed to obtain indirect estimates of average glottal airflow rates (L/sec) and average Psub (cm H20), along with simultaneous acoustic measures of F0 and SPL. These two main aerodynamic components (Psub and transglottal airflow) reveal indirect information about the underlying valving activity of the larynx (Stemple 2000).

Flow is a term used to describe the movement of a quantity (volume) of gas through a given area in a unit of time. As such, the rate of flow is also referred to as the volume velocity, and it can be quantified in liters or milliliters per second or per minute (Stemple 2000). Based on this definition, the glottal airflow rate will be the velocity of the air passing through the glottis during phonation (Baken et al. 2000). Airflow through the larynx is usually measured with a pneumotachometer (air rate meter). Airflow through the larynx can be estimated at the airway opening because mass airflow through the larynx is comparable with that at the airway opening. Thus airflow is an indicator of relative openness of the larynx and the extent to which it allows air to pass between the trachea and pharynx (Hixon et al. 2008).

Psub is defined as the air pressure below the glottis and it represents the energy available for creation of the acoustic signal of speech. It should be of appropriate magnitude and well-regulated for an efficient voice production. Inappropriate levels of Psub or inadequate pressure regulation can cause abnormal voice intensity or sudden shifts in F0, SLP or voice quality. Measurement of Psub is a useful tool for diagnosis and therapeutic procedures in the voice field (Baken et al. 2000). Several methods to estimate Psub have been developed. However, the most used method is the estimation of Psub from the Poral during production of voiceless
plosives. This method is not invasive. When the vocal folds are abducted and the oral airway closed for a sufficiently long period during the production of voiceless plosive, the Poral reaches its maximum. The magnitude of the Poral peak during plosive production can be considered, therefore, to be a good estimation of Psub. The typical protocol requires the subject to produce several repetitions of the syllable /pa/ at a certain rate. Poral is captured by a plastic tube and pressure transducer. Habitual speaking pitch and loudness are used, and syllables should have similar stress. Under these conditions, the peaks of Poral are considered to be equal to the Psub (Baken et al. 2000). Another pressure measure can be calculated when Psub and Poral are known: the Ptrans, which is defined as the difference between Psub and Poral. A positive Ptrans is always needed to produce airflow from the lungs to the lips. This airflow, in turn, will be capable to produce vocal folds oscillation.

1.7 Electroglottography

EGG is a noninvasive method to assess laryngeal movement, and more specifically, the vibratory behavior of the vocal folds during voice production. It was first introduced by Fabre in 1957 (Fabre. 1957). To obtain the EGG signal, two (or more) electrodes are placed at either side of the thyroid cartilage at the level of the vocal folds, and a low amperage frequency-modulated (0.3–5MHz) current is passed between them. Variations in the electrical conductance of the larynx are produced by variations in the contact area of vocal folds during the glottal cycle. This results in fluctuations in the current between the two electrodes (Baken et al. 2012). These variations of the electrical admittance are reflected in the EGG waveform (Baken et al. 2000; Kitzing et al. 2000). Reduction of the glottal space causes an increased overall conductance between the electrodes. Therefore, the
conductance increases when vocal folds make contact, which is interpreted as glottal closing. Conversely, a decrease in the conductance is interpreted as glottal opening (Titze, 1990).

Different phases in standard EGG signal are:

• An upward portion with a steep positive slope, due to a fast contacting phase of the vocal folds beginning in the lower margin up to the upper margin.
• A segment around the maximum peak, which represents the portion of maximal contact of the vocal folds.
• A less steep downward-sloping portion corresponds to the gradual loss of contact between the vocal folds. This segment may exhibit a “knee.”
• An open phase represented by a segment of the signal where there are several oscillations of small amplitude. (Titze, 1990).

Several quantitative parameters can be obtained from the EGG signal, such as CQ, open quotient (OQ), and F0, CQ being the most used in clinic and researches. This parameter is defined as the ratio of the duration of the contact phase to the entire glottal cycle period (Rothenberg et al. 1988; Titze 1990). Earlier investigations support the possible association between EGG CQ and the degree of vocal fold impact stress (collision during vocal fold vibration). Verdolini et al. (1998) conducted a study to observe the possible use of the EGG CQ as a noninvasive method to estimate vocal fold impact stress. Authors suggested that EGG CQ strongly correlates with the degree of glottal impact stress. The study was conducted on excised canine larynges. When impact stress increases (stronger collision during vibration), vocal folds also tend to stay together longer and hence CQ increases (Verdolini et al. 1998). Similar results have been obtained in a computer modelling study by Horacek et al. (2006). Glottal CQ has also been reported to distinguish some types of phonation, e.g. resonant voice from pressed
voice (Verdolini et al. 1998), breathy, normal and pressed voice (Kankare et al. 2012), and modal register and falsetto (Kitzing, 1982).

### 1.8 High speed digital imaging

Several visualization techniques of vocal fold vibration have been used in the diagnosis of laryngeal disorders and assessment of normal voice production. In order to capture aperiodic vibratory patterns of the vocal folds, High speed digital imaging (HSDI) has emerged as effective method of visualization (Patel et al. 2008).

HSDI system records images of sustained phonatory vocal fold vibration and it is designed to capture laryngeal motion at a rate of up to 10,000 frames per second, which is fast enough to study the actual phonatory vibrations of the vocal folds, allowing for the observation of aperiodic vibration (Patel. 2008). This system is not dependent on F0 detection for motion extraction, and initiates recording with the onset of phonation, therefore overcoming many of the disadvantages of stroboscopy (Wittenberg et al. 2000; Christopher et al. 2001).

HSDI technique has been applied in different kind of studies such as comparison with other visualization techniques of vocal fold vibration (Wittenberg et al. 2000; Christopher et al. 2001; Patel 2008), to investigate vocal fold vibratory characteristics in voice disorders (Lohscheller et al. 2008; Murugappan et al. 2009), to explore the physical mechanisms of vocal fold vibration during normal phonation (Doellinger et al. 2006; Bonilha et al. 2008; Doellinger, 2009; Freeman, 2012; Ahmad et al, 2012), comparison between normal and disordered voices (Bonilha et al. 2008; Inwald 2011), to identify new measures and methods to evaluate the physiology of vocal fold vibration (Yan, 2005; Deliyski, 2005; Orlikoff et al. 2009; Köster 1999; Zhang 2010; McDonnell et al, 2011; Krenmayr,
2012;), to explore physiology of singing voice (Lindestad et al. 2001; Echternach 2010), and to measure the effects of phonation exercises on voice production (Laukkanen et al. 2007; Granqvist et al. 2014).

A study designed to investigate the clinical value of HSDI compared to stroboscopy (Patel 2008), provided evidence that HSDI can be used to reveal physiologic fluctuations in the vibratory pattern of vocal folds related to some auditory-perceptual features of moderate to severe dysphonia. The data from this study confirm that analysis of aperiodic signals can be reliably performed with the HSDI. Moreover, authors suggested that the vibratory features of amplitude, mucosal wave, phase symmetry, glottal closure, vocal fold edge, vertical plane difference, phase closure, tissue pliability, and aperiodicity can be judged from HSDI to obtain proper interpretations of vocal function.

To obtain precise and objective information from HSDI videos, several techniques have been developed, the most used of them are Videokymography and Phonovibrography.

Videokymographic imaging shows the vibrations of the vocal folds in a single image. The technique of videokymography utilizes a special video camera that, besides registering the images of the whole vocal folds at standard video rates, can also scan a single line of the laryngeal image at a high-speed rate of about 8,000 images per second. Many successive images from this line are put together to form the resulting kymographic image, which shows the vibrations of the selected part of the vocal folds (Svec et al. 2007). The main limitation of videokymography is that this method is not capable to show the vibration patterns of the entire length of both vocal folds. A relatively new quantitative image-based method overcome this difficulty, the phonovibrography (Loescheller et al, 2007). Phonovibrography is a fast, robust, and precise image processing strategy to extract information of the vocal fold movements during phonation (Wittenberg et al. 2000). The method detects the vocal fold edges and transfers their movement into a static geometric
presentation (phonovibrograms), which can be visually analyzed through various calculations. Several glottal variables can be derived from high HSDI videos analyzed with phonovibrography (Echternach et al. 2013).

1.9 Laryngeal endoscopy

Endoscopic visualization methods are the most important and common tools to explore the laryngeal status. Visualization could be performed using a rigid endoscope through the oral cavity or a flexible endoscope through the nose. These methods are used either with a continuous or a stroboscopic light source.

Laryngeal videostroboscopy performed with a rigid endoscope is the most common method to assess laryngeal structures and to diagnose vocal fold pathologies in clinic. The principle of videostroboscopy is based on the fact that very rapid events, such as the phonatory opening and closing of the glottis cannot be followed by the eye. Talbot’s law states that the human eye is not capable to perceive more than five frames per second due to the fact each image captured by eye takes 0.2 seconds to reach the retina. In order to produce a visual illusion of slow-motion during vocal folds vibration, laryngeal videostroboscopy produces light flashes that are delivered slightly later in the vocal fold cycle than the one before. Therefore, the phase difference between the vocal fold cycle and the flash cycle steadily increases. When these illuminated different points of vocal fold cycles are joined together, a slow-motion version of sampled phonatory cycles is obtained. Thus stroboscopic flashes recreate cycles at a rate that is very much lower than the actual fundamental frequency. This apparent much lower vocal fold movement is possible to be followed by the eye (Baken et al, 2000).

Even though rigid laryngeal videostroboscopy is the most common method to assess laryngeal structures and to diagnose vocal fold pathologies in clinic,
nasofiberoscopy with continuous light source is a useful method that has some advantages over the rigid laryngoscopy: 1) it allows to explore not only laryngeal structures, but also nasal passage and pharynx; 2) it is capable to assess all laryngeal functions (respiration, swallowing, and phonation both in speaking and singing); 3) gag reflex is unlikely to be produced during the procedure; and 4) the patient is able to speak or sing almost normally. The fact that the mouth structures (tongue, jaw, teeth, etc.) are free during this endoscopic procedure, allows considering this exploratory method a useful tool in research to observe laryngeal and pharyngeal activity during regular speaking and singing voice tasks. One of the disadvantages of nasofiberoscopy may be that the nasal passage cannot be fully closed. This, in turn, may make the subjects to do some compensatory maneuvers which they would not otherwise do.

A number of investigations have been carried out thought naso laryngeal fibroscopy to explore pathological voices (Morrison et al. 1983; Rubin et al. 2006, Stager et al. 2000; Behrman et al. 2003), singing voice physiology (Yanagisawa et al. 1989; Pershall et al. 1987; Mayerhoff et al. 2014; Guzman et al. 2013, Guzman et al. 2015), and speaking voice physiology (Guzman et al. In press).

1.10 Aims of the study

The general aim of the present dissertation was to determine how aerodynamic variables, phonation, and vocal tract configuration are affected by different SOVTE.

The first study explored aspects related to the three subsystems in one single subject by multiple measures. Its aim was to investigate the vocal tract modifications and also the acoustic, aerodynamic and electroglottographic characteristics of the voice when comparing vocal exercising with two different
vocal tract impedances (during phonation into a glass tube and a stirring straw). We hypothesized that the glottis and/or the supraglottic behavior should adapt differently to different impedance of the vocal tract.

The second study concentrated only on glottal function (phonation), and it aimed to observe the influence of tube phonation into water on vocal fold vibration by using HSDI and phonovibrography. Two questions were attempted to answer: 1) Is there any influence of tube phonation into water on vocal fold vibration? and 2) Does immersion depth affect vocal fold vibration?

The third study focused on aerodynamic variables and glottal functions. Specifically, the purpose of this investigation was to determine the effect of different artificial lengthening of the vocal tract (phonation through a tube with the free end in water and in air) on air pressure variables and vocal fold adduction. Research question was: Do subjects with different voice conditions change differently the air pressure and vocal fold adduction variables during different artificial lengthening of the vocal tract?

The fourth study aimed to observe and compare the effect of different commonly used SOVTE on vertical laryngeal position (VLP), anterior-posterior (A-P) laryngeal compression and pharyngeal width in a group of subjects diagnosed with hyperfunctional dysphonia. Therefore, this study concentrated only on modification of vocal tract function (affecting resonance aspects of voice production).
2 MATERIALS AND METHODS

Article I. One single volunteer male subject (thirty-four years old) was assessed through CT, acoustic analysis, EGG, and aerodynamic measures, during and after tube phonation. Auditory perceptual analysis was also performed. The subject had seven years of experience using SOVTE as vocal training and warm-up exercises. No voice or hearing pathology was reported at the time of the experiment.

CT examination

Before CT examination, the subject was informed about the potential health hazards. During CT examination, the subject was asked to stay in supine position inside the CT machine, and he was required to produce the following phonatory tasks: (1) to sustain vowel [a:], (2) to phonate a sustained vowel-like sound into a glass tube (27 cm in length and 9 mm inner diameter) for fifteen minutes, and immediately after that, (3) to produce another sustained vowel [a:]. Subject was radiated only for few seconds during each phonatory task. Habitual loudness level and speaking pitch was used during all phonatory tasks. No biomechanical instruction (e.g. open your throat, close your mouth) were given to the subject, but he was asked to produce a stable sound with a good closure at the lips and to feel as strong as possible vibratory sensations on the front part of face and mouth. Fifteen minutes of vocal rest was produced after sustained vowel [a:] recorded after
exercising. Then, similar sequence was produced into a plastic stirring straw (2.5 mm inner diameter and 13.7 cm in length), during fifteen minutes. A sustained vowel [a:] was produced immediately after straw phonation. Each phonatory task was performed twice.

All CT samples were obtained using a Toshiba – Aquilion CT machine. Voltage 120 kV, Scan options were used as follows: helical CT, Time of the rotation 0.5 s, Slice thickness 0.5 mm, total number of slices: 510 during CT procedure. The subject was required to keep the same body posture in the CT machine during all examinations. A frame was used to keep the head position stable.

Distance (mm) and area (mm$^2$) measurements were obtained from midsagittal images. Moreover, cross-sectional areas were calculated from the transversal CT images for all phonatory tasks. All measurements were carried out using the software vPACS view, version 6.9.5 (Audioscan, Czech Republic). Distance measures from midsagittal samples (Figure 1) included 1) vertical length of the vocal tract (which is indicative of the vertical laryngeal position) measured as the distance between the lowest point of the odontoid process of the Atlas and the vocal folds following a vertical line, 2) horizontal length of the vocal tract measured as the distance between the lowest point of Atlas and the narrowest point between the lips, 3) lip opening measured as the distance between the lower edge of the upper lip and the upper edge of the lower lip, 4) jaw opening measured as the distance between the lowermost edge of the jawbone contour and the anterior end of the hard palate, 5) tongue dorsum height measured as the distance between the lowermost edge of the jaw bone and the uppermost point of the tongue dorsum, 6) oropharynx width measured as the distance between the lowest point of the second vertebra and the most posterior part of the tongue contour (for ensuring the same angle we used straight line from the anterior uppermost edge of the jawbone contour to the anterior lowest point of the second vertebra), 7) velum elevation
measured as the distance from the posterior upper edge of the hard palate and the anterior lowest point of the uvula, 8) hypopharynx width measured as the distance between the lowest point of the pharynx and the internal edge of the epiglottis following a line from the anterior uppermost edge of the jawbone contour to the lower point of the pharynx.

Figure 1: Distances (mm) measured in CT midsagittal images: 1) vertical length of the vocal tract, 2) horizontal length of the vocal tract, 3) lip opening, 4) jaw opening, 5) tongue dorsum height, 6) oropharynx width, 7) velum elevation, and 8) hypopharynx width.

Cross-sectional areas from the same midsagittal images (Figure 2) included: 1) oral cavity (A1) measured from the lips to the velum (up to the line connecting the lowermost edge of the jawbone contour and a break of declivity on the velum surface), 2) the pharyngeal region (A2) measured from the line ending A1 down to the horizontal line connecting the lower edge of the third vertebra and the lowermost edge of the jawbone contour, and 3) the epilaryngeal region (A3) measured from the line ending A2, down to the vocal folds.
Cross-sectional areas from the transversal CT samples (Figure 3), included: 1) the inlet to the lower pharynx (Ap) just above the collar of the epiglottis and 2) the outlet of the epilaryngeal tube (Ae) just below the collar of the epiglottis where the epilarynx and sinus piriformes form three separate tubes. The ratio between these two areas was also calculated (Ap/Ae). Areas from transversal CT images were chosen taking into account the bending of the vocal tract. A line from backbone to jawbone through the tip of arytenoids in a parallel way to the bottom line of the CT image was use to obtain Ae. A line just above to the tip of arytenoids, which was also parallel to the bottom line, was used to obtain Ap.

![Figure 2: Areas (mm²) measured from CT midsagittal images: oral cavity (A1), pharyngeal region (A2), and epilaryngeal region (A3).](image)

![Figure 3: Areas (mm²) measured in CT transversal images: the inlet to the lower pharynx (Ap) and the outlet of the epilaryngeal tube (Ae).](image)
All distance and area measures were based on previous studies performed with MRI. (Echternach et al 2010a; Echternach et al 2010b; Echternach et al 2011; Laukkanen et al 2010).

Acoustic, electroglottographic and aerodynamic recordings

While recording simultaneously the acoustic, EGG and aerodynamic signals, the subject was required to produce the following sequence of phonatory tasks:

1) To repeat several times (eight times approximately) the syllable [pa:] and to produce a sustained vowel [a:] following the last [pa:].
2) To phonate a sustained vowel-like sound into a traditional Finnish glass tube (27cm in length and 9 mm in inner diameter) during five minutes.
3) To repeat again the syllable [pa:] and sustain vowel [a] following the last [pa:].
4) Fifteen minutes of vocal rest (complete silence).
5) To repeat several times the syllable [pa:] and a sustained vowel [a:] following the last [pa:].
6) To phonate a sustained vowel-like sound into a plastic stirring straw of 2.5 mm inner diameter and 13.7 cm in length during five minutes.
7) To repeat again the syllable [pa:] and sustained vowel [a] following the last [pa:].

The same comfortable speaking pitch and loudness level was produced during the entire sequence. Pitch was auditorily monitored using an electronic keyboard by one of the experimenters. Sound pressure level was monitored during the series of consecutive syllables [pa:] by using a sound level meter (American recorder
technologies SPL-8810) (American recorder technologies, Inc. Simi Valley, CA) placed at a distance of 40 cm from the mouth. Slow time response and a Z frequency weighting were used. Subject was asked to feel vibratory sensations on the front part of face and mouth during tube phonation. This was subjectively controlled by the participant.

Acoustic, EGG and aerodynamic signals were recorded in a sound treated room. A condenser microphone (Behringer ECM 8000) at a distance of 40 cm from the mouth was used to record the audio signal. The mouth-to-microphone distance of 40 cm was used to make the recordings comparable with those made for several years in Speech and Voice Research Laboratory, University of Tampere. The choice of distance has been found to be best for spectral and also perceptual analyses, since shorter distances will provide the proximity effect and longer distances will increase the effect of room acoustics (Leino and Laukkanen). Poral was recorded as an estimate of Psub during the occlusion of the consonant [p:] in the syllable [pa:] and during shuttering of the outer end of the tube. A pressure transducer (PT-25, Glottal Enterprises, Syracuse, NY) connected to a plastic tube was used to capture Poral. The tube was inserted into the corner of the mouth, extending a few mm behind the lips, without touching the tongue or any other oral structure. The manometer MSIF2 (Glottal Enterprises, Syracuse, NY) was used.

A two-channel electroglottograph (Glottal Enterprises, Syracuse, NY) was used to record EGG signal. A 20 Hz high-pass filtering was also used to exclude slow variations in the signal amplitude, which could be due to articulatory movements. Before every capture, electrodes were cleaned with a slightly wet tissue, and a thin layer of conductive gel was applied (Mingograph electrode cream, Siemens-Elema AB). The quality of the EGG signal was monitored throughout the recordings with an oscilloscope.

Calibration of audio signal was performed for further sound pressure level measurements in dB (Z) using a 440 Hz tone. Calibration of air pressure signal was
carried out with a Glottal Enterprises calibrator, model MCU-4 (Syracuse, NY). Pressure and EGG signals were digitized and recorded simultaneously into two different channels using the Soundswell Signal Workstation (Hitech Development AB, Stockholm, Sweden). Sampling rate of 16 kHz was used. Audio signal was recorded at a sampling rate of 48 KHz with 16 bits quantization using a DAT recorder (Marantz PMD 671. Marantz, Mahwah, NJ).

From aerodynamic samples, the following variables were studied (Figure 4):

1) Psub (estimated from the maximum peak of the Poral during the occlusion of the consonant [p:] in the syllable [pa:]).
2) Psub during manual shuttering of the outer end of the tube.
3) Poral obtained during [a:] in the syllables and non-shuttered phase of tube phonation.
4) Ptrans calculated from the difference between Psub and Poral.

A Soundswell Signal Workstation (Hitech Development AB, Stockholm, Sweden) was used to measure the Poral values.

From EGG samples, mean value of CQ_{EGG} was obtained. The software EFxHist, version 1.5 (Mark Huckvale. University college of London, UK) was
used. Analysis was performed from the middle section of the each EGG sample. EFxHist software defines the contact phase using a criterion level of 50% from the peak to peak amplitude of the EGG signal (Figure 5).

![Figure 5: Calculation of CQ from the EGG waveform using a 50% peak to peak algorithm.](image)

Acoustic analysis of vowel samples captured before and after tube/straw phonation was performed using Praat software (version 5.2, Boersma & Weenink, 2008). FFT spectrum (spectral slice and spectrogram) and LTAS were applied. An approximation of the formant frequencies from F1 to F5 was obtained from FFT spectrum analysis. The formant frequencies were measured from the strongest peaks or in the middle of two adjacent equally strong peaks in spectral slices. To confirm formant values, wide band spectrograms (bandwidth 260 Hz) were used. There the formant frequencies were located in the middle of the bands with the strongest intensity. Calculation of frequency distances between F2-F1, F3-F2, F4-F3, and F5-F4 was also performed.

Several variables were obtained from LTAS analysis (Figure 6):

1) Singer’s/speaker’s formant energy (energy between 2500-4000 Hz),
2) Singing power ratio (the difference of energy between the highest peak around 0-2 KHz and the highest peak around 2-4 KHz).
3) Total SPL (energy between 50-6000 Hz).
4) F1 energy (spectral energy between 500-800 Hz). A frequency bandwidth of 25 Hz and Hanning window were used for LTAS analysis.

![Figure 6: Schema including all spectral measures obtained from each acoustic sample.](image)

Additionally, an auditory perceptual assessment to evaluate the voice quality of the samples was carried out by 4 blinded judges (1 man, 3 women). All of them were speech-language pathologists with at least eight years of experience in voice clinic and reported normal hearing. Acoustic samples recorded before and after tube/straw phonation were played in randomized pairs. Raters were asked to judge which sample in each pair was produced with a better voice quality or if there was no difference between them. The evaluation was performed in a sound-treated room using a high quality Audioengine loudspeaker (Audioengine, Kowloon, Hong Kong).

Article II. Eight volunteers (five males, three females) participated in the second study. The age range of the participants was 23-45 years. No voice or hearing problems were reported at the time of the experiment. Diagnosis of normal larynx
was confirmed through laryngoscopy. The subjects of the study included four trained singers, one subject with experience in speech training, and three subjects who had no voice training.

Two test sequences were registered. The first sequence included production of the following tasks during HSDI registration:

1) Sustained and stable vowel [i:].
2) A vowel-like sound into a flexible silicone (lax vox-like) tube (45 cm in length, 2 cm in inner diameter) whose free end was submerged 5 cm into recipient filled with water. Tube phonation was performed for five minutes.
3) Sustained and stable vowel [i:].

The second sequence consisted of:

1) Tube phonation with the tube submerged 5 cm into the water (taken from the previous sequence),
2) Tube phonation with the tube submerged 10 cm into the water
3) Tube phonation with the tube submerged 18 cm into the water.

Therefore, a total of 5 samples were obtained from each subject ([i:] pre tube, phonation during tube submerged 5 cm in water, [i:] post tube, phonation during tube submerged 10 cm in water, and phonation during tube submerged 18 cm in water). All phonatory tasks in both sequences were performed at the same comfortable pitch and vocal loudness that was chosen by the subjects in the first
One of the experimenters monitored pitch auditorily during voice production, and used a grand piano for reference.

HSDI registration during all phonatory tasks was performed via a rigid endoscope attached to a plastic mouth piece. The flexible tube was also attached to the same mouth piece. To ensure the depths of immersion, three marks were made on the water container at 5 cm, 10 cm, and 18 cm (Figure 7). Even though the subjects phonated for 5 seconds during each phonatory task, only one second (from the mid portion) of phonation was captured by the high speed camera.

![Figure 7: Experiental set-up during HSDI examination.](image)

Rigid laryngoscopic examinations were performed by two experienced laryngologists. All participants stayed in standing position during examinations, and no topical anesthesia was applied. A HRES-Endocam 5562 system (Fa. Wolf, Knittlingen, Germany) coupled with a 90° rigid endoscope (Fa. Wolf, Knittlingen, Germany) was used. The high speed camera allows recording up to 4000 frames per second with a pixel resolution of 256 x 256. A 300-W xenon light source (Auto-LP 5131, Richard Wolf, Knittlingen, Germany) was used for illumination. To prevent tissue overheating during the recordings, the duration of light exposure was kept at a minimum.
To extract information of the vocal fold movements during phonation, Phonovibrography was used (Lohscheller et al. 2007). Thousand frames (250 ms) from each sample of the high speed material were analyzed using the custom made Phonovibrogram Software Tool (Glottal Analysis Tools v5, Michael Döllinger and Denis Dubrovsky, Dept. of Phoniatrics and Pedaudiology, Medical School Erlangen, Germany). Only the mid and stable portions of each 250 ms sample (excluding the onset and offset) were analyzed. Segmentation of the glottis area was performed in a single frame and semi-automatically transferred to the other frames. Secondly, a phonovibrogram was established by construction of glottal length axis (Lohscheller et al, 2007). Figure 8 explains how phonovibrogram is established from a high speed sample.

![Figure 8](image_url)  
**Figure 8** (courtesy of Dr. Michael Doellinger): (A) A single image of an endoscopic high-speed recording is shown; the orientation is given from patients' view. Visible are the left and right vocal fold, the glottis (dark) and the vocal fold edges. (B) Generation of the Phonovibrogram: The vocal fold edges are detected and the distance between vocal fold edges and glottis axis are computed. Then, the left vocal fold edge is rotated by 180° around (P) - the posterior commissure. The distances between the vocal fold edges and the glottis axis are color encoded: brighter colors represent larger distances from the glottis axis. This is done for all single frames within a movie, yielding a two-dimensional picture representing the vocal fold vibrations (C). Here, a PVG is given for a human subject with normal phonation (i.e. periodic over time, left-right symmetric) over three cycles.

From all high speed samples, the following glottal variables were obtained.

- **Amplitude to length ratio (ALR):** ratio between vibrational amplitude at the center of the vocal folds and the length of the vocal folds.
- **Closed quotient (CQ):** ratio between closed phase and the entire glottal period.
• Closing quotient (ClQ): ratio between closing phase and entire glottal period.
• Fundamental frequency (F0): number of cycles of vocal fold vibration per second.
• Harmonic to noise ratio (HNR): relation between harmonic energy and noise energy.
• Spectral flatness (SF): spectral declination or spectral tilt (frequency range from 0 to 2.000 Hz). Flatness is maximal for white noise.
• Jitter %: frequency perturbation in percentage.
• Shimmer %: amplitude perturbation in percentage.

Only descriptive statistics were calculated for the variables, including median and interquartile range. We analyzed the trend of values for all parameters and phonatory tasks without using hypothesis testing to avoid lack of statistical power due to small sample size. All analyses and graphics were performed using Stata 13.1 (StataCorp, College Station, TX).

Article III. Informed consent was obtained from forty-five volunteer subjects for this study. All participants were asked to undergo rigid videostroboscopy by two experienced laryngologists, to confirm medical diagnosis. No topical anesthesia was used during endoscopic procedure. Twelve subjects had the diagnosis of healthy larynx and voice and no voice training was reported (6 males, 6 females). Nine subjects were diagnosed with healthy larynx and voice and voice training was reported (2 males, 7 females). Fourteen subjects were diagnosed with muscle tension dysphonia (MTD) (5 males, 9 females), and 10 subjects were diagnosed with unilateral vocal fold paralysis (paramedian vocal fold position) (2 males, 8 females). The average age was 24 (20-33) years for subjects with healthy voice and without voice training, 27 (21-37) years for subjects with healthy voice with voice training, 28 (23-35) years for subject with MTD, and 45 (31-56) years for subject
with vocal fold paralysis. Participants from all groups were native speakers of Spanish.

