Immediate and Short-term Effects of Straw Phonation in Air or Water on Vocal Fold Vibration and Supraglottic Activity of Adult Patients with Voice Disorders Visualized with Strobovideolaryngoscopy: A Pilot Study

Abstract

Purpose. The first purpose of this study was to investigate and compare the short-term effects *after* a semi-occluded vocal tract (SOVT) therapy session consisting of straw phonation (SP) in air or water on vocal fold vibration and supraglottic activity of adult patients with voice disorders, visualized with strobovideolaryngoscopy (SVL). The second purpose of this study was to investigate and compare immediate changes in the patients' vocal fold vibration and supraglottic activity *during* SP in air or water, visualized with SVL.

Methods. Twelve adult patients with voice disorders (8 women and 4 men, mean age 52 years) were assigned randomly to one of two study groups: SP in air or SP in water. Immediately before and after a therapy session of 15 min, participants underwent a rigid SVL to determine the short-term effects of the SP session. At the posttherapy examination, flexible SVL while performing SP was added to determine the effects occurring during SP. The visual-perceptual ratings were performed blindly and in random order by three laryngologists, using the Voice-Vibratory Assessment with Laryngeal Imaging rating form for stroboscopy.

Results. *Short-term effects after SP:* After the SP-in-air session, the supraglottic mediolateral compression decreased significantly. The SP-in-water session led to significantly increased left vibrational amplitude. *Immediate effects during SP:* During SP in air, a significantly increased left amplitude and mucosal wave, and significantly decreased mediolateral supraglottic activity, were found. SP in water tended to decrease the vibrational amplitude during performance of the task. A trend toward higher anteroposterior supraglottic compression was observed during both SP in air and water, being more prominent in the latter.

Conclusions. SP in water led to less false vocal fold adduction and consequently less hyperfunction. The small increment in anteroposterior supraglottic activity during SP in air and water might be related to epilarynx narrowing, an economic phenomenon associated with SOVT exercises. The effects on vibrational amplitude were rather ambiguous. The small reduction in amplitude during SP in water is expected to diminish vocal fold impact stress and therefore creates an ideal basis for voice therapy. The increment in amplitude and mucosal wave during SP in air might indicate insufficient supraglottic pressure to obtain the favorable effects of semi-occlusion. Whether or not the rise in amplitude after the SP-in-water session is due to voice efficiency or voice fatigue remains unknown. Future larger-scale investigation in subgroups of voice patients is needed to explore these hypotheses.

Keywords

efficacy, semi-occluded vocal tract therapy, straw phonation, vocal fold vibration, supraglottic activity, strobovideolaryngoscopy

1. Introduction

Semi-occluded vocal tract (SOVT) exercises are promising and widely used to achieve economic and efficient voice use [1-3]. They can be used in voice therapy for patients with dysphonia or as warm-ups for occupational voice users and elite voice performers [4-8]. The common feature of these exercises is a reduction in the cross-sectional area of the vocal tract during phonation [9]. Semi-occlusion can be formed by the articulators or by an assistive device, such as a straw inserted between the lips [1,6]. In the latter case, an artificial lengthening of the vocal tract is achieved [10]. Semi-occluding (and lengthening) the vocal tract increases the supraglottic pressure and inertive reactance, which reduces the phonation threshold pressure and facilitates voice initiation and self-sustained oscillation via a non-linear feedback mechanism [1,5,10-12]. In general, SOVT exercises elicit voice production that relies more heavily on non-linear source-filter interaction than on adductory stress to give the voice acoustic power [6]. Therefore, the exercises lead to more economic phonation, which minimizes voice injury [1-3]. The greater supraglottic pressure achieved by semi-occlusion results in reduced transglottic pressure (difference between sub- and supraglottic pressure) which leads to a relatively small vibrational amplitude, decreased glottal resistance and slightly separated vocal folds [1,8,13-19]. Because the impact stress on the vocal folds diminishes, phonation with high subglottic pressure and high pitch can be achieved with minimal risk of injury to the vocal fold mucosa, making SOVT exercises ideal for voice warm-up [1,8,13,14,15-19]. Moreover, a barely abducted or barely adducted vocal fold configuration is associated with a voice production that is neither breathy nor pressed; and, therefore, it is the target for both patients with laryngeal hyperfunction and those with hypofunction [11,20-22].

SOVT exercises can have either a single (i.e. only the vocal folds) or a double vibratory source. A secondary vibratory source at the distal part of the vocal tract produces fluctuating intraoral pressure that is hypothesized to create a ‘massage-like’ effect on the vocal folds and the vocal tract leading to muscle relaxation, improved blood circulation and more comfort [9,23-25]. For straw phonation (SP), the free end of the straw can be placed into air or water, creating a single or double vibratory source (i.e. the water bubbling), respectively.

Although SP has a long history in voice training and rehabilitation, empirical evidence dates from the last two decades [1,26]. Efficacy studies investigating the immediate effects of SP have focused mainly on aerodynamic, acoustic and electroglottographic (EGG) alterations [4,6,8-10,22,26-42]. Also, MRI and CT studies are available [15,43,44]. Surprisingly, only three studies explored SP-induced adjustments detectible via endoscopy or stroboscopy

[18,29,45]. These techniques are now widely available in clinics and could therefore serve as valuable tools in SOVT assessment and treatment [18]. Dargin et al. [18] investigated immediate changes in laryngeal and pharyngeal activity during SP in air (drinking straw), lip trills and tongue trills in four vocally healthy singers with flexible strobovideolaryngoscopy (SVL). The authors found high variability within and between subjects. SP led to increased mediolateral supraglottic constriction in subject one, no changes in subject two, improved phase closure and increased mucosal wave in subject three, and increased amplitude/mucosal wave and decreased laryngeal height adjustment in subject four. Limitations of this study were the small sample size, as well as the lack of group effects, blinded visual-perceptual ratings and inter/intrarater reliability results. Also, a sequence of three different SOVT exercises was performed in a non-randomized order by each subject, which could have led to a carry-over effect and bias of the results. Guzman et al. [45] also investigated the immediate effects of a non-randomized sequence of eight SOVT exercises (4 SP in air exercises with different straw diameter and lengths, 2 SP in water exercises with different straw diameter, hand-over-mouth and lip trills) with flexible laryngoscopy in 20 patients with hyperfunctional dysphonia. They found lowering of the larynx, widening of the pharynx and narrowing of the epilarynx during all exercises. Visual-perceptual ratings were performed blindly, and both inter (ICC = 0.79) and intrarater (ICC = 0.65 – 0.78) reliability were calculated. Costa et al. [29] did not find significant effects on glottal closure and vestibule constriction after 1 min of SP in air (stirring straw) in 48 subjects with or without vocal fold lesions, determined with rigid laryngoscopy. No information regarding vocal fold vibration was reported in the last two studies as no stroboscopy was used. Furthermore, none of these studies answers the important question of whether SP-induced laryngeal/pharyngeal changes remain when voice production returns to baseline. The study of Costa et al. [29] tried to address this question. However, 1 minute of SP might have been too short to detect any retention effects.

