The Laryngoscope

Using High-Speed Videoendoscopy to Analyze Laryngeal Closure Parameters During Normal Swallow

Rebecca J. Howell, MD; Amna S. Mira, MS, CF-SLP; Andres Llico, PhD; Victoria S. McKenna, CCC-SLP, PhD

Objective: This pilot study was designed to test the tolerability of a lower scope position and feasibility of custom-designed MATLAB graphical user interface (GUI) used to analyze playback review of laryngeal high-speed videoendoscopy (laryngeal HSV) during healthy volitional dry swallows. We hypothesized this method would conceptually provide time resolution for glottic closure events compared with standard (30 frames per second, fps), and enable a means to measure timing, sequence, and duration of laryngeal movements during swallowing not otherwise visualized.

Methods: A total of 14 healthy adults (4 male, 22–80 years) participated. We performed laryngeal HSV at 500fps. Measurements included: (i) feasibility and tolerability of the procedure; (ii) identification of a swallowing segment of interest (SOI) for the peak of the swallow; and (iii) description of laryngeal swallowing movements using a GUI.

Results: Fourteen subjects tolerated the procedure without discomfort and swallow images were able to be analyzed in 12. Using our GUI, mean SOI was 260 ms, yielding 130 frames for analysis (compared with seven in standard laryngoscopy). Vocal fold adduction, vocal fold medialization, and anterior–posterior arytenoid compression to the epiglottis prior to whiteout could be identified and sequenced.

Conclusion: Participants tolerated a low position of the endoscope during dry volitional swallows. The output of our GUI demonstrated a novel technique for identifying, describing, and sequencing a swallowing SOI. Future studies should investigate laryngeal closure and arytenoid positioning with a bolus and in a range of ages, genders, and etiologies in both healthy and abnormal populations to better understand swallowing physiology.

Key Words: deglutition, HSV, Laryngeal high-speed videoendoscopy, normal swallow, swallowing physiology.
Level of Evidence: NA

***Laryngoscope*, 00:1–7, 2024**

INTRODUCTION

Qualitative video analysis is the fastest and simplest review of executed motion too quick or too complex for the human eye, such as playback or stop-action viewing used widely in voice science. The aim of quantitative analysis is a detailed examination of movement patterns.[1,2] High-speed videoendoscopy (HSV) refers to the electronic recording of optical still images representing motion in contrast to strobed images (sequential sampling of pictures in rapid succession *perceived* as motion).[2,3] Laryngeal HSV has been well-described in visualization of mucosal wave patterns in the study of voice production.[2] Laryngeal HSV has also been used to categorize other quick laryngeal motions such as the kinematics in supraglottic phonation, abduction time, and vocal fold adduction in cough (albeit to a lesser extent).[3–5]

The only use to date of laryngeal HSV in swallowing describes visual recognition of anatomic structures during the pharyngeal phase of deglutition during flexible endoscopic evaluation of swallowing (FEES) in healthy volunteers by Aghdam et al.[6] We propose that the use of laryngeal HSV could visualize the critical timed sequence in glottic and supraglottic closure and arytenoid movement that would further our understanding of swallowing physiology. In this pilot study, we investigated glottic closure patterns during volitional dry swallows using a novel paradigm with flexible laryngeal HSV and placement of the scope in a low position toward the petiole of the epiglottis. We hypothesized that this paradigm would be well-tolerated and conceptually could provide time resolution for glottic closure events compared with standard clinical frame rate providing visualization of specific laryngeal action to allow for identification using a novel MATLAB graphical user interface (GUI).

METHODOLOGY

Participants

A total of 14 healthy adult participants (4 males, 10 females) with a mean age of 37.2 years (SD = 17.2),

From the Department of Otolaryngology-Head and Neck Surgery (R.J.H., A.S.M.), University of Cincinnati, Cincinnati, Ohio, U.S.A.; Department of Communication Sciences and Disorders (A.S.M., V.S.M.), University of Cincinnati, Cincinnati, Ohio, U.S.A.; College of Applied Medical Sciences (A.S.M.), King Saud Bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia; Department of Biomedical Engineering (A.L.), University of Cincinnati, Cincinnati, Ohio, U.S.A.; and the School of Communication Sciences and Disorders (V.S.M.), University of Central Florida, Orlando, Florida, U.S.A.

