ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Spectral Aggregate of the High-Passed Fundamental Frequency and Its Relationship to the Primary Acoustic Features of Adductor Laryngeal Dystonia

Katherine L. Marks,[a] Taylor F. Feaster,[a] Sarah Baker,[a] Manuel E. Díaz-Cádiz,[a] Philip C. Doyle,[b] and Cara E. Stepp[a]

[a] Boston University, MA [b] Stanford University, CA

ARTICLE INFO

Article History:
Received March 10, 2022
Revision received May 26, 2022
Accepted July 15, 2022

Editor-in-Chief: Peggy B. Nelson
Editor: Shaheen N. Awan

https://doi.org/10.1044/2022_JSLHR-22-00157

ABSTRACT

Objective: Currently, no clinically feasible objective measures exist that are specific to the signs of adductor laryngeal dystonia (LD), deterring effective diagnosis and treatment. This project sought to establish concurrent validity of a new automated acoustic outcome measure, designed to be specific to adductor laryngeal dystonia (AdLD): the spectral aggregate of the high-passed fundamental frequency contour (SAHf_o).

Method: Twenty speakers with AdLD read voiced phoneme-loaded (more symptomatic) and voiceless phoneme-loaded (less symptomatic) sentences. LD discontinuities (defined as phonatory breaks, frequency shifts, and creak), the acoustic ramifications of laryngeal spasms, were manually identified. The frequency content of the f_o contour was examined as a function of time, and content above 1000 Hz was summed to automatically calculate SAHf_o. Multiple linear regression analysis was applied to SAHf_o based on LD discontinuities and sentence type (voiced or voiceless phoneme-loaded).

Results: The regression model accounted for 41.1% of the variance in SAHf_o. Both the LD discontinuities and sentence type were statistically related to SAHf_o.

Conclusion: Results of this study provide evidence of concurrent validity. SAHf_o is an automatic outcome measure specific to acoustic signs of AdLD that may be useful to track treatment progress.

Laryngeal dystonia (LD), also known as spasmodic dysphonia, is a neurological focal dystonia of the laryngeal musculature (Blitzer et al., 2018). There are two main types of LD, differentiated by the muscles affected: adductor LD (AdLD) and abductor LD, with AdLD being the most common (Blitzer, 2010; Tisch et al., 2003). AdLD is characterized by involuntary spasms of the adductory laryngeal muscles during phonation, which result in phonatory breaks and other discontinuities, strained or strangulated voice quality (Creighton et al., 2015; Ludlow et al., 2008), and elevated vocal effort during phonation (Shoffel-Havakuk et al., 2019). Reduced intelligibility due to these laryngeal spasms compromises effective communication and negatively impacts quality of life and psychosocial well-being for individuals with AdLD (Bender et al., 2004; Braden et al., 2010; Hogikyan et al., 2001; Isetti et al., 2014; Rubin et al., 2004).

AdLD is currently a chronic, incurable disorder. Its typical treatment provides only temporary relief of symptoms (Creighton et al., 2015). Current gold standard treatment involves repeated injections of botulinum toxin (Botox) into the laryngeal muscles, which provides a temporary (3–6 months) relief in signs and symptoms. New,

---

Correspondence to Katherine L. Marks: kmarks6@bu.edu. ***Disclosure:*** *Cara E. Stepp has received consulting fees from Altec, Inc./Delsys, Inc., companies that focused on developing and commercializing technologies related to human movement. Stepp's interests were reviewed and are managed by Boston University in accordance with their conflict of interest policies. The other authors have declared that no other competing financial or nonfinancial interests existed at the time of publication.*

alternative treatments include off-label medications (Richardson et al., 2017), surgical approaches to anatomical remodeling of the larynx (Blitzer et al., 2018), and sodium oxybate (Rumbach et al., 2017; Simonyan et al., 2021). Experimental therapies such as vibrotactile and electrical stimulation approaches (Khosravani et al., 2019; Pitman, 2014), as well as noninvasive neuromodulation with repetitive transmagnetic stimulation or transcranial direct current stimulation (tDCS), are under investigation (Prudente et al., 2021). However, efficacy of these newer treatments is hampered by limited objective measures available for use in their validation.

Clinically available measures of vocal function in AdLD include patient-reported outcomes, laryngeal stroboscopy, listener perceptions of voice quality, and acoustic measures. Although patient-reported outcomes can provide some valuable screening information, they are susceptible to bias (e.g., placebo effects) and are not specific to the symptoms of AdLD, weakening their usefulness (Ludlow et al., 2018; Sackett, 1979; Yiu et al., 2021). Visualization of the larynx via laryngeal stroboscopy allows clinicians to rule out organic changes to the vocal folds and allows for subjective impressions about the movements of the larynx and features of vocal fold vibrations. However, based on simultaneous laryngeal electromyography and acoustics and accounting for electromechanical delay, spasm durations are often < 20 ms (Hillel, 2001), too fast to be imaged with clinical endoscopy systems, which are typically limited to 30 frames/s. High-speed videoendoscopy (HSV) records a higher rate of data (typically over 4,000 frames/s) and could, therefore, be used to characterize vocal fold adductory gestures (Díaz-Cádiz et al., 2019). However, HSV systems are not typically available in clinical settings due to high costs of the equipment, management and processing of large data files, and limitations on memory storage (Mehta & Hillman, 2012). Auditory-perceptual ratings can provide a sensitive index of function (Cannito et al., 2012), but only when averaged over many listeners (Eadie & Baylor, 2006; Shrivastav et al., 2005). Thus, this method of evaluation can provide valuable insight for scientific inquiries with access to large pools of listeners who can spend many hours in controlled environments, but it is not feasible for clinical purposes.

The relative promise of auditory-perceptual measures suggests potential in tapping the acoustic signal directly. Commonly reported outcomes include those related to fundamental frequency (f_o), perturbation measures, and phonation time (Adams et al., 1995; Truong et al., 1991; Watts et al., 2006; Wong et al., 1995). However, few clinically feasible automatic acoustic measures have been shown to be sensitive to AdLD (Moerman et al., 2015; Yanagida et al., 2018). Cepstral-spectral acoustic measures have shown promise for evaluating many voice disorders, but they are not sensitive to the disorder-specific features of AdLD (i.e., laryngeal discontinuities associated with laryngeal spasms). For example, the Cepstral Spectral Index of Dysphonia (CSID) is an automated acoustic measure that objectively estimates overall severity of dysphonia (Peterson et al., 2013), but not specifically features related to the primary signs of AdLD (i.e., features related to spasms). The CSID is calculated via an algorithm that incorporates measures from spectral and cepstral analyses and has been shown to vary with overall severity of voice across many disorders, despite differences in the underlying mechanisms of the different disorders (Awan et al., 2009; 2010; Peterson et al., 2013; Roy et al., 2014). Roy et al. (2014) evaluated the discriminant validity of CSID, finding that a CSID cutoff difference of 46.6 points had 66.7% sensitivity and 64.4% specificity based on task in differentiating AdLD from muscle tension dysphonia (MTD).[1] However, correlations between the CSID and listener-rated overall severity were much lower for AdLD than for MTD. This suggests that CSID may not capture the primary signs of AdLD, but instead may be capturing secondary features related to compensatory vocal effort (Buckley et al., 2020).

