The Laryngoscope

WILEY

ORIGINAL REPORT OPEN ACCESS

Automated Creak Identifies Laryngeal Dystonia During Conversational Speech

Daria A. Dragicevic[1] | Cara E. Sauder[2] | Daniel P. Buckley[1,3] | Laura E. Toles[4] | Lauren F. Tracy[3] | J. Pieter Noordzij[3] | Gregory A. Grillone[3] | Ted Mau[4] | John Paul Giliberto[5] | Tanya K. Meyer[6] | Albert L. Merati[6] | Tanya L. Eadie[2,6] | Cara E. Stepp[1,3,5] | Katherine L. Marks[1,7]

[1]Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts, USA | [2]Department of Speech and Hearing Sciences, University of Washington, Seattle, Massachusetts, USA | [3]Department of Otolaryngology—Head and Neck Surgery, Chobanian & Avedisian School of Medicine, Boston University, Boston, Massachusetts, USA | [4]Department of Otolaryngology—Head and Neck Surgery, Voice Center, University of Texas Southwestern Medical Center, Dallas, Texas, USA | [5]Department of Biomedical Engineering, Boston University, Boston, Massachusetts, USA | [6]Department of Otolaryngology—Head and Neck Surgery, University of Washington School of Medicine, Seattle, Washington, USA | [7]Department of Otolaryngology, Emory University School of Medicine, Atlanta, Georgia, USA

Correspondence: Daria A. Dragicevic (ddragic@bu.edu)

Received: 21 February 2025 | **Revised:** 12 November 2025 | **Accepted:** 12 December 2025

Keywords: spasmodic dysphonia | vocal fry | vocal hyperfunction

ABSTRACT

Objective: Creak is an acoustic feature that serves an umbrella term for irregularities in fundamental frequency. This study evaluated whether automated creak distinguished speakers with adductor laryngeal dystonia (AdLD), primary muscle tension dysphonia (pMTD), and those without voice disorders during conversational speech.

Methods: Two stimulus types (reading passage and conversational speech) were produced by 50 speakers with AdLD, 50 speakers with pMTD, and 50 control speakers. An automated creak detector was used to calculate the percentage (%) of creak in both stimuli. The effects of group, stimulus type, and their interaction on % creak were assessed. Receiver operating characteristic curve analyses were conducted and the area under the curve (AUC) was used to evaluate the effectiveness of the creak detector in distinguishing among groups.

Results: We found a statistically significant effect of group on % creak, but no effect of stimulus type on % creak, suggesting that creak distinguished speakers with AdLD, pMTD, and controls during both a reading passage and conversational speech. AUC values indicated acceptable discriminative ability in distinguishing AdLD speakers from both pMTD speakers and controls across both stimuli types.

Conclusion: Automated creak demonstrates promise as a discriminative feature for identifying differences between speakers with AdLD from pMTD and controls in both a reading passage and conversational speech.

Level of Evidence: 3.

1 | Introduction

Laryngeal dystonia (LD) is a focal dystonia with neurological origins [1]. Adductor laryngeal dystonia (AdLD), or spasmodic dysphonia, is the most common subtype, characterized by involuntary spasms of the adductor laryngeal muscles [2, 3]. These spasms cause phonatory breaks, pitch shifts, irregular fundamental frequency (f_o), as measured by spectral noise, low f_o,

Contents of this manuscript were presented at The 13th International Conference on Voice Physiology and Biomechanics, Erlangen, Germany on July 25, 2024.

aperiodicity [4–7], and creak [8]. Creak, an acoustic umbrella term for types of irregularities in f_o, can be characterized as low or irregular f_o, a constricted glottis with low glottal airflow indicated by lower values of H1–H2, damped glottal pulses, and/or the presence of subharmonics [9]. Over half of individuals with AdLD may experience co-occurring vocal tremor [10, 11]. AdLD symptoms hinder communication [12–14], voice-related quality of life [15–17], and work productivity [18–20].