Editor's Note: This Manuscript was accepted for publication on November 21, 2024.

Presented as Triological Candidate Thesis, Technology and Procedure Development.

The authors have no funding, financial relationships, or conflicts of interest to disclose.

Send correspondence to Rebecca J. Howell, Department of Otolaryngology Head & Neck Surgery, University of Cincinnati, Cincinnati, OH; Email: howellrb@ucmail.uc.edu

DOI: 10.1002/lary.31945

ranging from 22 to 80, were recruited for the study. Participants reported no current speech, language, hearing, voice, or swallowing problems (please see Table I for participant demographic information). Participants did not have any history of neurological conditions, and all were on regular diet with thin liquids. All participants scored within normal limits on the Voice Handicap Index (defined as score <30).[7] All participants scored within normal limits on the Eating Assessment Tool-10 (defined as score ≤2).[8] All participants further completed a standard clinical FEES (Olympus OTV-S190; PENTAX, ENF-VH) under halogen light by a certified speech-language pathologist. Trials of thin liquid (5, 10, 15 mL) and puree (5 mL) were assessed to verify normal swallowing ability, with all participants being classified as having normal swallowing physiology.

Participants were prospectively enrolled from the Spring to the Summer of 2023. Participants provided informed consent prior to participation (IRB no. 2022-0686).

Procedures

First, oxymetazoline was sprayed in the nasal passage and topical viscous lidocaine (2%) was applied to the nose using a cotton-tipped applicator. Lidocaine was not sprayed into the nose or applied to the throat to minimize theoretical impact on swallowing sensation. Next, a flexible endoscope (PENTAX, FNL10-RP3) was passed through the nasal cavity and into the nasopharynx to visualize the larynx. Participants were then instructed to "swallow your saliva" for up to five dry swallows, while the scope was placed in an inferior position toward the petiole of the epiglottis beyond the usual home position (Fig. 1). Video images were captured using a high-speed camera (Phantom, VEO-340 L; resolution of 512 × 512 pixels) at a frame rate of 500 frames per second (fps), allowing for a time resolution of 1 frame/2 ms. This frame rate was chosen because it has been shown to adequately quantify laryngeal adductory gestures of cough and fricative ab/adduction.[4,9] This frame rate allowed for up to 20 recordings for each participant, with each having a recording duration of ~9 s. The duration of each recording was determined by the resolution, frame rate, internal camera memory, and number of recordings chosen.

Measures

Tolerability and feasibility. Our analysis included reporting the total number of participants who tolerated the scope placement (tolerability) as well as the total number of participants whom we were able to adequately view their swallows (feasibility). In our study, *tolerability* refers to subjects' ability to complete examination tasks while having the scope positioned at the petiole of the epiglottis by self-report by individual subjects. *Feasibility* indicates the researcher's ability to view the laryngeal patterns before the whiteout using playback video recordings.

Swallow segment of interest. Playback video recordings were inspected by two speech-language pathologists independently and cropped to each specific swallowing event (VM, AM). We identified this as the swallow *segment of interest* (SOI), defined as the frames from the first adduction of the vocal folds, via identification of arytenoid movement medially, to the time of complete light reflection.[10] Next, these swallow events were entered into a custom-designed MATLAB graphical user interface (GUI). The MATLAB GUI was designed by the third author (AL), where he designed a special algorithm that allows researchers to analyze the high-speed video frame by frame and tag each laryngeal behavior (Fig. 2). Postprocessing of the data allowed for researchers to individually visualize and label each frame of the recorded

TABLE I.
Participant Demographic Information for all 14 Participants.

Subject ID	Age (Yrs)	Sex	Reflux*	VHI-Functional	VHI- Physical	VHI-Emotional	VHI-Total	EAT-10
P_01	36	Female	Yes	0	0	0	0	0
P_02	67	Female	Yes	2	0	0	2	1
P_03	80	Female	No	0	0	0	0	0
P_04	26	Female	No	0	0	0	0	0
P_05	22	Female	No	0	1	0	1	0
P_06	27	Male	Yes	0	0	0	0	0
P_07	28	Female	Yes	0	0	0	0	0
P_08	52	Male	No	0	0	0	0	0
P_09	24	Female	Yes	0	2	0	2	1
P_10	52	Female	Yes	0	0	0	0	0
P_11	30	Male	Yes	0	0	0	0	2
P_12	24	Male	Yes	0	0	0	0	0
P_13	24	Female	Yes	0	0	0	0	1
P_14	30	Female	No	0	0	0	0	0

EAT-10 = Eating Assessment Tool-10; ID = identification numbers; VHI = Vocal Handicap Index.
*Self-reported reflux symptoms.