Current methods of analyzing LD discontinuities (associated with laryngeal spasms in AdLD) appear more promising than automated measures (Cannito et al., 2012; Izdebski, 1984; Sapienza et al., 1999; Siemons-Luhring et al., 2009; Yanagida et al., 2018). However, these methods often require manual intervention by a trained technician, which makes them time-consuming and, thus, limits their clinical feasibility. Sapienza and colleagues published a series of papers describing the manual labeling of LD discontinuities, specifically phonatory breaks, frequency shifts, and aperiodic segments (Sapienza et al., 1998, 1999, 2000, 2002). They manually counted the number of phonatory breaks that were identified visually from the acoustic waveforms. Phonatory breaks were defined as any absence of voicing that lasted for more than 50 ms (Sapienza, 2002). The number of frequency shifts that occurred within each word or sustained vowel were also manually counted. A frequency shift was defined as a change of 50 Hz or more in f_o that occurred within 50 ms (Sapienza et al., 2002). Aperiodic segments were visually identified and counted from the acoustic waveform for selected words and were defined as segments consisting of nonrepetitive cycles (Sapienza et al., 2002). Overall, their data indicated that, for speakers with AdLD, the phonatory breaks that typically characterize AdLD speech were not the only notable characteristic that differentiated the speakers with MTD, but rather, it was a combination of phonatory breaks, f_o variation, and aperiodicity (Sapienza et al., 2000). Roy et al. (2008) also found that phonatory break analysis differentiated speakers with AdLD from speakers with MTD. Collectively, the studies

---

[1] Muscle tension dysphonia (MTD) is a voice disorder that can present similarly to AdLD. MTD is defined by Colton et al. (2006) as the presence of vocal hyperfunction in the absence of known pathology or neurological condition.

suggest that visual identification and manual counting of LD discontinuities provide a valid way to differentiate speakers with AdLD from speakers with MTD (Sapienza et al., 2000) and controls (Sapienza et al., 1999) and are sensitive to the effects of Botox injections (Sapienza et al., 2002). Although valid, the visual identification and manual counting of LD discontinuities is not clinically feasible, as it is time intensive.

Recently, Buckley et al. (2020) proposed a new outcome measure, which we term here the *spectral aggregate of the high-passed* f_o (SAHf_o). SAHf_o was automatically calculated by detection of fast transitions in f_o in voiced segments, indicative of frequency shifts, as well as discontinuities in f_o caused by removal of unvoiced segments (i.e., phonatory breaks). A model comprised of three relative fundamental frequency measures, designed to capture vocal effort and SAHf_o, designed to capture the laryngeal spasms, accounted for 85% of the variance in the auditory-perceptual evaluation of overall severity, which is promising evidence that an acoustic model could be used as an automated objective outcome measure specific to AdLD. The authors suggested that, based on the model results, the overall severity of dysphonia in AdLD is a result of both primary features (laryngeal spasms) and secondary (hyperfunctional) compensatory behaviors (Buckley et al., 2020). This reasoning was based on the empirical evidence supporting that relative fundamental frequency is sensitive to increased laryngeal tension and vocal effort (Lien et al., 2015; McKenna et al., 2016), but only the theoretical intent behind the development of SAHf_o. Thus, further investigation into the of the relationship between SAHf_o and the primary signs of AdLD is warranted.

The purpose of this study was to provide evidence of concurrent validity for a new outcome measure, SAHf_o, by determining whether it reflects the primary acoustic signs of AdLD (i.e., laryngeal spasms). Our long-term objective is to psychometrically validate SAHf_o as a clinically feasible outcome measure specific to AdLD, which will require additional validation studies that address different aspects of a psychometric evaluation, including discriminant validity between AdLD and MTD. This study is the first step. We hypothesized that SAHf_o would be positively associated with the number of laryngeal discontinuities in the acoustic signal, caused by adductory laryngeal spasms.

Method

Participants

Twenty adults (14 women, six men) were recruited as speakers for this study.[2] Speakers ranged in age from 55 to 69 years with a mean age for women of 63.1 years (standard deviation [*SD*] = 4.0) and for men of 59.8 years (*SD* = 3.3). All speakers had been diagnosed by the same board-certified, fellowship-trained laryngologist as having AdLD. The diagnosis was made based on independent assessments of an experienced laryngologist and an experienced voice-specialized speech-language pathologist (SLP). Discussion followed these independent assessments based on the information gathered. If either the physician or SLP identified any episodes that were consistent with abductor LD, or a mixed variety of LD, or potentially a hyperfunctional voice disorder, they were excluded. To further confirm the initial diagnostic impression of AdLD, a second follow-up visit was used as a confirmation of the initial diagnosis. If confirmation of AdLD by both the physician and SLP was consistent across the two visits and was further confirmed through auditory-perceptual assessment of voice recordings and videolaryngoscopy, the patient was invited to participate in the research study in the future. Patients were excluded if their age was outside the range of 35–80 years, or if they had a history of a breathing disorder or chronic obstructive pulmonary disease, history of head injury or stroke, observable evidence of vocal tremor or other sites of tremor (e.g., hand, head), gross disturbance in gait or balance, history of difficulty with mastication or swallowing, or history of surgery to the head and neck. Prior to the recording of voice samples, all participants provided informed consent in accordance with Western University's Research Ethics Board (REB #18588E). Each speaker's participation was voluntary, and none of the participants were compensated for their time.

The 20 recordings used in this study were obtained from a larger pool of patient recordings (n = 60) obtained from potential participants who were followed in this clinic. The recordings selected for inclusion in this study also represented a subsample from a larger series of participants who were seen for Botox treatment within the same clinic. All 20 speakers had been receiving ongoing Botox treatment for their AdLD. Speakers ranged in the level of AdLD severity. No speakers were excluded from participation based on severity or on task-dependent sign expression. The severity level for this group of speakers also was influenced by the timing of where each participant fell in their regular Botox treatment cycle. Therefore, all recordings took place before Botox injection at a time when voice abnormalities would be most prominent. The disorder had been present for at least 6 months at the time when voice and speech samples were gathered. A voice-specialized SLP rated the overall severity of voice for each sample using the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kempster et al., 2009). The SLP rerated approximately 20% of the recordings to determine intrarater reliability.

---

[2] Gender information was unavailable for this sample.

Instrumentation

All experimental samples were obtained by the same individual under consistent recording conditions within a sound-attenuated recording environment. For all 20 speakers, voice samples were obtained using a professional quality cardioid condenser microphone (SHURE PG81), a preamplifier, and Kay-Pentax Sona-Speech II software (Pine Brook), which was installed on a laptop computer (Toshiba Satellite P100). All samples were digitized at a 44.1-kHz sampling rate and stored for later editing.

Recording Procedure

During the recording procedure, all participants were seated comfortably with the microphone positioned at a fixed distance of 15 cm (microphone-to-mouth distance) using an adjustable microphone boom-arm. The experimental stimuli were displayed in large print on 8″ × 11″ cards positioned at eye level in front of the speaker. Once the microphone and card placement were established, instructions for the recording procedure were then provided.