Symptom severity in AdLD depends on task: connected speech is often more affected than sustained vowels [21]. Symptom severity is also greater during voiced-loaded speech than voiceless-loaded speech [22–25] and syntactically complex speech relative to syntactically simple speech [23]. AdLD symptoms can resemble primary muscle tension dysphonia (pMTD) [25–27], but differ by being task-specific, typically occurring in voiced-loaded tasks, whereas pMTD symptoms persist across tasks [21, 24, 27]. Despite these differences, differential diagnosis of AdLD remains challenging [28]. There is no single diagnostic test, and the task-specific nature of AdLD is often considered only when AdLD is already suspected [24, 29]. Differentiation of the two disorders is essential, as they respond differently to treatment options.

Creak is a common acoustic feature in speech [8]. Variations include prototypical creak (particularly at phrase boundaries) vocal fry, breathy creak, diplophonia, and extreme aperiodicity [9]. One study found that creak was significantly higher in AdLD compared to pMTD [30]. However, creak also occurs in typical speakers [31–34]. Creak is used to show expression and inflection or to indicate phrasal boundaries [35] and is more prevalent in informal contexts [36]. One small study found increased creak in conversational speech relative to a reading passage in controls [37]; whether this applies to AdLD remains unknown.

Marks et al. [30] found that creak distinguished AdLD from pMTD and from controls, using a standard reading passage in a small dataset ($n=16$ per group). One advantage they noted in the study is that creak appeared agnostic to phonemic stimuli. However, it is unclear whether findings would generalize to conversational speech, in which creak may be more common in typical speakers. In everyday speech, speakers typically use a more relaxed, conversational speaking style that includes more prosodic variation compared to reading [38, 39]. Effective clinical tools should be generalizable across contexts, such as speaking tasks [39]. Thus, this study aimed to evaluate the discriminative ability of automated creak estimates in differentiating among speakers with AdLD, speakers with pMTD, and speakers with typical voices during conversational speech compared to reading. We sought to validate the discriminative performance of creak in a larger sample size and during an ecologically valid speech task. We hypothesized that automated creak would have weaker discriminative accuracy between groups in conversation compared to reading aloud.

2 | Materials and Methods

2.1 | Participants

Fifty individuals with AdLD, 50 individuals with pMTD, and 50 individuals without voice disorders (controls) participated at one of the following sites: Boston University, the University of Washington, and the University of Texas Southwestern Medical Center. Participants provided written informed consent, per the Boston University Institutional Review Board (#2625).

Participants with AdLD and pMTD were included based on case history, comprehensive voice evaluation, and videostroboscopy by a board-certified otolaryngologist and a voice-specialized speech-language pathologist (SLP). Participants with AdLD were 42 females and 8 males ($M=58.5$ years, SD$=14.2$ years) and included individuals with co-occurring Meige syndrome ($n=1$), foot dystonia ($n=1$), tardive dyskinesia ($n=1$), and Barrett's esophagus ($n=1$). Otherwise, participants had no history of any other speech-related or neurological disorders. Ninety-four percent of participants with AdLD ($n=47$) were receiving botulinum toxin (Botox) injections. Recordings were conducted immediately prior to each participant's next scheduled Botox injection; injection intervals ranged from 8 to 40 weeks ($M=15$ weeks). Participants with pMTD and controls were matched to the AdLD group by sex and age within 2 years. Participants with pMTD were included based on a diagnosis of pMTD: hyperfunction in the absence of structural/phonotrauma or neurological condition. All control participants reported no history of voice, speech, hearing, or neurological disorders. All participants were speakers of American English. Full demographics are listed in Table 1.

Five blinded voice-specialized SLPs, with at least 6 years of experience working primarily with voice disorders ($M=9$ years), completed the overall severity (OS) of dysphonia section of the Consensus Auditory-Perceptual Evaluation of Voice [40] using conversational speech samples and the Rainbow Passage [41]. Samples were rated on a visual analog scale ranging from 0 to 100 mm, with textual anchors for mild at 10, moderate at 35, and severe at 72 mm; 20% of samples were repeated to assess intra-rater reliability. OS of dysphonia, averaged across SLPs, ranged from 6.43 to 83.58 in AdLD, 4.94 to 58.31 in pMTD, and 0.81 to 17.75 in control speakers. Intra- and inter-rater reliability were calculated using intraclass correlation coefficients (ICCs) [42]. Average intra-rater reliability ICC (A, 1) was 0.93 (SD$=0.06$; range: 0.83–0.97). Inter-rater reliability ICC (A, k) was 0.89.