Based on the literature reviewed above, there is a need for randomized clinical trials investigating group effects of SP in air or water on the laryngeal and pharyngeal activity of patients with voice disorders, both during the SOVT execution and after a therapy session, including stroboscopy, using a standardized, reliable, visual-perceptual evaluation protocol, blinded ratings and inter/intrarater reliability results. Therefore, the first purpose of our study was to investigate and compare the short-term effects *after* a SOVT therapy session consisting of SP in air or water on vocal fold vibration and supraglottic activity of adult patients with voice disorders, visualized with SVL. The second purpose of this study was to investigate and compare immediate changes in the patients' vocal fold vibration and supraglottic activity ***during*** SP in air or water, visualized with SVL.

2. Material and methods

This research was approved by the institutional review board of Drexel University College of Medicine (protocol number: 1908007341).

2.1 Participants

Adult patients who had a scheduled voice therapy session at the voice center of Philadelphia Ear, Nose and Throat Associates, Department of Otolaryngology – Head and Neck Surgery, Drexel University College of Medicine, were asked if they were willing to participate in the study. Smoking, pregnancy, mental health conditions, and physically limiting diseases that might interfere with study completion were selected as exclusion criteria. Twelve patients (8 women, 4 men) with a mean age of 52 years (SD 17.1, range 20-77 years) agreed to participate and provided written informed consent. They all had been diagnosed with dysphonia by an experienced otolaryngologist (R.T.S.). Information regarding the participants' demographics and voice-related history is summarized in Table I. There were no significant differences in age (Mann-Whitney U test, $p = 0.469$) or sex (Fishers Exact test, $p > 0.999$) between the two study groups.

2.2 Design

Participants were assigned randomly to one of two study groups: SP in air or SP in water. An online random number generator was used for this group assignment. Immediately before and after the therapy session, participants underwent rigid SVL to determine the short-term effects of the SP session on the vocal fold vibration and supraglottic activity. At the posttherapy examination, flexible SVL while performing SP was added to determine immediate effects occurring during SP.

2.3 SP voice therapy session

Each SP session was guided by one of three experienced, certified speech-language pathologists (B.R., J.P. or P.D.). The session lasted 15 min and consisted of either SP in air or SP in water. Subjects were asked to phonate through a straw in air or water while being instructed by the speech-language pathologist. A regular drinking straw (diameter 5mm, length 21cm) was used for both conditions. A water depth of 2 cm was used for the SP in water session, which was set by drawing a line on the straw. The exercises were performed in sitting position with the head upright. The therapy sessions were similar with respect to voice demand tasks (vowel phonation to continuous speech). The SP protocol is reviewed in Table II. To standardize therapeutic performance, all speech-language pathologists participated in a short training prior to the experiment.

2.4 Strobovideolaryngoscopy (SVL)

The SVL examinations were performed by a laryngology fellow (G.A.) or a senior otolaryngology resident (J.B.). Participants were examined in seated position, with or without administration of topical anesthesia. The KayPentax (Montvale, NJ) laryngeal stroboscope model 9400 was used for all examinations.

A rigid laryngoscope (model: KayPentax 9106) with stroboscopic light was used at both the pre- and posttherapy examinations. Participants were asked to produce a sustained vowel /i/ at a habitual pitch and loudness followed by a low-to-high glissando.

After the rigid SVL of the posttherapy examination, a flexible nasolaryngoscope (model: KayPentax VNL – 1170K) with stroboscopic light was introduced to determine the effects during SP. First, participants were asked to produce a sustained vowel /i/ at habitual pitch and loudness followed by a low-to-high glissando (phonatory condition without straw). After that, participants performed the same phonatory tasks during SP in air or water, depending on their group assignment (phonatory condition with straw). Participants were guided by their speech-language pathologist during these SP tasks.

2.5 Visual-perceptual ratings

After data collection, all video samples of the performed examinations were copied from the medical records, de-identified and randomized. A laryngology fellow (H.P.) and two senior otolaryngology residents (J.R. and M.B.) performed the visual-perceptual ratings using the Voice-Vibratory Assessment with Laryngeal Imaging (VALI) rating form for stroboscopy [46]. The following parameters were evaluated: glottal closure (complete, anterior gap, posterior gap, hourglass, spindle gap, irregular, or incomplete), amplitude (magnitude of lateral movement of the vocal folds, in %), mucosal wave (magnitude of lateral movement of the mucous membrane, in %), vertical level (on-plane, off-plane left lower, or off-plane right lower), nonvibratory portion (adynamic segments of tissue that appear stiff, in %), anteroposterior (AP) and mediolateral (ML) supraglottic activity (constriction of the supraglottic structures, rated 0 – 5 with the aid of concentric circles), free edge contour (normal, convex, concave, irregular, or rough), phase closure (open phase predominates, nearly equal, or closed phase predominates), phase symmetry (the degree of symmetry between the left and the right vocal folds in terms of opening and closing, in %), and regularity (consistency of averaged stroboscopic cycles, in %). For each parameter, the form provides a concise definition and high-quality graphic [46].

All raters were instructed about the use of the VALI form by the same investigator (G.A.) to establish a common understanding of all parameters. After that, each rater viewed and rated the video samples independently. Ratings were completed on a self-paced basis. Ten percent of the samples were randomly repeated to assess intrarater reliability. Audio signals were removed to avoid any bias of the participant's voice quality on the judges' ratings.

2.6 Statistical analyses

SPSS version 27 (SPSS Corporation, Chicago, IL, USA) was used for statistical analyses of the data. Analyses were conducted at $\alpha = 0.05$.

Interrater reliability was determined with a two-way mixed, consistency, average-measures intraclass correlation coefficient (ICC) for the continuous data [47,48]. For the nominal data, Cohen's kappa (K) was run for all rater pairs followed by computing the arithmetic mean of these estimates [47,49]. Intrarater reliability was determined with a two-way mixed, absolute agreement, single-measures ICC for the continuous data and with Cohen's K for the nominal data [47,48]. Only variables with an ICC or K of at least 0.40 (fair agreement, [50,51]) were kept for further analysis.

For the continuous data, the median scores of the three raters were used for further analysis. For the nominal data, the category that was selected by at least two of the three raters was retained for further analysis.

Linear mixed model analyses were performed to compare the groups over time (pre versus post) and phonatory condition (without straw versus with straw) on each continuous outcome measure using the restricted maximum likelihood estimation and scaled identity covariance structure. Time or Phonatory condition, Group, and Time or Phonatory Condition × Group interactions were specified as fixed factors. A random intercept for subjects was included. Model assumptions were checked by inspecting whether residuals were normally distributed. Within-group effects of time or phonatory condition were determined by posthoc pairwise comparisons. Marginal homogeneity tests were used to compare nominal data within groups.