All speakers were asked to produce a series of vocal tasks that comprised a standardized speech recording protocol. This included two productions of a sentence loaded with voiced phonemes: “Early one morning a man and a woman were ambling along a one-mile lane running near Rainy Island Avenue” (Dedo & Shipp, 1980), and two productions of a sentence loaded with voiceless phonemes: “He saw half a shape mystically cross the path about a half a block from her sister Kathy’s house,” adapted from Dedo and Shipp (1980). Pauses were provided between each vocal task or when requested by the speaker.

Prior to recording, explicit instructions were provided to each speaker to “Please read the materials to follow aloud using your typical voice and please use your typical pitch and loudness level. Do not worry if you have any disruption in your speech due to your voice disorder. We are seeking to obtain the most representative sample of your speech and we will only stop recording if you misread a stimulus item.” Recordings of each speaker’s sample stimuli were manually adjusted using Sona-Speech II to ensure that the signal input was not under- or overdriven. Once this setting was identified, it was continuously monitored and adjusted as needed across vocal tasks. All stimulus samples were assessed individually to ensure that no errors in production occurred (i.e., omitted or misspoken words). If an error occurred, the speaker was asked to repeat that sample.

Acoustic Analysis

Two trained technicians manually identified and labeled LD discontinuities from the acoustic waveforms and spectrograms in Praat (Boersma & Weenink, 2013) for each repetition for each sentence. Following criteria from Sapienza et al. (2000), phonatory breaks were defined as absence of voicing that typically occurs during a phonetic segment (example shown in Figure 1A). Frequency shifts were defined as a change of 50 Hz or more in a f_o that occurs within 50 ms (example shown in Figure 1B). *Creak* was used as an umbrella term for other types of irregularities in f_o following Keating et al. (2015; e.g., low and irregular f_o, multiply pulsed voice, aperiodic voice, or tense/pressed voice). Figure 1C displays one example of creak. Consensus judgments were made by a third technician (a voice specialized SLP) for instances in which the original two technicians did not agree (approximately 16% of all identified LD discontinuities). The percentage of LD discontinuities (both summed and individually) were calculated as the aggregate duration of LD discontinuities divided by the total duration of the sentence.

Consistent with the methods of Buckley et al. (2020), automated measures of SAHf_o were calculated using an instantaneous f_o estimator, Halcyon (Azarov et al., 2016) and custom MATLAB scripts (MATLAB, 2021). Figure 2 illustrates methods for calculating SAHf_o. First, the stimuli text was force-aligned to the audio using an open-source algorithm (P2FA; Yuan & Liberman, 2008), removing unvoiced sounds and short words (e.g., “of” and “a”). Stimuli were then processed by Halcyon, collecting a range of possible f_o candidates between 30 and 400 Hz over 1-ms sampling windows (see Figure 2, Panels A and B). The f_o contour was estimated by optimizing the most probable path across all f_o candidates and then resampled to correspond with the sampling rate of the original acoustic signal using linear interpolation. Spectrograms were calculated using the f_o contour as a function of time, with a 25-ms Hanning window and 90% overlap. The lower range was limited to 0 dB to make strong changes in f_o more evident against smaller fluctuations (i.e., close to the background low-energy level), as shown in Figure 2, Panels C and D. The time-average

Figure 1. Examples of a phonatory break (A), a frequency shift (B), and creak (C) identified via waveform and spectrogram of speech.

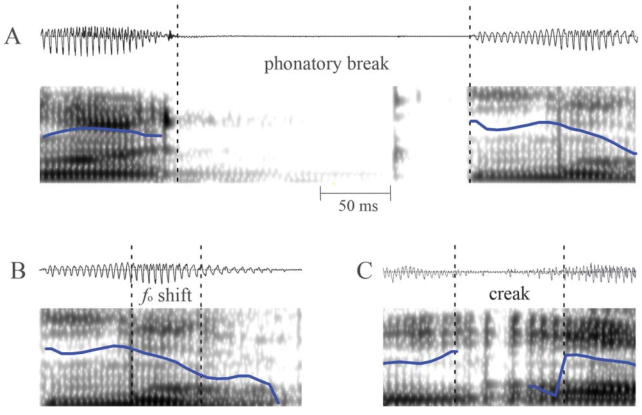
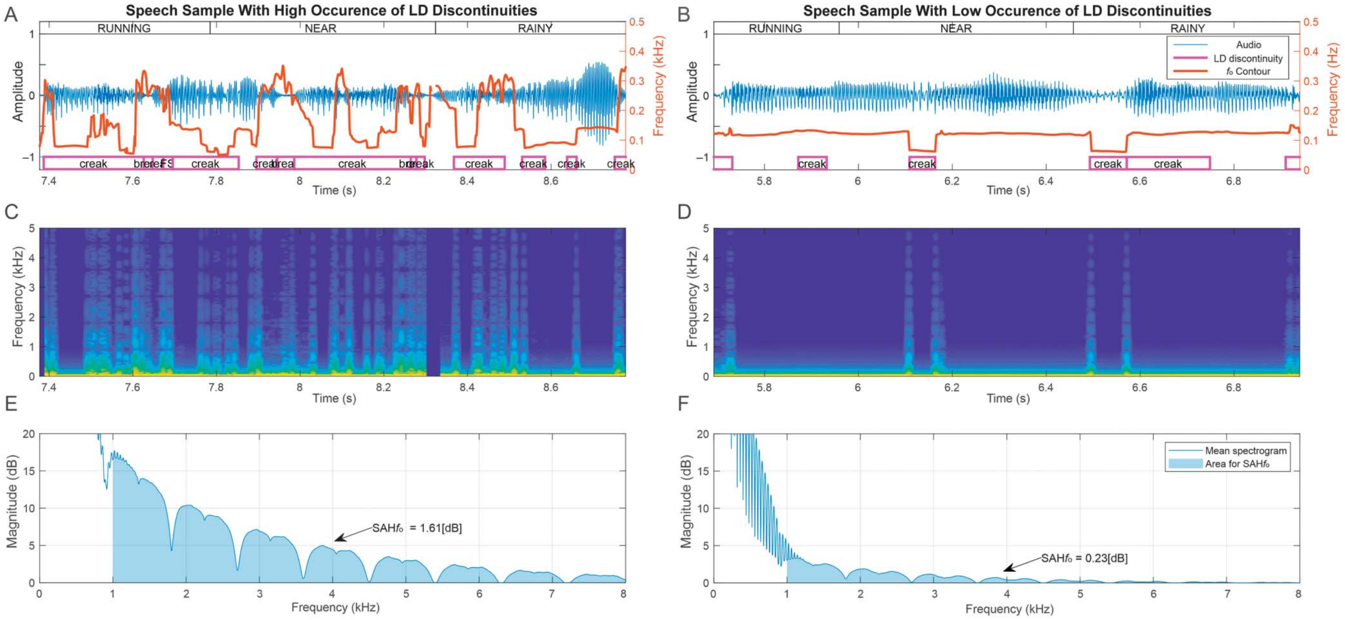


