Published in final edited form as:

J Voice. 2022 January ; 36(1): 27–33. doi:10.1016/j.jvoice.2020.02.024.

Retest reliability for complete airway interruption systems of aerodynamic measurement

Jim R. Lamb, Sarah A. Schultz, Austin J. Scholp, Emily R. Wendel, Jack J. Jiang

University of Wisconsin School of Medicine and Public Health, Department of Surgery, Division of Otolaryngology-Head and Neck Surgery, Madison, WI 53792

Abstract

Introduction: Measures of subglottal pressure (Ps), phonation threshold pressure (PTP), and laryngeal resistance (LR) can be used as indicators of vocal cord disorders. The gold standard non-invasive measurement uses labial interruption, which has been shown to have reliability inconsistencies. Mechanical interruption methods have demonstrated promise in measurement reliability. The goal of the present study is to compare retest reliability of labial and mechanical interruption methods.

Methods: 55 subjects aged 18–69 participated. 10 trials were performed for each method. For labial interruption, subjects produced five labial plosives at comfortable and quiet volumes. For mechanical interruption, subjects produced a sustained /α/ while a balloon valve interrupted phonation five times. 30 subjects completed a second study visit identical to the first approximately two weeks (15 days $\pm$ 3.76) after the first visit. Ps, PTP, mean airflow rate (MFR), and LR were determined for each subject and retest reliability for each was analyzed.

Results: The percent difference in measurement results for test-retest of Ps were 12.88% $\pm$ 10.15 for mechanical interruption and 27.56% $\pm$ 17.14 for labial interruption ($p = 0.0003$). The percent difference for PTP measurements were 21.46% $\pm$ 16.01 for mechanical and 17.04% $\pm$ 14.62 ($p = 0.3372$) for labial. Intra-subject coefficients of variation of Ps were 0.086 $\pm$ 0.046 for mechanical and 0.161 $\pm$ 0.078 for labial ($p < 0.0001$). For PTP, the coefficients were 0.177 $\pm$ 0.083 for mechanical and 0.186 $\pm$ 0.091 for labial ($p = 0.5402$). Lastly, for LR (Ps divided by MFR) the percent differences were 14.33% $\pm$ 10.06 for mechanical and 53.87% $\pm$ 43.19 for labial ($p < 0.0001$) with intra-subject variability of 0.115 $\pm$ 0.050 for mechanical and 0.287 $\pm$ 0.222 for labial ($p < 0.0001$).

Conclusions: Ps and LR measured using mechanical interruption showed more consistency for both retesting across separate study visits and intra-subject variability. PTP was similar in retesting and intra-subject variability. Continued work to improve mechanical interruption techniques is

Corresponding author: Jack J. Jiang, MD, PhD, 1300 University Ave, 2725 Medical Sciences Center, Madison, WI 53706, 608-265-9854, jjjiang@wisc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of Interest: Corresponding author Jack Jiang held a patent on the mechanical interruption device used in the present study. That patent has since expired and he has not received any financial benefit from it.

warranted as these methods offer higher reliability and consistency than the labial interruption methods.

Keywords

Retest reliability; Complete interruption; Subglottal pressure; Phonation threshold pressure; Laryngeal resistance; Mean airflow rate

1 - Introduction

Based on a 2012 study of a U.S. healthcare claims database, the American population has a dysphonia prevalence rate of 0.98% for the entire population, with the prevalence ranging from 0.6% for children ages 0–9 up to 2.5% for people age 70+ [1]. This indicates a need for capable diagnostic methods. Vocal pathologies are typically diagnosed with methods such as videostroboscopy. However, these types of visual assessment can be subjective [2]. Aerodynamic assessment offers quantitative measurements to perform more objective diagnosis. Complete aerodynamic assessment includes measures such as subglottal pressure (Ps), phonation threshold pressure (PTP), mean flow rate (MFR), and laryngeal resistance (LR).