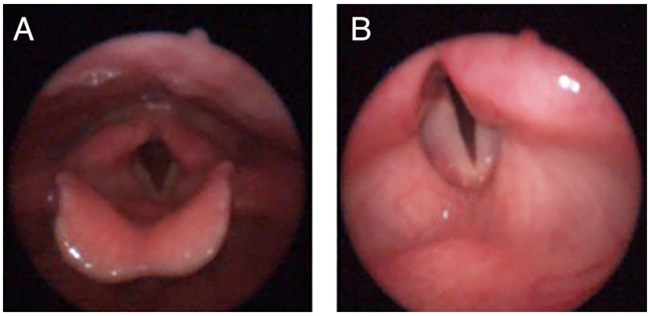

Fig. 1. (A). Example of the typical scope "home" position for FEES. (B). Example of the new, lower placement for laryngeal assessment during the swallow. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

swallow using the MATLAB GUI. This measure was reported in total number of frames. The output of the GUI allowed a graphic depiction of the number of frames for each laryngeal movement. Next, we provided a comparison between the number of frames provided during the swallow SOI from the high-speed camera to a calculation of the number of frames captured with a traditional clinical low-speed system at 30 fps.

Laryngeal swallowing movements. Using our custom-designed MATLAB GUI, we completed an analysis of identification of discrete swallow movements to investigate the timing of the observed swallowing events. Swallow events were identified and defined by the research team, which included two speech-language pathologists and one laryngologist.

RESULTS

Tolerability and Feasibility

A total of 14 participants tolerated the procedure for 100% tolerability. Coughing and/or gagging during the procedure were minimal and did not impact the ability to perform volitional dry swallows. In this study, we could not visualize laryngeal behaviors at the height of the swallow in 15% of participants (2 of 14). In the two cases that it was not possible, the scope was not positioned correctly either due to deflection out of the laryngeal vestibule (likely by the epiglottis) or due to rocking movement of the larynx. Of those participants who we could analyze, one to three swallows per participant could be analyzed using the swallow movements definitions described below. Therefore, feasibility was 86% or 12 of 14 participants.

Swallow Segment of Interest (SOI)

The *SOI* was defined as the height of the swallow, which included the last actions of (1) vocal fold medialization, (2) anterior–posterior movement, and (3) the whiteout. *Whiteout* (WO) is defined as the light reflection into the camera. The swallowing SOI was calculated for one dry swallow from each of the 12 participants whose swallows could be fully visualized. The average duration of the swallow SOI was 130 frames (SD = 78 frames, range = 70–317 frames), which is consistent with an average of 260 ms (SD = 157 ms) per swallow (Fig. 3). Calculation of the same SOI using a lower sampling rate comparable to clinical tools would have yielded an average of 7.8 frames (SD = 4.7 frames, range = 4–19 frames) for the same swallowing SOI. Figure 4 provides frame snapshots from a portion of the swallow SOI using 500 fps with HSV, whereas the same images at a simulated down-sampled rate at 30 fps (representing the current standard) would result in only two images.

Laryngeal Swallowing Movements

Using a single subject for feasibility of our custom MATLAB GUI, we identified discrete swallow movements and defined those movements for specific laryngeal gestures (Fig. 5).

Definitions:

1. Vocal fold (VF) abduction (Ab)—VF and arytenoids moving away from midline.
2. Vocal fold (VF) medialization (M)—The movement of VF and arytenoids towards the midline.
3. Vocal fold (VF) adduction full (AdFull)—The true VF are fully closed.
4. Arytenoid adduction (ARad) full—The arytenoids are fully closed, even if the true VF are NOT.
5. Anterior–Posterior (AP) movement (Mo)—AP movement (coming together or apart).
6. Anterior–Posterior (AP) compression full (Cfull)—the full AP closure between arytenoids and the petiole of the epiglottis.

We were able to visually identify distinct laryngeal gestures of arytenoid adduction and complete arytenoid closure, VF medialization and complete VF closure, initiation of anterior–posterior compression of the laryngeal surface of the epiglottis to the arytenoid, and the point of AP compression when view of the VFs are lost. Due to the scope position, we could not visualize epiglottic inversion but could infer the timing of epiglottic deflection based on scope displacement and ending in the whiteout (WO).