3. Results

3.1 Inter and intrarater reliability

Excellent interrater reliability was found for vertical level (K = 0.90) and supraglottic activity AP/ML (ICC = 0.76/0.85), good interrater reliability was found for mucosal wave right (ICC = 0.68) and phase symmetry (ICC = 0.73), and fair interrater reliability was found for glottal closure (ICC = 0.40), amplitude left (ICC = 0.51) and mucosal wave left (ICC = 0.47). The parameters amplitude right, nonvibratory portion right/left, free edge contour right/left, phase closure, and regularity were excluded from the analysis due to poor interrater reliability (K or ICC < 0.40) [50,51].

Intrarater reliability (K or ICC) ranged from 0.18 to 0.89 for rater 1, from 0.18 to 0.84 for rater 2 and from 0.15 to 0.82 for rater 3. Poor intrarater reliability was found for free edge contour right/left for rater 1, amplitude right and phase closure for rater 2 and non-vibratory portion left, free edge contour right and regularity for rater 3 (K or ICC < 0.40).

3.2 Short-term effects after SP

Short-term effects of the SP-in-air or water session determined with rigid SVL for the continuous data are presented in Table III. Linear mixed model analyses showed no significant Time × Group interactions for any of the outcome parameters, indicating no significant different evolution pre- to posttherapy between the SP-in-air and the SP-in-water group. Within group posthoc tests revealed significantly increased left amplitude in the SP in water group (+8, p = 0.049) and significantly decreased mediolateral supraglottic activity in the SP in air group (-0.5, p = 0.047). Marginal homogeneity tests showed no significant short-term effects of the SP in air or water session for the nominal data (p > 0.05, Table V).

3.3 Immediate effects during SP

Immediate effects during SP in air or water determined with flexible SVL for the continuous data are shown in Table IV. Linear mixed model analyses showed a significant Phonatory condition × Group interaction for amplitude left, mucosal wave left and mediolateral supraglottic activity. Within group posthoc tests revealed significantly increased left amplitude (+8, p = 0.049) and mucosal wave (+11, p = 0.047), and significantly decreased mediolateral supraglottic activity (-0.7, p = 0.045) in the SP-in-air group. Marginal homogeneity tests showed no significant immediate effects of SP in air or water for the nominal data (p > 0.05, Table V).

4. Discussion

First, this study aimed to investigate the short-term effects ***after*** a SOVT therapy session consisting of SP in air or water on vocal fold vibration and supraglottic activity of adult patients with voice disorders, visualized with SVL. Second, immediate changes occurring ***during*** the exercises were determined.

Due to lower transglottic pressure, reduced vibrational amplitude could be expected during SP [1,8,13-19,25]. A trend suggesting this phenomenon was noted in the SP in water group (non-significantly decreased left amplitude: mean decrease from 56% to 49%), which indicates lower vocal fold impact stress that offers promising vocal training and rehabilitation opportunities with minimal risk of phonotraumatic reactions [1,8,13-19]. The current results are consistent with two recent studies using high-speed rigid videolaryngoscopy during water resistance therapy with a flexible silicone tube [19,25]. Guzman et al. [19] found a lower amplitude-to-length ratio (ratio between the vibrational amplitude and the length of the vocal folds) in most of the healthy participants studied, and Laukkanen et al. [25] noted a diminished maximum glottal width and decreased glottal amplitude in a healthy, vocally trained male singer. Lack of significance in the current study might be due to the low sample size and/or the heterogeneity in voice diagnoses. Due to poor interrater reliability, it is unknown whether a similar decrease occurred in the right amplitude.

Surprisingly, opposite results were found for SP in air, with a significantly increased left vibrational amplitude (mean increase from 47% to 55%) and mucosal wave (mean increase from 57% to 68%) during the exercise. It can be hypothesized that a drinking straw in air did not create sufficient supraglottic pressure and got overpowered by the subglottic pressure produced by the participants, resulting in an increased transglottic pressure. Maxfield et al. [6] investigated the intraoral pressure produced by 13 different SOVT exercises. Based on their results, the authors assumed that straws with diameters larger than 3.5 mm in air might not produce sufficient intraoral pressure for efficient and effective voice therapy. Nevertheless, it should be noted that the increase in vibrational amplitude during SP in air was relatively small and remained within the expected range (40 – 60% [18]). Furthermore, the baseline amplitude was lower in this group. Dargin et al. [18] also found an increased amplitude and mucosal wave in one and an increased mucosal wave in two of the four vocally healthy singers during SP with a drinking straw in air.

For voice training and rehabilitation purposes, it is important that positive effects remain during non-SOVT phonation. After the SP in water session, however, the left amplitude was rated significantly higher compared to baseline (mean increment from 52% to 60%). The lower vocal fold impact stress found during the exercise seems not to be transferred to non-SOVT phonation after practicing for 15 min. A higher vibrational amplitude during and after SOVT exercises has been observed in earlier studies [52-54]. An increase in vibrational amplitude is often linked to more respiratory and glottal effort [55] and therefore might be related to voice fatigue. A good marker suggesting risk for voice fatigue seems to be a rise in phonation and collision threshold pressure [56-59]. Such rise has been found by Enflo et al. [60] after 2 min of water resistance therapy (glass tube, l = 27cm, ø = 9mm, 1-2 cm below water surface) in vocally healthy mezzo-sopranos. The authors speculate about whether these findings are a sign of voice fatigue or improved phonatory function. They suggest the latter by comparing their results to those of a vocal loading experiment [59]: the increments after water resistance therapy were less prominent, voice quality improved instead of worsened, and comfort was reported instead of fatigue. More phonatory comfort also has been reported by patients with dysphonia after a water resistance therapy program [24].