Subglottal pressure is the driving force behind phonation and PTP is the minimum Ps required to initiate vocal fold vibration [3]. Prior studies have shown that subjects with vocal pathologies such as vocal fatigue [4], benign mass lesions [5], and vocal fold dehydration [6] have higher PTP than healthy controls. Thus, a device capable of measuring these changes would have clinical relevance. Ps and PTP can be reliably measured through invasive methods such as tracheal puncture [7] or a transnasally inserted miniature pressure transducer catheter into the trachea [8], as well as noninvasive methods such as labial plosives [9], and mechanical interruption [10].

Currently, the labial method is the gold standard for clinical pressure measurements. This method requires the subject to perform labial plosives. The subject repeatedly says /pi/ or /pa/ and during the stop phase of /p/, the intraoral pressure equilibrates with Ps [9]. Ps is measured while speaking at a comfortable volume and PTP is measured while the subject phonates as quietly as they can, but without whispering. While this method is accurate [11], some studies indicate difficulty obtaining consistent measurements for pressure [12] and MFR [13]. The alternative method proposed in this study is a variation of the mechanical airflow interruption method presented by Hoffman *et al.* [14]. The mechanical interruption method uses a balloon valve to interrupt airflow and measure the subglottal pressure through a series of pressure equilibrations at the time of interruption. Previous studies performed by our lab have directly compared the labial method to mechanical methods with adults [15] as well as similar measurement reliability studies with children [16, 17]. The pediatric studies found that the reliability of pressure measurements was similar for mechanical and labial methods [16], but measurements of MFR and LR were more consistent for mechanical methods [17]. In Chapin *et al.* [15], mechanical interruptions were manually controlled and applied at 500 millisecond intervals. In contrast, our study randomly applied mechanical interruptions at a shortened interval of 250 milliseconds. Their study also briefly looked at

retest reliability. Ten subjects participated in a retest portion, in which 1–2 hours passed between taking two sets of measurements. Our study has extended the retest interval to two weeks, so that measurements are not being taken on the same day as each other.

Laryngeal resistance (LR) is a derived measurement that is obtained by dividing P_s by MFR. LR measurements are able to provide information on the geometric and biomechanical properties of the larynx and vocal folds [18]. Other glottal factors such as vocal fold adduction and roundness of the glottal entry and exit can affect the resulting LR [19]. Because LR is a derived measurement, it can be obtained by methods that measure both P_s and MFR, such as labial or mechanical methods. The above Chapin study found that intra-subject variability when measuring MFR was higher for the labial method than the mechanical method [15]. Thus, intrasubject variability was also present in the LR calculations.

This study aims to measure P_s , PTP, MFR, and LR using labial and mechanical interruption methods in adult subjects. The goal is to compare not only intra-subject variation for each of the parameters, but the test-retest reliability of the measurement methods. Previous comparisons have shown similarity in intra-subject variability for pressure measurements, but with mechanical interruptions having lower variability for MFR and LR measures. These studies have also found a difference in the P_s and PTP measures. This study will test if P_s and PTP measures will remain similar across a retest study. The difference in PTP measurement comes with a caveat because labial methods measure onset PTPT and mechanical methods measure offset PTP, with onset expected to be higher than offset [20]. This study will differ from previous retest reliability studies in two important ways. First, while previous studies have performed retest reliability between multiple sessions conducted in the same day, this study will wait approximately two weeks before retesting subjects [15]. Second, the balloon interruptions for the mechanical methods have been automated and shortened.

2 - Methods

2.1 - Subjects

This study was approved by the Institutional Review Board of the University of Wisconsin – Madison. Subjects provided informed written consent to participate. Fifty-five subjects (30 females, 25 males) who were at least 18 years old (mean age: 22.5 ± 5.0 years) were recruited via flyers. Each subject had no history of dysphonia and was a nonsmoker. Subjects performed labial and mechanical interruption. Each method was performed 10 times. Approximately two weeks (average time: 15.0 ± 3.8 days) after the first study activities, 30 subjects (18 females, 12 males; age 23.5 ± 10 years) came back for a second visit to perform the exact same methods as the first visit.