Additional unanticipated laryngeal movements were visualized and sequenced prior to the SOI defined as the preparation phase. A visual representation of laryngeal swallow movements during the preparatory phase and SOI output of the custom MATLAB GUI can be seen for the participant in Figure 5.

DISCUSSION

The current understanding of swallowing physiology can most reliably be quantified using either videofluoroscopy (VFSS) or flexible endoscopic evaluation of swallow (FEES). On VFSS, the anatomic location of the bolus defines penetration versus aspiration, but the true vocal folds are not actually visualized. Best practice in VFSS is captured at 30 fps.[11] Alternatively, the use of flexible endoscopes utilizes a continuous light source with

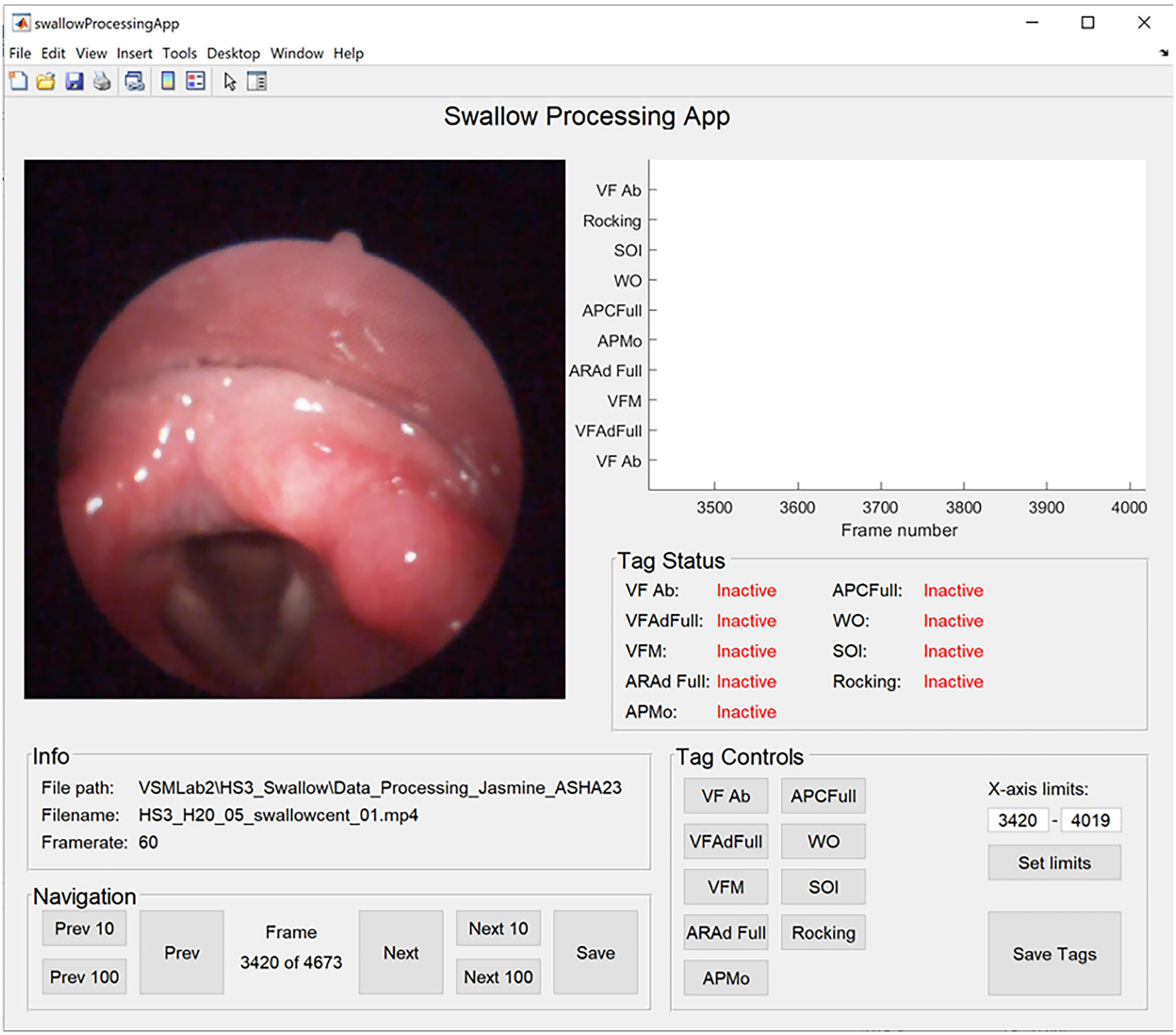

Fig. 2. Custom-designed MATLAB GUI Interface using one participant example. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