Whether the increment in vibrational amplitude after SP in water is a sign of voice fatigue or voice efficiency remains open for debate. Dargin et al. [18] stated that an increment in amplitude within the expected range (40 – 60%) might indicate more efficient vocal fold vibration, which consequently should be the case for the current increment (52 – 60%). Menezes et al. [52] also interpreted the increased vibrational amplitude after 3, 5 and 7 min of lip trill compared to baseline in vocally healthy men as a beneficial phonatory effect. However, voice fatigue cannot be excluded for several reasons: patients with dysphonia were included who might be more vulnerable, a straw was used instead of a tube providing extra flow resistance, and the session lasted 15 min which is a relatively long practice time in the SOVT literature. Laukkanen et al. [25] assume that water resistance therapy may cause symptoms of voice fatigue if performed for an excessively long time. Performance time has empirically been explored for SP in air in dysphonic women [61] and dysphonic children [34]. Both studies suggest an ideal duration between 3 and 5 min, as from 5 min on, some voice parameters deteriorated [34,61]. However, these findings are not fully generalizable to the current study due to differences in straw diameter (1.5 mm versus 5mm). Further, Menezes et al. [52] recommend a maximum of 3 min for vocally healthy women and 5 min for vocally healthy men when performing tongue trills, as sensations of discomfort increased with performance time. In dysphonic women, vocal tension tends to increase and voice quality decreased after 7 min of tongue trills [62]. With these findings in mind, it can be questioned whether practice time was too long in the current study. Nevertheless, it was the specific purpose of this research to investigate the effect of an SOVT therapy session, and 15 min is a realistic duration resembling clinical practice. Also, the increased vibrational amplitude could result from continuation of the rise in subglottic pressure needed to overcome the higher airflow resistance during SP in water [63-65]. The higher the resistance a semi-occlusion offers, the more subglottic pressure a subject needs to start and sustain phonation [27,32]. Subjects in the SP-in-water group perhaps simply maintained that high subglottic pressure after doing so for 15 min. It then seems logical that this was not the case for the other group as subjects needed lower subglottic pressure to overcome the semi-occlusion and resistance of the drinking straw in air [32]. This hypothesis was supported by recent results of Kang et al. [38,41] who found decreased phonation threshold pressure after SP with a drinking straw, implying vocal ease instead of fatigue. Due to the rather ambiguous effects of SP in air and water on the vibrational amplitude of patients with dysphonia, more research on this topic is advisable. Finally, it should be noted that amplitude and mucosal wave covary with pitch and loudness. Although subjects were asked to phonate at a habitual pitch and loudness, it is advisable to control these variables better in future research.

Based on earlier human efficacy and computational modeling studies, epilarynx tube (aryepiglottic) narrowing in the AP dimension is expected to occur during SOVT phonation [1,11,18,45]. A trend suggesting this increase was visible in both groups and was more prominent during SP in water. Guzman et al. [45] also found narrower AP laryngeal compression in the conditions with the highest airflow resistance, i.e. a drinking straw in water and a long stirring straw in air, compared with less resistive SOVT exercises. Although supraglottic compression often is associated with laryngeal hyperfunction, epilarynx tube narrowing can be beneficial for phonation by improving vocal economy (ratio of output to effort) [15,18,45,66-68]. It has been demonstrated that AP laryngeal compression is correlated highly with larynx lowering and pharynx widening, which are favorable phonatory configurations and frequent goals in voice therapy and singing pedagogy [45]. Vertical laryngeal position and pharyngeal area were not evaluated in the current study, but earlier SOVT studies included these measurements. Guzman et al. [45] found a lower larynx and wider pharynx visualized with flexible laryngoscopy during several SOVT exercises (including SP in air and water) in patients with hyperfunctional dysphonia. A lower larynx and wider hypopharynx also were measured with CT during and immediately after SP and tube phonation in air in a male classically trained singer [15]. Case studies of Laukkanen et al. [43] and Guzman et al. [15] reported widening of the lower pharynx over the epilarynx measured via MRI or CT in a healthy vocally trained female subject and a male classically trained singer during and immediately after SP in air. An increased ratio between pharyngeal and epilaryngeal tube opening contributes to the singer's and speaker's formant (i.e. a formant cluster between F3 and F5), which consequently leads to more output without increasing vocal effort, i.e. more vocal economy [15,18,45,66-68].

The ultimate goal of SOVT therapy is to maintain that higher vocal economy during non-SOVT phonation. Titze [1] highlighted the importance of the epilarynx in the process of transfer to non-SOVT phonation. The hypothesis is that the vocalist wants to hold on to the sensation of resistance and back pressure associated with a SOVT and therefore retains some epilarynx narrowing in the transfer process. In other words, the epilarynx could serve as an impedance matcher between the vocal folds and the vocal tract in trained speakers and singers, and SOVT exercises may assist in the awareness of this impedance matching [1,67,69]. Nevertheless, no retention of the rise in AP supraglottic activity after the SOVT session could be documented in the current study. This is not surprising as this transfer process is the result of frequent practice and might not be expected after only one therapy session [1,64]. In this study, no information is available on the ratio between pharyngeal and epilaryngeal tube opening, and therefore no definite conclusions about transfer and impedance matching after therapy can be made.

Although AP supraglottic compression can be a sign of economic voice use, authors and clinicians mostly agree that ML compression is a sign of hyperfunction and should therefore be avoided or eliminated [18,70]. Muscle tension dysphonia was diagnosed in 66.6% (8/12) of the participants with an equal distribution (4/6) between the two groups. Consequently, some ML compression was present at baseline, which seemed more prominent in the SP-in-air group (EM 2.0) than in the SP-in-water group (EM 1.0). A desired and significant decrease in ML compression was found both during and after SP in air. In other words, SP in air led to less false vocal fold adduction. It seemed that for this parameter, transfer to non-SOVT phonation did occur. Oliveira Maia et al. [71] also found less ML laryngeal compression immediately after high-pitched blowing SOVT exercise in women with and without voice complaints.

A similar decrease of ML compression did not occur in the SP in water group. On the contrary, a small non-significant increase was measured during the exercise. It can be hypothesized that this increase was a compensatory response to the extra resistance of the water. Earlier EGG studies suggested an increased contact quotient and consequently potentially higher impact stress during water resistance therapy in vocally healthy and dysphonic subjects [22,32]. Yet, other studies, suggest lower impact stress during water resistance therapy (tube phonation in water) by measuring an increased open quotient or a decreased closed quotient with high-speed videolaryngoscopy in vocally healthy [19,73] and dysphonic [64] subjects. A recent study by Laukkanen et al. [25] found slower closing of the glottis during the exercise in a healthy trained singer, again suggesting lower impact stress. These contradictory results can be due to variability in subjects (healthy versus dysphonic), assessment methods (EGG versus high-speed videolaryngoscopy) and water depths. Furthermore, in the current study, a straw instead of a tube was submerged under water. It is possible that the supraglottic resistance created by the combination of a small diameter and water submersion led to a rise in impact stress, and consequently some increase in ML compression. This is consistent with Guzman et al. [19,22,32] and Laukkanen et al. [55] who found higher closed quotients/smaller open quotients with deeper immersion depths compared with shallower immersion depths, smaller tubes compared with wider tubes, and longer tubes compared with shorter tubes. Independent of the rationale, it is unsure whether this small increase in ML compression during SP in water should be seen as something harmful as it was only temporary and not detectable during non-SOVT phonation after the therapy session. Notwithstanding, the hypothesis that SP in water might lead to less muscle tension due to the 'massage-like' effect of the water bubbling [9,23] was not supported by the current results. However, the smaller baseline values for ML compression in this particular group may not have allowed further decrease for this variable.