2.2 - Data Collection

All data were collected using a custom airflow interruption device similar to the device used in previous studies [16]. Data were recorded with a custom LabVIEW program. The device was made of a polyvinyl chloride (PVC) tube embedded with various sensors. An AKG

CK77 condenser microphone (Harman International Industries, Inc., Stamford, CT) was connected to a MP13 preamplifier (Rolls Co., Murray, UT) to record audio signals. Pressure and airflow measurements were obtained using a pneumotach amplifier (Series 1110, Hans Rudolph, Inc.). Pressure was measured through one port, while airflow was measured with two ports across a known resistance. A balloon valve (Series 9340, Hans Rudolph, Inc.) was attached to the distal end of the tube. A USB-6341 data acquisition board (National Instruments, Inc., Austin, TX) was used to sample data at 10 kHz. A digital flow meter (FMA-1610A, Omega Engineering, Stamford, CT) was used to calibrate dynamic airflow data and pressure was zeroed to the room's atmospheric pressure prior to data collection. A computer monitor provided visual feedback of SPL and airflow to help the subjects maintain consistency during trials (Figure 1).

2.3 - Labial Interruption

For labial interruptions, a 0.066 inch diameter intraoral tube was inserted into the pressure measurement port of the PVC tube. The distal end rested inside the subject's lips, sitting just on top of the tongue. A silicone mask was attached to the end of the PVC tube for the subject to place their face into. To measure Ps, subjects said /pɑ/ at a comfortable volume. For measuring PTP, subjects now said /pɑ/ as quietly as they could, but without going into a whisper. For each pressure measurement, 10 trials were performed, with the subject saying the plosive five times within each trial.

2.4 - Mechanical Interruption

During mechanical interruption, there was no intraoral tube present and the pressure measurement was taken from the port within the PVC tube, rather than inside of the subject's mouth. A silicone mouthpiece was attached to the distal end of the PVC tube. During airflow interruptions, subjects were instructed to plug their nose to prevent a pressure loss from air leaking out. Subjects produced a sustained /ɑ/ at a comfortable volume for approximately 10 seconds. During the 10 second trial, the balloon valve would inflate five times for 250 milliseconds to briefly cut off airflow. The measurement of Ps were taken at 150 milliseconds. The interruptions were automated by a computer that works on the order of gigahertz to ensure that interruption times were consistent in length. The 250 millisecond interruption is shorter than the 500 milliseconds reported previously [15], the measurement of Ps is taken at 150 milliseconds [14], indicating that the interruption shortening does not alter our measurement.

2.5 - Data analysis

Data were analyzed through custom LabVIEW and MATLAB programs. For the mechanical interruptions, Ps was measured as the pressure 150 milliseconds after interruption onset. PTP was calculated to be the difference between Ps and the pressure at the time of phonation offset (Figure 2). MFR and SPL were averaged during the intervals of constant phonation between trials. LR was then calculated as the mean Ps divided by the MFR for the given trial. For the labial interruptions, Ps and PTP were measured as the maximum pressure measured prior to phonation of each plosive for each respective trial type (Figure 3). MFR and SPL were averaged during the phonation of /pɑ/. LR was again calculated as the average Ps or PTP divided by the MFR for the entire trial.

2.6 - Statistical Analysis

Percent difference for each metric was calculated using the formula:

$$\% \text{ difference} = \frac{x_1 - x_2}{(x_1 + x_2)/2} \cdot 100\%.$$

This is taking the difference of the values divided by the average of the values and multiplying by 100%. Percent difference was used here in lieu of percent error because the results of the first visit or the second visit could be the “true” value. Means and coefficients of variations were calculated for all 55 subjects at the first study visit. Coefficients of variation were calculated by taking the metric mean and dividing by the standard deviation. To compare the results of each testing method, paired t-tests were performed with a significance level of $\alpha = 0.05$. Using the 30 subjects who completed both study visits, an intra-class correlation coefficient (ICC) for retest reliability was calculated for each variable using SPSS v25. The following scale was used to rank ICCs: Under 0.500 was poor reliability, between 0.500 and 0.700 was moderate reliability, between 0.700 and 0.900 was good reliability, and over 0.900 was excellent reliability.