or without video recording to quantitatively analyze swallowing motion, widely known as FEES.[12] Standard FEES are captured at a frame rate of 30 fps compared with laryngeal HSV, with frame rates up to 8000 fps. Glottic closure is *presumed* in the absence of visualized penetration or aspiration. Glottic closure is *presumed* to occur during the "whiteout" of the swallow where the passage of the bolus and movement of pharyngeal structures are not visualized due to the reflected light from the pharyngeal and laryngeal tissue to the endoscope. Furthermore, even if laryngeal visualization was possible, both VFSS and FEES lack the temporal resolution (adequate frame rate) to fully visualize quick laryngeal movements and view nuanced changes in glottic closure. It is also assumed that the closure pattern during swallowing is like closure patterns visualized during phonation.

Dynamic CT imaging has also been used to describe glottic closure during swallowing, finding a unique difference in true VF closure with viscosity. VF closure started earlier with thin liquids and lasted longer compared with thickened liquids by Inamoto et al.[13] Kendall et al. used VFSS to describe the onset of aryepiglottic fold closure occurring earlier, as the size of the bolus increased in healthy participants.[14] Van Daele et al. combined electromyographic and endoscopic analysis, to show arytenoid movement consistently preceded full glottic closure and cessation of the posterior cricoarytenoid muscle activity. During 89% of normal swallows, the glottis was partially open in the video frame before the bolus.[15] In this study, VF abduction occurs immediately prior to the SOI, which occurs over 260 ms or 130 frames (on average) prior to WO. Early arytenoid movement toward the epiglottis can

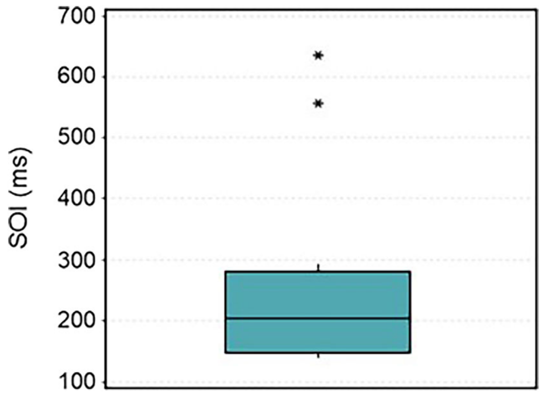

Fig. 3. Boxplot of the distribution of SOI (ms) for all 12 participants. Outliers are identified as asterisks. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

be viewed on standard (30 fps) videoendoscopy, yet laryngeal closure was assumed during the whiteout by Abe and Tsubahara.[16]

This pilot study sought to investigate the feasibility of a novel technique to further characterize quick laryngeal movements occurring during the height of swallowing, which correlates with the whiteout period of the pharyngeal swallow. In this study, we had 100% tolerability of a lower scope position, with 86% feasibility in playback review identifying a timed swallowing SOI in a small 14-subject healthy cohort. Using laryngeal HSV and a custom-designed MATLAB GUI, we could identify and begin to sequence specific laryngeal movements before (preparatory phase) and during the swallow. This investigation was directed to laryngeal movements associated with glottic closure and arytenoid motion during volitional dry swallows. Conceptually, this technique enabled visualization related to VF and concurrent arytenoid to epiglottis closure of the glottic inlet that is generally assumed but has yet to be fully visualized. Using our MATLAB GUI, we were able to describe differences between arytenoid movement and full arytenoid closure, VF medialization and full VF closure, and AP movement and AP closure. This study evaluated volitional dry swallows, which may be like reflexive pharyngeal swallow that normally occurs in less than 500 ms in normal swallowing physiology.[17] Our study found that the average duration of SOI was 260 ms, in this case yielding, 130 frames for analysis, which would have only yielded seven frames in either FEES or VFSS (30fps). Because the pharyngeal reflex is so rapid, it is known for poor inter-rater reliability in inexperienced raters.[18] The reliability of our custom GUI and the swallow events during the SOI will need to be further assessed in future studies.