SOVT exercises are expected to lead to barely abducted/adducted vocal fold, that is vocal folds that are just touching. Even in 1989, Sovijärvi et al. [73] assumed that water resistance therapy improved vocal fold closure. Improved glottal closure has been achieved by several SOVT exercises, including humming [74], high-pitched blowing [71], tongue trills [52], and lip trills [75]. The current study does not add to this list, as glottal closure remained stable for all subjects both during and after SP, except for one subject in the SP-in-air group (who evolved from an hourglass gap before to an anterior gap after therapy). However, it should be acknowledged that glottal closure was rated as a nominal variable, with a selection of one out of seven categories, as prescribed in the VALI-form [46]. Therefore, it is possible that minimal improvements in glottal closure were not detected. Future studies with larger sample sizes and specific subgroups of patients with a similar baseline glottal gap/closure should clarify this possibility.

While interpreting the current results, it should be noted that parameters were rated differently ('post' condition compared to 'without straw' condition) depending on whether a rigid or flexible endoscope was used. This might be due to the assumption that rigid SVL offers superior image quality and diagnostic ability over flexible SVL [76,77]. This superiority, together with the clinical routine of the researchers, was the reason why a rigid endoscope was used to determine the short-term effects of the SOVT therapy program. As SP can only be performed during flexible SVL, some discrepancy in video ratings for the two purposes of the study due to the use of different scopes cannot be excluded.

Another limitation of this study is that raters were not trained specifically to use the VALI form. The authors of the form [46] suggest a training of 2-2.5 hours with video examples of the end point for each feature. In the current study, the raters were instructed about the use of the VALI form and had access to the definitions and graphics for each parameter, although no video examples were provided. Poor inter and/or intrarater reliability for almost half (7/15) of the parameters might be a consequence of this limited training. However, in the study of Poburka et al. [46], the same number of parameters showed poor inter (glottal closure, vertical level, free edge contour right/left) or intrarater (mucosal wave right, supraglottic activity AP/ML) reliability, despite the form-specific training. The authors found higher reliability for the VALI form when high-speed videoendoscopy was used, which might be of importance for future research. Also, a consensus evaluation between raters could have provided reliable results for the parameters with poor interrater reliability. In general, visual-perceptual ratings remain subjective in nature. Although standardizing the evaluation is crucial, and forms as VALI are needed for both clinical and research aims, visual-perceptual ratings remain extremely difficult.

Final important limitations of the current study are the low sample size and the heterogeneity of voice diagnoses. Of course, the effect of different voice pathologies on voice production is not homogenous [64]. The oscillation pattern at baseline differed for each subject, which makes it difficult to find clear group effects of the interventions. Furthermore, it has been demonstrated previously that subjects react differently to a specific SOVT exercise and increased airflow resistance [4,6,18,19,55,65]. To obtain the most favorable therapeutic effects, it appears that the semi-occlusion must provide airflow resistance similar to that produced by the glottis itself, which in turn will depend on the voice pathology [6]. Therefore, future larger-scale efficacy studies on subgroups of patients are needed to establish diagnosis-specific SOVT treatment.

5. Conclusions

In conclusion, this study is unique as it is the first study that uses a randomized clinical trial to investigate group effects of SP in air or water on vocal fold vibration and supraglottic activity, determined by SVL. A trend of reduced vibrational amplitude was noticed during SP in water, which suggests decreased vocal fold impact stress and therefore creates an ideal basis for voice training and therapy. During SP in air, however, vibrational amplitude and mucosal wave increased, which might possibly indicate insufficient supraglottic pressure to obtain the favorable effects of semi-occlusion. After the therapy session, vibrational amplitude did not differ from baseline in the SP-in-air group but rose in the SP-in-water group. Whether or not this rise is due to voice efficiency or voice fatigue remains open for debate. Due to the rather ambiguous effects of SP in air and water on the vibrational amplitude of patients with dysphonia, more research on this topic is advisable.

SP in air led to a significantly reduced supraglottic mediolateral compression both during the execution and after the therapy session, indicating less false vocal fold adduction and consequently less hyperfunction. A trend toward higher anteroposterior supraglottic compression, which is expected to be due to desired epilarynx narrowing, was visible both during SP in air and SP in water, being more prominent in the latter. This finding was not noted after the session, which suggests lack of transfer. Future research should include a larger sample size and focus on specific subgroups of patients. Further, high-speed imaging and a consensus evaluation between raters might increase the validity of the visual-perceptual assessment. Importantly, long-term effects of SOVT therapy on vocal fold vibration and supraglottic activity require further attention to help guide SOVT protocols and therapy duration, and to define better the efficacy of SOVT.

References

1. Titze IR. Voice training and therapy with a semi-occluded vocal tract: Rationale and scientific underpinnings. *J Speech Lang Hearing Res.* 2006;49:448-459.
2. Gaskill CS, Quinney DM. The effect of resonance tubes on glottal contact quotient with and without task instruction: A comparison of trained and untrained voices. *J Voice.* 2006;26:e79-e93.
3. Croake DJ, Andreatte RD, Stemple JC. Immediate effects of the vocal function exercises semi-occluded mouth postures on glottal airflow parameters: A preliminary study. *J Voice.* 2017;31:245.e9-245.e14.
4. Dargin TC, Searl J. Semi-occluded vocal tract exercises: aerodynamic and electroglottographic measurements in singers. *J Voice.* 2015;29(2):155-164.
5. Kapsner-Smith MR, Hunter EJ, Kirkham K, et al. A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols. *J Speech Lang Hearing Res.* 2015;58:535-549.
6. Maxfield L, Titze I, Hunter E, Kapsner-Smith M. Intraoral pressures produced by thirteen semi-occluded vocal tract gestures. *Logop Phonatr Vocol.* 2015;40(2):86-92.
7. Mills R, Hays C, Al-Ramahi J, Jiang JJ. Validation and evaluation of the effects of semi-occluded face mask straw phonation therapy methods on aerodynamic parameters in comparison to traditional methods. *J Voice.* 2017;31:323-328.
8. Smith SL, Titze IR. Characterization of flow-resistant tubes used for semi-occluded vocal tract voice training and therapy. *J Voice.* 2017;31(1):113.e1-113.e8.
9. Andrade AP, Wood G, Ratcliffe P, et al. Electroglottographic study of seven semi-occluded exercises: LaxVox, straw, lip-trill, tongue-trill, humming, hand-over-mouth, and tongue-trill combined with hand-over-mouth. *J Voice.* 2014;28(5):589-595.
10. Conroy ER, Henninck TM, Awan SN, et al. Effect of variations to a simulated system of straw phonation therapy on aerodynamic parameters using excises canine larynges. *J Voice.* 2014;28(1):1-6.
11. Titze IR, Verdolini Abbott K. *Vocology: The Science and Practice of Voice Habilitation.* Salt Lake City, UT: National Center for Voice and Speech, 2012.
12. Guzman M, Jara R, Olavarria C, et al. Efficacy of water resistance therapy in subjects diagnosed with behavioral dysphonia: a randomized controlled trial. *J Voice.* 2017;31(3):385.e1-385.e10.
13. Titze IR. Phonation into a straw as a voice building exercise. *J Singing.* 2000;57(1):27-28.
14. Titze IR. How to use the flow-resistant straws. *J Singing.* 2002;58(5):429-430.
15. Guzman M, Laukkanen AM, Krupa P, et al. Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. *J Voice.* 2013;27(4):523.e19-523.e34.
16. Ogawa M, Hosokawa K, Yoshida M, et al. Immediate effects of humming on computed electroglottographic parameters in patients with muscle tension dysphonia. *J Voice.* 2014;28(6): 733-741.
17. Hampaia V, Laukkanen AM, Guzman MA, et al. Vocal fold adjustment caused by phonation into a tube: a double-case study using computed tomography. *J Voice.* 2015;29(6):733-742.
18. Dargin TC, DeLaunay A, Searl J. Semioccluded vocal tract exercises: changes in laryngeal and pharyngeal activity during stroboscopy. *J Voice.* 2016;30(3):377.e1-377.e9.
19. Guzman M, Laukkanen AM, Traser L, et al. The influence of water resistance therapy on vocal fold vibration: a high-speed digital imaging study. *Logop Phonatr Vocology.* 2017;42(3):99-107.
20. Verdolini K, Druker DG, Palmer PM, Samawi H. Laryngeal adduction in resonant voice. *J Voice.* 1998;12(3):315-327.
21. Berry DA, Verdolini K, Montequin DW, et al. A quantitative output-cost ratio in voice production. *J Speech Lang Hearing Res.* 2001;44:29-37.
22. Guzman M, Calvache C, Romero L, et al. Do different semi-occluded voice exercises affect vocal fold adduction differently in subjects diagnosed with hyperfunctional dysphonia. *Folia Phoniatr Logo.* 2015;67:68-75.
23. Guzman M, Acuña G, Pacheco F, et al. The impact of double source of vibration semioccluded voice exercises on objective and subjective outcomes in subjects with voice complaints. *J Voice.* 2018;32(6):770.e1-770.e9.