3 - Results

Mean PTP ($p < 0.0001$), MFR ($p = 0.0066$), and LR ($p = 0.0056$) varied across methods, unlike mean Ps ($p = 0.0740$) where mechanical and labial interruption had similar values (Table 1). Coefficients of variation for Ps ($p < 0.0001$), MFR ($p = 0.0004$), and LR ($p < 0.0001$) were lower for mechanical interruption than labial. The same was also true of percent difference for Ps ($p = 0.0003$), MFR ($p = 0.0045$), and LR ($p < 0.0001$). Measures for PTP were not different between mechanical and labial for both coefficient of variation and percent difference. Boxplots portraying the spread of the data are found in Figure 4. Results of reliability testing with the ICC are shown in Table 2. Mechanical Ps (0.755), MFR (0.845), and LR (0.684) all had good reliability, while mechanical PTP (0.655) had moderate reliability. Labial PTP (0.837) and MFR (0.717) had good reliability, but Ps (0.595) and LR (0.593) only had moderate reliability. The percent differences of SPL for mechanical ($2.32\% \pm 2.46$) and labial ($4.51\% \pm 2.60$) methods were small enough to indicate that subjects were vocally consistent across the two study visits. Thus, other aerodynamic comparisons can be made.

4 - Discussion

The goal of the present study was to analyze retest reliability of labial and mechanical interruption methods for measuring aerodynamic parameters over a time period of two weeks. This study agreed with previous studies and has also generated new findings. No significant differences were seen between the mean values of Ps, but significant differences were found in the variance of measurements of Ps, LR, and MFR. In an ideal scenario, both the percent difference and coefficient of variation would be zero, indicating the measurement was the exact same every single time. The smaller the variance measurements get, the more precision the measurement has.

The percent difference results and intra-subject coefficients of variation show the scale of differences in measurement methods. For measurements of Ps, MFR, and LR, the mechanical method had a significantly lower coefficient of variation than the labial method. This means that trial-to-trial variation during a given study visit was lower when performing measurements with the mechanical method. For the same three variables, the percent difference was also significantly lower for the mechanical method. Percent difference was used to compare the variability between study visits. Mechanical methods again proved to be more consistent. For PTP measurements, neither coefficient of variation or percent difference was better for either method.

Mechanical interruption produced superior intra-class correlation coefficient results for all metrics except PTP. This result is in agreement with the rest of our findings. Surprisingly, there was a discrepancy in ICC for PTP measurements, with the mechanical parameter being below the labial parameter. An outlier in our subject pool accounted for one-third of this variation in the percent difference measure for mechanical PTP (Figure 4). The cause of this subject's retest variability is unknown. When this subject was removed from the data set, the ICC for mechanical PTP became 0.779 (range 0.581–0.890) and labial PTP became 0.837 (range 0.682–0.920). Removing this outlier caused the mechanical interruption value to noticeably increase, but did not affect the labial measure. Thus, when the data were analyzed without the outlier the numbers better aligned with the remainder of the study. However, this discrepancy should be noted for future studies.

Due to the measurement devices and their corresponding techniques, the variability of MFR measurements with labial interruption were higher than that of mechanical interruption. Thus, LR also had high variability when measuring with labial interruption, as LR is a measurement derived from Ps and MFR. When performing the labial task, subjects said /pα/ into an anesthesia mask. Getting a perfect, airtight seal of the mask to the subject's face was difficult. A lack of a seal could affect airflow measurements, leading to higher variability. Pressure measurements were unaffected by this, as the pressure measurement occurred inside the subject's mouth via an intraoral tube for labial interruption.