Figure 2. Example of the analysis of the spectral aggregate of the high-passed fundamental frequency contour (SAHf_o) for two participants with adductor laryngeal dystonia (LD). Segments of the original speech sample (audio) and respective fundamental frequency (f_o) contour for an individual with (A) high occurrence of LD discontinuities and (B) low occurrence of LD discontinuities. To calculate SAHf_o, first, the stimuli text was force-aligned using an open-source algorithm (P2FA; Yuan & Liberman, 2008), removing unvoiced sounds and short words (e.g., "of" and "a"). Stimuli were then processed by Halcyon, collecting a range of possible f_o candidates between 30 and 400 Hz over 1-ms sampling windows. The f_o contour was estimated by optimizing the most probable path across all f_o candidates and then resampled to correspond with the sampling rate of the original acoustic signal using linear interpolation. Spectrograms were calculated using the f_o contour as a function of time, with a 25-ms Hanning window and 90% overlap. The lower range was limited to 0 dB to make strong changes in f_o more evident against smaller fluctuations (i.e., close to the background low-energy level). Panels C and D show the spectrogram representations of the f_o contours from Panels A and B, respectively. The time-average mean was calculated from the resulting spectrogram. Energy above 1000 Hz was aggregated, effectively high pass filtering the information. Finally, the accumulated quantity was normalized by the total number of frequency bins, resulting in a single estimate of SAHf_o per sentence. Panels E and F illustrate the aggregation over time of the previous spectrograms. The final feature calculation is illustrated by the filled-in areas in the frequency range above 1000 Hz, resulting in an SAHf_o value.



mean was calculated from the resulting spectrogram. Energy above 1000 Hz was aggregated, effectively high pass filtering the information, as illustrated in Figure 2, Panels E and F. Finally, the accumulated quantity was normalized by the total number of frequency bins, resulting in a single estimate of SAHf_o per sentence.

Statistical Analysis

Pearson correlation coefficient r was calculated for 20% of the sentences to assess intrarater reliability of the SLP auditory-perceptual ratings. LD discontinuity values and SAHf_o values were averaged per condition (voiced or voiceless phoneme-loaded sentences) per participant. Statistical analyses were conducted using *R* (R Development Core Team, 2010). First, the data set was assessed using histograms and box plots, which revealed one outlier for SAHf_o in the voiceless condition. This person's data from voiceless phoneme sentences were removed from the data set, as the value of SAHf_o was > 2 *SD*s from the mean. To assess the relationship between SAHf_o and the aggregate percentage of all LD discontinuities, a simple regression model was used to assess whether the aggregate percentage of LD discontinuities were associated with SAHf_o. Statistical significance was set a priori to .05. Standardized beta coefficients (B) values were used as estimates of effect size. To further explore the dataset and the individual contributions of the types of LD discontinuities, a multiple regression model was used to determine the amount of variance in SAHf_o that creak, phonatory breaks, and frequency shifts captured, controlling for condition (voiced or voiceless phoneme-loaded sentences). Based on the results of this model, we used two, post hoc paired t tests to compare LD discontinuities and SAHf_o between voiced phoneme-loaded sentences (more symptomatic) and voiceless phoneme-loaded sentences (less symptomatic).

Results

Overall severity ratings of dysphonia, LD discontinuities (aggregated and individually), and SAHf_o values per participant and per condition are displayed in Table 1. Intrarater reliability for the SLP rater was r = .75,

Table 1. Auditory-perceptual and acoustic characteristics of voice samples.

	Voiced phoneme-loaded sentences						Voiceless phoneme-loaded sentences					
Participant	Overall severity	LD discontinuities	Phonatory breaks	Frequency shifts	Creak	$\text{SAH}f_o$	Overall severity	LD discontinuities	Phonatory breaks	Frequency shifts	Creak	$\text{SAH}f_o$
F1	69.80	48.39	2.55	6.50	40.01	1.25	5.41	6.07	0.00	1.58	4.49	0.65
F2	78.63	42.97	23.17	3.06	17.26	1.23	68.10	20.86	11.58	3.64	5.64	1.37
F3	70.09	49.80	6.35	2.45	41.00	0.85	87.76	41.46	7.61	1.00	32.84	1.97
F4	54.70	29.44	2.64	6.09	20.97	0.50	26.77	19.54	0.00	3.41	16.12	0.93
F5	68.09	58.16	0.26	9.17	49.97	0.65	8.26	3.35	0.00	0.90	2.45	0.31
F7	67.67	36.28	11.51	5.06	19.71	1.55	65.81	31.00	5.20	2.37	23.43	1.90
F8	59.26	16.07	15.96	0.11	0.00	0.57	75.22	24.75	23.95	0.79	0.00	0.68
F9	53.85	29.49	14.72	0.32	14.46	1.10	25.93	17.05	4.15	1.23	11.67	1.14
F10	44.87	16.38	0.38	7.18	8.82	0.46	30.06	5.03	0.58	1.59	2.86	0.55
F11	73.79	14.39	4.21	4.15	6.02	0.75	67.79	20.88	0.00	4.36	16.51	1.61
F12	46.72	37.47	0.00	1.07	36.40	0.73	48.43	39.65	1.46	2.58	35.60	1.26
F13	56.27	51.64	0.22	0.32	51.11	0.89	45.87	25.43	2.98	2.39	20.06	1.56
F14	63.82	65.35	1.57	0.40	63.38	1.61	58.12	36.01	1.73	0.23	34.04	1.53
M1	42.74	21.04	0.56	0.08	20.40	0.29	43.59	33.42	9.67	0.10	23.66	0.76
M2	51.85	25.56	1.19	2.90	21.46	0.56	7.98	15.13	0.00	1.92	13.21	0.34
M3	36.47	25.66	0.00	0.00	25.66	0.23	12.25	21.21	3.95	0.00	17.25	0.47
M4	46.17	53.49	6.53	0.29	46.67	0.70	75.64	48.88	16.87	0.37	31.64	1.27
M5	40.74	12.57	9.08	0.00	3.48	0.08	41.32	12.96	6.07	0.00	6.88	0.30
M6	48.43	17.43	3.58	6.88	13.85	0.40	23.36	7.41	0.00	0.32	7.09	0.50
M	56.52	34.29	5.50	2.95	26.35	0.76	43.03	22.64	5.04	1.52	16.08	1.01
SD	12.35	16.52	6.58	3.03	18.12	0.43	25.86	13.12	6.55	1.32	11.57	0.55

Note. Overall severity was perceptually judged by a voice-specialized speech-language pathologist using the Consensus Auditory Perceptual Evaluation of Voice. Laryngeal dystonia (LD) discontinuities, phonatory breaks, frequency shift, and creak are displayed as percentages. Percent LD discontinuities were calculated as the aggregate duration of LD discontinuities divided by the total duration of the sentence. Displayed are OS ratings, % LD discontinuities (aggregate and individually), and spectral aggregate of the high-passed fundamental frequency contour ($\text{SAH}f_o$) values in dB per participant for the voiced phoneme-loaded sentences and the voiceless phoneme-loaded sentences. Mean is the average across participants, and *SD* indicates the standard deviation of the mean. F indicates female participant; M indicates male participant. Gender information was not available.

indicating good reliability. The mean overall severity of dysphonia was rated as more dysphonic in the voiced phoneme-loaded sentences (*M* = 56.5, *SD* = 12.4, range: 36.5–78.6) compared to the voiceless phoneme-loaded sentences, which were more variable across individuals (*M* = 43.0, *SD* = 25.9, range: 5.4–87.8). Across participants, the average percent duration of LD discontinuities for the voiced phoneme-loaded sentences was 35.2% (*SD* = 16.55). The average percent duration of LD discontinuities for the voiceless phoneme-loaded sentences was 23.6% (*SD* = 12.83). In both conditions, creak accounted for the largest percentage of LD discontinuities, followed by phonatory breaks and frequency shifts. The average SAHf_o value was 0.76 (*SD* = 0.41) for the voiced-loaded sentences and 1.11 (*SD* = 0.70) for the voiceless-loaded sentences.