The repetition of syllable [pa:] at a rate from 2.5 to 4 syllables per second, in habitual, comfortable speaking pitch and loudness, was recorded from all participants as the baseline condition. Then, a series of phonation samples on five semi-occluded vocal tract postures in a random order were produced by all subjects:

1) Drinking straw with the free end in air (5 mm in inner diameter and 25.8 cm in length).
2) Stirring straw with the free end in air (2.7 mm in inner diameter and 10.7 cm in length). 3) Silicon tube (Lax Vox-like) with the free end in air (10 mm in inner diameter and 55 cm in length).
4) Silicon tube (Lax Vox-like) with the free end submerged 3 cm below water surface.
5) Silicon tube (Lax Vox-like) with the free end submerged 10 cm below the water surface.

Three repetitions were performed for each phonatory task (including baseline condition). Therefore, a total of 810 samples were recorded (45 subjects x 6 postures x 3 repetitions). All the exercises were demonstrated to the participants by a trained speech-language pathologists before data collection. Participants were instructed to phonate with ease and to feel vibrations in the anterior face and mouth during all SOVTE. The use of a comfortable speaking pitch and loudness level was required during all phonatory tasks. Loudness level was perceptually controlled by experimenters. Pitch was also monitored by the experimenters using an electronic keyboard for a reference.
All samples were recorded digitally at a sampling rate of 22.1 KHz with 16 bits/sample quantization. Acoustic signal was recorded at a constant microphone-to-mouth distance of 20 cm, using a condenser microphone AKG CK 77 (AKG Acoustics, Vienna, Austria) integrated into the Phonatory Aerodynamic System (PAS). Acoustic samples were captured to obtain F0.

Aerodynamic and EGG devices were connected to a Computerized Speech Lab, Model 4500 (KayPENTAX, Lincoln Park, NJ), which in turn was connected to a desktop computer running a Real-Time aerodynamic and EGG analysis software (KayPENTAX, Model 6600, version 3.4, KayPENTAX, Lincoln Park, NJ). EGG signal was captured with an electroglottograph (KayPentax, model 6103, KayPENTAX, Lincoln Park, NJ). EGG signal quality was monitored using a real time oscillogram incorporated in the EGG software.

A Phonatory Aerodynamic System (PAS; KayPentax, model 4500, KayPENTAX, Lincoln Park, NJ) was used to collect aerodynamic data. Calibration of the air pressure was performed before data collection. The Poral was captured with a pressure transducer connected to a thin flexible plastic tube (13 cm in length). The tube was inserted into the corner of the mouth, extending a few millimeters behind the lips. Subjects were instructed not to touch the tube with any oral structure. To avoid air leakage through the nose, a nose clip was used for all participants during data collection.

A Real-Time aerodynamic and EGG analysis software (model 6600, version 3.4 KayPENTAX, Lincoln Park, NJ) was used to analyze all samples. From the middle section of each sample, the most stable part was analyzed.

From EGG, aerodynamic and acoustic samples, the following variables were obtained:

1) CQ from electroglottographic signal.
2) F0 from acoustic signal.
3) Psub estimated from the maximum peak of the Poral during the occlusion of the consonant [p:] in the syllable [pa:].

4) Psub estimated during manual shuttering of the outer end of the tube.

5) Poral obtained during vowel [a:].

6) Poral obtained during non-shuttered phase in tube phonation.

7) Ptrans obtained by calculating the difference between Psub and Poral.

8) Peak-to-peak values of Poral from samples where bubbling was produced (Figure 9).

9) Bubbling frequency from the oral pressure signal.

Criterion level of 25% from the peak to peak amplitude of the EGG signal was used for CQ analysis.

![Figure 9: Modulations in oral pressure caused by bubbling of water. The registration is from one participant while phonating into a silicon tube submerged 10 cm in water.](image)

All statistical analyses were performed using Stata® 13.1 (StataCorp, College Station, TX), p<0.05 was considered to be statistically significant, and all reported
p values were two-sided. Median and interquartile range (IQR) were used to describe numerical variables. Kruskal-Wallis test was used to compare phonatory task and vocal status separately. To observe the joint influence of phonatory task and vocal status in vocal parameters, a generalized multivariable linear model was performed. Separate subgroup analysis (Silicon tube submerged 3 and 10 cm into water) for minimal and maximal Poral was also performed using Wilcoxon test. Finally, linear correlation analysis using Pearson coefficient for overall correlation was used.

Study IV. Twenty-one subjects participated in the study (average age of them was 26 years, with a range of 20-28 years). All participants had the diagnosis of hyperfunctional dysphonia without any vocal fold lesions and reported at least one year of voice problems. None of them reported previous voice therapy, voice training or smoking habit. Laryngeal diagnosis was made through flexible laryngoscopy by a laryngologist with more than twenty years of experience in voice clinic.

Nasofiberoscopic procedure

Three laryngoscopic variables were evaluated (Figure 10): 1) VLP defined as the perceived distance between vocal folds and distal part of flexible endoscope, 2) pharyngeal width defined as the perceived area surrounded by the pharyngeal walls, and 3) anterior-posterior (A-P) laryngeal compression defined as the perceived distance between epiglottis and arytenoid cartilages. Laryngeal and pharyngeal activity was assessed through flexible laryngeal endoscopy (Olympus ENF type p4). The flexible laryngoscope was connected to a video camera (Sony DCX Ls1 Sintek) and a Richard Wolf LP 4200 light source. Digitalization of analogue samples was performed with a Pinnacle Studio HD 10 software. To allow a full view of the pharynx and larynx, the flexible endoscope was positioned right
below the tip of the uvula. This placement was maintained with the laryngologist’s finger, which fixed the fiberscope against the alar cartilage of the nose. Since observation of laryngeal height modifications and other laryngeal/pharyngeal configurations can be affected by movement of the fiberscope, a steady placement of the fiberscope is crucial. No topical nasal anesthesia was used during procedures.

![Figure 10: Laryngoscopic images showing how laryngoscopic variables were evaluated. Pharyngeal width (A), A-P laryngeal compression (B), and vertical laryngeal position (C).](image)

During flexible endoscopic examination, participants produced eight different SOVTE:

1) Phonation into a long-wide straw (6 mm inner diameter and 20 cm in length) with the free end in air.
2) Phonation into a long-narrow straw (3 mm inner diameter and 20 cm in length) with the free end in air.
3) Phonation into a short-wide straw (6 mm inner diameter and 10 cm in length) with the free end in air.
4) Phonation into a short-narrow straw (3 mm inner diameter and 10 cm in length) with the free end in air.
5) Phonation into a long-wide straw (6 mm inner diameter and 20 cm in length) submerged 3 cm below the water surface.
6) Phonation into a long-wide straw submerged 10 cm below the water surface.
7) Phonation using the hand-over-mouth technique.
8) Phonation with lip trill.

Plastic commercial drinking and stirring straws were used for tube phonation both with the free end in air and into the water. Two repetitions were performed for each phonatory task.

Each SOVTE was produced for 7 seconds and the entire sequence lasted for approximately fifteen minutes. Regular breathing was asked between tasks in order to avoid the effect of lung volume on the laryngeal height. Participants were required to phonate each exercise at three loudness levels: soft, moderate, and loud. Only moderate loudness was used during tube phonation into the water. Habitual speaking pitch was used to avoid variation in vertical laryngeal position. An electronic keyboard was used to give and control the pitch by one of the experimenters during the entire test sequence.

Three blinded laryngologists with more than 4 years of experience in working with voice patients assessed the VLP, A-P laryngeal compression, and pharyngeal width in each video sample. To avoid the possible effect of voice quality on the judges’ ratings, all audio signals were removed from samples. Judges rated the three endoscopic variables using a 5-point Likert scale. For VLP (1 = very high, 5 = very low), for A-P laryngeal compression (1 = very open, 5 = very narrow), and for pharyngeal width (1 = very narrow, 5 = very wide). The three judges participated in a 1-hour training session in videolaryngoscopy examination, to standardize the rating parameters and rating scales. Two one hour-sessions were necessary to assess all video samples by each rater. Fifteen percent of the samples were randomly repeated in order to assess the intra-rater reliability. Judges were not aware of these repetitions.

Descriptive statistics were calculated for the variables, including mean and standard deviation. A multivariate linear regression model was used to obtain an
intraclass correlation coefficient (ICC) to assess the judges’ concordance (intra and inter-rater agreement). Then, another multivariate linear regression model was performed considering VLP, pharyngeal width, and A-P laryngeal compression as outcomes, and phonatory tasks and loudness as predictive variables. Simple correlation analysis between vertical laryngeal position, pharyngeal width, and anterior-posterior laryngeal constriction was also conducted using Spearman’s rho. One-way ANOVA analysis for testing differences between phonatory task scores, was used. The analyses were performed using Stata® 12.1 (StataCorp. 2011. College Station, TX: StataCorp LP). A p-value < 0.05 was considered statistically significant.
3 RESULTS

Article I. Main findings observed for distance measures from midsagittal images were: 1) Lower VLP for both glass tube and straw, being more prominent during the later. This lowered laryngeal position remained after tube and straw phonation. 2) Another important change was detected in the velum position, which rose to seal the nasopharyngeal port during the tube and straw phonation. The higher velum position remained after tube and straw phonation. 3) A decrease in oropharynx width was observed when comparing samples before and after both straw and tube phonation. 4) Hypopharynx became wider during straw phonation compared to samples before. 5) A decrease during both tube and straw phonations was found for Jaw opening.

Main outcomes detected in area measurements were: 1) an increase for Ap area both during tube and straw compared to vowel phonation before exercise. More prominent changes were observed during straw. 2) The Ae area showed an increment during straw phonation. 3) The Ae area decreased after both tube and straw phonation. More prominent changes were observed after straw. 4) The ratio of areas Ap/Ae increased for all conditions compared to vowel phonation before them. 5) Oral cavity area (A1) decreased during both tube and straw phonation. 6) An increase was observed in pharyngeal region (A2) during both tube and straw compared to vowel phonation. 7) The epilaryngeal region (A3) became larger during both glass tube and straw phonation, being the change more prominent during straw.
Results from acoustic analysis revealed the following major changes: 1) There was an increment of the energy in the speaker’s/singer’s formant cluster region after straw 2) A decrease was observed in SPR when compared samples before and after both tube and straw. The difference was greater after straw. 3) Regarding formant frequencies, the major change was a decrease in F1 after both tube and straw. 4) The rest of formant frequencies also experimented a decrease; however, it was smaller than the change found in F1. 5) A formant cluster was observed between F3 and F4. They were closer to each other after exercising with both tube and straw. This cluster was located in 2588-2930 Hz and 2603-2966 Hz after tube and straw respectively.

Auditory perceptual assessment of voice samples before and after exercises revealed that all samples produced after straw were evaluated by the four judges as better in voice quality than the samples before. The same degree of agreement was not found for tube phonation.

Aerodynamic and EGG exploration showed: 1) $P_{sub}$ increased during straw phonation and this increment remained after exercise. 2) During glass tube, this variable did not demonstrate a clear change, but it became slightly higher after tube. 3) $P_{oral}$ also increased during both tube and straw, being the change greater during straw. 4) $P_{trans}$ decreased during both tube and straw. 4) There was a decrease of EGG CQ during both tube and straw. This change remained after tube and straw.

Article II. Results of this study are presented in two ways: 1) median and interquartile ranges for both sequences (Tables I and II). Recall that no hypothesis testing was used to avoid lack of statistical power due to small sample size. 2) trends of values for all parameters during both sequences. A trend among participants was considered when similar changes (compared to vowel phonation) were observed for the majority (five or more) of the subjects.
Table I: Medians and interquartile ranges for glottal area parameters pre, during and after tube phonation submerged 5 cm in water (sequence 1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pre 5cm</th>
<th>During 5cm</th>
<th>Post 5cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0</td>
<td>198.06  ±88.97</td>
<td>222.91 ±59.63</td>
<td>216.38 ±109.54</td>
</tr>
<tr>
<td>HNR</td>
<td>15.38 ±2.85</td>
<td>14.38 ±2.82</td>
<td>15.69 ±7.84</td>
</tr>
<tr>
<td>Jitter</td>
<td>3.73 ±3.91</td>
<td>3.23 ±1.72</td>
<td>2.19 ±1.83</td>
</tr>
<tr>
<td>Shimmer</td>
<td>0.29 ±0.21</td>
<td>0.39 ±0.48</td>
<td>0.45 ±0.49</td>
</tr>
<tr>
<td>SF</td>
<td>-20.26 ±3.75</td>
<td>-17.88 ±2.87</td>
<td>-20.00 ±1.89</td>
</tr>
<tr>
<td>CQ</td>
<td>0.22 ±0.05</td>
<td>0.26 ±0.15</td>
<td>0.23 ±0.13</td>
</tr>
<tr>
<td>ALR-R</td>
<td>0.08 ±0.04</td>
<td>0.07 ±0.05</td>
<td>0.07 ±0.03</td>
</tr>
<tr>
<td>ALR-L</td>
<td>0.11 ±0.05</td>
<td>0.09 ±0.05</td>
<td>0.09 ±0.04</td>
</tr>
</tbody>
</table>

Table II Medians and interquartile ranges for glottal area parameters during tube submerged 5 cm, tube submerged 10 cm, and tube submerged 18 cm (sequence 2).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>During 5cm</th>
<th>During 10 cm</th>
<th>During 18 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>F0</td>
<td>222.91 ±59.63</td>
<td>235.29 ±81.41</td>
<td>216.38 ±72.35</td>
</tr>
<tr>
<td>HNR</td>
<td>14.38 ±2.82</td>
<td>14.14 ±3.30</td>
<td>12.80 ±1.94</td>
</tr>
<tr>
<td>Jitter</td>
<td>3.23 ±1.72</td>
<td>4.09 ±0.89</td>
<td>3.71 ±4.18</td>
</tr>
<tr>
<td>Shimmer</td>
<td>0.39 ±0.48</td>
<td>0.39 ±0.16</td>
<td>0.49 ±0.30</td>
</tr>
<tr>
<td>SF</td>
<td>-17.88 ±2.87</td>
<td>-19.55 ±3.58</td>
<td>-17.95 ±5.08</td>
</tr>
<tr>
<td>CQ</td>
<td>0.26 ±0.15</td>
<td>0.25 ±0.11</td>
<td>0.18 ±0.13</td>
</tr>
<tr>
<td>ALR-R</td>
<td>0.07 ±0.05</td>
<td>0.07 ±0.02</td>
<td>0.08 ±0.05</td>
</tr>
<tr>
<td>ALR-L</td>
<td>0.09 ±0.05</td>
<td>0.10 ±0.05</td>
<td>0.12 ±0.03</td>
</tr>
</tbody>
</table>

For the first sequence (phonation pre, during and post tube submerged 5 cm into water), the following trends were observed during phonation into tube 5 cm in water: 1) CQ decreased (5 of 8 participants) compared to baseline condition; 2) HNR decreased (6 participants), 3) SF decreased (five participants); 4) ALR ratio (right and left vocal folds) decreased (5 and 6 participants, respectively), 5) F0 increased (6 participants); and 6) Jitter decreased (five participants). These results
suggest that tube phonation with the free end submerged 5 cm into the water is less tight, softer, a bit more noisy and yet more stable.

Main trends observed after phonation into tube 5 cm in water were: 1) an increase in CQ (6 participants), CIQ (5 participants), HNR (5 participants), and F0 (5 participants); and 2) a decrease in jitter (5 participants). The findings suggest that phonation was tighter, and voice was less noisy and more stable.

For the second sequence (tube submerged 5, 10 cm, and 18 cm into water), mayor changes were: 1) CQ decreased with tube 5 cm in water (5 participants); 2) CQ increased with tube 18 cm in water (six participants); 3) CIQ increased during tube 10 cm (6 participants); 4) HNR decreased for all tube depths (6 and 5 participants, respectively); 5) ALR decreased for all tube depths (6 and 5 participants, respectively); 6) Jitter decreased for tube 5 cm in water (5 participants); 7) Jitter increased for tube 10 cm in water (5 participants); 8) F0 increased for tube 5 cm and 10 cm in water; and 9) SF decreased during tube submerged 5, 10 and 18 cm below the water surface (five participants for each condition). These findings suggest that a less tight phonation is produced when tube is submerged 5 cm in water compared to the rest of depths of immersion. The results for deeper immersion are more difficult to interpret. Decreased ALR, HNR and SF seem to imply softer phonation (less collision between the vocal folds), while increased CIQ and CQ suggest the opposite.

Article III. Significant differences were found for the six phonatory conditions when compared to each other for all variables (p<0.05), except for F0. Moreover, an increase in Psub, Poral, Ptrans, and CQ was observed for all SOVTE produced for all vocal conditions when compared to the baseline. It is important to notice that most variables behaved in average in the same way regardless of the vocal status of the participants during phonatory tasks.

A strong positive correlation between Psub and Poral (r=0.82; p < 0.001) was found when taken all conditions together. When each group was studied separately,
similar results were evidenced. Correlation was observed between Psub and Poral for each group (vocal condition). For normal untrained participants: $r=0.86; p < 0.001$, for normal trained participants: $r=0.89; p < 0.001$, for subjects with muscle tension dysphonia: $r=0.79; p < 0.001$, and for subjects with vocal folds paralysis: $r=0.82; p < 0.001$.

Regarding maximum and minimum Poral during bubbling, results showed that when tube immersion depth was deeper, significantly higher values were observed for both maximum and minimum Poral. Nevertheless, no significant differences were found for the average peak-to-peak amplitude of Poral regarding immersion depth. No significant differences were observed between voice conditions either.

Article IV. A good intra-rater reliability was observed for each judge. Moreover, the three blinded judges obtained a high agreement (high inter-rater reliability) (ICC= 0.79 [0.66-0.87], $p<0.0001$).

When compared score averages for each laryngoscopic variable by phonatory task, all SOVTE were found to have a significant and differential effect ($p < 0.0001$). Thus, VLP, A-P laryngeal compression and pharyngeal width changed differently throughout the eight SOVTE. All semi-occlusions produced a decrease in VLP compared to the resting position (Figure 11). Phonation with tube into the water (10 and 3 cm below the surface) and phonation into a long-narrow tube produced the three lowest VLP. The same three phonatory tasks caused the widest pharynx and the narrowest A-P laryngeal compression.

Figure 11: Laryngoscopic images for baseline (A), during phonation into a long-wide straw (6 mm inner diameter and 20 cm in length) with the free end in air (B), and during phonation into a long-wide straw (6 mm inner diameter and 20 cm in length) with the free end submerged 10 cm in water.
Correlation analysis showed that VLP significantly correlates with pharyngeal width ($\rho=0.578; p<0.0001$) and A-P laryngeal constriction ($\rho=0.3364; p<0.0001$). Moreover, pharyngeal width significantly correlates with A-P laryngeal constriction ($\rho=0.18; p=0.001$).
Most physiologic voice therapy programs (e.g. Vocal Function Exercises, Resonant Voice Therapy, and Accent Method) are based on SOVTE. They consider that physiologic approach should affect simultaneously breathing, glottal and vocal tract functions during exercising. Therefore, the general aim of the present dissertation was to determine how the three levels involved in voice production are affected by different SOVTE. Different exploratory methods, subjects with different voice conditions, and voice conditions were used to accomplish this purpose.

4.1 Influence of semioccluded exercises on aerodynamic variables

Earlier investigations (Titze et al 2002; Horáček et al 2014; Radolf et al 2014; Granqvist et al. 2014) have shown changes in air pressure during SOVTE. However, these studies have included at the most two subjects and all of them have been vocally trained.

The third study included in the present dissertation sought to investigate the effect of phonation into tubes in air and submerged in water on air pressure variables and vocal fold adduction. This study is the first one including a large number of subjects (n=40) with four different voice conditions. Regarding air pressure variables, findings showed an increase in Psub and Poral compared to the
baseline condition during all SOVTE. In average, phonation with the silicon tube below the water surface and into the stirring straw in the air produced the highest values for Psub. Similar results were found in the single case study (Study I) of the present dissertation. There was an increase of Psub during straw phonation, and this increment remained after removing the straw. Poral also evidenced an increment during both tube and straw. Our results from Study I and III are in line to previous works regarding air pressure related variables. Recently, Radolf et al. (2014) in single case study reported that Psub also increased during voice production into a glass resonance tube submerged 10 cm below the water surface. The mean Poral also increased in both habitual comfortable and soft phonation, the increment being greater during the latter. Furthermore, Horáček et al. (2014), using a physical model of voice production, reported that when flow was kept constant, subglottic pressure was necessary to raise considerably more for phonation into a stirring straw in air compared to a resonance tube 10 cm in water.

Findings from earlier studies and ours suggest that the increased Psub could be a way to compensate the increased Poral, which in turn, is caused by the high degree of airflow resistance offered by certain semiclosures (e.g. glass tube into the water and narrow stirring straw in air). In other words, the increased Psub could reflect higher vocal effort used to compensate for increased vocal tract loading. This airflow resistance (vocal loading) varies depending on the diameter and length of the tube in air (Titze et al. 2002; Laukkanen et al. 2007; Horáček et al. 2014; Radolf et al. 2014; Amarante et al 2015), and the depth of immersion when tube is placed into the water (Horáček et al. 2014; Radolf et al. 2014; Amarante et al 2015). The strong positive correlation between Poral and Psub (r=0.8298; p < 0.001) found in the third study of the present dissertation, corroborate the physiologic interaction between these two aerodynamic variables.

An interesting therapeutic aspect could arise from the effect of SOVTE in air pressure measures. No matter of the voice condition or voice training, an
unconscious and slight abdominal and rib cage activation (movement) is produced when voice is being produced with some semi-occlusions of the vocal tract. Even though no objective evidence exists yet, clinical observations have been reported and patients indicate to feel abdominal and rib cage movements during exercises.

They also point out that abdominal and rib cage activation is greater when tube is submerged in water compared to tube phonation in air. Possibly, this activation is an abdominal and intercostal muscle reaction to increase Psub, which is necessary to overcome the increased Poral. Hence, this muscle activation should be proportional to the degree of Poral, and airflow resistance. Therefore, SOVTE could be an effective manner to train breath support (breathing technique whose aim is to reduce excessive laryngeal muscular effort during phonation) in voice therapy and training. For example, SOVTE could promote that subjects start recruiting respiratory muscles more during phonation, instead of just overusing adductor laryngeal muscles for raising vocal intensity.

Another interesting finding from the third study of this thesis is that subject groups representing all four voice conditions (normal trained, normal untrained, functional dysphonia and vocal fold paralysis) behaved quantitatively in a similar manner regarding air pressure variables. Tube in water and stirring straw in air presented the highest values in both Psub and Poral for all voice conditions. It seems that these variables are in general more dependent on the degree of airflow resistance than vocal status of participants.

Ptrans is obviously a variable expected to be modified during SOVTE. Titze et al. (2002) stated that when a semi-occlusion, such as tube phonation is produced, there should be a reduction in the Ptrans (which is considered to be the force driving vocal fold vibration), unless the subglottic pressure is raised. Results from the first study of this dissertation (single case study) showed that Ptrans decreased during both tube and straw. Contrary, in the study III, Ptrans was higher than baseline condition for all exercises. Outcomes from the latter could imply that
although both Poral and Psub increased, they do not change proportionally, i.e. Psub increases relatively more than Poral. A compensatory adjustment in order to sustain the phonatory airflow during phonation could be a possible explanation. Radolf et al. (2014) showed similar findings in for both tube and stirring straw phonation.

Two main aspects differentiate tube phonation when the free end of a certain tube is placed in air or into the water: 1) the degree of airflow resistance, and 2) the presence or absence of water bubbling. Clinical observations have suggested that bubbles produced during phonation in water may cause a sensation of relaxing massage on the oral, laryngeal and pharyngeal tissues. In the third study of the present dissertation, bubbling characteristics were analyzed. Results showed that the average bubbling frequency was 22 Hz (range 12-32 Hz), regardless of tube immersion depth and the vocal condition of the subjects. Water bubbling is reflected in oscillation of Poral (e.g. Enflo et al 2013; Granqvist et al 2014). Earlier works have reported that the bubbling frequency of pulsating Poral is 10-40 Hz (Radolf et al. 2014; Horáček et al. 2014; Granqvist et al. 2014). Bubbling frequency has been previously analyzed from the physical point of view. It has been stated that it depends on several factors, such as air flow, immersion depth, and the diameter of the orifice of the tube (Davidson and Amick 1956). These factors are likely responsible for differences found in frequency of bubbling when compared our results to previous findings.

No significant differences were found for the amplitude of Poral modulation (mean difference between Max. and Min. Poral) caused by bubbling when comparing tube immersion depths of 3 cm and 10 cm in water. These outcomes and Radolf et al. (2014) results, may suggest that larger modulation of the Poral is not necessarily caused by higher Psub. Therefore, a possible association between the degree of tube immersion and the degree of subjective massage-like effect (usually
reported by patients for 10 cm compared to 3 cm immersion depths in water) is unlikely when the flow is kept constant.

4.2 Influence of semioccluded exercises on glottal function

Glottal function is likely the most explored aspect during SOVTE. The second study of the present thesis analyzed vocal fold oscillatory patterns during phonation into a tube submerged at different depths into water. In general, it was demonstrated that vocal fold oscillation is affected by tube phonation. Even though some trends were observed, results suggest that tube phonation with the free end in water do not produce clear patterns but the effect depends on how a person reacts to the increased airflow resistance.

CQ obtained from HSDI samples was one of the dependent variables in the study III. Findings showed that tube phonation submerged 18 cm into water caused higher values of CQ than the other conditions and baseline for most subjects. On the other hand, phonation into 5 cm of showed a decrease in CQ compared to baseline condition. Comparable results were observed in the third study of the present dissertation. When compared EGG CQ during tube submerged 3 and 10 cm into water, the latter demonstrated higher values. Our results are concordant with a recent study by Guzman et al. (2015) Authors compared EGG CQ in eight different SOVTE in a large number of subject (n=80). Tube submerged 10 cm in water demonstrated higher values than tube submerged 3 cm in water for both subjects diagnosed with dysphonic voices and subjects with normal voice. Similarly, Radolf et al. (2014) compared EGG CQ between phonation into a glass resonance tube with the outer end in the air, tube submerged 2 cm in water, tube submerged 10 cm in water, and phonation into a very thin straw. Outcomes showed that CQ obtained the highest value with tube submerged 10 cm into water. It seems that the degree of
flow resistance, specifically the depth of immersion, played an important role in glottal adduction. A shallower depth tends to lead to a lower CQ, while deeper submersion tended to produce a higher the CQ. Possibly, when supraglottic load increases, a higher Psub and a compensatory glottal adduction are produced, no matter the vocal status of participants. It is important to notice that not only depth of immersion affect airflow resistances, but also the inner diameter and length of tube. Longer and narrower tubes offer more resistance to the airflow than shorter and wider ones due to frictional losses (Titze et al. 2002, Amarante et al. 2016). Laukkanen et al. (2007) in a HSDI study found higher CQ for longer tubes compared with shorter ones. Results from the study by Horáček et al. (2014) showed how tube diameter also affects glottal function.

If the degree of airflow resistance really influences the degree of vocal fold adduction, perhaps low flow resistance exercises should be used in subjects with vocal fold hyperadduction, whereas high flow resistance exercises should be used in patients with low vocal fold adduction (e.g. vocal fold paralysis or presbyphonia). This is supported by observations of Sovijärvi (1977; 1989), who has stated that tube submerged 1-2 cm below the water should be used for patients with hyperfunctional voice disorders, while a depth of 10-15 cm could be more appropriate for subjects with vocal fold paralysis.

Although, depth of immersion, inner diameter, and length of tubes could affect the degree of compensatory glottal adduction during exercises, it is also important to notice that instructions to the subjects is a relevant aspect that should be take into account during voice therapy or training as well. Usually, an easy and soft phonation is required from patients during voice therapy when using SOVTE.

It is interesting to note the possible relationship between CQ and air pressure measures. In our third study, phonatory tasks showing the highest values of CQ were tube in water and stirring straw in air. These exercises were also found to present the highest values of Psub and Poral. It seems that when supraglottic load
increased, a higher Psub and a compensatory glottal adduction were produced. Even though most earlier studies have demonstrated that the higher the depth of immersion, the higher is the CQ values, a recent HSDI investigation showed an increased OQ with increased water depth (2-6 cm) (Granqvist et al. 2014). Similar finding were reported in the article I from the present dissertation (i.e. increased resistance produced a decreased CQ = increased OQ). This discrepancy could be due to the fact the tube was submerged in shallower depth than depth used in the other studies. 5 cm, 10 cm and 18 cm were used in our high speed research. Perhaps, tube phonation with deeper immersion depths (> 6 cm) may thus require more adduction of the vocal folds.

Verdolini et al. (1998) suggested that CQ_{EGG} strongly correlates with the degree of glottal impact stress. Hence, considering that during high flow resistance exercises (e.g. tube into the water or phonation into narrow straw) Psub and CQ_{EGG} increased together, an increased Psub could be associated to a higher degree of vocal fold impact stress. Earlier modeling studies have explored this possible association (Jiang and Titze 1994; Horáček et al. 2007). In an excised larynx experiment, Jiang and Titze (1994) demonstrated that peaks of impact pressure were positively related to adduction of the vocal folds, elongation, and subglottal pressure. Maximum impact stress was found at the midpoint of the membranous vocal folds. Furthermore, Horáček et al. (2007), using an aeroelastic model of voice production showed that both impact stress and acceleration increased with Psub.

