Glottal Adduction and Subglottal Pressure in Singing

*,†Christian T. Herbst, ‡Markus Hess, ‡Frank Müller, *Jan G. Švec, and §,||Johan Sundberg, *Olomouc, Czech Republic, †Wien, Austria, ‡Hamburg, Germany, and §||Stockholm, Sweden

Summary: Previous research suggests that independent variation of vocal loudness and glottal configuration (type and degree of vocal fold adduction) does not occur in untrained speech production. This study investigated whether these factors can be varied independently in trained singing and how subglottal pressure is related to average glottal airflow, voice source properties, and sound level under these conditions. A classically trained baritone produced sustained phonations on the endoscopic vowel [i:] at pitch D4 (approximately 294 Hz), exclusively varying either (a) vocal register; (b) phonation type (from "breathy" to "pressed" via cartilaginous adduction); or (c) vocal loudness, while keeping the others constant. Phonation was documented by simultaneous recording of videokymographic, electroglottographic, airflow and voice source data, and by percutaneous measurement of relative subglottal pressure. Register shifts were clearly marked in the electroglottographic wavegram display. Compared with chest register, falsetto was produced with greater pulse amplitude of the glottal flow, H1-H2, mean airflow, and with lower maximum flow declination rate (MFDR), subglottal pressure, and sound pressure. Shifts of phonation type (breathy/flow/neutral/pressed) induced comparable systematic changes. Increase of vocal loudness resulted in increased subglottal pressure, average flow, sound pressure, MFDR, glottal flow pulse amplitude, and H1-H2. When changing either vocal register or phonation type, subglottal pressure and mean airflow showed an inverse relationship, that is, variation of glottal flow resistance. The direct relation between subglottal pressure and airflow when varying only vocal loudness demonstrated independent control of vocal loudness and glottal configuration. Achieving such independent control of phonatory control parameters would be an important target in vocal pedagogy and in voice therapy.

Key Words: Singing–Subglottal pressure–Airflow–Glottal adduction–Vocal register.

INTRODUCTION

In speech and singing, the quality of the produced sound can be influenced by neuronally controlled muscular adjustments in both vocal tract and larynx. Voice source control can be achieved by a speaker or singer by varying a number of physiological parameters: pitch (ie, the perceptual correlate of fundamental frequency), choice of vocal register (vocal fry, chest/modal, and falsetto), degree of vocal fold adduction, and intended vocal loudness. The last three of these parameters are particularly important to the voice source, that is, to the waveform of the glottal flow. Features like the duration of the closed phase and the amplitude and the skewing of the air flow pulse influence the abruptness of the cessation of glottal air flow, measured in terms of the maximum flow declination rate (MFDR).[1] The MFDR has a direct impact on the slope of the voice source spectrum and thus also the resulting voice timbre.[2,3]

A vocal register is a set or range of serial sounds that are similar in perception and produced by similar vocal fold vibratory patterns.[4] The two main registers for speech and singing are the chest (or modal) register and the falsetto register. Contraction or relaxation of the thyroarytenoid (TA, vocalis) muscle plays an important role for phonation in chest and falsetto registers, respectively.[5] Previous research has suggested that the TA muscle is responsible for adducting the membranous portion of the glottis,[6] and that this maneuver (termed membranous medialization) is related to the choice of vocal register,[7] thus possibly affecting the pressure-flow relationships during phonation.

Vocal fold adduction in the posterior (cartilaginous) portion of the glottis is mainly achieved by the lateral cricoarytenoid (LCA) and the interarytenoid muscles.[8] Insufficient vocal fold adduction leads to incomplete glottal closure and thus the absence of a closed phase during the glottal cycle. Glottal adduction may vary within the phonatory adductory range[9] when choosing a phonation quality along the *Breathy-Flow-Neutral-Pressed* dimension, where *Breathy* is characterized by a DC offset of the glottal flow (ie, the glottal flow never reaches zero during the glottal cycle), and *Pressed* is characterized by a lower peak flow amplitude than *Neutral* and *Flow*.[10] *Flow* phonation is characterized by being produced with the lowest degree of prephonatory glottal adduction that produces vocal fold closure; it is typically used as a baseline phonation type in classical singing and is apparently related to what is referred to as "resonant voice."[11,12] In healthy untrained subjects, incomplete glottal closure and thus a somewhat breathy voice timbre can be considered to be a normal laryngeal configuration,[13] particularly in females[14] and in children.[15]

Variation of vocal intensity can be controlled via two basic mechanisms[16]: at the laryngeal level (via the adduction of the vocal folds) and the respiratory level (by varying lung pressure). Stathopoulos and Sapienza[17] suggest that independent variation of these two control variables is not possible "because of the limitations inherent in natural speech production." This would imply that vocal loudness cannot be varied without introducing changes in the glottal configuration, presumably via changes of

Accepted for publication August 13, 2014.

From the *Voice Research Lab, Department of Biophysics, Faculty of Science, Palacký University Olomouc, Olomouc, Czech Republic; †Laboratory of Bio-Acoustics, Department of Cognitive Biology, University of Vienna, Wien, Austria; ‡Department of Voice, Speech and Hearing Disorders, University Medical Center Hamburg-Eppendorf, University of Hamburg, Hamburg, Germany; §Department of Speech, Music, and Hearing, School of Computer Science and Communication, KTH Voice Research Centre, Stockholm, Sweden; and the ||University College of Music Education Stockholm, Stockholm, Sweden.