24. Meerschman I, Van Lierde K, Ketels J, et al. Effect of three semi-occluded vocal tract therapy programmes on the phonation of patients with dysphonia: lip trill, water-resistance therapy and straw phonation. *Int J Lang Commun Disord.* 2019;54(1):50-61.
25. Laukkanen AM, Geneid A, Bula V, et al. How much loading does water resistance voice therapy impose on the vocal folds? An experimental human study. *J Voice.* 2020;34(3):387-397.
26. Story BH, Laukkanen AM, Titze IR. Acoustic impedance of an artificially lengthened and constricted vocal tract. *J Voice.* 2000;14(4):455-469.
27. Titze IR, Finnegan EM, Laukkanen AM, Jaiswal S. Raising lung pressure and pitch in vocal warm-ups: the use of flow-resistant straws. *J Singing.* 2002;58(4):329-338.
28. Sampaio M, Oliveira G, Behlau M. Investigation of the immediate effects of two semi-occluded vocal tract exercises. *Pro Fono.* 2008;20(4):261-266.
29. Costa CB, Costa LHC, Oliveira G, Behlau M. Immediate effects of the phonation into a straw exercise. *Braz J Otorhinolaryngol.* 2011;77(4):461-465.
30. Guzman M, Higueras D, Fincheira C, et al. Immediate acoustic effects of straw phonation exercises in subjects with dysphonic voices. *Logoped Phoniatr Vocol.* 2013;38(1):35-45.
31. Guzman M, Rubin A, Muñoz D, Jackson-Menaldi C. Changes in glottal contact quotient during resonance tube phonation and phonation with vibrato. *J Voice.* 2013; 27(3):305-311.
32. Guzman M, Castro C, Madrid S, et al. Air pressure and contact quotient measures during different semioccluded postures in subjects with different voice conditions. *J Voice.* 2015;30(6):759.e1-759.e10.
33. Robieux C, Galant C, Lagier A, et al. Direct measurement of pressures involved in vocal exercises using semi-occluded vocal tracts. *Logoped Phoniatr Vocol.* 2015;40(3):106-112.
34. Ramos LA, Gama ACC. Effect of performance time of the semi-occluded vocal tract exercises in dysphonic children. *J Voice.* 2017;31(3):329-335.
35. Mills RD, Rivedal S, DeMorett C, et al. Effects of straw phonation through tubes of varied lengths on sustained vowels in normal-voiced participants. *J Voice.* 2018;32(3):386.e21-386.e29.
36. Kang J, Scholp A, Tangney J, Jiang JJ. Effects of a simulated system of straw phonation on the complete phonatory range of excised canine larynges. *Eur Arch Otorhinolaryngol.* 2019;276(2):473-482.
37. Kang J, Xue C, Chou A, et al. Comparing the exposure-response relationships of physiological and traditional vocal warm-ups on eerodynamic and acoustic parameters in untrained singers. *J Voice.* 2019;33(4):420-428.
38. Kang J, Xue X, Lou Z, et al. The therapeutic effects of straw phonation on vocal fatigue. *Laryngoscope.* 2020;130(11):E674-E679.
39. Tangney J, Scholp A, Kang J, et al. Effects of varying lengths and diameters during straw phonation on an excised canine model. *J Voice.* 2021;35(1):85-93.
40. Manjunatha U, Nayak PS, Bhat JS. Can straw phonation be considered as vocal warm up among speech language pathologists? *J Voice.* In press.
41. Kang J, Xue C, Piotrowski D, et al. Lingering effects of straw phonation exercises on aerodynamic, electroglottographic, and acoustic parameters. *J Voice.* In press.
42. Keltz A, McHenry M. Steam and/or semi-occluded vocal tract exercise as morning vocal warm-up strategy. *J Voice.* In press.
43. Laukkanen AM, Horáček J, Krupa P, Švec J. The effect of phonation into a straw on the vocal tract adjustments and formant frequencies. A preliminary MRI study on a single subject completed with acoustic results. *Biomed Signal Proc.* 2012;37(2):75-82.
44. Guzman M, Miranda G, Olavarria C, et al. Computerized tomography measures during and after artificial lengthening of the vocal tract in subjects with voice disorders. *J Voice.* 2017;31(1):124.e1-124.e10.
45. Guzman M, Castro C, Testart A, et al. Laryngeal and pharyngeal activity during semioccluded vocal tract postures in subjects diagnosed with hyperfunctional dysphonia. *J Voice.* 2013;27(6):709-716.
46. Poburka BJ, Patel RR, Bless DM. Voice-vibratory assessment with laryngeal imaging (VALI) form: Reliability of rating stroboscopy and high-speed videoendoscopy. *J Voice.* 2017;31(4):513.e1-513.e14.

47. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8(1):23-34.
48. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropractic Med.* 2016;15:155-163.
49. Light RJ. Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin.* 1971;76(5):365-377.
50. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
51. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment.* 1994; 6(4):284-290.
52. Menezes MH, de Campos Duprat A, Olival Costa H. Vocal and laryngeal effects of voiced tongue vibration technique according to performance time. *J Voice.* 2005;19(1):61-70.
53. Schwarz K, Cielo CA. Vocal and laryngeal modifications produced by the sonorous tongue vibration technique. *Pró-Fono Revista de Atualização Científica.* 2009;21(2):161-166.
54. Cordeiro GF, Montagnoli AN, Tsuji DH. Comparison among phonation of the sustained vowel /ε/, lip trills, and tongue trills: The amplitude of vocal fold vibration and the closed quotient. *Otolaryngology.* InTech, Rijeka, Croatia, 2012.
55. Laukkanen AM, Pulakka H, Alku P, et al. High-speed registration of phonation-related glottal area variation during artificial lengthening of the vocal tract. *Logop Phonatr Vocol.* 2007;32:157-164.
56. Vilkman E, Lauri ER, Alku P, et al. Effects of prolonged oral reading on F_0, SPL, subglottal pressure and amplitude characteristics of glottal flow waveforms. *J Voice.* 1999;13(2):303-312.
57. Solomon NP, DiMattia MS. Effects of a vocally fatiguing task and systemic hydration on phonation threshold pressure. *J Voice.* 2000;14(3):341-362.
58. Chang A, Karnell MP. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *J Voice.* 2004;18(4):454-466.
59. Enflo L, Sundberg J. Vocal fold collision threshold pressure: An alternative to phonation threshold pressure? *Logop Phonatr Vocology.* 2009;34:210-217.
60. Enflo L, Sundberg J, Romedahl C, McAllister A. Effects on vocal fold collision and phonation threshold pressure of resonance tube phonation with tube end in water. *J Speech Lang Hearing Res.* 2013;56:1530-1538.
61. Paes SM, Behlau M. Dosage dependent effect of high-resistance straw exercise in dysphonic and non-dysphonic women. CoDAS. 2017;29(1):e20160048.
62. Menezes MHM, Ubrig-Zancanella MT, Cunha MGB, et al. The relationship between tongue trill performance duration and vocal changes in dysphonic women. *J Voice.* 2011;25(4):e167-e175.
63. Radolf V, Laukkanen AM, Horáček J, Liu D. Air-pressure, vocal folds vibration and acoustic characteristics of phonation during vocal exercising, part 1: measurement in vivo. *Eng Mech*. 2014; 21:53-59.
64. Echternach M, Raschka J, Kuranova L, et al. Immediate effects of water resistance therapy on patients with vocal fold mass lesions. *Eur Arch Oto-Rhino-L.* 2020;277:1995-2003.
65. Echternach M, Köberlein M, Gellrich D, et al. Duration of biodynamic changes associated with water resistance therapy. *Logop Phoniatr Vocol.* 2020;25:1-8.
66. Sundberg J. Articulatory interpretations of the singing formant. *J Acoust Soc Am.* 1974;55:838-844.
67. Titze IR, Story BH. Acoustic interactions of the voice source with the lower vocal tract. *J Acoust Soc Am.* 1997;101:2234-2243.
68. 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. *Logoped Phoniatr Vocol.* 2012;37:75-82.
69. Titze IR, Laukkanen AM. Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logoped Phonatr Vocol.* 2007;32:147-156.
70. Stager SV, Bielamowicz SA, et al. Supraglottic activity: evidence of vocal hyperfunction or laryngeal articulation. *J Speech Lang Hearing Res.* 2000;43(1):229-238.

71. Oliveira Maia MA, Oliveira Maia M, Côrtes Gama AC, Behlau M. Immediate effects of the high-pitched blowing vocal exercise. *J Soc Bras Fonoaudiol.* 2012;24(1):1-6.
72. Granqvist S, Simberg S, et al. Resonance tube phonation in water: High-speed imaging, electroglottographic and oral pressure observations of vocal fold vibrations - A pilot study. *Logoped Phoniatr Vocol*. 2015;40(3):113-21.
73. Sovijärvi A, Häyrinen R, et al. Äänifysiologisten kuntoutusharjoitusten ohjeita [Instructions for voice exercises]. Helsinki: Publications of Suomen Puheopisto, 1989.
74. De Bodt MS, Clement G, et al. The impact of phonation mode and vocal technique on vocal fold closure in young females with normal voice quality. *J Voice.* 2012;26:818e1-818e4.
75. Nam IC, Kim SY, et al. Effects of voice therapy using the lip trill technique in patients with glottal gap. *J Voice.* 2019;33(6):949.e11-949.e19.
76. Romak JJ, Heuer RJ, Hawkshaw MJ, Sataloff RT. In Sataloff RT. Professional Voice: The Science and Art of Clinical Care, 4th edition. San Diego, CA: Plural Publ; 2017:405-438.
77. Boles RW, Gao WZ, et al. Flexible versus rigid laryngoscopy: A prospective, blinded comparison of image quality. *J Voice.* In press.

Table I: participants' demographics and voice-related history

Group	Subject	Sex	Age	Occupation	Singing	Voice disorder/diagnosis	Hours of prior voice therapy	Familiar with SP prior to the experiment
SP in air	1	M	66	dentist	non-professional	reflux laryngitis Reinke's edema bilateral benign masses bilateral vocal fold paresis muscle tension dysphonia glottic insufficiency	11	yes
	2	F	55	teacher elementary school	no singing	left vocal fold cystic mass Reinke's edema laryngopharyngeal reflux right partial vocal fold paresis glottic insufficiency	4	no
	3	M	65	orthodontist	no singing	Reinke's edema laryngopharyngeal reflux multiple masses on left vocal fold left vocal fold paresis muscle tension dysphonia	2	no
	4	F	20	musical theater student	professional	right vocal fold cystic mass reflux laryngitis fluctuating vocal fold paresis R>L muscle tension dysphonia	7	yes
	5	F	45	physical therapist	non-professional	right vocal fold paresis bilateral benign masses left posterior granuloma sulcus vocalis glottic insufficiency	3	yes
	6	F	77	singer	professional	right VF weakness bilateral sulci reflux laryngitis bilateral local Reinke's edema	23	yes

						left benign vocal fold mass muscle tension dysphonia		
SP in water	7	F	55	singer voice teacher opera director	professional	right vocal fold mass edema Reinke's edema left vocal fold muscle tension dysphonia glottic insufficiency	19	yes
	8	F	36	teacher middle school	non-professional	reflux laryngitis Reinke's edema right vocal fold cyst alternating hypomobility (R>L) muscle tension dysphonia	10	no
	9	F	35	teacher middle school voice-over/ commercial representative	non-professional	left vocal fold mass bilateral vocal fold scar and stiffness laryngopharyngeal reflux unilateral vocal fold paresis Reinke's edema muscle tension dysphonia glottic insufficiency	3	yes
	10	M	37	motivational speaker pianist and singer	professional	reflux laryngitis left vocal fold cyst muscle tension dysphonia glottic insufficiency	53	yes
	11	M	65	courier hospital singer	professional	bilateral vocal fold masses bilateral sulci Reinke's edema right vocal fold scar laryngopharyngeal reflux presbylaryngis left vocal fold paresis glottic insufficiency	0	no
	12	F	66	nurse	non-professional	reflux laryngitis Reinke's edema left vocal fold paresis glottic insufficiency	4	no