The association between Psub and degree of impact stress has also an important clinical application that should be taken into account when choosing the right SOVTE depending on the perceptual, aerodynamic, electroglottographic and self-reported characteristic of voice.

One of the most frequently used SOVTE in voice clinic is tube or straw phonation with the free end in air. In our first study (single case study), CQ_{EGG}
decreased during both tube and straw. This change remained both after tube and after straw. However, no consistency has been observed in previous studies when measure CQEGG during tube phonation with the free end in air. (e.g. Laukkanen 1992; Titze et al. 2002; Gaskill et al. 2010; Gaskill et al. 2012; Guzman et al. 2013; Guzman et al. 2015). It is possible to hypothesize that perhaps vocal training status is a determinant variable that should be considered when comparing CQ during tube phonation and vowel phonation. Nevertheless, results from the third study of the present dissertation demonstrated no clear trend regarding vocal status of participants (normal voice without voice training, normal voice with voice training, functional dysphonia, and vocal folds paralysis). Therefore, it seems that glottal adduction during tube phonation in air is more dependent on the individual compensatory adjustments.

ALR was another glottal parameter analyzed in our HSDI study. In most cases, for all immersion depths, ALR evidenced a decrease during phonation into water. A decrease in Ptrans could be a possible explanation. Since from the biomechanics point of view a low ALR is associated to a low vocal fold impact stress, the possibility of vocal folds phonotrauma is expected to decrease when ALR is low. In a previous HSDI study, Laukkanen et al. (2007) analyzed vocal fold vibration using the amplitude to dynamic length (A-DL) ratio during tube phonation. An increased A-DL ratio was found for the longest tube, which according to authors, may suggest a raised vocal effort (increased Psub). In our HSDI study, a relatively high ALR was found for some subjects with tube 18 cm in water.

Another glottal variable related to the degree of vocal fold impact stress is the ClQ. This parameter is specifically related to the abruptness of vocal fold closure, which is also expected to influence the degree of vocal fold collision. The lower the ClQ value, the more abrupt the closure should be. Our results showed two trends. ClQ increased after 5 cm phonation compared to baseline condition and during phonation in the tube 10 cm in the water. Therefore, in these two cases the closing
phase was relatively longer compared to baseline condition. Laukkanen et al. (2007) observed that CIQ was lower for longer tubes compared to shorter ones. Since a long tube and/or narrow tube could be similar to a tube submerged deep under the water surface regarding the degree of airflow resistance, results from our study could be considered opposite to the outcomes by Laukkanen et al. (2007). Since mean values were not analyzed in our study and Laukkanen et al. (2007) study, individual variations due to different compensation strategies could be a suitable explanation for this discrepancy between these two previous studies.

Spectral flatness is defined as the spectral declination or spectral tilt. Previous studies using acoustic analysis of voice have reported that SOVTE produced a less steep spectral slope after exercising (Laukkanen 1992; Guzman et al. 2012; Guzman et al. 2013a, Guzman et al. 2013b) suggesting that semi-occlusions produce an increased spectral energy in the higher part of the spectrum. The second study of the present dissertation (HSDI investigation) showed no clear trends in samples obtained after tube phonation. Moreover, a decrease of spectral flatness during tube phonation for all immersion depths compared to baseline was observed. This implies a steeper spectral slope. The amplitude of higher harmonics is particularly sensitive to the speed at which the glottal airflow decreases at the end of the closing phase (Story et al. 2000). A lower SF seems to imply a smoother collision between the vocal folds (Fant 1960; Gauffin and Sundberg 1989).

Commonly, when a voice sample has a steeper spectral slope, i.e. less harmonic energy in the higher harmonics, there is also an increment of noise energy. HNR is the ratio between harmonic and noise energy. This variable was analyzed based on image analysis in our HSDI study. This parameter demonstrated an increase after tube phonation submerged into 5 cm of water compared to pre tube samples. Thus, more harmonic energy compared to noise was observed. Earlier studies, based on acoustic signal analysis (Paes et al. 2013; Guzman et al. 2011; Guzman et al. 2012)
have also reported a decreased noise energy after SOVTE in subjects diagnosed with functional dysphonia.

Perturbation measures, specifically Jitter and shimmer have also been assessed before and after semioccluded vocal tract exercises. In a study performed with school teachers diagnosed with slight dysphonia, both jitter and shimmer were found to be significantly lower after vocal exercises with stirring straw (Guzman et al. 2011). Moreover, a decrease in jitter and shimmer in normal-voiced participants after vocal warm-up using semiocclusions was reported by Barrichelo-Lindström et al. (2007). Results from our high speed study showed an interesting trend, jitter decreased during and after phonation into a tube submerged 5 cm in water for most participants. These results suggest that phonation into a tube in 5 cm water and after it has a stabilizing effect on glottal function.

4.3 Influence of semioccluded exercises on vocal tract

Two studies from the present dissertation addressed the influence of SOVTE on vocal tract configuration. The first study used CT, while the fourth study explored vocal tract changes through flexible transnasal laryngeal endoscopy. Main findings from both investigations are in general concordant.

VLP presented one of the major changes observed during semiocclusions. In our CT study, the subject demonstrated a lower LVP during both phonation into the traditional Finnish glass tube and stirring straw, the change being more prominent during the latter. Similar results were found in our flexible laryngoscopy study, all SOVTE produced a lower VLP compared to the resting position. Phonation with tube into the water (10 and 3 cm below the surface) and phonation into a long-narrow tube (a kind of a long stirring straw) produced the three lowest VLP. In both studies, the lowest VLP was found during exercises with the greatest airflow.
resistance (narrow straws and tube into water). Two earlier dual-channel electroglottograph investigations have reported similar results regarding VLP. Laukkanen et al. (1999) found a lower VLP compared to the resting position during occlusive consonants (also considered SOVTE). Moreover, Wistbacka et al. (2015) analyzed VLP during tube phonation with the free end in air and in water. EGG results showed that VLP was lower during tube in water, while it increased during phonation with the tube end in air. These findings confirm the fact that exercises with higher airflow resistance tend to lower the larynx.

Although, several studies have reported a lower VLP during SOVTE compared to resting position, the opposite effect of tube phonation on this dependent variable has also been demonstrated (Laukkanen et al. 1996; Laukkanen et al. 2012; Vampola et al.) Furthermore, two magnetic resonance imaging studies reported no changes on the VLP during phonation into a resonance tube and during voiced plosive consonants (Laukkanen 2010; Laukkanen, 2012). Additionally, Guzman et al. (In press), in a recent CT research performed in subjects with voice disorders, assessed VLP among other vocal tract features during phonation in commercial drinking and stirring straws. Even though larynx tended to be lower during both straws (the change being more prominent for the stirring straw), no significant differences were found compared to vowel phonation.

As it comes to different results concerning VLP, it is worth mentioning some possible underlying factors like individual differences. Differences in the subjects background education (e.g. singing vs speaking training) and also the exploratory method used (e.g. dual-channel EGG and imaging techniques) could be reasonable explanations for this discrepancy. Regarding exploratory methods, Laukkanen et al. (1999) in a study aimed to investigate the behavior of multi-channel EGG in VLP using videofluorography as a reference method, reported that results from EGG may be affected by e.g. anterior-posterior movements of the larynx, and even
opposite results can be obtained with EGG than what really happens based on videofluoroscopic findings.

It is commonly agree in voice clinic that a low VLP is desirable during voice exercises, since it has been linked to a healthy and relaxed way to produce voice. Therefore, semi-occlusions and lengthening of the vocal tract might have an important therapeutic effect if they really produce a laryngeal lowering. In Finland, a lowering of larynx is an important goal during tube phonation in water (water resistance therapy). Sovijärvi et al. (1989) stated that the positive outcomes of the resonance tube method are due to the efficient lowering of the larynx during exercising. It is important to notice that a lowered larynx during speaking is not a goal since it would cause an unnatural sound.

The lower VLP found in the first study of the present dissertation (CT study), is concordant to acoustic results obtained from FFT analysis. A decrease in the frequency of the first five formants was observed. In a previous study performed with MRI, similar results were observed in a male subject who lowered the larynx after vocal warm-up with spoken exercises (Laukkanen 2012). The acoustic theory of speech states that the formant frequencies depend on the length of the vocal tract and the cross-sectional shape of the vocal tract as a function of its length (Fant 1970; Kent 1993). Vocal tract length determines the average location of formant frequencies. In this regard, as length increases, the value of the formant frequency will decrease. Since a lower larynx increases the vocal tract length, all formant frequencies are expected to be lower. Lowering of formant frequencies, especially the first formant, could contribute to a greater source-filter interaction determined by the approximation of F1 to F0. Story et al. (2000) stated that phonating at a frequency at or near the first formant may allow for an efficient voice production that could possibly be associated with lower effort.

Originally, phonation in narrow tube with the free end in air or a resonance tube in water was recommended in clinic for patients with hypernasality since such
exercises were supposed to rise the velum (Gundermann 1977). Currently, velum position is another vocal tract feature being explored during SOVTE. Three previous studies have demonstrated that velum rose to seal the nasopharyngeal port during the tube phonation and remained elevated after exercising (Laukkanen et al. 2010; Vampola et al. 2011; Laukkanen et al. 2012). Our CT study showed exactly the same changes during and after phonation into both glass tube and stirring straw. In another CT study, Guzman et al (in press), demonstrated similar results in subjects with functional dysphonia during voice production in drinking and stirring straws. However changes were significant only during stirring straw, possible due to a higher degree of flow resistance.

Energy transfer is expected to increase with a proper velar closure by decreasing the damping caused by the nasal tract. This, in turn, should produce an increase in the total SPL (Laukkanen et al, 2010). However, acoustic data from the first study of the present dissertation did not demonstrate a greater total SPL after tube/straw phonation. Since the participant was asked to keep the same loudness level during all phonatory tasks, the lack of change in total SPL could be expected. The spectral energy of F1 region did not change during the sequence neither. Since energy of F1 supports most of total SPL, this is also expected.

An open and relaxed throat is another common goal of voice therapists and vocal coaches, in patients with excessive muscle tension and speech and singing training. Hence, a wide pharyngeal area during voice exercises should be desirable in voice therapy. Increment of pharyngeal area during all SOVTE was another change observed in the fourth study from the present dissertation. Phonation with tube into the water (10 and 3 cm below the surface) and phonation into a long-narrow tube produced the most prominent changes. Earlier studies have demonstrated similar outcomes on pharyngeal configuration (Laukkanen et al. 2010; Vampola et al. 2011; Laukkanen 2012). Moreover, in our CT study, several changes were observed during both glass resonance tube and narrow straw
phonation regarding pharyngeal configuration. The lower pharynx area, the middle pharyngeal region, and anterior-posterior width of the hypopharynx increased during exercising compared to vowel phonation prior to the exercises. All of these changes were larger during stirring straw than glass tube phonation. Vampola et al. (2011) in a CT investigation showed that the most dominant change in the vocal tract during phonation into a glass tube was caused by expansion of the cross-sectional area of the oropharynx. In another CT study performed in patients with voice disorders, a significant increment was found in hypopahrynx width during two types of plastic straws (drinking and stirring straws). Increase of the inlet to the lower pharynx and the epilaryngeal region was also observed during both straws, but no significant differences were reported (Guzman et al. In press). Additionally, correlations were found between inlet to the lower pharynx and epilaryngeal region for both type of straws, hypopharyngeal width and inlet to the lower pharynx for both straws, and hypopharyngeal width and epilaryngeal region for drinking straw (Guzman et al. in press).

The ratio between the inlet to the lower pharynx (Ap) and the outlet of the epilaryngeal tube (Ae) was also explored in the first study of the present dissertation (Study I). Sundberg (1974) stated that the epilaryngeal tube should act as a separate resonator (i.e. acoustically unlinked from the rest of the vocal tract) when the cross-sectional area in the pharynx (Ap) is at least six times wider than that of the laryngeal tube opening (Ae). When this relation is produced, the singer’s formant cluster (a prominent spectrum envelope peak near 3 kHz associated with the "ringing" voice quality) is likely to be produced. Acoustically, singing formant is produce by a clustering of third, fourth and fifth formants. Findings from our study I, demonstrated a clear tendency that the ratio Ap/Ae increases during and after both tube and straw phonation. The greatest Ap/Ae ratio was observed after straw phonation (5.5) which is close to the value suggested by Sundberg (1974) (6.0). Related to this, results from the study IV (flexible laryngoscopy) showed an
aryepiglottic narrowing (A-P laryngeal compression) during all SOVTE. This modification was more prominent during phonation with a tube submerged under the water (3 and 10 cm) and during phonation with a long-narrow tube, as occurred for the VLP and pharyngeal width. Similar findings have been obtained by Peltokoski et al. (2016) in a nasoendoscopy study. Epilaryngeal region narrowed more when the resonance tube was in water compared to resonance tube in air. In addition, two interesting correlations were found in the study IV of the present dissertation: 1) between pharyngeal width and A-P laryngeal compression, which means that when the pharynx widened, there was also a narrowing of the aryepiglottic sphincter. 2) Between VLP and A-P laryngeal compression, which indicates that when the larynx is lower, there was also a narrowing of the aryepiglottic sphincter. Regarding the latter correlation, Sundberg (1994) has suggested that a low VLP is a way to obtain an aryepiglottic narrowing.

Results from acoustic analysis in the first study of this dissertation (CT study) showed three interesting findings that could be related to results found regarding Ap/Ae ratio: 1) after straw phonation, the largest increase in the energy of the speaker’s/singer’s formant cluster region was found, 2) the singing power ratio (SPR) demonstrated the largest decrease after straw phonation, which implies a less steep spectral slope between the low harmonics and the harmonics located near to 3 kHz (singing formant region), 3) there was a clear formant cluster of F3 and F4 after exercising with both a glass tube and a straw, which was located in 2500-3000 Hz for vowel [a:] after exercising. In line with our acoustic findings, two earlier case studies have also showed an increment of the spectral energy in the singer’s/speaker’s formant region (Laukkanen et al. 2010; Laukkenen et al. 2012) after tube phonation. Additionally, two studies whose aim was to compare the effect on spectral energy distribution of SOVTE and open vowel exercising, showed an increased spectral energy in the higher part of the spectrum (2000-5000 Hz) after SOVTE compared to voice spectrum after vocalizations with open vocal
tract (Guzman et al. 2013a; Guzman et al. 2013b). Considering formant clustering after exercising with SOVTE, a significant decrease in the formant frequency distance after straw phonation was also found by Laukkanen et al (2010). Such formant clustering may have contributed to the changes mentioned above (lower SPR, and higher SPL between 2500-4000 Hz).

These acoustic data suggest a warm-up effect on voice spectrum after straw phonation, which could contribute to a more economic voice production. Vocal economy is defined as the ratio between voice output (decibels) and intraglottal impact stress (Kilopascal) (Verdolini et al. 1998; Berry et al. 2001). Since an increase of acoustic energy in the singer’s/speaker’s formant cluster region contribute to the perceived vocal loudness level, a lower SPR and a higher SPL between 2500-4000 Hz may help to project vocal sound without increasing laryngeal effort (recall that CQ decreased after straw phonation).

Another possible explanation of the energy increment observed in the singer’s formant region in our first study could be an increased aerodynamic-acoustic interaction of the vocal tract. This increment may be explained by an increased input impedance of the vocal tract due to the higher Ap/Ae ratio. This increment may be explained by an increased input impedance of the vocal tract due to the higher Ap/Ae ratio. Increased input impedance (Especially positive reactance at the fundamental frequency region) produces a faster cessation of the glottal flow and consequently a less tilting spectrum (Rothenberg 1986; Fant 1987, Story et al. 2000). According to the results by Vampola et al. (2011) the changes found after tube phonation resulted in an increase of total SPL, attributable to an increased reactance of the vocal tract after the exercise, which in turn was caused by an increment of Ap/Ae ratio. Specifically, the subject increased the vocal tract input reactance and make the vocal tract more inertive by expanding mainly the oropharynx rather than by reducing the epilaryngeal tube.
Moreover, an increased subglottic pressure after straw phonation could also explain a less steep spectral tilt and a higher SPL in the singer’s/speaker’s formant region. Nordenberg and Sundberg (2004) demonstrated that intensity is not linearly correlated to the spectral contour. Energy of all frequencies of the spectrum does not increase proportionally when there is an increase in total sound level. The gain in dB in the region of lower frequencies is lower than in region of high frequencies when total sound pressure level is increased (White 1998, Nordenberg et al. 2004).

Vocal tract changes observed in studies I and IV such as wider pharynx, lower larynx and raised velum should be expected to increase total vocal tract volume. In fact, two previous works have reported this modification during tube phonation. Vampola et al. (2011) revealed that the total volume increased 38.5% after phonation into the tube when compared to vowel phonation before tube. According to the authors the increase in volume was mostly due to transversal expansion of the vocal tract. Similar outcomes were found by Guzman et al. (In press). The total volume of the vocal tract was larger during both drinking straw and stirring straw compared to baseline, but it was significantly different only during use of the stirring straw. The main cause of this increase during stirring straw phonation is likely the same reason as explained by Vampola et al. (2011) study, an increase in transversal dimension of the vocal tract. Data by Guzman et al. (In press) showed a strong correlation of vocal tract total volume with the lower pharynx (A3) and the inlet to the lower pharynx (Ap) during straw exercises, suggesting an association between volume of the vocal tract and cross sectional areas. Even though changes in vertical length of vocal tract (i.e. lowering of the larynx and rising of the velum) may contribute to a larger total vocal tract volume during tube phonation, it seems that the main cause are the changes observed in cross sectional areas.

Increased total volume caused by all vocal tract changes described above is possibly due to the increased Poral during tube phonation. The increment on this parameter during SOVTE may have mechanically pushed the larynx down, the
pharyngeal walls laterally, and velum up. However, it is also possible that participants consciously closed the velum in order to get all the flow through the tube. Therefore, velum closure could also be made by active muscle functioning and not necessarily something that happens mechanically. Moreover, similar laryngeal muscle activation could cause the laryngeal lowering during tube phonation.

The first option mention above (increased volume is cause by increment in Poral) is supported by earlier studies showing that during occlusion at the lip and or artificial lengthening of the vocal tract, Poral raised compared to open vowel production (Titze et al. 2001; Amarante et al. 2016; Radolf et al. 2014; Maxfield et al. 2015). A recent investigation aimed to assess the static back pressure (Poral) and airflow for different tubes commonly used for voice therapy, reported that changes in the diameter of straws affect Poral considerably more compared with the same amount of relative change in length (Amarante et al. 2016). Furthermore, Maxfield et al. (2015) created a rank ordering by measuring the intraoral pressure produced by thirteen SOVTE commonly used in voice therapy. The highest values of Poral were found in straw submerged in water, raspberries, and stirring straw with the free end in air. These findings may explain the vocal tract differences found in our studies I and IV when compared different type of SOVTE. In the CT study, both Finnish glass tube and the thin stirring straw produce the same vocal tract changes, but all modifications were more prominent during the latter (it is narrower than the glass tube). In the flexible laryngeal endoscopy study, similar results were reported. Although, all exercises produced a wider pharynx, and lower larynx, more prominent changes were found during tube submerged into water (10 and 3 cm) and during narrow-long tube (type of stirring long straw). In addition, an interesting finding from our study IV may be linked to the impact of Poral on vocal tract changes. All dependent variables demonstrated the greatest degree of change when phonating with loud voice (Recall that three loudness levels were
requires to participants for all exercises). This result could be related to the higher
degree of Psub produced during artificial lengthening and occlusions of the vocal
tract when vocal intensity is higher. It seems that louder phonation caused by a
higher Psub is a factor to modify vocal tract shape. In this regard, loudness level
has also demonstrated in earlier works to be a factor influencing the degree of
vocal tract modifications in professional voice users (Yanagisawa et al. 1989;
Mayerhoff et al. 2014; Guzman et al. 2013, Guzman et al. in press). Data by
Mayerhoff et al. (2014) showed that both the degree of medial and A-P
compression were greater during loud phonation. These results were observed
during sustained vowel and connected singing productions by classically trained
singers. Guzman et al. (2013) found similar results. Authors reported that high
loudness produced the highest degree of laryngeal A-P compression, laryngeal
medial compression, and pharyngeal compression. Yanagisawa et al. (1989) found
greater A-P laryngeal compression when subjects performed loud phonation and
during the three loudest voice qualities: belting, twang, and opera.

4.4 Possible sources of error

Some possible sources of error from the studies included in the present dissertation
could be mentioned. The first study (CT investigation) only included one single
subject. Moreover, this subject was vocally trained and with several years of
experience using tubes and other SOVTE. However, most CT, acoustic and air
pressures measures findings are in line to earlier research and recent investigations
as well.

Due to small sample size, in the HSDI study (Study II) we only were able to
analyze the trend of values for all parameters and phonatory tasks without using
hypothesis testing to avoid lack of statistical power. Moreover individual variations
caused some difficulties in the interpretation of results. Additionally, in this study rigid laryngeal endoscopy was used. This could have affected voice production since tongue position during this exploratory method is not the natural position during regular speaking. Further research should include a larger sample size, statistical analysis and use flexible laryngeal endoscopy.

Even though our third study (air pressure and contact quotient measures study) included four groups of subjects representing four different voice conditions, it did not consider participants diagnosed with phonotraumatic organic dysphonia (e.g. vocal nodules, polyps, cysts, etc.). Likely, organic phonotraumatic lesions could cause different type of aerodynamic and glottal compensations during tube phonation.

Additionally, this study did not include glottal airflow measures. This variable is important since it is related to both glottal resistance and subglottic pressure.

One important aspect to be considered in the fourth study of the present dissertation if that non-objective measures were used. Flexible laryngeal endoscopy allows a full view of pharynx and larynx during examination, however, assessment is dependent on the rater’s experience. Moreover, the distance between the distal part of the fiber and laryngeal/pharyngeal structures could have been modified simply by the movement of the experimenter’s hand. To decrease the possible effect of these two sources of errors, we included otolaryngologists with more than 4 years of experience in voice disorders as judges. All of them had a laryngology fellow certification. Furthermore, inter and intra-rater reliability analysis was included. Moreover, the placement of flexible endoscope was fixed by securing the fiberscope against the alar cartilage of the nose with the laryngologist’s finger.
4.5 Future research

Even though several studies have been performed regarding semioccluded vocal tract exercises, there are still some issues that need to be addressed in further investigations. One important aspect that should be considered is the possible effect of semiocclusions on breathing function. To date an important number of studies have reported changes on air pressure measures (e.g. subglottic pressure) that somehow reflect some aspects of breathing function. Nevertheless, variables such as degree of thoracic and abdominal excursions during and after SOVTE, respiratory volumes, and respiratory muscle activity should be investigated.

Additionally, most studies on tube phonation have explored the effect during or immediately after exercises. However, to the best of our knowledge only three investigations have sought to know the efficacy of tube phonation protocols in subjects with voice disorders after several sessions of voice therapy. All of them have included a small sample size. In a recent randomized controlled trial study, Kapsner-Smith et al. (2015) found significant improvement in Voice Handicap Index scores compared to the control condition (no-treatment group) after a therapeutic program based on flow-resistant tube exercises. Moreover, significant improvement in roughness (from the CAPE-V scale) was found relative to control group. Additionally, two longitudinal studies have been designed using phonation into tubes submerged in water (water resistance therapy) in subjects with behavioral dysphonia, (Tapani 1992; Simberg et al, 2006) and one study was conducted to explore the effect of drinking straw phonation in air and the bilabial consonant /β:/ in a group of acting students diagnosed with muscle tension dysphonia (Guzman et al. 2012). More studies including a larger sample size with in different diagnosis and populations should be conducted in future research.
5 CONCLUSIONS

1. Overall, vocal exercises including semi-occlusions and artificial lengthening of the vocal tract have a simultaneous effect on phonation, vocal tract configuration and aerodynamic variables.

2. Most changes suggest that SOVTE are good tools for subjects with voice disorders and professional voice users. These changes seem to lead to a more economic voice production.

3. Subjects with different voice conditions (e.g. with voice disorders, healthy non-vocally trained, and healthy trained subjects) behave quantitatively and qualitatively in a similar manner under similar exercise’s conditions (i.e. flow resistance exercises).

4. The degree of flow resistance should be considered when choosing treatment exercises for patients with different types of diagnoses (e.g. hyperfunctional or hypofunctional dysphonia). High degree of flow resistance exercises (e.g. stirring straws and deep immersion when tube is submerged into the water) could be more beneficial for patients with vocal fold paralysis or presbyphonia. On the other hand, glottal function of patients with hyperaduction or vocal fatigue could be improved using exercises with low degree of flow resistance (e.g. traditional glass tube and shallow depths of immersion).
5. Aerodynamic aspects are mainly affected during SOVTE by increasing air pressure measures such as Poral and Psub, the latter being a compensation of the former. An increase in Poral, in turn, is caused by the airflow resistance offered by semi-occlusions.

6. Influence of SOVTE in phonation is mainly related to the degree of vocal fold collision and glottal adduction during exercising. This is mostly determined by the degree of resistance to the airflow that each exercise offers. High resistance seems to promote a greater vocal fold collision compared to a lower resistance exercise. Therefore, higher resistance exercises could be more beneficial for patients with diagnosis such as vocal fold paralysis or presbyphonia. On the other hand, glottal function of patients with hyperadduction or vocal fatigue could be improved using exercises involving lower airflow resistance.

7. Important therapeutic and training goals can be reached during SOVTE regarding vocal tract configuration. A lower larynx, a wider pharynx and a higher velum were produced during this type of exercises compared to vowel phonation, the change being more prominent when the airflow resistance is higher. These modifications are linked to positive changes in acoustic and perceptual output of voice.
ACKNOWLEDGMENTS

First of all, I would like to deeply thank Dr. Anne-Maria Laukkanen, my PhD advisor professor. Dr. Laukkanen helped me plan this thesis, revise its proposal, and monitor the whole process. Her continuous help and support were invaluable to this thesis and my career as a voice scientist. Her outstanding knowledge, generosity, and patience were very important to carry out all studies and writing process. She also motivated me to continue going forward in the science of voice field.

I am also extremely grateful to Dr. Matthias Echternach. Thanks him, his voice lab team, and his dedicated work and knowledge, the HSDI study was possible.

My sincere thanks go also to all scientists that collaborated in all steps in the CT study. Thanks to Dr. Jan Švec, Dr. Jaromír Horáček, and Dr. Petr Krupa.

Special thanks to my friend and colleague Dr. Ahmed Geneid. We share so many experiences together during the CT and HSDI studies in Czech Republic and Germany respectively. He is really a so talented professional.

I also express my sincere gratitude to Dr. Bernhard Richter and Dr. Louisa Traser for all their collaboration during the HSDI study in Germany.

I would like to also thank all the Chilean co-authors I had the pleasure to work with. Without their work, this thesis would not have been possible. Thanks for their support, understanding, and cooperation. Many thanks Dr. Cristian Olavarria, Dr. Miguel Leiva, Cristian Castro, Sofia Madrid, Dr. Daniel Muñoz, Dr. Alba Testart, and Elizabeth Jaramillo.
The study I was supported by the Academy of Finland, grant number 128095. In the Czech Republic, it was supported by the project GACR No. P101/12/1306 and the European Social Fund project no. CZ.1.07/2.3.00/20.0057.

The study II was supported by the Deutsche Forschungsgemeinschaft, grant Ec 409/1-1 and Ri 1050/4-1 in Germany. In Finland, it was supported by the Finnish Academy of Sciences (grant number 1128095).

Studies III and IV received no financial support.


Laukkanen AM, Horáček J, Havlík R. Case-study magnetic resonance imaging and acoustic investigation of the effects of vocal warm-up on two voice professionals. Logopedics Phoniatrics Vocolgy. 2012: 37, 75-82.


Lohscheller J, Eysholdt U, Toy H, Doellinger M. Phonovibrography: mapping high-speed movies of vocal fold vibrations into 2D-diagrams for visualizing and analyzing the underlying laryngeal dynamics. Institute of Electrical and Electronic Transaction on Medical Imaging. 2007.


Vocal Tract and Glottal Function During and After Vocal Exercising With Resonance Tube and Straw

*†Marco Guzman, ‡Anne-Maria Laukkanen, †Petr Krupa, †Jaromír Horáček, †*Jan G. Svec, and †‡#Ahmed Geneid, *Santiago, Chile, †Tampere and ‡Helsinki, Finland, ‡Brno, †Prague, and †‡Olomouc, Czech Republic, and #Ismailia, Egypt

Summary: Objective. The present study aimed to investigate the vocal tract and glottal function during and after phonation into a tube and a stirring straw.