Several limitations are clear: first is the use of volitional dry swallows used in this pilot project for initial safety and tolerability. The duration of SOI and timing of events may be different from reflexive swallows and will be affected by bolus consistency and volume. Both volitional and reflexive swallows would be best studied under simultaneous HSV and VFSS to determine the impact of the scope position on laryngeal physiology and to

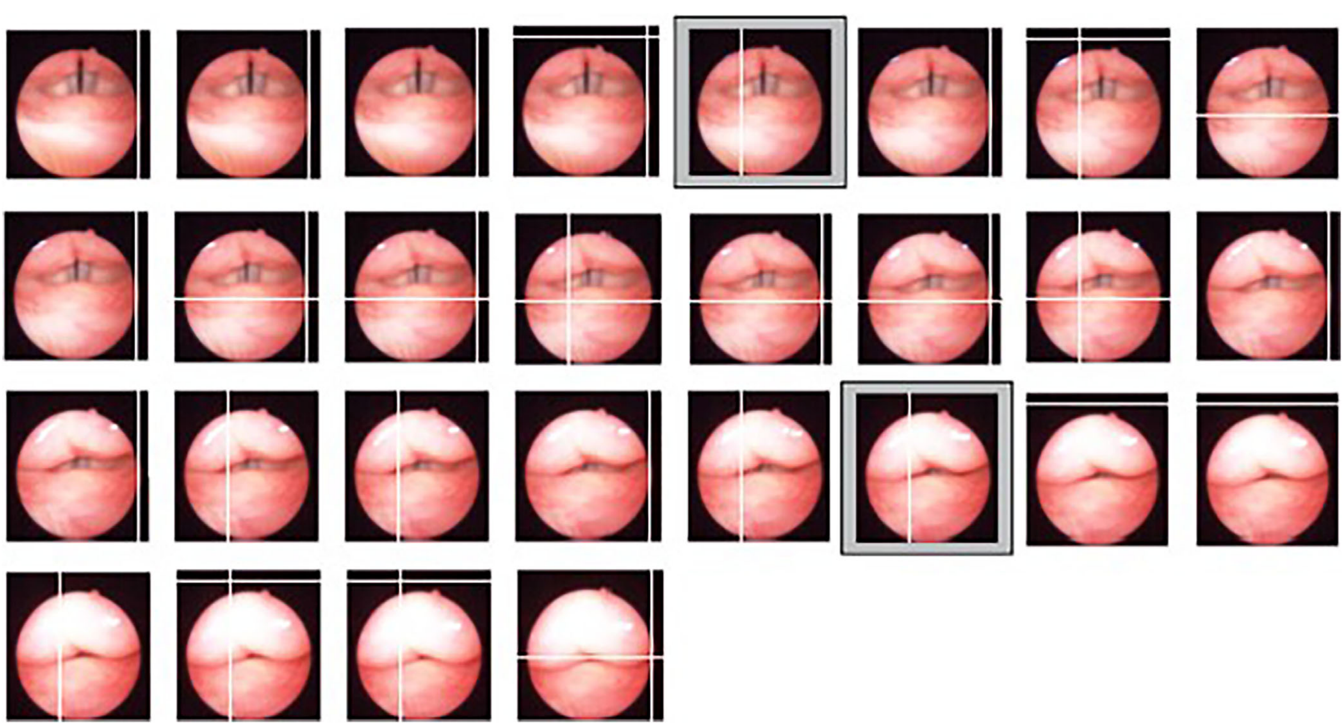

Fig. 4. Frames extracted from the swallow segment of interest (SOI) from the high-speed camera sampled at 500 fps. The images visualize true vocal fold medialization and anterior–posterior compression. If the same swallow event was captured at only 30 fps, there would only be two images captured. The highlighted (gray box) images are an example of two images captured by standard imaging. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