A multiple regression model revealed that the aggregate percentage of LD discontinuities and condition accounted for 41.1% of the variance in SAHf_o, which was statistically significant, $F(2, 35) = 13.9$; $p < .001$. LD discontinuities were positively associated with SAHf_o, $B = .21$; $SE = .021$, $t(35) = 4.88$ $p < .001$, when controlling for sentence type. SAHf_o was .49 dB higher in the voiceless condition, controlling for LD discontinuities, $B = .49$, $SE = .13$, $t(35) = 3.67$, $p > .001$. A second multiple regression model assessed the contributions of each type of LD discontinuity (creak, phonatory breaks, and frequency shifts) to SAHf_o, controlling for sentence type. The overall regression was significant, accounting for 47.8% of the variance, $F(4, 33) = 7.55$, $p < .001$. Table 2 contains a summary of the regression results. Sentence type, phonatory breaks, and creak were all associated with SAHf_o ($p < .05$).

Table 2. Multiple regression results of laryngeal dystonia (LD) discontinuity contributions to spectral aggregate of the high-passed fundamental frequency contour (SAHf_o).

Variables	SAHf_o (dB)	Standardized regression coefficients (*B*)
Intercept	–0.11 (–0.51, 0.27)	—
Voiceless sentences	0.55* (0.27, 0.84)	.56
% Creak	0.02* (0.01, 0.03)	.63
% Phonatory breaks	0.03* (0.01, 0.05)	.40
% Frequency shifts	0.06* (–0.00, 0.11)	.28
Observations	38	
R^2	0.48	
Adjusted R^2	0.41	
Residual standard error	0.39 (*df* = 33)	
F statistic	7.55* (*df* = 4, 33)	

Note. Results of multiple regression model for the association between creak, phonatory breaks, frequency shifts, and SAHf_o (dB). Statistical significance indicated by * as $p > .05$. 95% CI are in parentheses. Standardized regression coefficients, which serve as estimates of effect sizes, are listed for each predictor variable. Em dash indicates no value.

Post hoc paired *t* tests revealed a statistically significant difference of mean percent LD discontinuities in the voiced and voiceless phoneme-loaded sentences, $t(19) = 3.15$, $p = .005$, 95 CI [3.90, 19.36]. The mean difference of LD discontinuities was 11.63 percentage points higher in the voiced condition. The difference of mean SAHf_o in the voiced and voiceless sentences was also statistically significant, $t(19) = -2.89$, $p = .009$, 95 CI [−0.61, −0.10], $d = -.63$; however, mean SAHf_o was lower in the voiced-loaded sentences, which differs from the result for the reference measure of percent LD discontinuities.

Discussion

The purpose of this study was to evaluate the concurrent validity of a new outcome measure, SAHf_o, by determining whether SAHf_o reflects the primary acoustic signs of AdLD (i.e., laryngeal spasms). Our long-term objective in this line of research is to validate SAHf_o as an automated and clinically feasible outcome measure specific to AdLD. Toward that goal, the objective of this study was to provide evidence of concurrent validity of SAHf_o as an outcome measure in AdLD. Results of the simple regression model demonstrated a statistically significant relationship between SAHf_o and LD discontinuities, the primary acoustic features caused by laryngeal spasms. Furthermore, LD discontinuities that were manually identified from the acoustic signal accounted for 41% of the variance in SAHf_o. This finding provides promising evidence of concurrent validity; SAHf_o is associated with the primary acoustic features that are specific to AdLD speech (i.e., LD discontinuities).

An unanticipated finding of this study was that SAHf_o is sensitive to phoneme-specific stimuli. Since the primary signs of AdLD are particularly elicited by production of voiced phonemes, this finding yields important information for the design of future studies using SAHf_o. Laryngeal spasms typically occur more frequently during voiced phonemes than during voiceless phonemes in AdLD (Erickson, 2003; Roy et al., 2007), leading to relatively more LD discontinuities. As expected, there were significantly more LD discontinuities in the voiced phoneme-loaded sentences compared to the voiceless phoneme-loaded sentences, consistent with evidence in the literature (Sapienza et al., 1999, 2000, 2002). Additionally, overall severity of voice, although rated by a single SLP, was generally rated more dysphonic in the voiced condition compared to the voiceless condition, with greater variability in the voiceless condition: a finding consistent with Roy et al. (2007) and Erickson (2003). This trend varied among individuals, as described in Table 1. In contrast, SAHf_o was found to be higher in the voiceless phoneme-loaded sentences than the voiceless phoneme-loaded sentences, indicating a greater occurrence of fast transitions in the f_o contour. This finding suggests that

SAHf_o may be impacted by phonemic context. SAHf_o captures abrupt changes in f_o in the signal that excludes voiceless sounds. Thus, despite the underlying changes in vocal function of the participants (wherein they were likely experiencing fewer spasms during the voiceless phoneme-loaded sentences), the SAHf_o values were higher in the voiceless phoneme-loaded sentences, likely due to the more frequent opportunities for abrupt changes in f_o to occur. This means that the effects of LD discontinuities were masked by the abrupt changes that naturally occur in the voiced segments of the voiceless phoneme-loaded sentences. Based on this finding, SAHf_o values should only be compared on identical or similar text material (i.e., SAHf_o values for voiced phoneme-loaded sentences should only be compared to values from voiced phoneme-loaded sentences) in future related studies. Regardless, use of voiced phoneme-loaded sentences may be all that is necessary for monitoring treatment progress of speakers with AdLD, as speakers have the most difficulty with voiced phonemes.

Clinical and Research Implications

SAHf_o was found to be specific to the primary acoustic signs of AdLD, which offers preliminary evidence that SAHf_o could be employed clinically to support diagnosis and assessment of AdLD; however, further psychometric evaluation is warranted before SAHf_o can be implemented clinically. SAHf_o is the first automated acoustic measure specific to the signs of AdLD, filling a significant gap in its assessment. However, SAHf_o may not be used in conjunction with differential assessments due to the issues described with interpreting differences in SAHf_o between voiced and voiceless phoneme-loaded sentences. If calculated from voiced stimuli only, SAHf_o could potentially be used to assess individuals' responses to treatment, which could potentially aid in determining Botox dosing or dosing schedule. SAHf_o also has the potential to be used to validate new treatment modalities that are currently being developed (Blitzer et al., 2018; Khosravani et al., 2019; Pitman, 2014; Prudente et al., 2021; Richardson et al., 2017; Rumbach et al., 2017; Simonyan et al., 2021). Furthermore, the implementation of SAHf_o as an objective outcome for AdLD could reduce reliance on subjective measures that are prone to error, such as auditory-perceptual methods that are only reliable when averaged over many listeners (Eadie & Baylor, 2006). However, additional psychometric analyses of SAHf_o are necessary to provide evidence of criterion, construct, and discriminant validity as well as sensitivity to change, before SAHf_o can be implemented clinically.