Methods. A male classically trained singer was assessed. Computerized tomography (CT) was performed when the subject produced [a:] at comfortable speaking pitch, phonated into the resonance tube and when repeating [a:] after the exercise. Similar procedure was performed with a narrow straw after 15 minutes silence. Anatomic distances and area measures were obtained from CT midsagittal and transversal images. Acoustic, perceptual, electroglottographic (EGG), and subglottic pressure measures were also obtained.

Results. During and after phonation into the tube or straw, the velum closed the nasal passage better, the larynx position lowered, and hypopharynx area widened. Moreover, the ratio between the inlet of the lower pharynx and the outlet of the epilaryngeal tube became larger during and after tube/straw phonation. Acoustic results revealed a stronger spectral prominence in the singer/speaker’s formant cluster region after exercising. Listening test demonstrated better voice quality after straw/tube than before. Contact quotient derived from EGG decreased during both tubé and straw and remained lower after exercising. Subglottic pressure increased during straw and remained somewhat higher after it.

Conclusion. CT and acoustic results indicated that vocal exercises with increased vocal tract impedance lead to increased vocal efficiency and economy. One of the major changes was the more prominent singer’s/speaker’s formant cluster. Vocal tract and glottal modifications were more prominent during and after straw exercising compared with tube phonation.


INTRODUCTION

Semi-occluded vocal tract setting has been extensively used by speech pathologists and voice trainers as therapeutic and training exercises, respectively. Various types of tubes have been used to perform these voice exercises. One of them is the traditional Finnish glass tube, “resonance tube,” 26–28 cm in length and 8–9 mm in inner diameter. 1,2 A more accessible option is commercial plastic drinking straws. Moreover, Titze 3,4 has proposed the use of stirring straws, shorter and thinner plastic straws, which are commonly used to stir coffee in the United States. Resonance tube phonation into the water (water resistance therapy) also has a long tradition as a therapeutic tool. 1,5

Several benefits have been attributed to vocal exercises involving tube phonation or other semi-occluded vocal tract postures, such as y-buzz, 4 tongue trill, lip trill, and voiced bilabial fricative [β:]. Some of these benefits are an increase in vocal tract impedance, specifically resulting in changes in the inertive reactance, 7–11 which may be favorable to voice production by decreasing phonation threshold pressure 10 and by increasing skewing of the glottal flow waveform (faster cessation of the flow). 9,10 The vocal tract impedance can affect the voice source function in two ways (1) through an acoustic-aerodynamic interaction and (2) through a mechano-acoustic interaction. 7,12

In the former, the shape of the glottal flow pulse is affected by the acoustic pressures in the vocal tract. 7,8 When fundamental frequencies (F0s) are below the first formant, the skewing of the flow pulse is increased by supraglottic acoustic pressures compared with the glottal area, so that the airflow is suppressed at glottal opening and maintained during the glottal closing phase. This increased skewing of the glottal flow waveform leads to strengthening of the higher harmonics (less spectral tilt), increase in sound pressure level (SPL), 9,13,14 and to a more resonant voice quality, that is, vibratory sensations in the frontal part of face, alveolar ridge, and head area with easy voice production. 15 This is reflected in a brighter and louder sound. Titze and Sundberg 16 pointed out that because the skewing of the glottal airflow signal is one of the determinants of vocal intensity, the source-filter interaction can be used to increase intensity rather than vibrational amplitude, thus avoiding an increase in the vocal fold impact stress. 14

The second way that glottal source function can be affected by vocal tract impedance is the mechano-acoustic interaction of the vocal tract pressures and the vocal folds. 7,8,17,18 This
occurs when the increased vocal tract inertance affects the vocal fold vibration itself (through affecting the pressure above and inside the glottis). Specifically, the inertive reactance lowers the phonation threshold pressure (the subglottal pressure required to barely initiate and sustain phonation). A low phonation threshold pressure would produce an ease of phonation (decrease in perceived phonatory effort).

Not only an anterior semi-occlusion or lengthening of the vocal tract can produce this favorable condition but also a narrowing in the lower vocal tract. Titze and Story reported that when the epilaryngeal tube is quite narrow, making the input impedance to the vocal tract comparable to the glottal impedance, there is an important source-filter interaction. The inertance (the positive component of the reactance) of the vocal tract facilitates vocal fold vibration by lowering the oscillation threshold pressure. When the narrowed epilaryngeal tube is combined with a wide pharynx, two relevant effects occur: (1) the vocal tract is inertive for a wide range of \( P_0 \) values and (2) a formant cluster between \( F_3 \) and \( F_5 \) is produced (singer’s or speaker’s formant), which is desirable to increase the vocal loudness without an increase in vocal effort.

An increased supraglottal pressure and consequently an elevation of the intraglottal pressure have also been reported during semi-occlusions. When the vocal tract constriction is tight or the tube added to the vocal tract is narrow or long, it increases supraglottal air pressure which tends to separate the vocal folds if not compensated by raising subglottic pressure and adduction. Separation of the vocal folds (decreased adduction), in turn, would explain a decreased contact quotient (CQ). Several studies have reported a change in the relative closed time of the glottis on an electroglosstographic (EGG) signal (eg, CQ) when semi-occlusion is compared with vowel phonation. Only some of them have shown a decreased CQ.

Different methods have been used to obtain data of the effects of these vocal exercises. Electroglosstography, acoustics, modeling, and simulation studies are the most dominant. In addition, some studies have been carried out using aerodynamic, electromyographic, endoscopic, and radiological measurements. Most of them have examined the changes of vocal folds parameters during and/or after semi-occluded exercises. Nevertheless, few researches have reported outcomes related to vocal tract configuration changes (shape modification) as an effect of the semi-occlusion or artificial lengthening of the vocal tract.

In this regard, in a recent investigation conducted by Laukkonen et al., one female participant was assessed with magnetic resonance imaging (MRI) and acoustic analysis before and after straw phonation. Midsagittal area of the vocal tract increased and the velar closure improved during and after straw exercising. Furthermore, the ratio of the transversal area of the lower pharynx over that of the epilarynx increased both during and after the straw. Acoustic changes showed a higher total SPL and also more energy in the speaker’s formant cluster. The distances between \( F_4 \) and \( F_3 \) and \( F_5 \) and \( F_4 \) demonstrated a decrease. Authors suggested that the use of straw vocal exercises helps to produce a speaker’s formant cluster, which increases loudness and thus improves vocal economy. Similar MRI and acoustic findings were observed in another study designed to observe changes in voice production after warm-up for two professional voice users. One of the participants used semi-occlusions ([B:], [m:]), and closed vowels [y:] and [u:]).

Vampola et al. examined the vocal tract shape in a female subject before, during, and after phonation into a tube using computerized tomography (CT) and finite element models (FEMs) to study changes in vocal tract input impedance. Results indicated that the phonation into a tube causes changes in the vocal tract, which remain also when the tube is removed. Authors observed tightened velopharyngeal closure and enlarged cross-sectional areas of the oropharyngeal and oral cavities during and after the tube phonation when phonating on vowel [a:]. FEM calculations revealed an increased input inertance (inertive or positive reactance) of the vocal tract, especially in the frequency range from 2500 to 4000 kHz, and increased acoustic energy radiated out of the vocal tract after the tube phonation.

To date, no studies have been addressed to assess the possible different training effects produced by two different types of artificial lengthening on the vocal tract setting. Few earlier studies have reported differences in glottal source changes when using two or more different semi-occlusions. The present study therefore aimed at investigating the vocal tract modifications and also the acoustic, aerodynamic, and EGG characteristics of the voice when comparing vocal exercising with two different vocal tract impedances (during phonation into a glass tube and a stirring straw). We hypothesize that the glottis and/or the supraglottic behavior should adapt differently to different load impedances of the vocal tract.

**METHOD**

**CT scanning**

CT was carried out in Surgical Clinic, Department of Imaging and Radiology in Brno, Czech Republic. The CT images were acquired using a Toshiba-Aquilion CT machine. The CT imaging parameters used to provide images of the vocal tract were voltage 120 kV, scan option: helical CT, time of the rotation 0.5 seconds, slice thickness 0.5 mm, and total number of slices: 510. The examination was performed for one male classically trained singer (34 years) who has 7 years of experience using tube phonation and other semi-occlusions as vocal training and warm-up exercises. The subject did not report any known voice or hearing pathology at the time of the experiment. In supine position inside the CT machine, the subject was asked to produce the following phonatory tasks: (1) to sustain vowel [a:], (2) to phonate a sustained vowel-like sound into a glass tube (27 cm in length and 9 mm inner diameter) for 15 minutes, and, immediately after that, (3) to produce another sustained vowel [a:]. All phonations were carried out at habitual loudness level and speaking pitch. The participant was required to produce a stable sound with a good closure at the lips and to feel as strong as possible vibratory sensations on the alveolar ridge, face, and head areas during tube phonation. After 15 minutes of complete silence (vocal rest), phonation into a plastic coffee straw (2.5 mm inner diameter and 13.7 cm in length) was
performed for 15 minutes. Immediately after that, the participant was asked to produce another sustained vowel [a:]. This straw was chosen because it was used in a previous study. The subject was scanned two times while producing each phonatory task. He was asked to adopt a relaxed posture in the CT scanner and exactly the same body and head position was kept during the entire CT procedure. The head position was mechanically fixed in a frame during all experiments. The participant was a volunteer and he was informed about the potential health hazards of the CT examination.

CT image analysis
Ten CT midsagittal images (five phonatory tasks × two repetitions) were chosen to perform a series of distance measurements (mm). Anatomic distances (Figure 1) of interest included (1) vertical length of the vocal tract (which is indicative of the vertical laryngeal position [VLP]) measured as the distance between the lowest point of the odontoid process of the Atlas and the vocal folds following a vertical line, (2) horizontal length of the vocal tract measured as the distance between the lower edge of the upper lip and the upper edge of the lower lip, (4) jaw opening measured as the distance between the lowermost edge of the jawbone contour and the anterior end of the hard palate, (5) tongue dorsum height measured as the distance between the lowermost edge of the jawbone and the uppermost point of the tongue dorsum, (6) oropharynx width measured as the distance between the lowest point of the second vertebra and the most posterior part of the tongue contour (for ensuring the same angle, we used straight line from the anterior uppermost edge of the jawbone contour to the anterior lowest point of the second vertebra), (7) velum elevation measured as the distance from the posterior upper edge of the hard palate and the anterior lowest point of the uvula, and (8) hypopharynx width measured as the distance between the lowest point of the pharynx and the internal edge of the epiglottis following a line from the anterior uppermost edge of the jawbone contour to the lower point of the pharynx. All CT measurements were performed using the software vPACS view, Version 6.9.5 (Audio-scan, Prague, Czech Republic).

Moreover, three cross-sectional areas (mm²) were measured from the same 10 midsagittal CT images (Figure 2). The areas were (1) oral cavity (A1) measured from the lips to the velum (up to the line connecting the lowermost edge of the jawbone contour and a break of declivity on the velum surface), (2) the pharyngeal region (A2) measured from the line ending A1 down to the horizontal line connecting the lower edge of the third vertebra and the lowermost edge of the jawbone contour, and (3) the epilaryngeal region (A3) measured from the line ending A2 down to the vocal folds.

Additionally, from the transversal CT images, two areas (mm²) were measured37,38 (Figure 3): (1) the inlet to the lower pharynx (Ap) just above the collar of the epiglottis and (2) the outlet of the epilaryngeal tube (Ae) just below the collar of the epiglottis where the epilarynx and sinus piriformis form three separate tubes. The ratio between these two areas was also calculated (Ap/Ae). Areas from transversal CT images were chosen taken into account the bending of the vocal tract. A line from backbone to jawbone through the tip of arytenoids in a parallel way to the bottom line of the CT image was use to obtain Ae. A line just above the tip of arytenoids, which was also parallel to the bottom line, was used to obtain Ap.

Recordings
The recordings were performed separately from the CT measurements in a sound-treated room. Three signals were
digitized: audio, subglottic (oral) pressure, and EGG signal. The audio signal was recorded by a condenser microphone (Behringer ECM 8000; Behringer, Behringer City, China) at a distance of 40 cm from the mouth. An estimate of the subglottic pressure was recorded from the oral pressure during the occlusion of the consonant [p:] in the syllable [pa:] and by shuttering the outer end of the tube. This pressure was captured with a pressure transducer connected to a thin plastic and flexible tube. The tube was inserted into the corner of the mouth, extending a few millimeters behind the lips, without touching the tongue or any other oral structure. The manometer MSIF2 (Glottal Enterprises, Syracuse, NY) was used.

The EGG signal was recorded with a two-channel electroglottograph (EG2; Glottal Enterprises) using a 20 Hz high-pass filtering (to exclude slow variations in the signal amplitude, which could be due to articulatory movements). The electrodes were cleaned with a slightly wet tissue, and a thin layer of conductive gel was applied (Mingograph electrode cream; Siemens-Elema AB, Munich, Germany). They were positioned near the laminae of the thyroid cartilage and secured with a velcro strip, which was wrapped around the participant’s neck as tightly as possible to prevent any movement of electrodes throughout the data collection. The quality of the EGG signal was monitored throughout the recordings with an oscilloscope.

Audio signal was calibrated for further SPL measurements in dB (Z) using a 440 Hz tone. Air pressure signal was calibrated using a Glottal Enterprises calibrator, model MCU-4. Pressure and EGG signals were digitized and recorded simultaneously into different channels at a sampling rate of 16 kHz using the Soundswell Signal Workstation (Hitech Development AB, Stockholm, Sweden). Audio signal was recorded at a sampling rate of 48 kHz with 16 bits quantization using a DAT recorder (Marantz PMD 671; Marantz, Mahwah, NJ).

Figure 3. Areas (mm²) measured in CT transversal images: the inlet to the lower pharynx (Ap) and the outlet of the epilaryngeal tube (Ae).

Phonatory tasks
EGG and subglottic pressure signals were recorded during the repetition of the syllable [pa:] at habitual loudness and comfortable pitch before performing the vocal exercises. After this sequence, the subject phonated a sustained vowel-like sound into the glass tube (27 cm in length and 9 mm in inner diameter) as a vocal exercise for 5 minutes using his habitual pitch and loudness. He was asked to produce a stable sound with a good closure at the lips and to feel perceptible vibratory sensations on the alveolar ridge, face, and head areas during phonation. Immediately after tube phonation exercises, the participant produced again the same series of consecutive syllable [pa:] and sustained vowel [a] following the last [pa:], using the same pitch and loudness level to assess the eventual effect of tube phonation after the tube exercising. Because subglottic pressure is one of the important factors that affect SPL, one of the experimenters monitored the SPL during the series of consecutive [pa:] syllable with a sound level meter (American recorder technologies SPL-8810) (American Recorder Technologies, Inc., Simi Valley, CA) positioned at a distance of 40 cm from the mouth. Z frequency weighting and slow time response were used for monitoring SPL. The SPL choice was made by the subject in pre-exercising recording (at comfortable loudness). Then, these pre-exercising made free choices became the targets for postexercising samples. An electronic keyboard was used to give and control the pitch, which was monitored auditorily by the participant and one of the experimenters. The recording captured during this consecutive syllable [pa:] was saved as the sample after tube phonation. After 15 minutes of complete silence (vocal rest), exactly the same entire procedure was performed with a plastic stirring straw of 2.5 mm inner diameter and 13.7 cm in length.

Acoustic, air pressure, and EGG data analyses
To compare the samples recorded before and after tube and straw phonations, acoustic measures were made using Praat software (Version 5.2; Boersma and Weenink, 2008, University of Amsterdam, Amsterdam, The Netherlands). Long-term average spectrum (LTAS) and Fast Fourier Transformation (FFT) spectrum (spectral slice and spectrogram) were used. From LTAS analysis, the following variables were assessed: total SPL (energy between 50 and 6000 Hz), F1 energy (spectral energy between 500 and 800 Hz), singer/speaker’s formant energy (energy between 2500 and 4000 Hz), and singing power ratio (SPR) (the difference of energy between the highest peak around 0–2 kHz and the highest peak around 2–4 kHz). A frequency bandwidth of 25 Hz and Hanning window was used for LTAS analysis. FFT spectrum was performed to obtain
an approximation of the formant frequencies from F1 to F5. The formant frequencies were measured from the strongest peaks or in the middle of the two adjacent equally strong peaks in spectral slices (average from each sample). Wide band spectrograms (bandwidth 260 Hz) were used for a comparison. There the formant frequencies were located in the middle of the bands with the strongest intensity. Frequency distances between formants F2-F1, F3-F2, F4-F3, and F5-F4 were then calculated.

Every sample captured before, during, and after exercising into both tube and straw were analyzed to obtain the average subglottic pressure (estimated from the maximum peak of the oral pressure during the occlusion of the consonant [p:] in the syllable [pa:] and during manual shuttering the outer end of the tube), oral pressure (obtained during nonshuttered phase), and the glottal CQ. Transglottal pressure was also calculated from the difference between subglottic and oral pressures. A Soundswell Signal Workstation (Hitech Development AB) was used to calculate the oral pressure values. EGG CQ (the ratio of the duration of the “contact phase” to the entire glottal period) was obtained with the software EFxHist, Version 1.5 (Mark Huckvale, University College of London, UK) from the middle section of the each EGG sample. EFxHist software defines the contact phase using a criterion level of 50% from the peak-to-peak amplitude of the EGG signal.

Auditory-perceptual assessment

To evaluate the voice quality of the samples, a perceptual analysis was conducted with four listeners (one man and three women). This group of blinded judges consisted of speech-language pathologists with more than 8 years of experience in voice training and rehabilitation. Samples recorded before and after tube/straw phonation were played in randomized pairs. Raters were required to judge which sample in each pair was produced with a better voice quality or if there was no difference between them. Listeners could replay each sample as many times as they wanted before making their decision and moving on to the next sample. The evaluation was performed in a sound-treated room using a laptop computer and a high-quality Audioengine loudspeaker (Audioengine, Kowloon, Hong Kong). The listeners were located at approximately 2 m from the loudspeaker. All the listeners reported normal hearing.

RESULTS

CT distance measurements

Mean of the two repeated scans for the anatomic distances (mm) calculated from the midsagittal images of the vocal tract obtained from the CT measurements performed before, during, and after tube/straw phonations are presented in Figure 4. Changes were observed in the vertical length (which is indicative of the VLP). It increased during tube phonation (8%) and even more during straw phonation (21%) (Figures 5 and 6). This lowered laryngeal position remained after tube and straw phonations. The most evident change in laryngeal height occurred when comparing phonations before and during straw phonation (19 mm of difference, 21%). Because during the tube and straw phonations the horizontal length is determined by the length of the tube and straw, respectively, this distance was only measured before and after. Small changes were observed in the horizontal length (less than 2% in all samples). An important modification was observed in the velum position, which rose to seal the nasopharyngeal port during the tube and straw phonations (Figures 5 and 6). In both the tube and straw phonations, the uvula rose 7 mm (35%). Although after tube and straw phonations, the velum was not as high as during them, the higher uvula position remained as compared with the position before tube and straw phonations. When comparing oropharynx width before and after straw phonations, there was about 37% of decrease from 8.7 to 5.65 mm, respectively. This difference was also present after the tube phonation (25%), but it was not as clear as compared with

FIGURE 4. Mean of the two repeated scans for anatomic distances (mm) calculated from the midsagittal images of the vocal tract obtained from the CT measurements performed before, during, and after tube and straw phonations. BT, before tube; DT, during tube; AT, after tube; DS, during straw; AS, after straw.
the sample after the straw phonation. Hypopharynx became wider during straw phonation (38%). During and after tube phonation, negligible changes were found in the hypopharynx width. Jaw opening also showed a decrease during both tube and straw phonations (15% and 13%, respectively). However, this change was probably a direct consequence of the straw and tube presence between lips. Jaw opening did not change substantially after tube/straw phonation. Small changes were revealed in tongue dorsum height for all sample types (less than 7%), except for the sample during straw (13%). No clear differences were demonstrated in lip opening throughout the sequence. Because during tube and straw phonations the lip opening is determined by the diameter of the tube and straw, respectively, this distance was only measured before and after. Additionally, midsagittal images also show a more frontal tongue position in both during and after tube/straw phonation. This change was more evident during that after exercising (Figures 5 and 6).

The mean differences between phonation before compared with phonation during and/or after tube/straw were greater than the differences between repetitions for vertical length of vocal tract, velum elevation, pharyngeal width, jaw opening, and tongue dorsum height (Table 1). Therefore, one may suggest that the reported variations between before, during, and after tube/straw phonations are true differences.

CT area measurements
Mean of the two repeated scans for the cross-sectional areas (mm²) measured from the midsagittal and transversal images of the vocal tract obtained from the CT scannings before, during, and after tube/straw phonations are presented in Figure 7. Ap area increased both during tube and straw phonations compared with vowel phonation before them. This increment was clearly much larger during straw (91%) than glass tube (9%) phonation (Figure 8). Negligible changes were observed when comparing Ap between vowel phonations before and after tube and straw phonations (0.2% and 3.5%, respectively). The Ae area became larger during straw phonation (65%) (Figure 9), but no clear changes were observed comparing the phonations before and during tube phonation (2%). The same area showed a decrease after tube (16%) and even more after straw phonation (23%) (Figure 10). The ratio of areas Ap/Ae measured from the transversal CT images showed a clear tendency: during tube, the ratio increased by 7.9%; after tube, by 19.4%; during straw, by 16%; and after straw, by 35%. A change was observed in oral cavity area (A1), which decreased during both tube and straw phonations (37% and 29%, respectively) (Figures 5 and 6). Although the oral cavity remained smaller after tube and straw phonations compared with the vowel production before them, this change (1% and 9% for tube and straw, respectively) was not as substantial as the modification during them. Because the presence of tube and straw phonations might affect the oral cavity area, the changes during exercising should be taken with caution. The pharyngeal region (A2) showed the same increase (15%) during tube and straw phonations as compared with vowel production before exercising, but it decreased after tube and straw phonations compared with phonation before them (20% and 27%, respectively). The epilaryngeal region (A3) became larger during both glass tube (13%) and straw (73%) phonations, the change being clearly much larger during

FIGURE 5. Midsagittal images of the vocal tract. Before tube (left), during tube (middle), and after tube phonation (right). The two most dominant changes are the higher velum and the lower laryngeal position in both during and after tube phonations.

FIGURE 6. Midsagittal images of the vocal tract. Before straw (left), during straw (middle), and after straw (right). Likewise as in Figure 2, the two most evident changes are the higher velum and the lower laryngeal position in both during and after straw phonations. Additionally, the hypopharynx is much wider during straw phonation. The straw is not visible due to the fact that it is made of a very thin walled and a very soft material, which is out of the CT scan possibilities.
<table>
<thead>
<tr>
<th>Acoustic Parameters</th>
<th>Before Tube and Straw 1 (dB)</th>
<th>Before Tube and Straw 2 (dB)</th>
<th>Difference Before 1 and 2</th>
<th>After Tube 1 (dB)</th>
<th>After Tube 2 (dB)</th>
<th>Difference After 1 and 2</th>
<th>After Straw 1 (dB)</th>
<th>After Straw 2 (dB)</th>
<th>Difference After 1 and 2</th>
<th>Mean After Tube 1 (dB)</th>
<th>Mean After Straw 1 (dB)</th>
<th>Difference Mean Before/After Tube</th>
<th>Difference Mean Before/After Straw</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPL at 40 cm</td>
<td>75.8</td>
<td>75.9</td>
<td>0.1</td>
<td>76.1</td>
<td>75.9</td>
<td>0.2</td>
<td>76.9</td>
<td>75.3</td>
<td>1.6</td>
<td>75.9</td>
<td>76.0</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Energy SF cluster</td>
<td>27.5</td>
<td>27.1</td>
<td>0.4</td>
<td>27.6</td>
<td>26.6</td>
<td>1.0</td>
<td>30.3</td>
<td>29.4</td>
<td>0.9</td>
<td>27.9</td>
<td>27.1</td>
<td>0.8</td>
<td>1.9</td>
</tr>
<tr>
<td>SPR</td>
<td>19.6</td>
<td>19.7</td>
<td>0.1</td>
<td>17.7</td>
<td>16.7</td>
<td>1.0</td>
<td>14.4</td>
<td>16.0</td>
<td>1.6</td>
<td>19.6</td>
<td>17.2</td>
<td>2.4</td>
<td>4.4</td>
</tr>
<tr>
<td>F1 energy</td>
<td>49.8</td>
<td>50.0</td>
<td>0.2</td>
<td>48.4</td>
<td>47.3</td>
<td>1.1</td>
<td>49.1</td>
<td>48.9</td>
<td>0.2</td>
<td>49.9</td>
<td>47.9</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Formant Frequencies</td>
<td>Before Tube and Straw 1 (Hz)</td>
<td>Before Tube and Straw 2 (Hz)</td>
<td>Difference Before 1 and 2</td>
<td>After Tube 1 (Hz)</td>
<td>After Tube 2 (Hz)</td>
<td>Difference After 1 and 2</td>
<td>After Straw 1 (Hz)</td>
<td>After Straw 2 (Hz)</td>
<td>Difference After 1 and 2</td>
<td>Mean After Tube 1 (Hz)</td>
<td>Mean After Straw 1 (Hz)</td>
<td>Difference Mean Before/After Tube</td>
<td>Difference Mean Before/After Straw</td>
</tr>
<tr>
<td>F1</td>
<td>707</td>
<td>697</td>
<td>10</td>
<td>624</td>
<td>614</td>
<td>10</td>
<td>588</td>
<td>624</td>
<td>36</td>
<td>702</td>
<td>619</td>
<td>606</td>
<td>83.0</td>
</tr>
<tr>
<td>F2</td>
<td>1143</td>
<td>1060</td>
<td>83</td>
<td>1008</td>
<td>1007</td>
<td>1</td>
<td>1010</td>
<td>1008</td>
<td>2</td>
<td>1102</td>
<td>1009</td>
<td>94.0</td>
<td>93.0</td>
</tr>
<tr>
<td>F3</td>
<td>2691</td>
<td>2681</td>
<td>10</td>
<td>2587</td>
<td>2589</td>
<td>2</td>
<td>2608</td>
<td>2598</td>
<td>10</td>
<td>2686</td>
<td>2588</td>
<td>98.0</td>
<td>83.0</td>
</tr>
<tr>
<td>F4</td>
<td>3200</td>
<td>3197</td>
<td>3</td>
<td>2940</td>
<td>2920</td>
<td>20</td>
<td>2962</td>
<td>2961</td>
<td>1</td>
<td>3199</td>
<td>2930</td>
<td>269.0</td>
<td>233.0</td>
</tr>
<tr>
<td>F5</td>
<td>3761</td>
<td>3636</td>
<td>125</td>
<td>3361</td>
<td>3346</td>
<td>15</td>
<td>3332</td>
<td>3355</td>
<td>23</td>
<td>3699</td>
<td>3354</td>
<td>3344</td>
<td></td>
</tr>
<tr>
<td>Formant Distances</td>
<td>Before Tube and Straw 1 (Hz)</td>
<td>Before Tube and Straw 2 (Hz)</td>
<td>Difference Before 1 and 2</td>
<td>After Tube 1 (Hz)</td>
<td>After Tube 2 (Hz)</td>
<td>Difference After 1 and 2</td>
<td>After Straw 1 (Hz)</td>
<td>After Straw 2 (Hz)</td>
<td>Difference After 1 and 2</td>
<td>Mean After Tube 1 (Hz)</td>
<td>Mean After Straw 1 (Hz)</td>
<td>Difference Mean Before/After Tube</td>
<td>Difference Mean Before/After Straw</td>
</tr>
<tr>
<td>F2-F1</td>
<td>436</td>
<td>363</td>
<td>73</td>
<td>384</td>
<td>393</td>
<td>9</td>
<td>422</td>
<td>384</td>
<td>38</td>
<td>399.5</td>
<td>388.5</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>F3-F2</td>
<td>1548</td>
<td>1621</td>
<td>73</td>
<td>1579</td>
<td>1582</td>
<td>3</td>
<td>1598</td>
<td>1590</td>
<td>8</td>
<td>1584.5</td>
<td>1580.5</td>
<td>1594</td>
<td>10.5</td>
</tr>
<tr>
<td>F4-F3</td>
<td>509</td>
<td>516</td>
<td>7</td>
<td>353</td>
<td>331</td>
<td>22</td>
<td>354</td>
<td>363</td>
<td>9</td>
<td>512.5</td>
<td>342</td>
<td>170</td>
<td>154</td>
</tr>
<tr>
<td>F5-F4</td>
<td>561</td>
<td>439</td>
<td>22</td>
<td>421</td>
<td>426</td>
<td>5</td>
<td>370</td>
<td>394</td>
<td>33</td>
<td>500</td>
<td>423.5</td>
<td>382</td>
<td>77</td>
</tr>
<tr>
<td>F5-F3</td>
<td>1070</td>
<td>955</td>
<td>115</td>
<td>774</td>
<td>757</td>
<td>17</td>
<td>724</td>
<td>757</td>
<td>33</td>
<td>1012.5</td>
<td>765.5</td>
<td>740.5</td>
<td>247</td>
</tr>
<tr>
<td>Differences Between Formant Distances</td>
<td>Before/After Tube 1 (Hz and %)</td>
<td>Before/After Tube 2 (Hz and %)</td>
<td>Difference Before/After Tubes 1 and 2</td>
<td>Before/After Straw 1 (Hz and %)</td>
<td>Before/After Straw 2 (Hz and %)</td>
<td>Difference Before/After Straws 1 and 2</td>
<td>Mean Before/After Tube 1 (Hz and %)</td>
<td>Mean Before/After Straw 1 (Hz and %)</td>
<td>Difference Mean Before/After Tube and %</td>
<td>Difference Mean Before/After Straw and %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2-F1</td>
<td>52 (11.9)</td>
<td>-30 (8.3)</td>
<td>22</td>
<td>14 (3.2)</td>
<td>-21 (5.8)</td>
<td>7</td>
<td>11 (11.9)</td>
<td>-3.5 (4.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3-F2</td>
<td>-31 (2.0)</td>
<td>39 (2.4)</td>
<td>8</td>
<td>-50 (3.2)</td>
<td>31 (1.9)</td>
<td>19</td>
<td>4 (2.2)</td>
<td>-9.5 (2.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4-F3</td>
<td>156 (36.6)</td>
<td>185 (35.9)</td>
<td>29</td>
<td>155 (30.5)</td>
<td>153 (29.7)</td>
<td>2</td>
<td>170.5 (36.25)</td>
<td>154 (30.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5-F4</td>
<td>140 (24.9)</td>
<td>13 (2.9)</td>
<td>127</td>
<td>140 (24.9)</td>
<td>13 (2.9)</td>
<td>127</td>
<td>76.5 (13.9)</td>
<td>76.5 (13.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F5-F3</td>
<td>296 (27.7)</td>
<td>198 (20.7)</td>
<td>128</td>
<td>346 (32.3)</td>
<td>198 (20.7)</td>
<td>148</td>
<td>247 (24.2)</td>
<td>272 (26.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
straw. Epilaryngeal region demonstrated small changes after tube (2.4%) and after straw (5%) (Figures 5 and 6). The mean differences between phonation before compared with phonation during and after tube/straw were greater than the differences between repetitions for most of the area measurements. More detailed numerical information about CT distances and area measurements are presented in the Appendix.