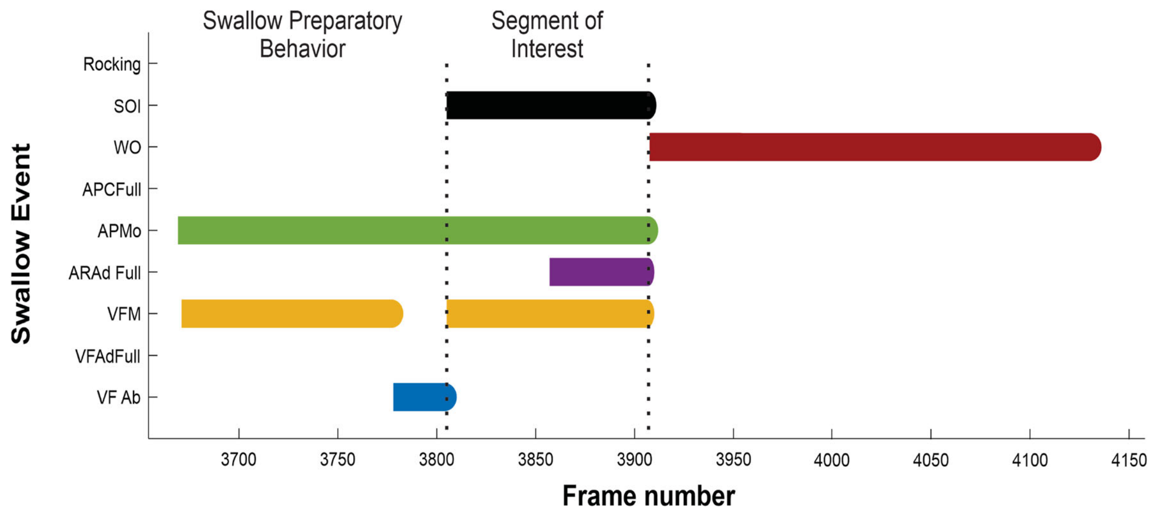

Fig. 5. MATLAB GUI Output example of the swallow preparation and segment of interest (SOI) with specific laryngeal movements identified. WO = whiteout; Vocal fold (VF) abduction (Ab)—VF and arytenoids moving away from midline; Vocal fold (VF) medialization (M)—the movement of VF and arytenoids toward the midline; Vocal fold (VF) adduction full (AdFull)—the true VF are fully closed; Arytenoid adduction (ARad) full—the arytenoids are fully closed, even if the true VF are NOT; Anterior–Posterior (AP) movement (Mo)—AP movement (coming together or apart); Anterior–Posterior (AP) compression full (Cfull)—the full AP closure between arytenoids and the petiole of the epiglottis. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

understand whether the scope placement increases penetration/aspiration risk due to physical obstruction. Furthermore, 15% of the participants were unable to be visualized, and this may be due to neck position, epiglottic deflection of the scope, or light/speed of camera. Although the use of 500 fps was initially used in prior laryngeal kinematic studies, it does not preclude the utilization of 1000 or 2000 fps with the current capability of laryngeal HSV.

Our results present promising pilot data, although it does not have immediate clinical applicability. First, bolus consistency influences the timing of VF closure,[13] warranting further exploration focusing on bolus-driven swallows using the novel technique and is beyond the scope of this pilot study. Second, the use of high-speed imaging may not be accessible clinically, although the novel approach is meant to document information that is clinically relevant but undocumented. While the novel scope position could be implemented clinically with a halogen light, we hypothesize that the rapid laryngeal behaviors documented here would not be visualized with low-speed imaging.

CONCLUSION

The placement of an endoscope into the laryngeal vestibule during a volitional dry swallow was 100% tolerable in this small pilot study. Using laryngeal HSV (500 fps) with a custom MATLAB GUI, we demonstrate 86% feasibility in identifying a timed swallow SOI. This is a promising and novel technique to sequence and time rapid laryngeal motions such as glottic and supraglottic compression not otherwise visualized by either FEES or VFSS in standard (30 fps) dysphagia evaluation. Future studies should investigate laryngeal closure and arytenoid positioning with a bolus present, and in a range of ages, genders, and etiologies in both healthy and abnormal populations to better understand swallowing physiology.