Address correspondence and reprint requests to Christian T. Herbst, Voice Research Lab, Department of Biophysics, Faculty of Science, Palacký University Olomouc, 17 listopadu 12, Olomouc 771 46, Czech Republic. E-mail: herbst@ccrma.stanford.edu

Journal of Voice, Vol. 29, No. 4, pp. 391-402
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.08.009

the degree of vocal fold adduction. Such alterations of laryngeal configuration would have an impact on the relationship between subglottal pressure and glottal airflow. In particular, variation of vocal fold adduction would affect the glottal flow resistance,[18–20] which can be estimated in analogy to Ohm's law as the quotient of subglottal pressure divided by the average airflow.[21]

Several studies showed that glottal flow resistance is not kept stable when vocal loudness is changed in normophonic subjects: Stathopoulos and Sapienza[17] found a significant effect of vocal intensity level on laryngeal airway resistance in untrained adult females and males. Isshiki[16] reported that glottal flow resistance increased as a function of vocal loudness for low- and medium-speaking fundamental frequencies in one normophonic adult male. The same study reported no change of glottal flow resistance as a function of vocal loudness in high (falsetto) voice. However, in that case, glottal flow resistance was about twice as high compared with modal voice, suggesting an already maximally adducted glottis. Sulter and Wit[22] reported that glottal flow resistance increased with increasing intensity in both males and females, which they ascribe to "the increment in adductory forces."

This notion is corroborated by an electromyography study conducted by Baker et al,[23] who found that muscular activity in the LCA muscle, ie, a main adductory muscle,[8] tended to increase as a function of vocal loudness in normophonic females. Further support is provided by Södersten et al, who showed with videostrobolaryngoscopic evidence that the degree of posterior adduction varies as a function of vocal intensity in healthy middle-aged women.[14]

In contrast, trained singers seem to be able to vary glottal adduction independently of other parameters.[7,11,24] To the best of our knowledge, the possible independence of intended vocal intensity and adductory settings as phonatory control parameters in singing has not yet been experimentally documented. In this single-subject study, such an attempt is made by means of direct measurement of subglottal pressure and glottal airflow, complemented by simultaneous observation of the laryngeal configuration via endoscopy and videokymography (VKG).

In particular, two questions will be evaluated, related to intentional muscular voice production control and physical voice source parameters: (1) Can vocal loudness be varied independently of glottal configuration, ie, the type (cartilaginous adduction and membranous medialization/vocal register) and the degree of vocal fold adduction? (2) How is subglottal pressure related to average glottal airflow and sound level in singing, when changing (a) cartilaginous adduction; (b) vocal register; and (c) vocal loudness, while keeping the other physiological control parameters constant?

METHODS

Subject and phonatory tasks

A single subject (author C.T.H.) participated in this study. He is a 40-year-old semiprofessional baritone, who would be classified as category 5.5 (local/community–Church Soloist) in the singer taxonomy proposed by Bunch & Chapman.[25] He had served as a subject also in previous investigations, where he demonstrated the ability to reliably and independently control cartilaginous (ie, posterior) adduction and register function.[7,24]

The subject produced a series of phonations where he varied one of three basic phonatory parameters at a time, keeping the other two as constant as possible. The three parameters were (a) vocal loudness (ie, a subjective quantity, mainly representing pulmonary effort), from soft to loud and back; (b) phonation type along the dimension from *Breathy* to *Pressed*, and back; and (c) register, changing from falsetto to chest register, or vice-versa. The latter two parameters were termed cartilaginous (posterior) adduction and membranous medialization, respectively, in a previous publication, and it has been shown that they can be varied independently.[7]

In total, 79 phonations were produced on the endoscopic vowel [i:] at a target pitch of D4 (fundamental frequency approximately 294 Hz). The accuracy of the phonatory maneuvers was assessed by postinspection of the endoscopic video material. The criteria were (a) that the glottal configuration remained unchanged during variation of vocal loudness, and (b) that it matched previous observations as regards the subject's behavior in falsetto and chest register and in *Breathy* versus *Pressed* phonation.[7,24] Based on this analysis, three phonations had to be removed from the data corpus, because the subject accidentally changed the degree of vocal fold adduction when varying vocal loudness. Only periodic or nearly periodic portions of the phonations, classified as type I signals in the Titze system[26] were selected for analysis; for this, signal classification was based on spectrograms obtained from the *Praat* software (www.fon.hum.uva.nl/praat/), applying a window length of 0.1 seconds and a dynamic range of 60 dB.

Data acquisition

Data acquisition was performed at the Department of Phoniatrics and Pedaudiology, University Medical Center Hamburg-Eppendorf, with the setup described in a recent publication.[27] Oral airflow was recorded using a circumferentially vented pneumotachograph system (Glottal Enterprises, Syracuse, NY), which the subject pressed firmly against his face. An ENF-VH endoscope (Olympus Europa SE & Co. KG, Hamburg, Germany) was inserted through an airtight seal in the flow mask[28] so as to allow the acquisition of VKG data during phonation. VKG recordings were made at the point of maximum vocal fold vibratory amplitude along the glottal axis, with a second generation VKG camera system[29] from Cymo, Groningen, The Netherlands, using a Xenon XL-M 180 light source (ILO electronic GmbH, Quickborn, Germany).

Subglottal pressure (P_{Sub}) was recorded with an SPC-320 pressure transducer (Millar Instruments Ltd, Houston, TX) introduced percutaneously via a cannula into the space between the subject's thyroid and cricoid cartilages, about 1.5 to 2 cm below the glottis. The approximate frequency response of this transducer was assessed by comparing it with that of a 1″ B&K 4144 microphone, showing a frequency response within ±0.2 dB in the frequency range of 20 to 2048 Hz.