Although the pressure measurements themselves were unaffected by mask seal, higher variability was still seen in the labial method for Ps measurements. This variability likely stemmed from human inconsistencies, as labial interruptions were subject controlled, whereas mechanical interruptions were machine controlled. Even with visual feedback, it was difficult for subjects to consistently create a labial plosive at the same vocal intensity every single time. Subjects showed much less variability in vocal intensity during the mechanical trials when the computer controlled the interruptions.

The most surprising finding of this study was that PTP remained consistent during the labial interruption methods. As previously stated, PTP is typically an inconsistent measure due to the difficulties associated with labial interruption. However, in this study, PTP was the only metric where labial measurements had a smaller percent difference of measurement between the two study visits. Although the PTP percent difference was lower, the coefficient of variation for intra-subject variability was still higher for labial interruption.

It is important to recognize that labial and mechanical interruption methods are measuring different forms of PTP. While the labial method measures onset PTP, the mechanical method measures offset PTP. This observation explains the mechanical method's significantly lower mean PTP, as sustaining phonation requires less energy than initiating it [20]. Even these fundamental differences between mechanical and labial interruption did not lead to any significant differences in variation or percent difference. With the exception of the difference in mean measurement, no other conclusions can be made about the relationship between mechanical and labial interruption from this study. Future research into the hysteresis relationship of onset to offset PTP is warranted.

5 - Conclusion

Mechanical interrupters continue to be more reliable and more consistent than their labial counterparts, when measuring aerodynamic vocal parameters. Thus, future studies should promote the use of mechanical aerodynamic measurement techniques in clinical settings. The differences in measurement of laryngeal resistance found in this study indicate the need for further research into the diagnostic capabilities of this measure. Additionally, future work should aim to determine the relationship between onset and offset phonation threshold pressure.

Acknowledgements

This study was funded by NIH grant number R01 DC008153 from the National Institute on Deafness and other Communicative Disorders.

References

- [1]. Cohen SM, Kim J, Roy N, Asche C, and Courey M. 2012. "Prevalence and causes of dysphonia in a large treatment-seeking population." *Laryngoscope* 122 (2): 343–348. 10.1002/lary.22426. [PubMed: 22271658]
- [2]. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, Van De Heyning P, Remacle M, and Woisard V. 2001. "A basic protocol for functional assessment of voice pathology especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques." *Eur Arch Otorhinolaryngol* 258: 77–82. 10.1007/s004050000299. [PubMed: 11307610]
- [3]. Titze IR. 1992. "Phonation threshold pressure: A missing link in glottal aerodynamics." *J Acoust Soc Am* 91 (5): 2926–2935. 10.1121/1.402928. [PubMed: 1629485]
- [4]. Solomon NP, and DiMattia MS. 2000. "Effects of a Vocally Fatiguing Task and Systemic Hydration on Phonation Threshold Pressure." *J Voice* 14 (3): 341–362. 10.1016/S0892-1997(00)80080-6. [PubMed: 11021502]
- [5]. Jiang JJ, Stern J, Chen H, and Solomon NP. 2004. "Vocal Efficiency Measurements in Subjects with Vocal Polyps and Nodules: A Preliminary Report." *Ann Otol Rhinol Laryngol* 133:277–282. <https://doi.org/10.1177%2F000348940411300404>.
- [6]. Verdolini K, Titze IR, and Druker DG. 1990. "Changes in Phonation Threshold Pressure with Induced Conditions of Hydration." *J Voice* 4(2):142–151. 10.1016/S0892-1997(05)80139-0.
- [7]. Isshiki N. 1964. "Regulatory Mechanism of Voice Intensity Variation." *J Speech Hear Res* 7: 17–29. 10.1044/jshr.0701.17. [PubMed: 14130741]
- [8]. Kitzing P, and Lofqvist A. 1975. "Subglottal and oral air pressures during phonation - preliminary investigation using a miniature transducer system." *Med Bio Eng* 13: 644–648. 10.1007/BF02477320. [PubMed: 1186324]