Individual Contributions of LD Discontinuities

The highest average percentage of LD discontinuity type was creak, which was defined as an umbrella term for other types of irregularities in f_o, following Keating et al. (2015; e.g., low and irregular $f_{o,}$ multiply pulsed voice, aperiodic voice, or tense/pressed voice). The average percentage of creak was 26.4% in voiced phoneme-loaded sentences, a level that was consistent with the 23.3% reported in Sapienza et al. (2002), and substantially higher than the 0.05% reported in Sapienza et al. (2000). It is worth noting, however, that the stimuli between our study and the two Sapienza studies were different: Instead of voiced phoneme-loaded sentences, Sapienza et al. (2000, 2002) used 15 words from voiced segments of the Rainbow Passage (Fairbanks, 1960). Sapienza's definition of aperiodicity was much stricter than Keating et al.'s (2015) definition of creak adopted in this study. This approach was pursued due to the technicians' observations of discontinuities that were perceived to be associated with laryngeal spasms, but did not fall into any of the three categories described by Sapienza. The average percentage of phonatory breaks (5.5% in the voiced phoneme-loaded sentences) was somewhat higher than Sapienza et al.'s (2002) findings of 2.3%, but comparable to their finding of 6.5% in 2000. Results of this study included a lower percentage of frequency shifts (2.9% in the voiced phoneme-loaded sentences) than Sapienza et al.'s (2002) findings of 6.2%, but is comparable to their finding of 2.2% (Sapienza et al., 2000). In terms of their association with SAHf_o, phonatory breaks and creak were positively associated with SAHf_o, whereas frequency shifts were not statistically associated with the measure. Further investigation with use of control groups (individuals without voice disorders and individuals with MTD) is warranted to determine if SAHf_o is specific to LD.

Limitations

The relationship between LD discontinuities and auditory-perceptual ratings was out of the scope of the current investigation. Thus, overall severity of dysphonia was rated by only one voice-specialized SLP in this study. Although the voiced phoneme-loaded sentences were generally rated as more severe than the voiceless phoneme-loaded sentences, it is important to consider potential bias from the use of only one rater. Nevertheless, that finding is consistent with studies that employed multiple raters (Erickson, 2003; Roy et al., 2007).

It should be noted that the voiceless phoneme-loaded sentence used in this study was altered from the original Dedo and Shipp (1980) sentences, which resulted in the voiceless phoneme-loaded sentence having six fewer syllables than the voiced phoneme-loaded sentence. However, because SAHf_o is calculated by aggregating the spectra normalized by frequency bins, SAHf_o should not have been affected by this discrepancy in the number of syllables per sentence.

Future Directions

More work is needed to fully validate and psychometrically evaluate SAHf_o before it can be used clinically. Future work may address the association of SAHf_o and the primary signs of AdLD visualized via high-speed videoendoscopy of the larynx, offering further support for concurrent validity. Sensitivity to change could be assessed in additional work comparing the effect sizes of change in vocal function (e.g., pre- and post-Botox treatment) between SAHf_o, HSV measures, and auditory-perceptual ratings. SAHf_o also needs to be assessed in terms of construct validity: the extent to which the measure accurately assesses what it intended to assess. Specifically, it will be critical to determine its discriminant validity.

A significant issue in treating AdLD is its frequent misdiagnosis. AdLD is often mistaken for MTD, a hyperfunctional voice disorder with increased global laryngeal tension, but without laryngeal spasms (Roy et al., 2005). Differential diagnosis can present challenges even for experts due to the need to rely on patient-reported symptoms (Shoffel-Havakuk et al., 2019) and auditory-perceptual evaluation (Ludlow et al., 2018). Consequently, patients endure an average of 4–5 difficult years after first seeking medical help for their voice symptoms (Creighton et al., 2015; de Lima Xavier & Simonyan, 2019) and see an average of 3.95 physicians (Creighton et al., 2015) before receiving an accurate diagnosis of laryngeal dystonia. Therefore, determining whether SAHf_o can distinguish between AdLD and MTD is critical future work. If SAHf_o can be used to support a differential diagnosis, it may help to reduce the time it takes for individuals with AdLD to receive appropriate treatment and return to their daily voicing activities.

Conclusions

Results of this study provided essential evidence of concurrent validity for an automated acoustic measure, SAHf_o, which was found to reflect the primary acoustic signs of AdLD. This finding demonstrates the potential of SAHf_o as a clinically feasible, automated outcome measure to support the diagnosis and assessment of AdLD. Further psychometric evaluation is required before SAHf_o can be implemented clinically, particularly in terms of criterion, construct, and discriminant validity as well as sensitivity to change.

Data Availability Statement

Data for this study are unavailable due to the identifiable nature of voice and speech recordings.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders (Grants DC015570 awarded to Cara Stepp and DC020349 awarded to Katherine Marks) and Boston University Sargent College Student Research Grant (awarded to Katherine Marks).

References

Adams, S. G., Hunt, E. J., Irish, J. C., Charles, D. A., Lang, A. E., Durkin, L. C., & Wong, D. (1995). Comparison of botulinum toxin injection procedures in adductor spasmodic dysphonia. *The Journal of Otolaryngology, 24*(6), 345–351.

Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics, 23*(11), 825–841. https://doi.org/10.3109/02699200903242988

Awan, S. N., Roy, N., Jetté, M. E., Meltzner, G. S., & Hillman, R. E. (2010). Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: Comparisons with auditory-perceptual judgements from the CAPE-V. *Clinical Linguistics & Phonetics, 24*(9), 742–758. https://doi.org/10.3109/02699206.2010.492446

Azarov, E., Vashkevich, M., & Petrovsky, A. (2016). *Instantaneous pitch estimation algorithm based on multirate sampling*. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Bender, B. K., Cannito, M. P., Murry, T., & Woodson, G. E. (2004). Speech intelligibility in severe adductor spasmodic dysphonia. *Journal of Speech, Language, and Hearing Research, 47*(1), 21–32. https://doi.org/10.1044/1092-4388(2004/003)

Blitzer, A. (2010). Spasmodic dysphonia and botulinum toxin: Experience from the largest treatment series. *European Journal of Neurology, 17*(Suppl. 1), 28–30. https://doi.org/10.1111/j.1468-1331.2010.03047.x

Blitzer, A., Brin, M. F., Simonyan, K., Ozelius, L. J., & Frucht, S. J. (2018). Phenomenology, genetics, and CNS network abnormalities in laryngeal dystonia: A 30-year experience. *The Laryngoscope, 128,* S1–S9. https://doi.org/10.1002/lary.27003

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* (Version 5.3.39). http://www.praat.org