Acoustic analysis results

Table 1 shows the results of acoustic analysis of vowel [a:] pre- and posttraining with tube and straw for LTAS and FFT. Some of the LTAS measures (SPL and F1 energy) did not change substantially after tube and straw phonations as compared with phonation before training (less than 1.6 dB; 4%). The major change was observed in SPR, which decreased after both tube (2.5 dB; 12%) and straw (4.4 dB; 22%) phonations, the difference being larger after straw. The energy in the speaker/singer’s formant cluster region increased in average 2.5 dB after straw (Figure 11). The mean differences between phonation before compared with phonation after tube/straw phonation were greater than the differences between repetitions for SPR and the energy in the speaker/singer’s formant cluster region (Table 1). Therefore, one may suggest that the reported variations between before and after tube/straw phonations are true differences.

Formant frequencies values from F1 to F5, distances between F2-F1, F3-F2, F4-F3, F5-F4, and F5-F3 before and after training with tube and straw, and the differences in hertz and percentage of the formant frequency distances are also showed in Table 1. The most evident change related to formant frequencies in the different phonations is found in F1 values. It decreased after both tube and straw by 12% and 14%, respectively. Changes in the rest of the other formant frequencies measured were smaller than the change found in F1 (8% for F2, 4% for F3, 8% for F4, and 9% for F5). It is interesting to note that all formant frequencies of vowel [a:] became lower after both tube and straw training. From Table 1, it is possible to point out that the clearest change was found in F4-F3 distance difference. F3 and F4 were closer to each other after exercising with both tube and straw. The difference was in average 170 Hz (36.2%) and 154 Hz (30.1%) after tube and straw, respectively. This formant cluster between F3 and F4 was located in 2588–2930 Hz and 2603–2966 Hz after tube and straw, respectively.

![FIGURE 7](image7.png)

**FIGURE 7.** Mean of the two repeated scans for the cross-sectional areas (mm$^2$) calculated from the midsagittal and transversal images of the vocal tract obtained from the CT measurements performed before, during, and after tube and straw phonations. For definitions of the areas, Ae, Ap, A1, A2, and A3, recall Figures 2 and 3. BT, before tube; DT, during tube; AT, after tube; DS, during straw; AS, after straw.

![FIGURE 8](image8.png)

**FIGURE 8.** Transversal images of the vocal tract. Before tube/straw (left), during tube (middle), and during straw (right). Ap increased during both tube and straw phonations compared with vowel phonation. This increment was clearly larger during straw phonation than during glass tube phonation.
The average $F_0$ measured throughout the sequence was 113 Hz with a range of 111–115 Hz.

**Auditory-perceptual analysis**

Table 2 displays the results of the listening test. Data show that there is a high degree of agreement among listeners for straw phonation samples. All the samples produced after straw exercises were evaluated by the four judges as better in voice quality than the samples before. Assessment for tube phonation did not show the same degree of agreement found in straw phonation recordings.

**Air pressures and CQ**

Subglottic pressure, oral pressure, transglottal pressure, and CQ values before, during, and after tube and straw phonations are presented in Table 3. There was an increase of subglottic pressure during straw phonation (18%), and this increment remained after removing the straw (7%). Subglottic pressure value did not demonstrate a significant change during tube phonation and became higher after tube phonation (6%). There was also an increase of oral pressure during both tube and straw phonations. Transglottal pressure decreased during both tube and straw phonations. Oral pressure increased 17 times more during a straw than a tube phonation, and transglottal pressure decreased about two times more during a straw phonation.

Glottal CQ decreased during both tube and straw phonations 4.5% and 20%, respectively. This change remained both after tube (8.3%) and after straw phonations (9.2%). Figure 12 shows representative EGG waveforms from the samples before, during, and after straw phonation.

**DISCUSSION**

One of the major differences observed between samples before and both during tube and during straw phonations is the increase in the vertical length of the vocal tract, which reflects a lower VLP and also a larger overall vocal tract length. This change remained during vowel production after tube and straw phonations. Findings from the acoustic analysis are concordant with the CT results. FFT showed a decrease in the frequency of the first five formants. According to the acoustic theory of speech, the formant frequencies depend on the length of the vocal tract and the cross-sectional shape of the vocal tract as a function of its length. Vocal tract length determines the average location of formant frequencies. In this regard, as length increases, the value of the formant frequency will decrease. In a previous study performed with MRI, similar results were observed in a male subject who lowered the larynx after vocal warm-up with spoken exercises. On the other hand, in two MRI studies performed with a female participant, no changes were reported in the VLP neither after straw phonation nor after other semi-occluded vocal tract exercises ([ß:], [m:], [y:], and [u:]). It is important to point out that both the male subject from the present study and the male subject from the earlier MRI investigation have classical singing voice technique, whereas the female participant has a long experience in speaking voice training. Because the classical singing is produced with a “covered” sound which implies a longer vocal tract due to a lower VLP and lip protrusion, it is possible to question whether the lowered VLP demonstrated by our subject is due to the effect of tube phonation or not.
An important aspect that may be regarded is that the tube or straw may make the perception of vibratory sensations stronger and incite the subjects to change the vocal tract and glottal setting in such a way that it is prone to intensify and pertain these sensations. Therefore, the adjustments could be modified according to the previously learned voice patterns, which are regulated by the voice use type (speech or singing or different styles of singing). According to Titze, the origin of vibratory sensations in resonant voice could be due to the efficiency on the energy conversion process at the glottis (from aerodynamic to acoustic energy). When this process is efficient, the vibrations are distributed all over the face and head areas. On the other hand, when the energy conversion process is poor, the vibrations will remain mostly in the laryngeal area.15

In earlier investigations performed with other assessment methods, there is evidence for both the effects, laryngeal lowering and laryngeal raising compared with the resting laryngeal position. Some studies have evidenced a lower VLP during semi-occluded tasks,40,44 whereas others have reported the opposite.23,34 Because people diagnosed with hyperfunctional dysphonia commonly present a high VLP due to the abnormal contraction of suprahyoid muscles,46 semi-occlusions and lengthening of the vocal tract might have an important therapeutic effect if they really produce a laryngeal lowering. In this regard, according to Sovijärvi et al.,47 the positive outcomes of the resonance tube method are due to the efficient lowering of the larynx and the firming of the vibration of the vocal folds. The author pointed out that the length of the tube should be chosen according to the lowering of the larynx during the tube phonation.1–48

Later also Simberg and Laine5 stated that to choose the tube that best enhances the lowering of the larynx during the phonation is an important factor. Both Sovijärvi and Simberg and Laine used tubes submerged into the water (water resistance therapy). It is important to mention that tube phonation has not only been used in cases of hyperfunction as mentioned but also in patients with hypofunctional dysphonia.5,47 As it comes to the amount of laryngeal lowering, Sovijärvi49 remarked that it may just be some millimeters or rather avoidance of raising the larynx, so the aim of the method was not to cause a similar laryngeal lowering as in classical singing.

The velum position was another important vocal tract modification in the present study. It rose to seal the nasopharyngeal port during and after the tube and straw phonations. Our results are concordant with previous investigations carried out with MRI and CT.37–39 A suitable explanation for this change could be the increased oral pressure produced during semi-occluded vocal tract postures.3,19 The oral pressure measured in the present study during both tube and straw phonations was higher than during vowel production.

An improved energy transfer by decreasing the damping caused by the nasal tract and hence an increase in the total SPL may be expected with a proper velar closure.35 Nevertheless, our findings did not show an SPL increment after tube/straw phonation. This lack of change is probably because the participant was required to produce the same loudness level during vowel production before and after exercising to avoid spectral changes merely caused by intensity increments. Related to this, the energy of F1 did not show important changes.

### TABLE 2.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Before Tube</th>
<th>After Tube</th>
<th>No Difference</th>
<th>Before Straw</th>
<th>After Straw</th>
<th>No Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Sample 2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

The numbers indicate the amount of listeners regarding which sample in each pair was produced with a better voice quality or if there was no difference between them. Listeners n = 3 in total.
during the sequence either. This is not surprising because the spectral energy of F1 region supports most of the overall SPL. Because velum elevation has been found, through MRI examination, to be a common modification in classical singers during singing,\textsuperscript{50–53} it could be argued that the higher velum position in the present study may be due to the classical vocal background of our participant. Nevertheless, previous radiological studies have demonstrated the same modification when using artificial vocal tract lengthening with a tube in subjects who have an extensive experience teaching speaking voice technique.\textsuperscript{37,38} Therefore, phonation into a narrow tube may be a useful tool for the treatment of hypernasality as previously reported for using a narrow tube in air or a resonance tube in water.\textsuperscript{4,54}

CT transversal images demonstrated a clear tendency that the ratio Ap/Ae increases during and after both the tube and straw phonations. Ap/Ae has been suggested to be an important factor for the singer’s formant cluster production (a prominent spectrum envelope peak near 3 kHz associated with the “ringing” voice quality). Sundberg\textsuperscript{55} suggested that when the cross-sectional area in the pharynx is at least six times wider than that of the laryngeal tube opening, the epilaryngeal tube is acoustically unlinked from the rest of the vocal tract acting as a separate resonator. Therefore, an extra formant would be added to the vocal tract transfer function. If additionally the sinus of Morgagni is wide, this extra formant would be tuned between the third and fourth formants. Furthermore, if the sinus piriformis are wide, the fifth formant would lower down to about 3 kHz.\textsuperscript{55} These three things can be reached by lowering of the larynx.

In the present study, the greatest Ap/Ae ratio was observed after straw phonation (5.5), which is close to the value suggested by Sundberg (6).\textsuperscript{55} Interestingly, acoustic results of the present study showed the largest increase in the energy of the speaker/singer’s formant region after straw. Additionally, the SPR showed the largest decrease after straw phonation as well. Recall that SPR is a spectral measurement created originally for quantifying the singer’s formant.\textsuperscript{56} Therefore, if the singer/speaker’s formant region demonstrates more energy after exercising, it is expected that SPR value will be lower. Thus, it is possible to state that the outcomes obtained from CT are in line with acoustic analysis results. These data suggest a clear immediate effect on spectral characteristics after straw phonation, specifically an increased spectral prominence in the singer’s/speaker’s formant region and also a change in the spectral slope declination (ie, less steep slope). Values of SPR after straw represent an increased energy in the higher harmonics of the spectrum compared with the lower ones. The increase of SPL in the singer/speaker’s formant region has also been reported after tube phonation in earlier works.\textsuperscript{37–39} Furthermore, two studies\textsuperscript{30,31} whose aim was to compare the effect on spectral energy distribution of semi-occluded vocal exercises and open vowel exercising demonstrated that sustained vowel production after phonation into stirring straws

<table>
<thead>
<tr>
<th>TABLE 3. Subglottic Pressure, Oral Pressure, Transglottal Pressure, and CQ Before, During, and After Tube and Straw Phonations</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQ (%)</td>
</tr>
<tr>
<td>Before tube and straw 1</td>
</tr>
<tr>
<td>Before tube and straw 2</td>
</tr>
<tr>
<td>During tube 1</td>
</tr>
<tr>
<td>During tube 2</td>
</tr>
<tr>
<td>After tube 1</td>
</tr>
<tr>
<td>After tube 2</td>
</tr>
<tr>
<td>During straw 1</td>
</tr>
<tr>
<td>During straw 2</td>
</tr>
<tr>
<td>After straw 1</td>
</tr>
<tr>
<td>After straw 2</td>
</tr>
<tr>
<td>Mean before tube and straw</td>
</tr>
<tr>
<td>Mean during tube</td>
</tr>
<tr>
<td>Mean after tube</td>
</tr>
<tr>
<td>Mean during straw</td>
</tr>
<tr>
<td>Mean after straw</td>
</tr>
<tr>
<td>Mean before tube and straw</td>
</tr>
<tr>
<td>Mean during tube</td>
</tr>
<tr>
<td>Mean after tube</td>
</tr>
<tr>
<td>Mean during straw</td>
</tr>
<tr>
<td>Mean after straw</td>
</tr>
</tbody>
</table>

FIGURE 12. Representative EGG waveforms from the samples before (top), during (middle), and after (bottom) straw phonation. Horizontal axis is time and vertical is impedance (increasing downward).
and vocal function exercises (voice program which involves semi-occlusions) produced more spectral energy increase in the higher part of the spectrum (2000–5000 Hz) than vocalizations with open vocal tract setting. The acoustic differences between tube and straw phonations found in the present study (more increase of SPL in the singer/speaker’s formant after straw than tube) seem to be in line with the results of the listening test. All samples after straw phonation were rated as representing a better voice quality by all the listeners. Assessment for tube phonation did not show the same degree of agreement as found in straw phonation recordings.

As mentioned above, the singer/speaker’s formant, in general terms, can mainly be explained as a resonatory phenomenon arising from a clustering of third, fourth, and fifth formants. According to our findings, there was a clear cluster of F3 and F4 after exercising with both a glass tube and a straw. This formant cluster of F3 and F4 was located in 2500–3000 Hz for vowel [a:] after exercising. F5 also showed a decrease, and the frequency difference between F5 and F4 was smaller after tube/straw phonation as compared with vowel production before them. However, this change was not as clear as the change in F4-F3. A significant decrease in the formant frequency distance after straw phonation was also found by Laukkanen et al. This formant cluster may have contributed to the changes mentioned above (lower SPR and higher SPL between 2500 and 4000 Hz).

According to Sundberg, the larynx lowering, typical of male classical singing, seems to be a way to obtain a high ratio between the cross-sectional area of the low pharynx and the epilaryngeal tube opening and drawing the higher formants F3-F5 closer to each other. Therefore, the lowered VLP observed during and after both tube and straw phonations in the present study might be the physiological cause of the strengthening of the singer’s formant cluster range showed by our participant. However, if a low VLT is desirable to produce a singer’s formant, how can it be explained that without a laryngeal lowering it is also possible to observe a formant cluster in vocally trained speakers? There are some differences between the singer’s and speaker’s formant cluster. The former is usually located between 2 and 3 kHz, whereas the latter is produced between 3 and 4 kHz. The lower frequency of this peak in singers may be related to lowering of the larynx. On the other hand, the VLP change does not seem to be the main cause of the speaker’s formant. Previous studies suggested that a speaker’s formant could be obtained through a slight narrowing of the epilaryngeal region, widening of the back of the mouth cavity, and narrowing of the front part of it. The modeling results by Leino et al. suggest that these two modifications may be responsible for the peak at 3.5 kHz showed in male actors. In addition, there is empirical evidence that vocal tract changes other than laryngeal lowering are able to produce a spectral prominence and a brilliant singing voice quality.

Regarding the increase of SPL in the singer/speaker formant region, it is important to point out that not only a formant clustering could be the explanation for this change but it could also be explained by increased input impedance of the vocal tract, for example, due to the higher Ap/Ae ratio. Higher input impedance produces a faster cessation of the glottal flow and consequently a less tilting spectrum and a more resonant voice quality (as found in the present study after straw) and easy voice production. Therefore, both a resonance strategy (formant cluster, possibly only due to changes in the vocal tract length and thus formant frequencies) and increased aerodynamic-acoustic interaction of the vocal tract might contribute to a louder voice without increasing the vocal effort. These resonance and acoustic-aerodynamic changes are desirable effects for both vocal warm-up and voice training. A change in the lower vocal tract may simultaneously cause these both. In addition, an increase of subglottic pressure after straw phonation (as found in the present study) would also be a possible explanation for somewhat higher SPL in the singer/speaker’s formant region. An SPL rise of 1 dB is typically associated with a 1.5 dB difference near 3000 Hz in the LTAS. However, in the present study, SPL was controlled to be the same before and after exercising, and there was an increase of just 0–1.1 dB between the samples before and after exercising, whereas the energy of the singer’s formant cluster increased up to 2.8 dB after the straw.

The CQ values decreased during both tube and straw phonations and remained lower after them as compared with vowel phonation before them. This change was larger for straw than tube phonation. Several studies have reported a change in CQ when semi-occlusion is compared with vowel phonation. However, only some of them have demonstrated a decrease. There is no clear pattern regarding the effect of semi-occlusion or lengthening of the vocal tract on CQ. From the theoretical point of view, an important effect of straw phonation, voiced bilabial fricative and also tube phonation when the tube is submerged into water is the resistance against airflow. This will produce an increased supraglottal pressure, which consequently will cause a decrease of the transglottal air pressure (the force driving vocal fold vibration) as found in our data during both tube and straw and an elevation of the intraglottal pressure if not compensated by increasing subglottic pressure and addition. Increased intraglottal air pressure, in turn, would produce a lower CQ. Titze stated that a semi-occluded setting (especially during phonation into a thin straw) may produce a slight separation of the vocal folds and hence a smaller CQ. If this really occurs, it is possible to assume that occlusions of the vocal tract might be beneficial to decrease the vocal fold impact stress. When the vocal folds impact stress increases (due to more intense vocal folds collision), the CQ tends to rise and vice versa. However, it is not a strictly linear phenomenon.

Another possible explanation for the decreased CQ during and after tube/straw phonation could be the low VLP. Earlier studies have shown that a low VLP has also been correlated to low glottal CQ. Therefore, it seems to be that lower VLP is associated with less glottal adduction. This phonatory effect might be due to an abductive component of tracheal pull when the larynx is low.

During straw phonation, not only the lowest CQ but also the highest subglottic pressure across the three stages of the sequence (before, during, and after) was found. Increased subglottic pressure has also been reported in a previous study.
carried out with stirring straw phonation. Earlier research has shown that impact stress increases with lung pressure, that is, the impact stress is higher at higher lung pressure values. However, in the present study, a high subglottic pressure was observed seemingly without increment of impact stress. This assumption is based on the fact that low CQ was obtained during straw. What drives the vocal folds is the transglottal pressure, that is, the difference between the subglottal and the supraglottal pressure. As mentioned above, a downstream occlusion increases the supraglottal air pressure and reduces the transglottal air pressure, which in turn reduces the additive force. Occlusion of the vocal tract makes it possible to lower the transglottal pressure (which is desirable) even if the subglottal pressure is high. In this regard, Titze et al mentioned that with narrow straws, high subglottal pressure can be reached without a high risk of vocal fold impact stress.

Interesting modifications were also observed in the pharyngeal region. First, the Ap area increased during both straw and tube phonations compared with vowel phonation before them (Figure 8). Second, and related to the first one, anterior-posterior width of the hypopharynx became wider during straw phonation (Figure 6). Moreover, the pharyngeal region area (A2) showed an increase during both tube and straw phonations as compared with vowel production before exercising (Figures 5 and 6), and finally, the epilaryngeal region area (A3) became larger during both glass tube and straw phonations (Figures 5 and 6). All these modifications were larger during straw than tube phonation. An important therapeutic application could arise from these observations: the widening of the pharyngeal area in people with vocal hyperfunction. Vocal hyperfunction is considered as excessive and/or imbalanced use of muscular forces, which is characterized by excessive laryngeal and extralaryngeal tension. One of the common features treated by voice therapists in patients diagnosed with excessive muscle tension is the relaxation of the laryngo-pharyngeal muscles and widening of the vocal tract, especially in the pharyngeal area. The relaxation of this area is also an important goal of classical singing and speaking voice pedagogy. Different exercises to promote an open throat sensation have been one of the most used tools to produce freedom or lack of excessive tension in the area of the throat and a better tone quality in both normal and pathologic voices. From our finding, it is possible to state that straw phonation might be an option to produce the same effect. Results from other CT and MRI studies support our findings related to the modification of the pharyngeal area with tube or straw exercises. The increased oral pressure during straw may push the pharyngeal tissues and maybe it also then has the possibility to aid muscle relaxation in the pharyngeal area.

Different from earlier studies, the present investigation did not show neither an increase in jaw opening nor an increase in the lip opening after tube/straw phonation. The previously cited reports were obtained from a female subject who has mainly speech training experience, whereas our participant owns a classical singing voice background. This fact could be a suitable explanation for the discrepancy in the results. Because jaw opening and lip opening seem to be directly related in classical singing technique, the lack of changes in both together (jaw and lip opening) in the present study appears reasonable. Nevertheless, this relation (the fact that lip and jaw may go together) should not be taken as a rule for all cases. In earlier studies using tube phonation, it was possible to see lowering of the jaw, which was not directly related to lip opening. A similar movement relationship has been also observed between jaw opening and tongue dorsum height. They vary independently. Likewise, our outcomes showed negligible changes for tongue dorsum height after both tube and straw phonations. Moreover, no important changes were observed after tube and straw phonations in A1 area (oral cavity). These results differ from previous MRI and CT studies where oral cavity was found larger after exercises with straw and tube, respectively. Last, but not least, midsagittal images also showed a more frontal tongue position during and after tube/straw phonations, which is in line with earlier studies. This would imply that the pharyngeal constriction was reduced; however, this implication did not occur in all CT samples in the present study. The change in the tongue configuration could also have affected some vowel formants, especially F2. A more frontal tongue position is supposed to raise F2; however, this acoustic change was not observed in our data.

CONCLUSION

Our results suggest that both straw and tube phonations may have vocal training and vocal warm-up effects, one of the major being an increased spectral prominence in the singer/speaker’s formant cluster region. This may be explained by an increase of the ratio between the cross-sectional area of the low pharynx and the area of the epilaryngeal tube opening. The lower laryngeal position observed in the CT images could be a reasonable explanation for this finding. Moreover, listening test demonstrated a better voice quality after straw/tube than before exercising, hence, this finding seem to be in line with the acoustic and anatomic results. Therapeutic effects could also arise from our data such as a better velar closure (as treatment for hypernasality), a widening of the pharyngeal region (as treatment for patients with laryngeal and extralaryngeal muscle hypertension), and possibly a laryngeal lowering (as an exercise for subjects with high larynx position due to vocal hyperfunction). Some of these therapeutic effects have been described earlier. Because the changes in the vocal tract, CQEGG, subglottic pressure, acoustic parameters, and auditory-perceptual findings were greater during and after exercising with the straw than with the tube, it is reasonable to state that more prominent changes are obtained when the vocal tract input impedance during exercises is higher. Considering all the observed effects in the present study, it would be possible to state that vocal exercises with increased vocal tract impedance assist in increasing vocal efficiency and economy (more loudness without an increase of vocal loading due to
increased vocal fold collision). In the present study, only immediate results after exercising were obtained, long-term effects still remain to be studied.

REFERENCES

43. Miller DG, Schutte HK. Toward a definition of male ‘head’ register, passag. 157–170.
44. Miller DG, Schutte HK. Toward a definition of male ‘head’ register, passag. 157–170.


Spieß G. *Methodische Behandlung der nervösen Aphonie und einiger anderer Stimmstrungen.* *Arch Laryngol Rhinol.* 1899;9:368–376. [in German].

## APPENDIX

### CT Distances (mm) and Cross-Sectional Areas (mm²)

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Before Tube 1</th>
<th>Before Tube 2</th>
<th>During Tube 1</th>
<th>During Tube 2</th>
<th>After Tube 1</th>
<th>After Tube 2</th>
<th>During Straw 1</th>
<th>During Straw 2</th>
<th>After Straw 1</th>
<th>After Straw 2</th>
<th>Mean Before</th>
<th>Mean During Tube</th>
<th>Mean After Tube</th>
<th>Mean During Straw</th>
<th>Mean After Straw</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical length</td>
<td>88</td>
<td>87.9</td>
<td>93</td>
<td>95</td>
<td>95</td>
<td>93</td>
<td>106</td>
<td>107</td>
<td>95</td>
<td>96</td>
<td>87.95</td>
<td>94</td>
<td>94</td>
<td>106.5</td>
<td>95.5</td>
</tr>
<tr>
<td>Horizontal length</td>
<td>91.9</td>
<td>90.6</td>
<td>90.9</td>
<td>90.8</td>
<td>91.9</td>
<td>91.6</td>
<td>91.25</td>
<td>90.85</td>
<td>91.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lip opening</td>
<td>9.7</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>9.6</td>
<td>9.6</td>
<td>9.5</td>
<td>9.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaw opening</td>
<td>76.1</td>
<td>83.1</td>
<td>65.9</td>
<td>68.5</td>
<td>75.8</td>
<td>76.6</td>
<td>68.7</td>
<td>69</td>
<td>78.7</td>
<td>79.5</td>
<td>79.6</td>
<td>79.6</td>
<td>79.6</td>
<td>76.2</td>
<td>75.5</td>
</tr>
<tr>
<td>Tongue dorsum height</td>
<td>64.7</td>
<td>64.2</td>
<td>61.8</td>
<td>62.1</td>
<td>65.5</td>
<td>66</td>
<td>57.3</td>
<td>55.5</td>
<td>68.4</td>
<td>68.9</td>
<td>64.45</td>
<td>61.95</td>
<td>65.75</td>
<td>56.4</td>
<td>68.65</td>
</tr>
<tr>
<td>Oropharynx width</td>
<td>8.8</td>
<td>8.6</td>
<td>8</td>
<td>8.1</td>
<td>7</td>
<td>6.3</td>
<td>6</td>
<td>6.5</td>
<td>5.2</td>
<td>6.1</td>
<td>8.7</td>
<td>8.05</td>
<td>6.65</td>
<td>6.25</td>
<td>5.65</td>
</tr>
<tr>
<td>Velum elevation</td>
<td>21.6</td>
<td>19.5</td>
<td>13.9</td>
<td>13.4</td>
<td>16.8</td>
<td>17.4</td>
<td>13.7</td>
<td>13.4</td>
<td>15.8</td>
<td>16.6</td>
<td>20.55</td>
<td>13.55</td>
<td>17.1</td>
<td>13.55</td>
<td>16.2</td>
</tr>
<tr>
<td>Hypopharynx width</td>
<td>18.7</td>
<td>18.2</td>
<td>19.5</td>
<td>19.5</td>
<td>19</td>
<td>18.7</td>
<td>25.8</td>
<td>19.5</td>
<td>18.7</td>
<td>18.5</td>
<td>18.9</td>
<td>26</td>
<td>19.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (mm²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outlet of the epilaryngeal tube (Ae)</td>
<td>114</td>
<td>109</td>
<td>110</td>
<td>117</td>
<td>95</td>
<td>91</td>
<td>187</td>
<td>181</td>
<td>84</td>
<td>87</td>
<td>111.5</td>
<td>113.5</td>
<td>93</td>
<td>184</td>
<td>85.5</td>
</tr>
<tr>
<td>Inlet to the lower pharynx (Ap)</td>
<td>460</td>
<td>444</td>
<td>495</td>
<td>497</td>
<td>498</td>
<td>404</td>
<td>881</td>
<td>849</td>
<td>471</td>
<td>465</td>
<td>452</td>
<td>496</td>
<td>451</td>
<td>865</td>
<td>468</td>
</tr>
<tr>
<td>Oral cavity (A1)</td>
<td>1577</td>
<td>1489</td>
<td>958</td>
<td>964</td>
<td>1478</td>
<td>1353</td>
<td>1083</td>
<td>1087</td>
<td>1367</td>
<td>1393</td>
<td>1533</td>
<td>961</td>
<td>1415.5</td>
<td>1085</td>
<td>1380</td>
</tr>
<tr>
<td>Pharyngeal region (A2)</td>
<td>429</td>
<td>444</td>
<td>522</td>
<td>489</td>
<td>377</td>
<td>313</td>
<td>416</td>
<td>594</td>
<td>312</td>
<td>317</td>
<td>436.5</td>
<td>505.5</td>
<td>345</td>
<td>505</td>
<td>314.5</td>
</tr>
<tr>
<td>Epilaryngeal region (A3)</td>
<td>547</td>
<td>533</td>
<td>608</td>
<td>616</td>
<td>551</td>
<td>555</td>
<td>933</td>
<td>938</td>
<td>547</td>
<td>587</td>
<td>540</td>
<td>612</td>
<td>553</td>
<td>935.5</td>
<td>567</td>
</tr>
</tbody>
</table>
The influence of water resistance therapy on vocal fold vibration: a high-speed digital imaging study

Marco Guzman\textsuperscript{a,b}, Anne-Maria Laukkanen\textsuperscript{c}, Louisa Traser\textsuperscript{d}, Ahmed Geneid\textsuperscript{e}, Bernhard Richter\textsuperscript{d}, Daniel Muñoz\textsuperscript{f} and Matthias Echternach\textsuperscript{d}

\textsuperscript{a}Department of Communication Sciences and Disorders, University of Chile, Santiago, Chile; \textsuperscript{b}Department of Otolaryngology, Las Condes Clinic, Santiago, Chile; \textsuperscript{c}Speech and Voice Research Laboratory, School of Education, University of Tampere, Tampere, Finland; \textsuperscript{d}Institute of Musicians’ Medicine, Freiburg University Medical Center, Freiburg, Germany; \textsuperscript{e}Department of Otorhinolaryngology and Phoniatrics—Head and Neck Surgery, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland; \textsuperscript{f}Department of Otolaryngology, University of Chile, Santiago, Chile

\textbf{ABSTRACT}

\textbf{Purpose:} This study investigated the influence of tube phonation into water on vocal fold vibration.