- [9]. Smitheran JR, and Hixon TJ. 1981. "A clinical methods for estimating laryngeal airway resistance during vowel production." *J Speech Hear Disord* 46: 138–146. 10.1044/jshd.4602.138. [PubMed: 7253590]
- [10]. Bard MC, Slavit DH, McCaffrey TV, and Lipton RJ. 1992. "Noninvasive technique for estimating subglottic pressure and laryngeal efficiency." *Ann Otol Rhinol Laryngol* 101: 578–582. <https://doi.org/10.1177%2F000348949210100706>. [PubMed: 1626903]
- [11]. Lofqvist A, Carlborg B, and Kitzing P. 1982. "Initial validation of an indirect measure of subglottal pressure during vowels." *J Acoust Soc Am* 72 (2): 633–635. 10.1121/1.388046. [PubMed: 7119286]
- [12]. Hertegard S, Gauffin J, and Lindestad P. 1995. "A Comparison of Subglottal and Intraoral Pressure Measurements during Phonation." *J Voice* 9 (2): 149–155. 10.1016/S0892-1997(05)80248-6. [PubMed: 7620537]
- [13]. Hoffman MR, Rieves AL, Surender K, Devine EE, and Jiang JJ. 2013. "Evaluation of Auditory and Visual Feedback for Airflow Interruption." *J Voice* 27 (2): 149–154. 10.1016/j.jvoice.2012.10.002. [PubMed: 23280384]
- [14]. Hoffman MR, Baggott CD, and Jiang JJ. 2009. "Reliable Time to Estimate Subglottal Pressure." *J Voice* 23 (2): 169–174. 10.1016/j.jvoice.2007.09.005. [PubMed: 18207358]
- [15]. Chapin WJ, Hoffman MR, Rieves AL, and Jiang JJ. 2011. "Comparison of Labial and Mechanical Interruption for Measurement of Aerodynamic Parameters." *J Voice* 25 (3): 337–341. 10.1016/j.jvoice.2010.01.004. [PubMed: 20189755]
- [16]. Hoffman MR, Scholp AJ, Hedberg CD, Lamb JR, Braden MN, McMurray JS, and Jiang JJ. 2019. "Measurement Reliability of Phonation Threshold Pressure in Pediatric Subjects." *Laryngoscope* 129: 1520–1526. 10.1002/lary.27418. [PubMed: 30408173]
- [17]. Scholp AJ, Hedberg CD, Lamb JR, Hoffman MR, Braden MN, McMurray JS, and Jiang JJ. 2019. "Measurement Reliability of Laryngeal Resistance and Mean Flow Rate in Pediatric Subjects." *J Voice* In press. 10.1016/j.jvoice.2019.02.005.
- [18]. Netsell R, Lotz WK, DuChane AS, and Barlow SM. 1991. "Vocal Tract Aerodynamics during Syllable Productionss: Normative Data and Theoretical Implications." *J Voice* 5 (1): 1–9. 10.1016/S0892-1997(05)80157-2.
- [19]. Scherer RC, and Guo CG. 1990. "Effect of vocal fold radii on pressure distributions in the glottis." *J Acoust Soc Am* 88 (S1): S150. 10.1121/1.2028673.
- [20]. Lucero JC. 1999. "A theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offsett." *J Acoust Soc Am* 105(1):423–431. 10.1121/1.424572. [PubMed: 9921668]

Figure 1.

Schematic of experiment. The set-up is mainly the same for mechanical and labial interruptions. For mechanical interruptions, a silicon mouthpiece is used. For labial interruptions, a facemask is used and an intraoral tube extends from the tube's pressure port into the subject's mouth.

Figure 2.

Example of analysis from a mechanical interruption. Audio and pressures traces are shown. T_1 is the beginning of the interruption. T_2 is the time of the audio cut-off and P_1 is subtracted from P_s to measure PTP . T_3 is a 75 millisecond marker. T_4 marks 150 milliseconds and is where the supraglottal pressure is measured and equilibrated to subglottal pressure.

Figure 3.