Braden, M. N., Johns, M. M., III., Klein, A. M., Delgaudio, J. M., Gilman, M., & Hapner, E. R. (2010). Assessing the effectiveness of botulinum toxin injections for adductor spasmodic dysphonia: clinician and patient perception. *Journal of Voice, 24*(2), 242–249. https://doi.org/10.1016/j.jvoice.2008.08.003

Buckley, D. P., Cadiz, M. D., Eadie, T. L., & Stepp, C. E. (2020). Acoustic model of perceived overall severity of dysphonia in adductor-type laryngeal dystonia. *Journal of Speech, Language, and Hearing Research, 63*(8), 2713–2722. https://doi.org/10.1044/2020_JSLHR-19-00354

Cannito, M. P., Doiuchi, M., Murry, T., & Woodson, G. E. (2012). Perceptual structure of adductor spasmodic dysphonia and its acoustic correlates. *Journal of Voice, 26*(6), 818.e5–13. https://doi.org/10.1016/j.jvoice.2012.05.005

Colton, R. H., Casper, J. K., & Leonard, R. J. (2006). *Understanding voice problems: A physiological perspective for diagnosis and treatment*. Lippincott Williams & Wilkins.

Creighton, F. X., Hapner, E., Klein, A., Rosen, A., Jinnah, H. A., & Johns, M. M. (2015). Diagnostic delays in spasmodic dysphonia: A call for clinician education. *Journal of Voice, 29*(5), 592–594. https://doi.org/10.1016/j.jvoice.2013.10.022

de Lima Xavier, L., & Simonyan, K. (2019). The extrinsic risk and its association with neural alterations in spasmodic dysphonia. *Parkinsonism & Related Disorders, 65,* 117–123. https://doi.org/10.1016/j.parkreldis.2019.05.034

Dedo, H. H., & Shipp, T. (1980). *Spastic dysphonia: A surgical and voice therapy treatment program*. College-Hill.

Díaz-Cádiz, M., McKenna, V. S., Vojtech, J. M., & Stepp, C. E. (2019). Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy. *Journal of Speech, Language, and Hearing Research, 62*(6), 1685–1706. https://doi.org/10.1044/2019_JSLHR-S-18-0405

Eadie, T. L., & Baylor, C. R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice, 20*(4), 527–544. https://doi.org/10.1016/j.jvoice.2005.08.007

Erickson, M. L. (2003). Effects of voicing and syntactic complexity on sign expression in adductor spasmodic dysphonia. *American Journal of Speech-Language Pathology, 12*(4), 416–424. https://doi.org/10.1044/1058-0360(2003/087)

Fairbanks, G. (1960). *Voice and articulation Drillbook (Vol. 2)*. Harper and Row.

Hillel, A. D. (2001). The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *The Laryngoscope, 111*(4, Pt. 2, Suppl. 97), 1–47. https://doi.org/10.1097/00005537-200104001-00001

Hogikyan, N. D., Wodchis, W. P., Spak, C., & Kileny, P. R. (2001). Longitudinal effects of botulinum toxin injections on voice-related quality of life (V-RQOL) for patients with adductory spasmodic dysphonia. *Journal of Voice, 15*(4), 576–586. https://doi.org/10.1016/S0892-1997(01)00060-1

Isetti, D., Xuereb, L., & Eadie, T. L. (2014). Inferring speaker attributes in adductor spasmodic dysphonia: Ratings from unfamiliar listeners. *American Journal of Speech-Language Pathology, 23*(2), 134–145. https://doi.org/10.1044/2013_ajslp-13-0010

Izdebski, K. (1984). Overpressure and Breathiness in spastic dysphonia: An acoustic (LTAS) and perceptual study. *Acta Oto-Laryngologica, 97*(3–4), 373–378. https://doi.org/10.3109/00016488409131003

Keating, P. A., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *ICPhS, 2015*(1), 2–7.

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: Development of a standardized clinical protocol.

Khosravani, S., Mahnan, A., Yeh, I.-L., Watson, P. J., Zhang, Y., Goding, G., & Konczak, J. (2019). Atypical somatosensory-motor cortical response during vowel vocalization in spasmodic dysphonia. *Clinical Neurophysiology, 130*(6), 1033–1040. https://doi.org/10.1016/j.clinph.2019.03.003

Lien, Y.-A. S., Michener, C. M., Eadie, T. L., & Stepp, C. E. (2015). Individual monitoring of vocal effort with relative fundamental frequency: Relationships with aerodynamics and listener perception. *Journal of Speech, Language, and Hearing Research, 58*(3), 566–575. https://doi.org/10.1044/2015_JSLHR-S-14-0194

Ludlow, C. L., Domangue, R., Sharma, D., Jinnah, H. A., Perlmutter, J. S., Berke, G., Sapienza, C., Smith, M. E., Blumin, J. H., Kalata, C. E., Blindauer, K., Johns, M., Hapner, E., Harmon, A., Paniello, R., Adler, C. H., Crujido, L., Lott, D. G., Bansberg, S. F., . . . Stebbins, G. (2018). Consensus-based attributes for identifying patients with spasmodic dysphonia and other voice disorders. *JAMA Otolaryngology–Head & Neck Surgery, 144*(8), 657–665. https://doi.org/10.1001/jamaoto.2018.0644

Ludlow, C. L., Hoit, J., Kent, R., Ramig, L. O., Shrivastav, R., Strand, E., Yorkston, K., & Sapienza, C. M. (2008). Translating principles of neural plasticity into research on speech motor control recovery and rehabilitation. *Journal of Speech, Language, and Hearing Research, 51*(1), S240–S258. https://doi.org/10.1044/1092-4388(2008/019)

MATLAB. (2021). *Version 9.10.0 (R2021a)*. The MathWorks Inc.

McKenna, V. S., Heller Murray, E. S., Lien, Y.-A. S., & Stepp, C. E. (2016). The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *Journal of Speech, Language, and Hearing Research, 59*(6), 1283–1294. https://doi.org/10.1044/2016_JSLHR-S-15-0406

Mehta, D. D., & Hillman, R. E. (2012). Current role of stroboscopy in laryngeal imaging. *Current Opinion in Otolaryngology & Head and Neck Surgery, 20*(6), 429–436. https://doi.org/10.1097/MOO.0b013e3283585f04

Moerman, M., Martens, J. P., & Dejonckere, P. (2015). Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia: A compilation of own research. *Logopedics, Phoniatrics, Vocology, 40*(1), 24–29. https://doi.org/10.3109/14015439.2014.936497

Peterson, E. A., Roy, N., Awan, S. N., Merrill, R. M., Banks, R., & Tanner, K. (2013). Toward validation of the cepstral spectral index of dysphonia (CSID) as an objective treatment outcomes measure. *Journal of Voice, 27*(4), 401–410. https://doi.org/10.1016/j.jvoice.2013.04.002

Pitman, M. J. (2014). Treatment of spasmodic dysphonia with a neuromodulating electrical implant. *The Laryngoscope, 124*(11), 2537–2543. https://doi.org/10.1002/lary.24749