\textbf{Method:} Eight participants were analyzed via high-speed digital imaging while phonating into a silicon tube with the free end submerged into water. Two test sequences were studied: (1) phonation pre, during, and post tube submerged 5 cm into water; and (2) phonation into tube submerged 5 cm, 10 cm, and 18 cm into water. Several glottal area parameters were calculated using phonovibrograms.

\textbf{Results:} The results showed individual differences. However, certain trends were possible to identify based on similar results found for the majority of participants. Amplitude-to-length ratio, harmonic-to-noise ratio, and spectral flatness (derived from glottal area) decreased for all tube immersion depths, while glottal closing quotient increased for 10 cm immersion and contact quotient for 18 cm immersion. Closed quotient decreased during phonation into the tube at 5 cm depth, and jitter decreased during and after it.

\textbf{Conclusion:} Results suggest that the depth of tube submersion appears to have an effect on phonation. Shallow immersion seems to promote smoother and more stable phonation, while deeper immersion may involve increased respiratory and glottal effort to compensate for the increased supraglottal resistance. This disparity, which is dependent upon the degree of flow resistance, should be considered when choosing treatment exercises for patients with various diagnoses, namely hyperfunctional or hypofunctional dysphonia.

\textbf{Introduction}

Water resistance therapy includes phonation of a sustained vowel sound into a tube with the distal end submerged in water. The therapeutic process consists of several steps occurring during sessions throughout a period of weeks. At the beginning of water resistance therapy, the patient uses a limited pitch range for the first week(s) of training (1). Gradually, the patient starts to use a more varied intonation such as glides and simple intervals in a glissando mode. The patient is asked to keep the phonation stable and to follow a normal and comfortable breathing pattern in all exercises. Optimal body posture and breath control are also important aspects in water resistance therapy (1).

Two main versions of water resistance therapy have been used: (1) Phonation into a traditional Finnish resonance tube (made of glass, 24–28 cm in length with an 8–9 mm inner diameter) that is submerged in a bowl of water (2); and (2) The Lax Vox technique, which involves phonation into a flexible silicone tube (35 cm in length with an inner diameter of 9–12 mm) that is submerged into a water-filled bottle (3). In both versions of water resistance therapy, the tubes are kept approximately 1 mm between the teeth, with the lips rounded so that no air leaks from the mouth (1). In the glass tube method, tubes are kept 1–2 cm or 5–15 cm below the surface of the water, depending on the patient’s needs. With the Lax Vox method, the participant is instructed to keep the tube 1–7 cm in water for voice therapy (3). In both techniques, the patient is asked to feel vibratory sensations in the anterior areas of the facial tissues. According to Titze (4), vibration sensations on these areas are related to the efficiency of the energy conversion process at the glottis.

Although water resistance therapy requires regular practice for at least several weeks, most relevant studies have been performed using tube phonation for a short period of time (seconds or minutes). Some of these studies have been designed to investigate changes in the glottal function or aerodynamic measures of phonation (5,6), while others have looked for changes in the vocal tract configuration (7).

The possible effect of resonance tube phonation in water on phonation threshold pressure and collision threshold pressure (CTP) was studied by Enflo et al. (6). CTP was found to be higher and the voice quality was perceived to be
better after the exercise. The perceptual changes were more prominent in singers who did not practice singing daily and in singers who had less experience in singing in general (6).

According to the recent *in vivo* measurements by Radolf et al. (8) the mean oral pressure increased about 4 times in habitual comfortable phonation and about 9 times in soft phonation, when the subject was phonating into a resonance tube inserted 10 cm in water. The subglottic pressure doubled in normal phonation and quadrupled in soft phonation, thus the subject compensated for the increase in supralaryngeal resistance. Fundamental frequency (F0) decreased 11–15 Hz. Comparable results were obtained by Horáček et al. (5) using a physical model of voice production. Flow rate was set constant in modeling. Oral pressure increased 10 times with a resonance tube 10 cm in water, and subglottic pressure had to be increased 1.4 times to keep the flow rate constant. Flow resistance increased 1.5 times. F0 remained constant. Larger changes were observed for soft phonation, and F0 decreased 38 Hz (19%).

Bubbles produced during phonation through a tube in water generate a pulsating oral pressure at a frequency of 15–40 Hz (5,8). Therefore, phonation into water may cause a sensation of massage on the laryngeal and pharyngeal tissues. A massage-like effect with reduction of muscle hypertension could be desirable in patients with voice disorders, especially in subjects diagnosed with laryngeal and pharyngeal hyperfunctionality. It has been hypothesized that a massage-like effect during phonation into water could also increase blood flow in the vocal folds. This assumption is supported by evidence from the field of sports medicine indicating that massage increases blood flow in muscles and skin (9–11). Other semi-occluded exercises such as tongue and lip trills may have a comparable effect due to the oscillation of oral pressure produced by tongue and lip vibration (12,13). Even though the massage-like sensation of bubbling via tube phonation into water has been clinically reported, to date there are no data supporting the hypothesis of a massage-like effect on the vocal folds or vocal tract.

Enflo et al. (6) stated that during phonation in the resonance tube in water the water bubbles generate oscillations of oral pressure, EGG, and audio signals. Specifically, oscillation of values of oral pressure modifies the transglottal pressure (which drives the vocal folds) and this, in turn, produces changes in the EGG signal amplitude. The results of the *in vivo* study by Radolf et al. (8) showed about 2–4 times higher peak-to-peak variation in oral pressure with the tube immersed 10 cm in water, compared to phonation on [u:] (larger increase for soft phonation). Contact quotient, measured from the EGG signal, increased 11% in habitual phonation. According to the modeling experiments by Horáček et al. (5) phonation at conversational loudness resulted in 2.2 times larger peak-to-peak pressure variation and 1.6 times larger glottal amplitude variation with the tube submerged 10 cm in water, compared to phonation on [u:]. Thus, the modulation of oral pressure during bubbling (tube phonation into water) seems to have an effect on vocal fold oscillation and possibly on vocal fold tissues. A recent high-speed imaging study, designed to investigate tube phonation, reported modulations of vocal fold vibration and EGG signal due to back pressure when the tube was held in water. Increased mean value of open quotient with increasing water depth was also reported (14).

Earlier investigations support the association between EGG contact quotient and the degree of vocal fold impact stress. When impact stress increases (stronger collision between the vocal folds during vibration), the vocal folds also tend to stay together for longer intervals, thus increasing the value of contact quotient or decreasing the value of open quotient (15).

Guzman et al. (7) evaluated via flexible laryngoscopy the effect of eight different semi-occluded vocal tract postures on vertical laryngeal position (VLP), pharyngeal constriction, and laryngeal compression in subjects diagnosed with hyperfunctional dysphonia. All semi-occluded techniques produced a lower VLP, narrower aryepiglottic opening, and a wider pharynx than in a resting position. More prominent changes were obtained with tube phonation into water 3 cm and 10 cm deep compared to the other semi-occluded exercises that did not involve water. Sovijärvi et al. (16) assumed that the positive outcomes of phonation into a resonance tube in water are due to the efficient lowering of the larynx and an improved vocal fold closure. Sovijärvi stated that the length of the tube and depth into the water should be chosen so that a clear lowering of the larynx would occur during the exercise. The goal after the exercise, however, is normal voicing with a neutral (not lowered) larynx (17–19).

Water resistance exercising precipitates changes in self-assessment of voice. Paes et al. (20) studied the immediate effects of it on teachers with behavioral dysphonia. The Finnish resonance tube immersed 2 cm into water was used. Significantly greater phonatory comfort and improved perceptual voice quality after the exercises were reported by the subjects. Less spectral noise in the acoustic signal and lower fundamental frequency were also observed (20).

Only two longitudinal studies have been conducted using the water resistance therapy. Positive results have been obtained in the treatment of behavioral dysphonia (21,22). Perceptual assessment and results from a questionnaire on the occurrence of vocal symptoms revealed significant changes in the treatment group compared with the control group (21).

Several visualization techniques of vocal fold vibration have been used in the diagnosis of laryngeal disorders and in the investigation of normal voice production. High-speed imaging (HSI) has emerged as an effective method of visualization during the last decades (22). Commercial HSI systems record images of sustained vocal fold vibration at the rate of 2,000 to 10,000 frames per second, which is fast enough to capture the actual phonatory vibrations of the vocal folds (23,24). For research, however, it was shown that using flexible endoscopy high-speed recordings are possible with a frame rate up to 20,000 frames per second, which allows more detailed analysis (24). To obtain precise and objective information from HSI videos, several methods have been developed. Phonovibrography is a fast, robust, and precise image-processing strategy to extract information of the vocal fold movements during phonation (25). The method detects the vocal fold edges and transfers their movement...
into a static geometric presentation (phonovibrograms). Several glottal variables can be calculated from phonovibrograms (26).

The present study aimed to observe the influence of tube phonation into water on vocal fold vibration by using high-speed digital imaging and phonovibrography. Specifically, we attempted to answer two questions: (1) Is there any influence of tube phonation into water on vocal fold vibration; and (2) Does immersion depth affect vocal fold vibration? We hypothesize that: (1) phonation into water causes modulation in glottal variables; (2) the modulations imply a decrease in the average values of closed quotient, amplitude-to-length ratio, perturbation, noise-to-harmonic ratio, and spectral flatness; (3) the average parameter change observed during water resistance exercising remains in vowel phonation afterwards; (4) the effect is stronger when the immersion is deeper than 5 cm; and (5) deeper than 10 cm immersion in water would promote more glottal compensation for increased airflow resistance (more adducted focal folds and increase of closed quotient), and therefore higher impact stress than shallower immersion.

Methods
Participants and phonatory tasks
Eight volunteer participants (five male and three female) were analyzed using high-speed digital imaging. All participants reported normal voice and hearing at the time of the experiment. The health of the larynx was ascertained through laryngoscopy. Four participants were trained singers; one participant reported experience in speech training, and three had no voice training. The age range of the participants was 23–45 years.

Two test conditions were studied:

1. The sequence for the first condition was: phonation pre, during, and post tube submerged 5 cm into water. In this sequence, subjects were asked to produce a sustained and stable vowel [i:] before and after phonation into the tube. Each individual phonation sample was produced for approximately 5 seconds. Successive sequences of tube phonation were produced for 5 minutes to gain a possible training effect. A flexible silicone (Lax Vox-like) tube (45 cm in length, 2 cm in inner diameter) was used. Participants were asked to phonate at a comfortable pitch and vocal loudness in all three tasks.

2. The sequence for the second condition consisted of three 5-second tube phonation samples with the tube submerged 5 cm (taken from the previous sequence), 10 cm, and 18 cm into water. Participants were asked to use the same comfortable pitch produced in the first sequence for these trials.

Samples of the vocal fold vibration were obtained pre, during, and post tube phonation via a rigid endoscope attached to a plastic mouth piece. The flexible tube was also attached to the same mouth piece (Figure 1). The mouth piece was maintained between the rounded lips, so that no air would leak from the mouth; the free end of the tube was kept submerged in the water as an extension of the vocal tract. Three marks were made on the water container at 5 cm, 10 cm, and 18 cm, to help in maintaining the appropriate depth of immersion. A grand piano was used to give the pitch, which was auditorily monitored by one of the authors (M.G.). Only one recording per phonatory task was performed for each participant. A total of five samples were obtained from each participant (pre tube, phonation during tube submerged 5 cm, post tube, phonation during tube submerged 10 cm, and phonation during tube submerged 18 cm). In each sample, participants phonated for approximately 5 seconds. However, only 1 second of phonation was captured by the high-speed camera. The captured samples were obtained from the mid-portion of each phonatory task.

Instrumentation
Data collection was performed in a room typically used for clinical laryngoscopic assessment of voice. Laryngeal endoscopic procedure was performed using a HRES-Endocam 5562 system (Fa. Wolf, Knittlingen, Germany) coupled with a 90° rigid endoscope (Fa. Wolf, Knittlingen, Germany). This system also includes software for digital storing of the recordings. The high-speed camera allows recording of 4,000 frames per second with a pixel resolution of 256 × 256. For illumination, a 300-W xenon light source (Auto-LP 5131, Richard Wolf, Knittlingen, Germany) was used. To prevent tissue overheating during the recordings, the duration of light exposure was kept at a minimum. No calibration of the images could be performed. Laryngoscopic procedures were performed by a laryngologist and a phoniatrician, both coauthors of this study (M.E. and A.G.). Participants were required to stay in standing position during examinations, and no topical anesthesia was used.

Image processing
Phonovibrography was used to extract information of the vocal fold movements during phonation. For each sample, 1,000 frames (250 ms) of the high speed material were analyzed using the custom-made Phonovibrogram Software Tool (Glottal Analysis Tools v5, Michael Dollinger and Denis
Glottal area variables

The following glottal variables were derived from all high-speed registrations as described in the literature (28). These parameters were chosen due to the fact that most of them have been used in earlier studies. Moreover, these parameters do not require calibration of the images from pixels into area.

- Amplitude-to-length ratio (A-LR): ratio between vibrational amplitude at the center of the vocal folds and the length of the vocal folds.
- Closed quotient (CQ): ratio between closed phase and the entire glottal period.
- Closing quotient (CQ): ratio between closing phase and entire glottal period.
- Fundamental frequency (F0): number of cycles of vocal fold vibration per second.
- Harmonic-to-noise ratio (HNR): relation between harmonic energy and noise energy.
- Spectral flatness (SF): spectral declination or spectral tilt (frequency range from 0 to 2,000 Hz). Flatness is maximal for white noise.
- Jitter%: frequency perturbation in percentage.
- Shimmer%: amplitude perturbation in percentage.

It is worth noting that, even though HNR, SF, jitter, and shimmer are often calculated from acoustic signals, in the present study these parameters were calculated from glottal area variation.

Only descriptive statistics were calculated for the variables, including median and interquartile range. We analyzed the trend of values for all parameters and phonatory tasks without using hypothesis testing to avoid lack of statistical power due to small sample size. All analyses and graphics were performed using Stata 13.1 (StataCorp, College Station, TX, USA).

Results

Tables 1 and 2 show median and interquartile ranges for sequence 1 and 2, respectively. Moreover, results are presented below in terms of the observed trends among participants, i.e. when similar changes (compared to vowel phonation) were observed for the majority (five or more) of the participants. Figures 2 and 3 summarize the trends for sequence 1 and 2, respectively.

First sequence: phonation pre, during, and post tube submerged 5 cm into water

During phonation into a tube submerged in water to a depth of 5 cm, CQ decreased (5 of 8 participants) compared to baseline condition, HNR decreased (6 participants), spectral flatness decreased (5 participants), A-LR (right and left vocal folds) decreased (5 and 6 participants, respectively), F0 increased (6 participants), and jitter decreased (5 participants). Thus phonation seemed to be less tight, softer, a bit more noisy, and yet more stable. After phonation into tube 5 cm in water there was an increase in CQ (6 participants), CIQ (5 participants), HNR (5 participants), and F0 (5 participants) and a decrease in jitter (5 participants). Thus, it seems that phonation was tighter, and the voice was less noisy and more stable.

Second sequence: tube submerged 5 cm, 10 cm, and 18 cm into water

CQ decreased with the tube 5 cm in water (5 participants), increased during phonation into a tube submerged in water to a depth of 18 cm (6 participants), CIQ increased during phonation into a tube submerged in water to a depth of 10 cm (6 participants), HNR and A-LR decreased for all tube depths (6 and 5 participants, respectively), jitter decreased for the tube 5 cm in water (5 participants) and increased for the tube 10 cm in water (5 participants), and F0 increased for the tube 5 cm and 10 cm in water. SF decreased during tube submerged 5 cm, 10 cm, and 18 cm below the water surface (5 participants for each condition). These findings seem to imply less tight phonation for the tube 5 cm in water, while the results for deeper immersion are more difficult to interpret. Decreased A-LR, HNR, and SF seem to imply softer phonation (less collision between the vocal folds), while increased CIQ and CQ suggest the opposite.
Discussion

This study analyzed vocal fold oscillatory patterns during phonation into a tube submerged at different depths into water. In general, it was demonstrated that vocal fold oscillation is affected by tube phonation. This has been noted before when phonation through a tube in the air has been studied, e.g. using air pressure and electroglottographic registration (29), by calculating the electroglottographic contact quotient (30,31), and by applying high-speed filming (32).

In line with earlier results (31,32), our results showed that participants behaved differently and thus no statistically significant differences were observed in any parameters studied. This suggests that water resistance exercises do not give automatically certain results but the effect depends on how a person reacts to the increased airflow resistance. Certain trends were observed, though.

A-LR was in most cases lower during phonation into water, for all immersion depths. A decrease in A-LR could be expected during semi-occlusions because of the decrease in transglottal pressure (the driving force for phonation). A low A-LR is expected to imply a low impact stress on vocal folds. Therefore, a phonotraumatic reaction would be unlikely. To the best of our knowledge, only one earlier study used a similar measure, the amplitude-to-dynamic length (A-DL) ratio, in investigating the effects of tube phonation. An increased A-DL ratio was found for the longest tube, which, according to the authors, may suggest a raised vocal effort (increased subglottic pressure) (32). In the present study, a relatively high A-LR was found for some subjects with the tube 18 cm in water.

To date, no clear patterns have been evidenced regarding contact quotient from electroglottography (CQ_EGG) signal during and after semi-occlusions (30,31,33–39). Some studies have found a decrease in CQ_EGG during and after semi-occluded exercises, while others have reported an increase. An interesting trend was detected for CQ (closed quotient obtained from high-speed samples) in the present study. Tube phonation into 18 cm of water showed higher values of CQ than the other conditions and baseline. Phonation into 5 cm of water in turn demonstrated a decrease in CQ compared to baseline condition. It is possible to speculate that the effect of vocal exercises on CQ depends on the degree of flow resistance (e.g. depth of immersion) and on the subjects’ reactions to it (38). From our findings, it seems that a shallower depth tends to lead to a decrease in CQ, while a deeper submersion tends to increase the glottal CQ. Similarly, Laukkanen et al. (32) in a high-speed imaging study found higher CQ for longer tubes compared with shorter ones. Results from the study by Horáček et al. (5) showed how tube diameter also affects glottal function. Since depth of immersion, and tube length and inner diameter all affect the amount of flow resistance, it seems plausible to expect that they also affect CQ. An increase of CQ could

Figure 2. Trends for sequence 1 (phonation pre, during, and post tube submerged 5 cm into water).
result from increased adduction, or it could be a passive consequence of increased subglottic pressure (5). It is most likely that in speech production these variables co-vary and the glottis reacts reflexively to the information from various sensory receptors. As mentioned before, the importance of CQ during tube phonation is due to the fact that there is a relationship between CQ and vocal fold impact stress (15).

Two trends were observed for CIQ. It increased after 5 cm phonation compared to baseline condition and during phonation in the tube 10 cm in the water. In other words, the closing phase seemed to be relatively longer compared to baseline condition. Different changes were observed by Laukkanen et al. (32). CIQ values were smaller for longer tubes. As mentioned before, a long and/or narrow tube could be equivalent to a tube submerged deep under the water surface because of the high flow resistance. Since a low CIQ is related to the abruptness of vocal fold closure, one may expect that high flow resistance exercises (i.e. long and/or narrow tubes or deep immersion) cause a higher vocal fold impact stress compared to shallower depth of immersion. However, this was not observed in our data.

Harmonic-to-noise ratio was also examined in the present study, based on image analysis. This parameter demonstrated an increase after tube phonation submerged into 5 cm of water compared to pre tube samples. Thus, more harmonic energy compared to noise was observed. Decreased noise after semi-occlusion exercises has been reported in previous works, based on acoustic signal analysis (20,42,43). This change has been observed after a few minutes of exercises (20), and after several weeks of treatment (43). However, the subjects of these studies were patients with hyperfunctional dysphonia.

Related to harmonic energy, spectral flatness decreased in tube phonation for all immersion depths compared to baseline. Therefore, it is possible to state that our data show a steeper spectral slope for all conditions compared to baseline. Recall that SF in the present study was derived from glottal area variation. No trend was observed in the present study in samples obtained after tube phonation. Earlier studies on acoustic spectra have shown a less steep spectral slope after phonation with different semi-occluded vocal tract exercises (2,39,44,45), suggesting that semi-occlusions produce an increased spectral energy in the higher part of the spectrum. The amplitude of higher harmonics is particularly sensitive to the speed at which the glottal airflow decreases at the end of the closing phase (46). A lower SF during phonation through a tube in water seems to imply a smoother collision between the vocal folds (47).

The effect of the depth of immersion on perturbation measures was also analyzed. There was an interesting trend
that jitter decreased during and after phonation into a tube submerged 5 cm in water for most participants. A number of investigations have used acoustic measures to observe the effect of semi-occluded exercises. However, only few of them have utilized perturbation measures. Jitter and shimmer were recently assessed before and after stirring straw phonation in a group of school teachers with slight dysphonia (42). Both parameters were found to be significantly lower after performing vocal exercises. Furthermore, Barrichelo-Lindström et al. (48) reported a decrease of jitter and shimmer in normal-voiced participants after vocal warm-up using Y-buzz, another semi-occluded exercise. It is possible to speculate that lower perturbation values may reflect better sensory-motor control (49) or greater activity in laryngeal muscles if one increases subglottic pressure and adduction during and after tube phonation (50–53). A negative correlation has been found between perturbation measures and F0 and SPL. Earlier studies have suggested that a higher F0 and SPL would imply higher muscle activity, and this, in turn, would reduce perturbation of F0 and amplitude (49–52).

Earlier investigations have reported that tube phonation and other semi-occluded exercises may also modify F0. Our data showed higher values for most conditions compared to baseline. Most previous studies regarding semi-occluded vocal tract exercises have reported a decrease in F0 during exercising (2,20,29,40–43). Others have indicated an increase (13), no change (5), or no clear trend. According to two previous studies, F0 shifts depended on the length of the tube and depth of immersion. Laukkanen et al. (32) found lower F0 values for longer tubes compared to shorter ones. Furthermore, Horaček et al. (5) observed (for soft phonation) that F0 was lower with a resonance tube in water and with a thin straw. Therefore, it seems that F0 is related to the degree of airflow resistance.

It is important to mention that F0 results could depend on the instructions given in the recording process. In several studies, participants have been instructed to keep the same pitch before and after exercising in order to make the results comparable (as it was required in the present study). Moreover, considering that all earlier studies have demonstrated changes of only a few Hz, these data should be taken with caution because it likely has no clinical impact and it can be only a physical modification. Additionally, no F0 changes have clearly remained after exercising.

It seems that choosing the most appropriate semi-occlusion (e.g. length/diameter of the tube and depth of immersion) for each subject would promote the best vocal function. The results suggest that there is a tendency for smoother collision between the vocal folds and increased stability in phonation through a tube in water with a smaller immersion depth. A decrease in SF, HNR, and A-LR observed in most cases in tube phonation for all immersion depths suggests a smoother collision between the vocal folds. The increase in CQ for deep immersion depth, however, suggests that the possibility for increased vocal loading cannot be excluded when deep immersion depth is used. The desired vocal goal always depends on what the starting-point is. Our results suggest that a shallow immersion may be suited for patients with hyperfunctional voice disorders, while deep immersion may be a good option for patients with hypofunctional voice disorder. Thus, our results are in line with the common clinical tradition (see e.g. Simberg and Laine) (1). Even though technical characteristics of semi-occlusions (e.g. length or inner diameter of tube) are important, instructions and guidance on how to perform the exercise are also crucial when these exercises are used in voice rehabilitation and training.

As mentioned above, it has been found that CQ correlates with impact stress. However, during phonation into a tube in water the bubbling of water affects the vocal fold vibration. Therefore, variation in CQ values can be also present. Further studies are needed to know how well a mean CQ and other mean values of glottal area variables characterize vocal fold vibration and impact stress during water resistance therapy.

Conclusion

Even though subjects showed individual variation for most glottal area parameters, data from the present study suggest that tube phonation into water causes changes in glottal variables. In most cases the amplitude-to-length ratio, harmonic-to-noise ratio, and spectral flatness (derived from glottal area) decreased, which seems to imply a smoother collision between the vocal folds. However, the effect does not necessarily remain in vowel phonation after tube phonation. Findings obtained from shallow immersion suggest a favorable mechanical and physiological interaction in terms of decreased closed quotient and reduced jitter. Deeper immersion, instead, showed higher closed quotient. This may imply an increased respiratory and glottal effort (higher subglottic pressure and more adducted vocal folds) to compensate for the strongly increased supraglottal resistance. Thus, the impact stress during phonation through a tube into water may increase during deep immersion. The results suggest that this disparity of the effects, which is dependent upon the degree of flow resistance, should be considered when choosing treatment exercises for patients with different types of diagnoses, namely hyperfunctional or hypofunctional dysphonia. Deeper immersion could be more beneficial for patients with vocal fold paralysis or presbyphonia. On the other hand, glottal function of patients with hyperadduction or vocal fatigue could be improved using shallower depths.

Acknowledgements

Special thanks to Denis Dubrovsky and Michael Döllinger for providing the phonovibrogram software.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

In Germany, this study was supported by the Deutsche Forschungsgemeinschaft [grant Ec 409/1-1 and Ri 1050/4-1]. In Finland, it was supported by the Finnish Academy of Sciences [grant number 1128095].
References


17. Sovijärv A. Aänisyöksilogaista ja artikulaatiotekniikasta [On voice physiology and articulatory technique]. Helsinki, Finland: Department of Phonetics, University of Helsinki press; 1966. [In Finnish]


Air Pressure and Contact Quotient Measures During Different Semiclosed Postures in Subjects With Different Voice Conditions

*†Marco Guzmán, ‡Christian Castro, *Sofía Madrid, §Christian Olavarria, §Miguel Leiva, ¶Daniel Muñoz, †Elizabeth Jaramillo, and **Anne-Maria Laukkanen, *†,‡,§,¶Santiago, Chile; ‡Valparaiso, Chile; and **Tampere, Finland

Summary: Objective. The purpose of this study was to investigate the effect of phonation into tubes in air and tubes submerged in water on air pressure variables and vocal fold adduction in subjects with different voice conditions. Methods. Forty-five participants representing four vocal conditions were included: (1) subjects diagnosed with normal voice and without voice training, (2) subjects with normal voice with voice training, (3) subjects with muscle tension dysphonia, and (4) subjects with unilateral vocal fold paralysis. Participants phonated into different kinds of tubes (drinking straw, 5 mm in inner diameter; stirring straw, 2.7 mm in inner diameter; silicon tube, 10 mm in inner diameter) with the free end in air and in water. Aerodynamic, acoustic, and electroglottographic signals were captured simultaneously. Mean values of the following variables were considered: glottal contact quotient (CQ) measured by electroglotograph, fundamental frequency, subglottic pressure (Psub), oral pressure (Poral), and transglottal pressure. Results. All exercises had a significant effect on Psub, Poral, transglottal pressure, and CQ (P < 0.05). Phonation into a 55-cm silicon tube submerged 10 cm in water and phonation into a stirring straw resulted in the highest values for CQ, Psub, and Poral compared with baseline (repetition of syllable [pa:]). Key Words: Tube phonation–Vocal semiocclusion exercises–Water resistance therapy–Functional dysphonia–Vocal fold paralysis.