Table II: SP protocol

1.	Introducing eutonic posture and costo-abdominal breathing
2.	Costo-abdominal breathing through the straw without phonation (3 times)
3.	SP on /o/ vowel: habitual pitch and loudness (3 times)
4.	SP on /o/ vowel with phonatory breaks on one exhalation: /o/-/o/-/o/
5.	SP on /o/ vowel: glissando (3 times) – low – high – high – low – low – high – low – high – low – high
6.	SP while "reading" phrases – using prosodic patterns without articulation – followed by reading the same phrases without the straw *Why not?* *Where are you going?* *Why not one?* *What time is it?* *Would you like one?* *What's your name?* *What a wonderful world.* *Whatever you want.* *Walk through the woods.* *Wherever you want to go.*
7.	SP while "reading" the *Rainbow Passage* – using prosodic patterns without articulation – followed by reading the same text without the straw
8.	SP while "singing" *Happy Birthday* – using prosodic patterns without articulation – followed by singing the same song without the straw
9.	SP during "spontaneous speech" – answering questions *(e.g. What are your plans for the weekend?)* – using prosodic patterns without articulation – followed by telling the same story without the straw

Table III: Short-term effects *after* the SP in air or water session on vocal fold vibration and supraglottic activity: continuous data (rigid SVL results)

		Time				Linear Mixed Model		
		Pre		Post		**Time**	**Group**	**Time × Group**
Outcome	**Group**	*EM*	*95%CI*	*EM*	*95%CI*	*p-value*	*p-value*	*p-value*
Amplitude left (%)	Air	54	[43, 65]	56	[45, 67]			
	Water	52	[41, 62]	60	[50, 70]	0.133	0.900	0.338
Mucosal wave right (%)	Air	52	[31, 73]	59	[38, 81]			
	Water	52	[31, 73]	50	[29, 71]	0.581	0.718	0.400
Mucosal wave left (%)	Air	66	[50, 82]	64	[48, 80]			
	Water	50	[35, 65]	55	[40, 70]	0.779	0.193	0.516
Supraglottic activity AP	Air	1.7	[0.3, 3.0]	1.8	[0.5, 3.2]			
	Water	1.5	[0.1, 2.9]	1.3	[0.0, 2.7]	>0.999	0.696	0.605
Supraglottic activity ML	Air	2.0	[1.1, 2.9]	1.5	[0.6, 2.2]			
	Water	1.0	[0.1, 1.9]	0.8	[0.0, 1.8]	0.110	0.181	0.401
Phase symmetry (%)	Air	78	[51, 105]	74	[47, 101]			
	Water	73	[48, 98]	70	[45, 95]	0.654	0.784	0.967

EM: estimated mean, CI: confidence interval, AP: anteroposterior, ML: mediolateral

*****: indicates a significant fixed effect ($p < 0.05$)

: indicates a significant within-group effect: posthoc pairwise comparison 'pre' vs 'post' ($p < 0.05$)

Table IV: Immediate effects *during* SP in air or water on vocal fold vibration and supraglottic activity: continuous data (flexible SVL results)

		Phonatory condition				**Linear Mixed Model**		
		Without straw		With straw		**Phonatory condition**	**Group**	**Phonatory conditon × Group**
Outcome	**Group**	*EM*	*95%CI*	*EM*	*95%CI*	*p-value*	*p-value*	*p-value*
Amplitude left (%)	Air	47	[39, 54]	55	[47, 63]			
	Water	56	[48, 64]	49	[41, 57]	0.712	0.836	0.037*****
Mucosal wave right (%)	Air	63	[49, 77]	67	[53, 81]			
	Water	66	[51, 81]	66	[51, 81]	0.729	0.908	0.729
Mucosal wave left (%)	Air	57	[51, 64]	68	[61, 74]			
	Water	71	[64, 79]	67	[60, 75]	0.253	0.102	0.024*****
Supraglottic activity AP	Air	1.5	[0.8, 2.2]	1.8	[1.1, 2.6]			
	Water	1.4	[0.6, 2.2]	2.0	[1.2, 2.8]	0.225	0.929	0.724
Supraglottic activity ML	Air	2.2	[1.3, 3.0]	1.5	[0.7, 2.3]			
	Water	1.2	[0.3, 2.1]	1.6	[0.7, 2.5]	0.550	0.439	0.035*****
Phase symmetry (%)	Air	77	[58, 96]	76	[56, 97]			
	Water	74	[53, 95]	65	[44, 86]	0.510	0.557	0.541

EM: estimated mean, CI: confidence interval, AP: anteroposterior, ML: mediolateral

*****: indicates a significant fixed effect ($p < 0.05$)
: indicates a significant within-group effect: posthoc pairwise comparison ‘without straw’ vs ‘with straw’ ($p < 0.05$)

Table V: Results of the nominal data for each subject in each condition (pre and post SP session, without or with straw)

			Time		**Marginal homogeneity test**	**Phonatory condition**		**Marginal homogeneity test**
Outcome	**Group**	**Subject**	Pre	Post	*p-value*	Without straw	With straw	*p-value*
Glottal closure	Air	1	hourglass	anterior gap	0.317	complete	/	>0.999
		2	hourglass	hourglass		posterior gap	posterior gap	
		3	/	/		complete	complete	
		4	hourglass	hourglass		posterior gap	posterior gap	
		5	/	incomplete		posterior gap	/	
		6	hourglass	hourglass		complete	/	
	Water	7	hourglass	hourglass	>0.999	hourglass	/	>0.999
		8	posterior gap	posterior gap		posterior gap	posterior gap	
		9	posterior gap	posterior gap		posterior gap	posterior gap	
		10	spindle gap	spindle gap		/	/	
		11	complete	complete		complete	complete	
		12	posterior gap	posterior gap		posterior gap	/	
Vertical level	Air	1	on-plane	on-plane	>0.999	on-plane	on-plane	>0.999
		2	on-plane	on-plane		on-plane	on-plane	
		3	/	/		on-plane	on-plane	
		4	on-plane	on-plane		on-plane	/	
		5	on-plane	on-plane		on-plane	on-plane	
		6	on-plane	on-plane		on-plane	on-plane	
	Water	7	on-plane	on-plane	>0.999	on-plane	on-plane	>0.999
		8	on-plane	on-plane		on-plane	on-plane	
		9	on-plane	on-plane		on-plane	on-plane	
		10	on-plane	on-plane		/	/	
		11	/	on-plane		on-plane	on-plane	
		12	on-plane	on-plane		on-plane	on-plane	

/: missing data due to insufficient visibility and/or disagreement between raters (three different categories rated)