Prudente, C. N., Chen, M., Stipancic, K. L., Marks, K. L., Samargia-Grivette, S., Goding, G. S., Green, J. R., & Kimberley, T. J. (2021). Effects of low-frequency repetitive transcranial magnetic stimulation in adductor laryngeal dystonia: A safety, feasibility, and pilot study. *Experimental Brain Research, 240,* 561–574. https://doi.org/10.1007/s00221-021-06277-4

R Development Core Team. (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Richardson, S. P., Wegele, A. R., Skipper, B., Deligtisch, A., Jinnah, H., & Investigators, D. C. (2017). Dystonia treatment: Patterns of medication use in an international cohort. *Neurology, 88*(6), 543–550. https://doi.org/10.1212/WNL.0000000000003596

Roy, N., Gouse, M., Mauszycki, S. C., Merrill, R. M., & Smith, M. E. (2005). Task specificity in adductor spasmodic dysphonia versus muscle tension dysphonia. *The Laryngoscope, 115*(2), 311–316. https://doi.org/10.1097/01.mlg.0000154739.48314.ee

Roy, N., Mauszycki, S. C., Merrill, R. M., Gouse, M., & Smith, M. E. (2007). Toward improved differential diagnosis of adductor spasmodic dysphonia and muscle tension dysphonia. *Folia Phoniatrica et Logopaedica, 59*(2), 83–90. https://doi.org/10.1159/000098341

Roy, N., Mazin, A., & Awan, S. N. (2014). Automated acoustic analysis of task dependency in adductor spasmodic dysphonia versus muscle tension dysphonia. *The Laryngoscope, 124*(3), 718–724. https://doi.org/10.1002/lary.24362

Roy, N., Whitchurch, M., Merrill, R. M., Houtz, D., & Smith, M. E. (2008). Differential diagnosis of adductor spasmodic dysphonia and muscle tension dysphonia using phonatory break analysis. *The Laryngoscope, 118*(12), 2245–2253. https://doi.org/10.1097/MLG.0b013e318184577c

Rubin, A. D., Wodchis, W. P., Spak, C., Kileny, P. R., & Hogikyan, N. D. (2004). Longitudinal effects of Botox injections on voice-related quality of life (V-RQOL) for patients with adductory spasmodic dysphonia: Part II. *Archives of Otolaryngology—Head & Neck Surgery, 130*(4), 415–420. https://doi.org/10.1001/archotol.130.4.415

Rumbach, A. F., Blitzer, A., Frucht, S. J., & Simonyan, K. (2017). An open-label study of sodium oxybate in spasmodic dysphonia. *The Laryngoscope, 127*(6), 1402–1407. https://doi.org/10.1002/lary.26381

Sackett, D. L. (1979). Bias in analytic research. In M. A. Ibrahim (Ed.), *The case-control study consensus and controversy* (pp. 51–63). Elsevier.

Sapienza, C. M., Cannito, M. P., Murry, T., Branski, R., & Woodson, G. (2002). Acoustic variations in reading produced by speakers with spasmodic dysphonia pre-Botox injection and within early stages of Botox injection. *Journal of Speech, Language, and Hearing Research, 45*(5), 830–843. https://doi.org/10.1044/1092-4388(2002/067)

Sapienza, C. M., Murry, T., & Brown, W. S. (1998). Variations in adductor spasmodic dysphonia: Acoustic evidence. *Journal of Voice, 12*(2), 214–222. https://doi.org/10.1016/s0892-1997(98)80041-6

Sapienza, C. M., Walton, S., & Murry, T. (1999). Acoustic variations in adductor spasmodic dysphonia as a function of speech task. *Journal of Speech, Language, and Hearing Research, 42*(1), 127–140. https://doi.org/10.1044/jslhr.4201.127

Sapienza, C. M., Walton, S., & Murry, T. (2000). Adductor spasmodic dysphonia and muscular tension dysphonia: Acoustic analysis of sustained phonation and reading. *Journal of Voice, 14*(4), 502–520. https://doi.org/10.1016/S0892-1997(00)80008-9

Shoffel-Havakuk, H., Marks, K. L., Morton, M., Johns, M. M., III, & Hapner, E. R. (2019). Validation of the OMNI vocal effort scale in the treatment of adductor spasmodic dysphonia. *The Laryngoscope, 129*(2), 448–453. https://doi.org/10.1002/lary.27430

Shrivastav, R., Sapienza, C. M., & Nandur, V. (2005). Application of psychometric theory to the measurement of voice quality using rating scales. *Journal of Speech, Language, and Hearing Research, 48*(2), 323–335. https://doi.org/10.1044/1092-4388(2005/022)

Siemons-Luhring, D. I., Moerman, M., Martens, J. P., Deuster, D., Muller, F., & Dejonckere, P. (2009). Spasmodic dysphonia, perceptual and acoustic analysis: Presenting new diagnostic tools. *European Archives of Oto-Rhino-Laryngology, 266*(12), 1915–1922. https://doi.org/10.1007/s00405-009-0995-7

Simonyan, K., Barkmeier-Kraemer, J., Blitzer, A., Hallett, M., Houde, J. F., Kimberley, T. J., Ozelius, L. J., Pitman, M. J., Richardson, R. M., Sharma, N., & Tanner, K. (2021). Laryngeal dystonia: Multidisciplinary update on terminology, pathophysiology, and research priorities. *Neurology, 96*(21), 989–1001. https://doi.org/10.1212/WNL.0000000000011922

Tisch, S. H., Brake, H. M., Law, M., Cole, I. E., & Darveniza, P. (2003). Spasmodic dysphonia: Clinical features and effects of botulinum toxin therapy in 169 patients—An Australian experience. *Journal of Clinical Neuroscience, 10*(4), 434–438. https://doi.org/10.1016/s0967-5868(03)00020-1

Truong, D. D., Rontal, M., Rolnick, M., Aronson, A. E., & Mistura, K. (1991). Double-blind controlled study of botulinum toxin in adductor spasmodic dysphonia. *The Laryngoscope, 101*(6), 630–634. https://doi.org/10.1288/00005537-199106000-00010

Watts, C., Nye, C., & Whurr, R. (2006). Botulinum toxin for treating spasmodic dysphonia (laryngeal dystonia): A systematic Cochrane review. *Clinical Rehabilitation, 20*(2), 112–122. https://doi.org/10.1191/0269215506cr931oa

Wong, D. L., Irish, J. C., Adams, S. G., Durkin, L. C., & Hunt, E. J. (1995). Laryngeal image analysis following botulinum toxin injections in spasmodic dysphonia. *The Journal of Otolaryngology, 24*(1), 64–68.

Yanagida, S., Nishizawa, N., Hashimoto, R., Mizoguchi, K., Hatakeyama, H., Homma, A., & Fukuda, S. (2018). Reliability and validity of speech evaluation in adductor spasmodic dysphonia. *Journal of Voice, 32*(5), 585–591. https://doi.org/10.1016/j.jvoice.2017.06.022

Yiu, Y., Baylor, C. R., Bamer, A. M., Shelly, S., Klein, A. M., Garrett, C. G., & Pitman, M. J. (2021). Validation of the communicative participation item bank as an outcome measure for spasmodic dysphonia. *The Laryngoscope, 131*(4), 859–864. https://doi.org/10.1002/lary.28897

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America, 123*(5), 3878. https://doi.org/10.1121/1.2935783