BIBLIOGRAPHY

1. Pueo B. High-speed cameras for motion analysis in sports science. *J Hum Sport Exerc*. 2016;11(1):53-73. https://doi.org/10.14198/jhse.2016.111.05.
2. Deliyski DD, Hillman RE, Mehta DD. Laryngeal high-speed videoendoscopy: rationale and recommendation for accurate and consistent terminology. *J Speech Lang Hear Res*. 2015;58(5):1488-1492. https://doi.org/10.1044/2015_JSLHR-S-14-0253.
3. Zacharias SRC, Myer CM, Meinzen-Derr J, Kelchner L, Deliyski DD, de Alarcón A. Comparison of videostroboscopy and high-speed videoendoscopy in evaluation of supraglottic phonation. *Ann Otol Rhinol Laryngol*. 2016;125(10):829-837. https://doi.org/10.1177/0003489416656205.
4. Diaz-Cadiz M, McKenna VS, Vojtech JM, Stepp CE. Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy. *J Speech Lang Hear Res*. 2019;62(6):1685-1706. https://doi.org/10.1044/2019_JSLHR-S-18-0405.
5. Chhetri DK, Neubauer J, Sofer E. Posterior cricoarytenoid muscle dynamics in canines and humans. *Laryngoscope*. 2014;124(10):2363-2367. https://doi.org/10.1002/lary.24742.
6. Aghdam MA, Ogawa M, Iwahashi T, Hosokawa K, Kato C, Inohara H. A comparison of visual recognition of the laryngopharyngeal structures between high and standard frame rate videos of the fiberoptic endoscopic evaluation of swallowing. *Dysphagia*. 2017;32(5):617-625. https://doi.org/10.1007/s00455-017-9803-5.
7. Jacobson B, Johnson A, Grywalsky C. The voice handicap index (VHI): development and validation. *Am J Speech Lang Pathol*. 1997;6:66-70.
8. Belafsky PC, Mouadeb DA, Rees CJ, et al. Validity and reliability of the eating assessment tool (EAT-10). *Ann Otol Rhinol Laryngol*. 2008;117(12):919-924.
9. McKenna VS, Llico AF, Howell RJ. Vocal fold kinematics for volitional and reflexive cough in healthy adults using high-speed flexible laryngoscopy. Presented at: American speech-language and hearing association convention; 2022; New Orleans, LA.
10. Mohd Khairuddin KA, Ahmad K, Mohd Ibrahim H, Yan Y. Analysis method for laryngeal high-speed videoendoscopy: development of the criteria for the measurement input. *J Voice*. 2021;35(4):636-645. https://doi.org/10.1016/j.jvoice.2019.12.005.
11. Bonilha HS, Blair J, Carnes B, et al. Preliminary investigation of the effect of pulse rate on judgments of swallowing impairment and treatment recommendations. *Dysphagia*. 2013;28(4):528-538. https://doi.org/10.1007/s00455-013-9463-z.
12. Langmore SE, Kenneth SMA, Olsen N. Fiberoptic endoscopic examination of swallowing safety: a new procedure. *Dysphagia*. 1988;2(4):216-219. https://doi.org/10.1007/BF02414429.

13. Inamoto Y, Saitoh E, Okada S, et al. The effect of bolus viscosity on laryngeal closure in swallowing: kinematic analysis using 320-row area detector CT. *Dysphagia*. 2013;28(1):33-42. https://doi.org/10.1007/s00455-012-9410-4.
14. Kendall KA, McKenzie S, Leonard RJ, Gonçalves MI, Walker A. Timing of events in normal swallowing: a videofluoroscopic study. *Dysphagia*. 2000; 15(2):74-83. https://doi.org/10.1007/s004550010004.
15. van Daele DJ, McCulloch TM, Palmer PM, Langmore SE. Timing of glottic closure during swallowing: a combined electromyographic and endoscopic analysis. *Ann Otol Rhinol Laryngol*. 2005;114(6):478-487. https://doi.org/10.1177/000348940511400610.
16. Abe H, Tsubahara A. Observation of arytenoid movement during laryngeal elevation using videoendoscopic evaluation of swallowing. *Dysphagia*. 2011;26(2):150-154. https://doi.org/10.1007/s00455-010-9285-1.
17. Kim Y, McCullough GH, Asp CW. Temporal measurements of pharyngeal swallowing in normal populations. *Dysphagia*. 2005;20(4):290-296. https://doi.org/10.1007/s00455-005-0029-6.
18. Lee JT, Park E, Hwang J, Jung T, Park D. Machine learning analysis to automatically measure response time of pharyngeal swallowing reflex in videofluoroscopic swallowing study. *Sci Rep*. 2020;10(1):14735-4. https://doi.org/10.1038/s41598-020-71713-4.