INTRODUCTION

Voice exercises that raise the airflow resistance of the vocal tract compared with open vowels are widely used in voice therapy and training. They are commonly called semioccluded vocal tract exercises (SOVTEs). These exercises include phonation on voiced fricatives, nasals, lip and tongue trills, or during sudden covering of the mouth by hand or finger kazoo and phonation into different tubes with the outer end either freely in the air or immersed in water. In voice training of subjects with normal voice, semiocclusion exercises have been used to improve voice quality and increase loudness in an effortless way. Especially, phonation into tubes with the outer end immersed in water has been used in voice therapy both to reduce hypertension and, in contrast, to increase adductory activity in patients with hypofunctional voice production.1–5

From the physical point of view, resistance is defined as the impediment to flow. Mathematically, resistance is a ratio between mean pressure and flow. At a given potential (difference of pressure), a lower flow will be produced when there is a high resistance. Therefore, flow is inversely proportional to resistance and proportional to potential.6 Resistance is higher when a constriction in the vocal tract is narrower. Longer and narrower tubes offer more resistance to the airflow than shorter and wider ones because of frictional losses.7,8 Moreover, when a tube is submerged in water, an extra resistance is added due to hydrostatic pressure, which is dependent on the depth of immersion and also on the angle of immersion in water.9–10

Titze et al7 measured the resistance of different kinds of tubes and straws that are commercially available and thus easy to obtain. From the results, it can be seen that a so-called resonance tube (7.5 mm in inner diameter, 30 cm in length)11 offers very little flow resistance, whereas a narrow straw (2.0 mm in inner diameter, 11.5 cm in length) in the United States used to stir coffee, offered 56 times more resistance than a resonance tube and 37 times more than a drinking straw (6.0 mm in inner diameter, 19.5 cm in length), at a range of flow which is typical for ordinary speech production. In a recent paper, Amarante et al8 studied the oral pressure (Poral; so-called back pressure) to flow relation for 10 different tubes that are commonly used for voice exercises. These tubes were studied with the outer end free in the air. Additionally, the resonance tube and a silicone lax vix tube (10 mm in inner diameter, 35 cm in length) were studied with the outer end immersed 1–7 cm in water. As expected, the results showed that a change in tube diameter affects back pressure remarkably more than a change in tube length. An increase in flow increased back pressure, especially for thinner straws. When a wider tube is immersed in water, the Poral needs to overcome the pressure generated by the water depth before flow can start. Otherwise, for a tube in water, the back pressure seems...
to be mainly determined by the immersion depth and not much by changes in flow.8

Titze et al7 stated that when airflow resistance is increased at the lips, the intraoral (supraglottal) pressure is positive (compared with atmospheric pressure), and this in turn would reduce the transglottal pressure (Ptrans; difference between subglottic and Poral) while the person is phonating, unless the subglottic pressure (Psub) is raised.7 The higher the resistance a tube offers, the more Psub a subject needs to generate to start and sustain phonation.7 The effect of different airflow resistances of the vocal tract on air pressure and on vocal fold vibration has been studied with physical model of voice production.10 According to the results, when flow was kept constant, Psub was necessary to raise considerably more for phonation into a stirring straw in air compared with a resonance tube 10 cm in water. Larger changes in Psub were observed for soft phonation compared with habitual loudness of speech. Peak-to-peak variation in Poral was largest for phonation into the tube in water.10

Several studies have investigated voice production of human subjects during and immediately after SOVTEs.7,9,12–14 Maxfield et al11 created a rank ordering by measuring the intraoral pressure produced by 13 semiocclusions commonly used in voice therapy. The highest values of Poral were evidenced in straw submerged in water, lip trills, and stirring straw with the free end in the air. According to the results by Radolf et al10 obtained from one trained female speaker, the Poral increased in phonation into a resonance tube and a stirring straw most when the resonance tube was 10 cm in water. During voice production into a resonance tube submerged 10 cm below the water surface, the increase in mean Poral (compared with vowel phonation) was about four times in habitual comfortable phonation and about nine times in soft phonation. This illustrates the fact that Poral is affected by flow and thus reflects not only supraglottic load but also Psub and adduction. Results by Radolf et al10 showed that Psub tended to increase relatively more than Poral, and thus Ptrans was higher in the SOVTEs compared with vowel phonation. Contact quotient (CQ) measured by electroglottograph (EGG) increased. In tube 10 cm in water, Ptrans decreased and CQ increased, which seems to imply increased adduction as compensation. A trained male singer in the study by Guzman et al13 in turn, showed decreased Ptrans and CQ both for phonation into a resonance tube and for phonation into a stirring straw in the air. Titze et al7 estimated Psub and registered EGG signal for two trained singers (one classically trained amateur tenor and one female professional pop-jazz singer) phonating at four pitches into seven different straws in the air. The results showed individual differences in the behavior of the subjects when phonating into tubes. Similarly, Gaskill and Erickson reported different changes in CQEGG for vocally untrained subjects phonating into a resonance tube in the air.15

Phonation into a tube in water results in bubbling of the water. Water bubbles produce a varying back pressure, which in turn modulates the Poral and gives an effect that has been described as a massage-like effect.15,16 The Poral modulations cause rhythmic modulations in vocal fold contact and glottal area and also on vertical position of the larynx.9,15,16

Changes in aerodynamic variables and vocal fold vibration have also been observed immediately after SOVTEs. In some studies, the changes followed those observed during the exercises,13 whereas in others, no consistent pattern was found within participants or across exercises.17

To date, to the best of our knowledge, all studies measuring voice production in humans during semiocclusions have been carried out with vocally normal participants and a small sample size. There is no evidence of whether subjects with different vocal skills and vocal fold status differ from each other in their reactions toward altered aerodynamic conditions of phonation. The purpose of this study was to investigate the effect of different artificial lengthening of the vocal tract (phonation through a tube with the free end in water and in air) on air pressure variables and vocal fold adduction. Research question was: Do subjects with different voice conditions change differently the air pressure and vocal fold adduction variables during different artificial lengthening?

Based on previous data and clinical observations, we hypothesize that semioclusion with narrow straws and with tube with the free end in water results in more prominent increment (1) in air pressures and (2) in CQ compared with other exercises, and that (3) changes in pressure and CQ are affected by voice condition. Specifically, we hypothesize that trained subjects are able to compensate for changes in supraglottic load and thus keep Ptrans and CQ the same, which would keep phonation stable and unaltered. Vocally untrained patients and subjects with vocal fold paralysis may try to compensate more for changes in supraglottic load and thus Ptrans and CQ would increase. Patients with muscle tension dysphonia (MTD) may tend to “give in” for increased supraglottic load, and thus decreased Ptrans and CQ are hypothesized.

METHODS

Participants

Forty-five participants were included in this study. They were chosen to represent four groups: (1) subjects diagnosed with normal voice and without voice training (n = 12; six males, six females), (2) subjects with normal voice and with voice training (n = 9; two males, seven females), (3) subjects diagnosed with MTD (n = 14; five males, nine females), and (4) subjects diagnosed with unilateral vocal fold paralysis (n = 10; two males, eight females). Vocal fold position in all participants with vocal fold paralysis was paramedian (ie, the glottal closure was not complete). The average age was 24 (20–33) years for subjects with normal voice and without voice training, 27 (21–37) years for subjects with normal voice with voice training, 28 (23–35) years for subjects with MTD, and 45 (31–56) years for subjects with vocal fold paralysis. Participants from all groups were native speakers of Spanish. This study was reviewed and approved by the University of Chile, Faculty of Medicine Review Board. Informed consent was obtained from all participants.

Phonatory tasks

Prior to aerodynamic and EGG assessment, all participants were asked to undergo rigid videostroboscopy to confirm medical
diagnosis. Laryngoscopic examinations were performed by two experienced laryngologists at the University of Chile Hospital. No topical anesthesia was used during endoscopic procedure.

Firstly, baseline recording was made for all participants. It consisted of repetition of syllable [pa:] at a rate from 2.5 to four syllables per second, in habitual, comfortable speaking pitch and loudness. Thereafter, participants were asked to select and produce a series of five semioccluded vocal tract postures in a random order: (1) drinking straw with the free end in the air (5 mm in inner diameter and 25.8 cm in length), (2) stirring straw with the free end in the air (2.7 mm in inner diameter and 10.7 cm in length), (3) silicon tube with the free end in the air (10 mm in inner diameter and 55 cm in length), (4) silicon tube with the free end submerged 3 cm below water surface, and (5) silicon tube with the free end submerged 10 cm below the water surface. During the different postures, the subjects were phonating. They were asked to maintain the same pitch as they used in the baseline samples. An electronic keyboard was used to cue the pitch, which was also auditorily monitored by the experimenters. Loudness level was perceptually controlled by experimenters. When performing the semiocclusion exercises, the participants were instructed to phonate with ease voice and to sense vibrations in the anterior face and the mouth. Before data collection, all the exercises were demonstrated to the participants by trained speech-language pathologists. Three repetitions were performed for each phonatory task. This resulted in 810 samples in total (45 subjects × six postures × three repetitions). Each entire recording session took approximately 30 minutes.

**Equipment and data collection**

Data were collected in the Voice Research Laboratory at University of Chile. Aerodynamic and EGG signals were captured simultaneously during all phonatory tasks. Aerodynamic data were collected with a Phonatory Aerodynamic System (PAS; model 4500, KayPENTAX, Lincoln Park, NJ). EGG data were obtained with an EGG (model 6103, KayPENTAX). Both aerodynamic and EGG systems were connected to an interface (Computerized Speech Lab, Model 4500, KayPENTAX), which in turn was connected to a desktop computer running a Real-Time EGG Analysis (Model 6600, version 3.4, KayPENTAX). To obtain fundamental frequency (F0), acoustic signal was recorded at a constant microphone-to-mouth distance of 20 cm, using a condenser microphone AKG (AKG Acoustics, Vienna, Austria) integrated into the PAS. All samples were recorded digitally at a sampling rate of 22.1 KHz with 16 bits/sample quantization.

At the beginning of the examination, participants were asked to sit comfortably upright on a chair. Two EGG electrodes were attached over the thyroid cartilage by means of a lightweight elastic band, which was comfortably wrapped around the participant’s neck as tightly as possible to prevent any movement of the electrodes throughout the data collection. Quality of EGG signal was monitored from the real-time oscillogram incorporated in the EGG software. To obtain optimal electrical conductivity, the electrodes were cleaned with a slightly wet tissue, and a thin layer of conductive gel was applied (Spectra 360 electrode gel, Parker Laboratories, New Jersey).

The Poral was captured with a pressure transducer (incorporated into the PAS) connected to a thin flexible plastic tube (13 cm in length). The tube was inserted into the corner of the mouth, extending a few millimeters behind the lips, without touching the tongue or any other oral structure. Calibration of the air pressure was performed according to the manufacturer’s instructions. A nose clip was used for all participants during data collection to avoid air leakage through the nose. Oral airflow leakage was controlled by one of the experimenters by placing the hand near the participant’s lips during each phonatory task. If some leakage was suspected, the task was repeated with a firmer closure at the lips. Figure 1 shows the experimental settings.

---

**FIGURE 1.** Experimental settings using tube in air (a) and tube into the water (b).
The EGG, acoustic, and pressure signals were analyzed to obtain the mean values for glottal CQ, F0, Psub (estimated from the maximum peak of the Poral during the occlusion of the consonant [p:] in the syllable [pa:] and during manual shuttering of the outer end of the tube), Poral (obtained during vowel [a:] in the syllable and during nonshuttered phase in tube phonation), and Ptrans (difference between subglottic and Poral).

Because during phonation into water, bubbling of water causes modulations in Poral,15,16 maximum and minimum Porals were measured from samples where bubbling was produced (silicon tube submerged 3 and 10 cm below water). Peak-to-peak Poral was manually obtained (Figure 2). In addition, bubbling frequency was measured from the pressure signal. All samples were analyzed with Real-Time EGG Analysis (model 6600, version 3.4, KayPENTAX). Only the most stable part from the middle section of each sample was analyzed. Criterion level of 25% from the peak-to-peak amplitude of the EGG signal was used for CQ analysis.

Statistical analysis
Numerical variables were described by median and interquartile range, and compared by phonatory task and vocal status separately using Kruskal-Wallis test. A generalized multivariable linear model to observe the joint influence of phonatory task and vocal status in vocal parameters was fitted. Separate subgroup analysis (silicon tube submerged 3 and 10 cm into water) for minimal and maximal Poral was also performed using Wilcoxon test. Finally, linear correlation analysis using Pearson coefficient for overall correlation was used. All analyses were performed using Stata 13.1 (StataCorp, College Station, TX), P < 0.05 was considered to be statistically significant, and all reported P values were two sided.

RESULTS
Table 1 shows the comparison between score averages (obtained from Kruskal-Wallis test) by phonatory task for each variable. P values indicate that the six phonatory conditions differed significantly from each other for all variables (P < 0.05), except for F0. Figures 3–6 show the distribution of values for...
CQ, Psub, Poral, and Ptrans for all groups and in all semioccluded postures.

Results from the multivariate linear regression model including F0, CQ, Psub, Poral, and Ptrans as outcomes and phonatory task as predictive variable are shown in Tables 2–5. Each table includes data from each voice condition separately. In general, all semioccluded postures produced an increase in Psub, Poral, Ptrans, and CQ compared with the baseline for all vocal conditions. In addition, during semiocclusion exercises, most variables behaved in average in the same way regardless of the vocal status of the participants.

Correlation analysis (considering all voice conditions together) demonstrated a strong linear relation between Psub and Poral ($r = 0.82; P < 0.001$). Similar results were obtained when each group was studied separately. The only correlation found for all groups was between Psub and Poral. Correlation for normal untrained participants: normal untrained participants: $r = 0.86; P < 0.001$, for normal trained participants: $r = 0.89; P < 0.001$, for subjects with MTD: $r = 0.79; P < 0.001$, and for subjects with vocal folds paralysis: $r = 0.82; P < 0.001$.

Tables 6 and 7 display values of maximum and minimum Poral during phonation into the tube in water. As can be expected, both maximum and minimum Poral were significantly higher when tube immersion depth was deeper (Table 6). However, the average peak-to-peak amplitude of Poral did not differ significantly with immersion depth. No significant differences were observed between voice conditions either (Table 7).

**DISCUSSION**

All semioccluded postures produced an increase in Psub and Poral compared with the baseline condition (syllable production). In average, phonation with the silicon tube below the water surface produced the highest values for Poral (Table 1). The same phonatory tasks also showed the highest values for Psub. Poral and

---

**TABLE 1.** Median and Interquartile Range of Variables in Phonatory Tasks

<table>
<thead>
<tr>
<th>F0 (Hz)</th>
<th>Men</th>
<th>Women</th>
<th>CQ (%)</th>
<th>Psub (cm H2O)</th>
<th>Poral (cm H2O)</th>
<th>Ptrans (cm H2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>128.27 (17.50)</td>
<td>215.13 (32.45)</td>
<td>54.14 (10.20)</td>
<td>7.28 (2.53)</td>
<td>0.22 (0.14)</td>
<td>6.90 (2.42)</td>
</tr>
<tr>
<td>Drinking straw</td>
<td>130.37 (9.14)</td>
<td>208.38 (23.46)</td>
<td>56.43 (9.66)</td>
<td>10.62 (4.73)</td>
<td>2.84 (2.54)</td>
<td>7.20 (3.11)</td>
</tr>
<tr>
<td>Stirring straw</td>
<td>127.63 (7.94)</td>
<td>160.55 (90.49)</td>
<td>57.50 (8.56)</td>
<td>14.47 (5.57)</td>
<td>6.75 (4.44)</td>
<td>6.58 (2.93)</td>
</tr>
<tr>
<td>Silicon tube in air</td>
<td>129.45 (18.20)</td>
<td>215.97 (50.60)</td>
<td>54.61 (13.26)</td>
<td>10.32 (4.33)</td>
<td>0.55 (0.82)</td>
<td>9.38 (3.85)</td>
</tr>
<tr>
<td>Silicon tube in water (3 cm)</td>
<td>134.91 (19.12)</td>
<td>222.04 (62.15)</td>
<td>54.99 (13.40)</td>
<td>12.05 (3.60)</td>
<td>32.94 (7.67)</td>
<td>8.31 (3.56)</td>
</tr>
<tr>
<td>Silicon tube in water (10 cm)</td>
<td>143.91 (42.08)</td>
<td>207.64 (50.03)</td>
<td>59.44 (11.09)</td>
<td>17.82 (3.42)</td>
<td>38.03 (7.01)</td>
<td>9.07 (3.49)</td>
</tr>
</tbody>
</table>

**FIGURE 4.** Distribution of values for all groups separately in all semioccluded postures for subglottic pressure.
$P_{\text{sub}}$ correlated strongly ($r = 0.8298; P < 0.001$). The stirring straw itself is known to offer a higher flow resistance than a (resonance) tube submerged 10 cm in water.\textsuperscript{7,8,10} However, in the present study, the highest $P_{\text{oral}}$ and $P_{\text{sub}}$ were measured for the tube submerged 10 cm in water. A similar observation has been presented by Radolf et al\textsuperscript{9} (for one subject). Such a result may be due to a leakage at the lips (overblowing) when phonating into the stirring straw. Another reason could be that people tend to reduce airflow during phonation into a straw, as phonation may feel uncomfortably hard when $P_{\text{oral}}$ increases too much.

In the present study, multivariate linear regression model showed that subject groups representing all four voice conditions (normal trained, normal untrained, MTD, and vocal fold paralysis) behaved similarly regarding air pressure variables. Tube

**FIGURE 5.** Distribution of values for all groups separately in all semioccluded postures for oral pressure.

**FIGURE 6.** Distribution of values for all groups separately in all semioccluded postures for transglottal pressure.
### TABLE 2.
Mean Parameter Values, Beta Coefficients, and Standard Errors From Multivariate Linear Regression Model for Normal Untrained Participants

<table>
<thead>
<tr>
<th>Gender</th>
<th>F0 (Hz)</th>
<th>CQ (%)</th>
<th>Psub (cm H2O)</th>
<th>Poral (cm H2O)</th>
<th>Ptrans (cm H2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>42.74</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Task</td>
<td></td>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Drinking straw</td>
<td>119.27***</td>
<td>53.80***</td>
<td>9.43***</td>
<td>2.9**</td>
<td>6.48***</td>
</tr>
<tr>
<td>Stirring straw</td>
<td>88.97*</td>
<td>55.38***</td>
<td>12.52***</td>
<td>5.86***</td>
<td>6.57***</td>
</tr>
<tr>
<td>Silicon tube in air</td>
<td>133.17***</td>
<td>49.19***</td>
<td>9.01***</td>
<td>0.48 (−1.55; 2.52)</td>
<td>7.80***</td>
</tr>
<tr>
<td>Silicon tube in water (3 cm)</td>
<td>121.40***</td>
<td>48.82***</td>
<td>12.40***</td>
<td>30.67***</td>
<td>8.34***</td>
</tr>
<tr>
<td>Silicon tube in water (10 cm)</td>
<td>131.23***</td>
<td>55.74***</td>
<td>18.5***</td>
<td>37.28***</td>
<td>9.71***</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001.

### TABLE 3.
Mean Parameter Values, Beta Coefficients, and Standard Errors From Multivariate Linear Regression Model for Normal Trained Participants

<table>
<thead>
<tr>
<th>Gender</th>
<th>F0 (Hz)</th>
<th>CQ (%)</th>
<th>Psub (cm H2O)</th>
<th>Poral (cm H2O)</th>
<th>Ptrans (cm H2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>−10.98 (−96.70; 74.73)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Task</td>
<td></td>
<td></td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Drinking straw</td>
<td>162.16**</td>
<td>54.61***</td>
<td>9.71***</td>
<td>3.3***</td>
<td>6.23***</td>
</tr>
<tr>
<td>Stirring straw</td>
<td>137.57 (−13.01; 288.14)</td>
<td>58.24***</td>
<td>13.29***</td>
<td>7.14***</td>
<td>5.99***</td>
</tr>
<tr>
<td>Silicon tube in air</td>
<td>178.24***</td>
<td>55.94***</td>
<td>9.56***</td>
<td>1.01 (−6.72; 2.68)</td>
<td>8.46***</td>
</tr>
<tr>
<td>Silicon tube in water (3 cm)</td>
<td>178.78***</td>
<td>49.81***</td>
<td>10.67***</td>
<td>30.22***</td>
<td>6.56***</td>
</tr>
<tr>
<td>Silicon tube in water (10 cm)</td>
<td>176.44***</td>
<td>57.53***</td>
<td>16.75***</td>
<td>37.15***</td>
<td>7.77***</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001.
### TABLE 4.
Mean Parameter Values, Beta Coefficients, and Standard Errors From Multivariate Linear Regression Model for Participants With Muscle Tension Dysphonia

<table>
<thead>
<tr>
<th>Task (ref. baseline)</th>
<th>F0 (Hz)</th>
<th>CQ (%)</th>
<th>Psub (cm H2O)</th>
<th>Poral (cm H2O)</th>
<th>Ptrans (cm H2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task (ref. baseline)</td>
<td>−28.80 (−74.62; 16.85)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Drinking straw</td>
<td>182.25*** (121.24; 243.26)</td>
<td>57.18*** (43.76; 70.61)</td>
<td>10.55*** (8.25; 12.84)</td>
<td>2.96*** (35.47; 38.83)</td>
<td>7.63*** (5.12; 10.15)</td>
</tr>
<tr>
<td>Stirring straw</td>
<td>181.29*** (94.26; 268.31)</td>
<td>57.64*** (43.71; 71.57)</td>
<td>14.71*** (12.33; 17.10)</td>
<td>7.52*** (5.99; 9.05)</td>
<td>6.67*** (4.51; 9.73)</td>
</tr>
<tr>
<td>Silicon tube in air</td>
<td>191.88*** (130.87; 252.89)</td>
<td>56.28*** (42.85; 69.70)</td>
<td>11.01*** (8.70; 13.29)</td>
<td>0.63 (−.84; 2.11)</td>
<td>10.34*** (7.83; 12.85)</td>
</tr>
<tr>
<td>Silicon tube in water (3 cm)</td>
<td>187.54*** (126.53; 248.56)</td>
<td>54.84*** (41.46; 68.32)</td>
<td>13.51*** (11.22; 15.81)</td>
<td>34.17*** (32.69; 35.64)</td>
<td>9.82*** (7.69; 11.95)</td>
</tr>
<tr>
<td>Silicon tube in water (10 cm)</td>
<td>189.89*** (131.54; 247.80)</td>
<td>58.15*** (44.73; 71.58)</td>
<td>18.51*** (16.21; 20.80)</td>
<td>39.21*** (36.73; 39.68)</td>
<td>10.70*** (8.57; 12.83)</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001.

### TABLE 5.
Mean Parameter Values, Beta Coefficients, and Standard Errors From Multivariate Linear Regression Model for Participants With Vocal Fold Paralysis

<table>
<thead>
<tr>
<th>Task (ref. baseline)</th>
<th>F0 (Hz)</th>
<th>CQ (%)</th>
<th>Psub (cm H2O)</th>
<th>Poral (cm H2O)</th>
<th>Ptrans (cm H2O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task (ref. baseline)</td>
<td>42.22 (−45.23; 129.68)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Drinking straw</td>
<td>212.69* (43.44; 381.93)</td>
<td>60.72*** (42.28; 79.16)</td>
<td>13.89*** (10.27; 17.51)</td>
<td>5.48* (1.20; 9.76)</td>
<td>8.19*** (4.08; 12.31)</td>
</tr>
<tr>
<td>Stirring straw</td>
<td>168.73* (35.27; 302.26)</td>
<td>62.67*** (42.96; 82.39)</td>
<td>19.07*** (14.75; 23.40)</td>
<td>10.19** (4.67; 15.72)</td>
<td>6.03*** (3.06; 13.68)</td>
</tr>
<tr>
<td>Silicon tube in air</td>
<td>167.01** (68.21; 265.80)</td>
<td>56.08*** (37.64; 74.52)</td>
<td>11.73*** (7.92; 15.54)</td>
<td>0.9 (−3.61; 5.41)</td>
<td>10.89*** (6.56; 15.23)</td>
</tr>
<tr>
<td>Silicon tube in water (3 cm)</td>
<td>218.04*** (138.55; 297.52)</td>
<td>62.42*** (45.03; 79.80)</td>
<td>14.49*** (10.87; 18.10)</td>
<td>35.49*** (31.21; 39.77)</td>
<td>9.79*** (6.47; 13.10)</td>
</tr>
<tr>
<td>Silicon tube in water (10 cm)</td>
<td>175.20*** (100.02; 250.38)</td>
<td>67.44*** (50.94; 83.93)</td>
<td>19.85*** (16.23; 23.47)</td>
<td>43.91*** (39.63; 48.19)</td>
<td>9.08*** (5.77; 12.40)</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001
in water and stirring straw in air presented the highest values in both Psub and Psupra for all voice conditions. It seems that these variables are in general more dependent on the degree of airflow resistance than vocal status of participants.

Earlier investigations have shown changes in vocal tract configuration during SOVTEs. These changes may result from changes in Poral. Guzman et al reported that during glass tube and stirring straw phonation, the hypopharyngeal area widened, the larynx position lowered, and the velum closed the nasal passage better compared with open vowel phonation. Better closure of the nasal passage may contribute to increase in Poral. All changes were more prominent during stirring straw in the air (high degree of airflow resistance) than in Finnish glass tube in the air (low resistance). In another study, where eight different semioccluded exercises were explored, all of them produced a lower vertical laryngeal position, narrower aryepiglottic opening, and a wider pharynx than resting position. More prominent changes were obtained with a tube in the water (10 and 3 cm) and narrow tube in the air.

In the present study, Ptrans was higher than baseline condition and significantly different throughout all semioccluded postures (except for silicon tube in the air) (Tables 1–5). This implies that although both Poral and Psub increase, they do not change proportionally. Psub increases relatively more than Poral. This seems to be due to a compensatory adjustment to sustain the phonatory airflow during phonation. The same trend was seen in Radolf et al for both tube and stirring straw phonation. Regarding voice condition, no clear pattern was observed in our data for Ptrans in the multivariate linear regression model. It is worth noticing, however, that distribution of Ptrans was very wide and narrow tube in the air.

In the present study, CQ was overall significantly greater for all semioccluded postures than for the baseline (syllable production). Phonatory tasks that showed in average the highest values of CQ were tube 10 cm in water and stirring straw in air (Table 1 and Figure 3). Recall that these exercises were also found to present the highest values of Psub and Poral. It seems that when supraglottic load increased, a higher Psub and a compensatory glottal adduction were produced, and that the changes were related to the degree of airflow resistance that the semiocclusions offered. Similar results have been obtained by Radolf et al and Guzman et al for CQEGG and by Guzman et al for CQ derived from high-speed registration, even though the effect in the latter was not statistically significant. Opposite results have been found by Granqvist et al who in their high-speed study reported (for two subjects) increased open quotient with increased water depth (2–6 cm). Guzman et al studied depths of 5 cm, 10 cm, and 18 cm. Phonation with deeper immersion depths (>6 cm) may thus require more adduction of the vocal folds.

If CQ is really affected by airflow resistance as also shown previously, it should be considered that patients with low vocal fold adduction (eg, vocal fold paralysis or presbyphonia) could be treated using high-flow resistance exercises. Instead, lower flow resistance should be applied in subjects with vocal fold hyperadduction. In this regard, clinical recommendations were proposed by Sovijärvi: Tube submerged 1–2 cm below the water could be used for patients with hyperfunctional voice disorders, whereas tube submerged in a depth of 10–15 cm could be more appropriate for subjects with vocal fold paralysis. These recommendations have been followed in clinical practice of water resistance therapy in Finland.

During tube phonation into water, the water bubbling frequency in our data showed an average of 22 Hz, with a range of 12–32 Hz, regardless of tube immersion depth and the vocal condition of the subjects. The amplitude of Poral modulation (mean difference between max. and min. Poral) caused by bubbling did not show significant differences either when comparing tube immersion depths of 3 cm and 10 cm in water.