Example of analysis for a labial interruption. SPL (sound pressure level), audio, and pressure traces are shown. T_1 is the start of the labial interruption. T_2 is where the maximum pressure occurs prior to phonation. For trials at comfortable volume, this represents P_s , and for trials at quietest possible volume, this represents PTP. (The dotted black line near the top is not relevant to this analysis)

Figure 4.

A series of boxplots comparing percent differences and coefficients of variation (CoV) across mechanical and labial methods. Top Left) Pressure CoV; Top Right) Pressure percent difference; Bottom Left) MFR and LR CoV; Bottom Right) MFR and LR percent difference. Each box represents the data between the first (q1) and third (q3) quartiles and the solid line in each box represents the median. Each marker outside the whiskers represents an outlier, determined as being greater than $q3 + 2.7 * (\text{standard deviation}) * (q3 - q1)$. There were only outliers above the upper bound.

Table 1.

Summary statistics for each method and variable. Values presented as mean $\pm$ standard deviation. MFR and LR mean and CoV have 54 subjects because there was a problem with airflow data collection for one subject; the pressure readings were unaffected.

	Parameter	n	Mechanical	Labial	t value	P value
	Mean (cmH ₂ O)	55	6.18 $\pm$ 1.56	6.81 $\pm$ 2.45	-1.822	0.0740
Ps	CoV (unitless)	55	0.086 $\pm$ 0.046	0.162 $\pm$ 0.077	-6.330	<0.0001 **
	% Difference (unitless)	30	12.88 $\pm$ 10.15	27.56 $\pm$ 17.14	-4.132	0.0003 **
	Mean (cmH ₂ O)	55	2.52 $\pm$ 0.99	3.49 $\pm$ 1.23	-4.497	<0.0001 **
PTP	CoV (unitless)	55	0.177 $\pm$ 0.083	0.185 $\pm$ 0.090	-0.616	0.5402
	% Difference (unitless)	30	21.46 $\pm$ 16.01	17.04 $\pm$ 14.62	0.976	0.3372
	Mean (L/min)	54	13.49 $\pm$ 5.31	11.15 $\pm$ 5.05	2.831	0.0066 **
MFR	CoV (unitless)	54	0.122 $\pm$ 0.054	0.225 $\pm$ 0.196	-3.762	0.0004 **
	% Difference (unitless)	30	16.82 $\pm$ 12.24	44.38 $\pm$ 43.87	-3.079	0.0045 **
	Mean (cmH ₂ O/L/min)	54	0.508 $\pm$ 0.232	1.187 $\pm$ 1.724	-2.886	0.0056 **
LR	CoV (unitless)	54	0.115 $\pm$ 0.050	0.287 $\pm$ 0.222	-5.400	<0.0001 **
	% Difference (unitless)	30	14.33 $\pm$ 10.06	53.87 $\pm$ 43.19	-4.912	<0.0001 **

* indicates significant difference at $\alpha = 0.05$

** indicates significant difference at $\alpha = 0.01$

CoV = coefficient of variation; Ps = subglottal pressure; PTP = phonation threshold pressure; MFR = mean flow rate; LR = laryngeal resistance

Table 2.

Intra-class correlation coefficients for each method and variable. The lower and upper bounds indicate the 95% confidence interval. All calculations were done using SPSS v25.

Parameter	Mechanical			Labial		
	ICC	Lower	Upper	ICC	Lower	Upper
Ps	0.755 [#]	0.547	0.875	0.595 [^]	0.305	0.784
PTP	0.655 [^]	0.391	0.820	0.837 [#]	0.685	0.919
MFR	0.845 [#]	0.701	0.923	0.717 [#]	0.485	0.854
LR	0.684 [#]	0.434	0.836	0.445 [^]	0.107	0.691

Reliability classifications: Excellent (ICC > 0.900);

[#] Good (0.900 > ICC > 0.700);

[^] Moderate (0.700 > ICC > 0.500); Poor (ICC < 0.500)

Ps = subglottal pressure. PTP = phonation threshold pressure. MFR = mean flow rate. LR = laryngeal resistance.