### TABLE 6.
**Median and Interquartile Range of Maximum, Minimum, and Difference Between Maximum and Minimum of Poral Considering Depth of Tube Immersion in Water (Wilcoxon Test)**

<table>
<thead>
<tr>
<th></th>
<th>Silicone Tube in Water (3 cm)</th>
<th>Silicone Tube in Water (10 cm)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. Poral (cm H2O)</td>
<td>5.58 (1.43)</td>
<td>10.67 (2.47)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Min. Poral (cm H2O)</td>
<td>2.01 (1.20)</td>
<td>6.69 (1.40)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Diff. Poral (cm H2O)</td>
<td>3.64 (1.32)</td>
<td>3.86 (1.19)</td>
<td>0.2066</td>
</tr>
</tbody>
</table>

### TABLE 7.
**Median and Interquartile Range of Maximum and Minimum Poral for Tube in Water Considering Different Groups (Both Immersion Depths Together)**

<table>
<thead>
<tr>
<th></th>
<th>Normal Untrained</th>
<th>Normal Trained</th>
<th>Dysphonia</th>
<th>Paralysis</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. Poral (cm H2O)</td>
<td>8.32 (5.75)</td>
<td>7.17 (5.18)</td>
<td>7.08 (4.34)</td>
<td>10.47 (6.60)</td>
<td>0.2509</td>
</tr>
<tr>
<td>Min. Poral (cm H2O)</td>
<td>4.46 (5.46)</td>
<td>3.74 (5.15)</td>
<td>3.05 (4.05)</td>
<td>6.42 (4.75)</td>
<td>0.1121</td>
</tr>
</tbody>
</table>
Based on the results by Radolf et al. and our results, this may suggest that a higher Psub does not necessarily imply a larger modulation of the Poral. Therefore, a stronger massage-like effect (usually reported by patients for 10 cm compared with 3 cm immersion depths in water) most likely cannot be associated with the degree of tube immersion when the flow is kept constant.

Our study has certain limitations. Because only 45 participants were included in the present study and they were divided in four different groups, it was not possible to perform variance analysis, and it was not possible to consider division by gender or type of voice training (speaking or singing voice). Moreover, more types of commonly used tubes could have been included such as the traditional Finnish glass tube. Further investigations should assess whether changes in air pressures and CQ remain over a longer period of time (long-term effect). In addition, patients with phonotraumatic vocal fold lesions should also be included in future studies. Finally, airflow measurements should also be considered.

CONCLUSION

During semiocclusion exercises, most variables behaved in average in the same way regardless of the vocal status of the participants. This implies the importance of proper instructions in voice therapy, as the expectations of therapy outcome naturally are different for patients with different types of voice disorders.

REFERENCES

Laryngeal and Pharyngeal Activity During Semioccluded Vocal Tract Postures in Subjects Diagnosed With Hyperfunctional Dysphonia

*Marco Guzman, †Christian Castro, ‡Alba Testart, §Daniel Muñoz, and ||Julia Gerhard, Accepted for publication May 17, 2013.

Summary: High vertical laryngeal position (VLP), pharyngeal constriction, and laryngeal compression are common features associated with hyperfunctional voice disorders. The present study aimed to observe the effect on these variables of different semioccluded vocal tract postures in 20 subjects diagnosed with hyperfunctional dysphonia. During observation with flexible endoscope, each participant was asked to produce eight different semioccluded exercises: lip trills, hand-over-mouth technique, phonation into four different tubes, and tube phonation into water using two different depth levels. Participants were required to produce each exercise at three loudness levels: habitual, soft, and loud. To determine the VLP, anterior-to-posterior (A-P) compression, and pharyngeal width, a human evaluation test with three blinded laryngologists was conducted. Judges rated the three endoscopic variables using a five-point Likert scale. An intraclass correlation coefficient to assess intrarater and interrater agreement was performed. A multivariate linear regression model considering VLP, pharyngeal width, and A-P laryngeal compression as outcomes and phonatory tasks and intensity levels as predictive variables were carried out. Correlation analysis between variables was also conducted. Results indicate that all variables differ significantly. Therefore, VLP, A-P constriction, and pharyngeal width changed differently throughout the eight semioccluded postures. All semioccluded techniques produced a lower VLP, narrower aryepiglottic opening, and a wider pharynx than resting position. More prominent changes were obtained with a tube into the water and narrow tube into the air. VLP significantly correlated with pharyngeal width and A-P laryngeal compression. Moreover, pharyngeal width significantly correlated with A-P laryngeal compression.

Key Words: Semiocclusion—Vocal tract—Voice therapy—Hyperfunction—Dysphonia—Aryepiglottic narrowing—Vertical laryngeal position—Pharyngeal width

INTRODUCTION

It is generally agreed among clinicians and voice scientists that the vertical laryngeal position (VLP) is an important aspect of voice production in both normal and pathological voices.\(^1\)\(^2\) It seems that several factors affect the VLP, such as phonetic features,\(^3\)\(^4\) lung volume,\(^5\) voice technique,\(^6\)\(^7\) pitch control,\(^6\) respiration technique,\(^8\) and vocal loudness.\(^9\)

A high laryngeal position is commonly associated with voices that have a strong component of muscle tension, especially in patients diagnosed with hyperfunctional voice disorders. Commonly, the abnormally high tension in extrinsic laryngeal muscles may cause a high position of the larynx.\(^10\)\(^11\)\(^12\) Therefore, a lowering of the elevated larynx is usually an important goal in clinical voice therapy and classical singing pedagogy.\(^13\)\(^14\)\(^15\)\(^16\)\(^17\)\(^18\)\(^19\)\(^20\)\(^21\)\(^22\)\(^23\)\(^24\) Several vocal exercises have been reported as useful therapeutic and training tools to lower the larynx. The yarn-sight technique is one of the most popular among voice pathologists and voice teachers.\(^14\) Other exercises are the prolonged consonant /br/,\(^15\) soft and aspirate vocal onset,\(^16\) and laryngeal manipulation.\(^17\)\(^18\)\(^19\)

VLP has both acoustic and physiological implications. An upward movement of the larynx from its resting position shortens vocal tract length, which raises all formant frequencies; this, in turn, produces a brighter vocal quality.\(^20\)\(^21\) The low position of the larynx produces the opposite acoustic effect. The VLP also has important effects on the biomechanical properties of the vocal folds. A high VLP stiffens the vocal fold tissues, therefore increasing fundamental frequency and potentially changing the folds’ vibratory pattern. Furthermore, high VLP usually facilitates a tight vocal fold adduction as part of the valving laryngeal function for airway protection.\(^20\)\(^21\) Moreover, Titze\(^25\) reported that vocal folds are likely to be thicker when the larynx is lowered. Thus, the cover of vocal folds loosens and the medial surfaces make a better glottal closure. When this occurs, a greater maximum flow declination rate is produced, which contributes to the increased vocal intensity without additional vocal effort.

Another common feature treated by voice therapists in patients diagnosed with muscle tension is the relaxation and opening of the pharyngeal area. This is also an important goal of singing pedagogy. Exercises to produce an open throat have been one of the most used tools to produce freedom or lack of tension in the area of the throat, resulting in a lack of constriction and a better voice quality in both normal and pathological voices.\(^13\)\(^23\)\(^24\) Most teachers include the use of the open throat technique as an important feature in singing training, especially in classical singing. The purpose of these types of exercises is described by voice trainers to be a way of maximizing pharyngeal space and/or achieving abduction of the ventricular folds.\(^13\) Titze\(^25\) as well as Titze and Story\(^26\)
described a “wide pharynx” as an acoustic enhancement to the first formant and to the overall sound. An open throat production has perceptually been described as a rounded, free, effortless, and warm sound.27

Supraglottic activity refers to the movements and configurations of structures above the vocal folds. There are two types of supraglottic activity: (1) anterior-to-posterior (A-P) laryngeal compression (aryepiglottic narrowing), which occurs when the arytenoid cartilages approximate the petiole of the epiglottis and (2) medial constriction, which refers to addition of the false vocal folds.28,29 Supraglottic activity has been commonly classified as a sign of nonorganic hyperfunctional dysphonia by clinicians.30 In addition, for many years, the development of several benign lesions on the vocal fold surface has been assumed to be related to hyperfunctional behavior or phonotrauma.31 On the other hand, some studies show that supraglottic activity could be present in subjects with normal voice.28,29,32 In fact, both A-P and medial compression have been found to be normal and even desirable laryngeal behaviors in singing33–36 and speaking among professional voice users.37

The present study aimed to observe and compare the effect of eight semioccluded vocal tract postures on VLP, A-P laryngeal compression, and pharyngeal width in a group of subjects diagnosed with hyperfunctional dysphonia.

METHODS

Participants

This study was approved by the research ethics committee at the School of Communication Disorders of the University of Valparaiso, Chile. Informed consent was obtained from 28 adult subjects (19 women and 9 men). The average age of this subject set was 26 years, with a range of 20–28 years old. Inclusion criteria for this study included (1) no previous voice therapy or voice training and (2) diagnosis of hyperfunctional dysphonia without any vocal fold lesions. Individuals with a history of smoking were excluded from this study. Although 28 subjects were recruited, seven of them did not meet the inclusion criteria. Therefore, only 21 were included in the analysis. Participants were asked to undergo flexible laryngoscopy to corroborate laryngeal diagnosis and to perform the phonatory tasks. The initial diagnosis was made by a laryngologist with more than 20 years of experience in voice disorders. All participants reported at least 1 year of voice problems.

Laryngoscopic procedure

At the beginning of the examination, participants were asked to sit upright in a comfortable chair. Assessment of the laryngeal and pharyngeal activity was carried out through endoscopic examination with a flexible fiberoptic endoscope (Olympus ENF type p4; Olympus, Center Valley, PA) connected to a video camera (Sony DCX-LS1 Sintek; Sony Corporation, New York, NY) and a Richard Wolf LP 4200 light source (Richard Wolf Medical Instrument Corporation, Vernon Hills, IL). Analog images were digitalized with Pinnacle Studio HD 10 software (Corel Corporation, Fremont, CA), and views were monitored on a color television monitor (Sony SSM-20L120). All examinations were performed without topical nasal anesthesia. The flexible endoscope was placed directly below the tip of the uvula, allowing a full view of the pharynx and larynx. This placement was fixed by securing the fiberscope against the alar cartilage of the nose with the laryngologist’s finger. A steady placement of the fiberscope is crucial because observation of laryngeal height adjustments and other laryngeal configurations can be affected by movement of the fiberscope. For the purposes of this study, three aspects were observed during laryngoscopic procedure: VLP, pharyngeal width, and A-P laryngeal compression.

Phonatory tasks

During the observation with the flexible endoscope, each participant was required to produce four different semioccluded vocal tract exercises at habitual speaking pitch: lip trill, hand-over-mouth technique, tube phonation into air, and tube phonation into water. All subjects performed these phonatory tasks twice. Participants were asked to perform tasks at their habitual speaking pitch to avoid variation (the effect of pitch on) in VLPs. Participants were asked to produce each exercise at three loudness levels: soft, moderate, and loud. Tube into water exercises were performed at only moderate loudness. The tube phonation tasks were performed using four different types of plastic commercial drinking (wide) and stirring (narrow) straws. Each participant was instructed to hold the straw with one hand, straight out from the mouth. The straw was maintained a few millimeters between the rounded lips, so that no air would leak from the mouth; the free end was kept either in the air or submerged under the water as an extension of the vocal tract. Careful control of pitch throughout the entire sequence was performed. An electronic keyboard was used to give and control the pitch, which was monitored auditorily. Pitch control is relevant because it may influence both laryngeal and pharyngeal activities. Each phonatory task was produced for a minimum of 7 seconds, and subjects were required to breathe normally between tasks to avoid the effect of lung volume on the laryngeal height. Before each semioccluded task, participants returned to a phonatory resting position to obtain baseline measures.

Each complete assessment session was accomplished in approximately 15 minutes with the following protocol:

1. Phonation into a long-wide tube (6 mm of inner diameter and 20 cm in length).
2. Phonation into a long-narrow tube (3 mm of inner diameter and 20 cm in length).
3. Phonation into a short-wide tube (6 mm of inner diameter and 10 cm in length).
4. Phonation into a short-narrow tube (3 mm of inner diameter and 10 cm in length).
5. Phonation into a long-wide tube submerged 3 cm below the water surface.
6. Phonation into a long-wide tube submerged 10 cm below the water surface.
7. Phonation using the hand-over-mouth technique.
8. Phonation with lip trill.

The order of the tasks in the protocol was not randomized.
Visual evaluation
To determine the VLP, A-P compression, and pharyngeal width, we conducted a human evaluation test with three blinded judges (2 men, 1 woman; mean age of 46 years with a range of 42–50 years). All judges were laryngologists with more than 4 years of experience in voice disorders. All audio signals were removed from video samples before performing the assessment to avoid the possible effect of voice quality on the judges’ ratings. To standardize the rating parameters and rating scales, the three judges participated in a 1-hour training session in videolaryngoscopic examinations. Video samples from each subject were played to the judges, and they were instructed to rate the three endoscopic variables using a five-point Likert scale; for VLP (1 = very high and 5 = very low), A-P laryngeal compression (1 = very opened and 5 = very narrow), and pharyngeal width (1 = very narrow and 5 = very wide). The evaluation was performed in a quiet room. Ratings were completed in two sessions by all the raters. Each session lasted no more than 1 hour. All raters reported normal or corrected-to-normal vision. Fifteen percent of the samples were randomly repeated to assess the intrarater reliability. Judges were not aware of these repetitions.

Statistical analysis
Descriptive statistics were calculated for the variables, including mean and standard deviation. A multivariate linear regression model was used to obtain an intraclass correlation coefficient (ICC) to assess the judges’ reliability (intrarater and interrater agreement) controlled by phonatory task and vocal loudness. Then, another multivariate linear regression model considering VLP, pharyngeal width, and A-P laryngeal compression as outcomes as well as phonatory tasks and its intensity as predictive variables (and its interactions if exist) was performed. Simple correlation analysis using Spearman rho between VLP, pharyngeal width, and A-P laryngeal constriction was also conducted. One-way analysis of variance for test differences between phonatory task scores was used. The analysis was performed using Stata 12.1 (StataCorp LP, College Station, TX) software. An alpha of .05 was used for the statistical procedures.

RESULTS
Table 1 shows the results from the intrarater reliability analysis. A good intrarater concordance was demonstrated for each judge. Moreover, the three blinded judges obtained a high agreement (intrarater reliability) (ICC = 0.79 [0.66–0.87], P < 0.0001).

Table 2 and Figure 1 display the comparison between score averages by phonatory task for each variable (outcome). P values indicate that all variables were found to have a significant effect, and all of them differ significantly from each other (P < 0.0001). Therefore, VLP, A-P laryngeal compression, and pharyngeal width changed differently throughout the eight semioccluded postures.

Results from the multivariate linear regression model including VLP, pharyngeal width, and A-P laryngeal constriction as outcomes as well as phonatory task and its intensity as predictive variables (and its interactions if they exist) are shown in Table 3.

VLP significantly correlates with pharyngeal width (rho = 0.578; P < 0.0001) and A-P laryngeal constriction (rho = 0.3364; P < 0.0001). Furthermore, pharyngeal width significantly correlates with A-P laryngeal constriction (rho = 0.18; P = 0.001).

DISCUSSION
The present study aimed to observe the effect of eight semioccluded vocal tract postures on VLP, A-P laryngeal compression, and pharyngeal width in a group of subjects diagnosed with hyperfunctional dysphonia. This is the first study designed to compare the effect of a large number of semioccluded vocal exercises and different loudness levels on pharyngeal and laryngeal activities. Result revealed that the effect on these variables is statistically significant throughout all phonatory tasks.

All semioccluded postures produced a decrease in VLP compared with the resting position. Phonation with tube into the water (10 and 3 cm below the surface) and phonation into a long-narrow tube produced the three lowest VLPs. Interestingly, the same three phonatory tasks caused the widest pharynx and the narrowest A-P laryngeal compression. In fact, the correlation analysis demonstrated a high correlation between all these dependent variables.

Sovijärvi et al38 stated that one of the most relevant effects of tube phonation is the lowering of the larynx. The degree of this effect would be related to the length of the tube.39–41 The same author pointed out that the goal is not necessarily to reach a very low larynx but to avoid a high VLP, especially in subjects with hyperfunctional laryngeal activity.42

Earlier investigations have demonstrated similar effects of tube phonation on VLP. In a computerized tomography study, Guzman et al43 reported lowering of the larynx during both glass resonance tube and stirring straw phonation. This change remained during vowel production after tube and straw. In a videofluorographic and dual-channel electroglottographic registration, Laukkanen et al44 found a lower VLP compared with the resting position during other semioccluded vocal tract postures. The opposite effect of tube phonation on VLP has also been demonstrated.45,46 Furthermore, two recent magnetic resonance imaging studies reported no changes on the VLP during phonation into a resonance tube and during voiced plosive consonants.47,48

It is important to highlight that no previous studies have reported the effect of semioccluded postures on VLP in subjects.

<table>
<thead>
<tr>
<th>Judge</th>
<th>ICC (95% Confidence Interval)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.71 (0.61–0.86)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>2</td>
<td>0.78 (0.67–0.88)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>3</td>
<td>0.65 (0.54–0.79)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

TABLE 1. Intrarater Reliability Analysis
diagnosed with nonorganic hyperfunctional dysphonia. A high VLP is commonly presented in patients with this vocal condition. Because VLP significantly decreased in our participants, it can be concluded that semioccluded postures, especially tube submerged into the water and phonation into long-narrow tube (stirring straw), are suitable therapeutic tools for people with high VLP because of laryngeal muscle tension.

Muscle tension may be reflected in the pharyngeal activity as well, more specifically in pharyngeal narrowing. Findings from the present study revealed that all phonatory tasks produced a widening of the pharynx during exercising. Earlier studies have demonstrated similar outcomes on pharyngeal configuration. Several changes were observed during both glass resonance tube and narrow straw phonation in a recent study. The lower pharynx area, middle pharyngeal region, and A-P width of the hypopharynx increased during exercising compared with vowel phonation before the exercises. All these changes were larger during straw than tube phonation. Moreover, in a computed tomography and finite-element modeling investigation, the most dominant change in the vocal tract during phonation into the tube was caused by expansion of the cross-sectional area of the oropharynx. An increase in the area of the junction between the pyriform sinuses and epilaryngeal tube was also observed. Hence, lengthening of the vocal tract may also have a positive effect on pharyngeal configuration in patients with hyperfunction. It may also be used as a vocal training exercise in normal individuals.

Two possible explanations could be reasonable for the effect on VLP and pharyngeal width of phonation with a tube submerged under the water and a long-narrow tube (stirring straw). First, oral pressure increases during these types of exercises. In a previous study, acoustic impedance and mean supraglottal resistance were raised by phonating into different tubes (different in length and inner diameter) in the air and submerged under the water. The results showed that the oral pressure is higher when phonating into narrow straws (stirring straws) than wide straws (drinking straws) and even higher when straws are submerged under the water. Furthermore, the deeper the straw is under the water the higher the oral pressure. These findings are in line with the results reported by Titze. Guzman et al also reported increased oral pressure during resonance tube and stirring straw phonation, being greater in the latter. Interestingly, in the present study, the lowest VLP and widest pharynx were observed during the most resistive phonatory tasks: phonation into a straw submerged under the water (3 and 10 cm) and phonation into a long-narrow tube (stirring straw). The other semioccluded postures also produced the same effect but with a lower degree. The increased oral pressure during semioccluded exercises may have directly pushed the larynx down and the pharyngeal walls laterally. Therefore, the first possible explanation would be a mechanical effect. The second explanation of the effect on VLP and pharyngeal widening of semiocclusions is the muscle relaxation. The pushing effect accomplished by oral pressure may produce a relaxation of the laryngeal and pharyngeal musculature, and this in turn may have produced a lowering of the larynx and widening of the pharynx.
Another interesting result from the present study was the aryepiglottic narrowing (A-P compression) found during all phonatory tasks. As occurred for the VLP and pharyngeal width, the A-P compression was also more prominent during phonation with a tube submerged under the water (3 and 10 cm) and during phonation with a long-narrow tube. Aryepiglottic narrowing or A-P laryngeal compression has been described as both a sign of laryngeal hyperfunction and also as a good and desirable feature in voice performers. Sundberg reported that the aryepiglottic narrowing contributes to the formation of the singer’s formant, a cluster of the third, fourth, and fifth formants. This acoustic characteristic, typical in trained singers, may help singers obtain a louder and brighter voice quality because of a high concentration of acoustic energy around 3 kHz. These findings were later confirmed by Titze and Story.

It not surprising that in the present study, VLP and A-P laryngeal compression were correlated. In this regard, Sundberg has suggested that aryepiglottic narrowing can be reached by lowering of the larynx. According to the author, a low VLP is a way to obtain a high ratio between the cross-sectional area of the low pharynx and the epilaryngeal tube opening, which is the necessary setting to produce the singer’s formant cluster. Nevertheless, this is not the only way to reach the high ratio. Earlier investigations have demonstrated that a spectral prominence near 3 kHz could also be obtained by other vocal tract strategies.

A high correlation between pharyngeal width and A-P laryngeal compression was demonstrated in the present study as well. This means that when the pharynx widened, there was also a narrowing of the aryepiglottic sphincter. A greater ratio of inlet to the pharynx over the outlet of the epilarynx has previously been reported using magnetic resonance imaging and computerized tomography examinations when using artificial lengthening of the vocal tract. This increased ratio would help to the formation of the singer/speaker’s formant. Related to this, in the three previous cited studies where a greater ratio was obtained, a more evident spectral prominence around 2.5 kHz (singer’s formant) was also reported. Therefore, it is feasible to assume that vocal tract semiocclusions may contribute to a high concentration of spectral energy at the singer’s formant region because of an increment of the ratio between pharyngeal and epilaryngeal tube openings.

According to Titze and Story, when a narrowed epilarynx (produced by an aryepiglottic narrowing) is combined with a wide pharynx, the acoustic load of the vocal tract is inertive for all possible values of fundamental frequencies. This produces strong interactions between the source and the filter. Specifically, the inertance of the vocal tract facilitates vocal fold vibration by lowering the oscillation threshold pressure. This effect may also be important for people with hyperfunctional dysphonia.

In the present study, all dependent variables demonstrated the greatest degree of change when phonating with loud voice. No significant changes were observed during soft and moderate loudness for pharyngeal width and A-P laryngeal compression. For VLP, all loudness levels caused a significantly lower larynx. However, as mentioned previously, the greatest change was seen during loud voice production. This is an interesting result that could be related to the degree of oral pressure produced during artificial lengthening and occlusions of the vocal tract. Therefore, this independent variable should be considered when using these types of exercises to modify the laryngeal and pharyngeal activities. In an earlier study, Yanagisawa et al. found similar results with regard to the effect of loudness on supraglottic activity. The authors reported that more aryepiglottic narrowing was obtained when loudness level increased across different singing voice qualities.

Because the order of the tasks in the protocol was not randomized, one could suspect that in fact only the first task had an effect, which then persisted across the other tasks. However,
<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients (95% Confidence Interval)</th>
<th>t</th>
<th>P Value</th>
<th>Variables</th>
<th>Coefficients (95% Confidence Interval)</th>
<th>t</th>
<th>P Value</th>
<th>Variables</th>
<th>Coefficients (95% Confidence Interval)</th>
<th>t</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intensity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft</td>
<td>0.21 (0.11, 0.28)</td>
<td>3.11</td>
<td>0.007</td>
<td>Soft</td>
<td>0.11 (–0.05, 0.18)</td>
<td>1.27</td>
<td>0.091 (NS)</td>
<td>Soft</td>
<td>0.10 (–0.08, 0.17)</td>
<td>1.13</td>
<td>0.238 (NS)</td>
</tr>
<tr>
<td>Habitual</td>
<td>0.23 (0.09, 0.38)</td>
<td>3.18</td>
<td>0.002</td>
<td>Habitual</td>
<td>0.14 (–0.01, 0.29)</td>
<td>1.84</td>
<td>0.067 (NS)</td>
<td>Habitual</td>
<td>0.11 (–0.04, 0.28)</td>
<td>1.42</td>
<td>0.157 (NS)</td>
</tr>
<tr>
<td>Loud</td>
<td>0.38 (0.24, 0.53)</td>
<td>5.20</td>
<td>&lt;0.001</td>
<td>Loud</td>
<td>0.41 (0.25, 0.56)</td>
<td>5.30</td>
<td>&lt;0.001</td>
<td>Loud</td>
<td>0.30 (0.13, 0.46)</td>
<td>3.60</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Phonatory task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-wide tube</td>
<td>0.37 (0.19, 0.55)</td>
<td>3.60</td>
<td>&lt;0.001</td>
<td>Long-wide tube</td>
<td>0.57 (0.25, 0.90)</td>
<td>3.69</td>
<td>&lt;0.001</td>
<td>Long-wide tube</td>
<td>0.78 (0.45, 1.11)</td>
<td>4.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Long-narrow tube</td>
<td>0.41 (0.20, 0.62)</td>
<td>3.90</td>
<td>&lt;0.001</td>
<td>Long-narrow tube</td>
<td>0.41 (0.19, 0.62)</td>
<td>3.75</td>
<td>&lt;0.001</td>
<td>Long-narrow tube</td>
<td>0.26 (0.03, 0.50)</td>
<td>2.27</td>
<td>0.023</td>
</tr>
<tr>
<td>Short-wide tube</td>
<td>–0.25 (–0.20, –0.04)</td>
<td>–2.40</td>
<td>0.017</td>
<td>Short-wide tube</td>
<td>–0.28 (–0.50, –0.06)</td>
<td>–2.60</td>
<td>0.010</td>
<td>Short-wide tube</td>
<td>–0.01 (–0.24, 0.21)</td>
<td>–0.13</td>
<td>0.894 (NS)</td>
</tr>
<tr>
<td>Short-narrow tube</td>
<td>0.20 (–0.001, –0.41)</td>
<td>1.95</td>
<td>0.052</td>
<td>Short-narrow tube</td>
<td>0.25 (0.03, 0.47)</td>
<td>2.31</td>
<td>0.022</td>
<td>Short-narrow tube</td>
<td>0.23 (0.004, 0.47)</td>
<td>2.01</td>
<td>0.045</td>
</tr>
<tr>
<td>Tube into the water (3 cm)</td>
<td>0.28 (–0.017, –0.59)</td>
<td>1.85</td>
<td>0.065 (NS)</td>
<td>Tube into the water (3 cm)</td>
<td>0.59 (0.27, 0.91)</td>
<td>3.69</td>
<td>&lt;0.001</td>
<td>Tube into the water (3 cm)</td>
<td>0.57 (0.23, 0.91)</td>
<td>3.30</td>
<td>0.001</td>
</tr>
<tr>
<td>Tube into the water (10 cm)</td>
<td>0.81 (0.50, 1.11)</td>
<td>5.21</td>
<td>&lt;0.001</td>
<td>Tube into the water (10 cm)</td>
<td>1.07 (0.75, 1.39)</td>
<td>6.63</td>
<td>&lt;0.001</td>
<td>Tube into the water (10 cm)</td>
<td>0.81 (0.47, 1.15)</td>
<td>4.67</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hand over mouth</td>
<td>–0.23 (–0.44, –0.03)</td>
<td>–2.25</td>
<td>0.025</td>
<td>Hand over mouth</td>
<td>0.03 (–0.18, 0.24)</td>
<td>0.29</td>
<td>0.773 (NS)</td>
<td>Hand over mouth</td>
<td>–0.06 (–0.29, 0.16)</td>
<td>–0.54</td>
<td>0.593 (NS)</td>
</tr>
<tr>
<td>Lip trill</td>
<td>–0.49 (–0.70, –0.28)</td>
<td>–4.65</td>
<td>&lt;0.001</td>
<td>Lip trill</td>
<td>–0.44 (–0.66, –0.22)</td>
<td>–4.04</td>
<td>&lt;0.001</td>
<td>Lip trill</td>
<td>0.12 (–0.10, 0.36)</td>
<td>1.07</td>
<td>0.285 (NS)</td>
</tr>
</tbody>
</table>

**Abbreviation:** NS, not significant.
this is unlikely because of the degree of the effected change throughout the sequence in all participants. For instance, the task that showed the most prominent effect in all variables (tube submerged 10 cm into the water) was performed sixth in the sequence. Additionally, lip trills, which were carried out at the end of the sequence, showed the lowest degree of change in two of the three independent variables.

The present study may have some limitations. First, all the participants were diagnosed with hyperfunctional dysphonia, and none of the subjects with hypofunctional dysphonia were included. A different effect could be likely between these two groups. Second, only the short-term effect was assessed. Possible long-term changes remain unknown. Finally, quantitative imaging may be able to provide more accurate information regarding the VLP, pharyngeal width, and A-P compression during semioccluded postures.

CONCLUSION

VLP, A-P laryngeal compression, and pharyngeal width can be modified by semioccluded vocal tract exercises in subjects diagnosed with nonorganic hyperfunctional dysphonia. A low larynx, narrow aryepiglottic opening, and wide pharynx may be reached by using these types of exercises. Phonation into a tube submerged under the water and a stirring straw produce more prominent changes than the other examined semioccluded postures. Loud voice productions also demonstrated a greater degree of change than soft and moderate loudness levels. Two possible explanations arise for these findings: an increase in oral pressure (mechanical effect) and/or a relaxation of the laryngeal and pharyngeal musculature because of semiocclusions. The observed effect is only short term, and the retention of the effect remains unknown.

REFERENCES