As part of the current study, participants with AdLD were also evaluated for vocal tremor using a combined sample of sustained vowel production, a voice-loaded sentence, and a voiceless-loaded sentence based on a validated screening tool for assessing LD subtypes and the presence of co-occurring tremor [43]. Although the screening tool was originally developed for use via telephone, in this study, we used the same acoustic recordings that were analyzed for creak. For each sample, the blinded voice-specialized SLP rated vocal tremor as present or absent and marked whether they were certain. For the three participants for whom the primary SLP was uncertain, two additional blinded voice-specialized SLPs performed ratings and the majority opinion determined the final tremor distinction. Based on these ratings, 36% of participants with AdLD presented with co-occurring tremor, roughly consistent with the previously reported co-occurrence of tremor in AdLD [10, 11].

2.2 | Procedure

Acoustic recordings were obtained using a headset microphone placed 45 degrees from the midline and 7 cm from the lips and

TABLE 1 | Demographic characteristics and CAPE-V overall severity ratings for participants with adductor laryngeal dystonia (AdLD), muscle tension dysphonia (PMTD), and controls.

Group		AdLD	PMTD	Control
Age, *M* (SD), years		*58* (14)	*56* (12)	*59* (14)
Age range, years		25–80	24–76	24–79
Sex, *n*	Female	42	42	42
	Male	8	8	8
Race, *n*	Asian	1	2	0
	Black or African American	2	15	5
	White	44	23	43
	More than one race	1	2	1
	Unknown/not reported	2	8	1
Ethnicity, *n*	Hispanic or Latino	2	6	2
	Not Hispanic or Latino	44	33	47
	Not reported or prefer not to answer	4	11	1
Overall severity of dysphonia, *M* (SD)	Rainbow Passage	40.6 (15.4)	16.2 (8.8)	8.0 (3.4)
	Conversational speech	38.3 (18.1)	16.0 (9.9)	8.5 (3.6)

Note: The mean (*M*), standard deviation (SD), and age range are reported for each group. Sex, race, and ethnicity are presented as number of individuals per group (*n*). Mean (*M*) and standard deviation (SD) values are also provided for overall severity of dysphonia, as measured by the consensus auditory–perceptual evaluation of voice (CAPE-V), for each group and each stimulus type.

digitized at 44.1 kHz. Among AdLD participants, 14% were recorded in a sound-treated room and 86% in a quiet room; of the pMTD participants, 26% were recorded in a sound-treated room and 74% in a quiet room; all controls were recorded in a sound-treated room. There was no difference in % creak for any of the groups based on the recording environment. Two stimulus types were used: the first paragraph of the Rainbow Passage and conversational speech. Prompts were provided to elicit the conversational speech, including: “How did you get here today?; What did you do last weekend?” The first 20 s of each recording were analyzed to ensure equal data across participants and stimulus types.

The percentage (%) of creak was calculated via an automated open-source algorithm [44] implemented in a custom MATLAB script [30]. The algorithm was trained to detect at least three patterns of creak: highly irregular temporal characteristics, fairly regular temporal characteristics with strong secondary excitation peaks, and fairly regular temporal characteristics without strong secondary excitation peaks [44]. The % creak was calculated per speaker per stimulus type, defined as the total time creak occurs in the voiced segments divided by the total duration of the voiced segments.

2.3 | Statistical Analysis

An analysis of covariance (ANCOVA) was performed to test the effects of group (AdLD, pMTD, controls), stimulus type (Rainbow Passage, conversational speech), and their interactions on % creak. Tremor and OS were included as covariates to account for the presence of tremor in the AdLD group and the varying levels of OS across the groups. Effect sizes for significant effects were calculated as partial eta squared (η_p^2) with 95% confidence intervals (CI). Values were interpreted as small (0.01), medium (0.09), and large (0.25) [45]. Post hoc comparisons between groups were made using Tukey's tests of multiple comparisons. Effect sizes for significant differences were calculated using Cohen's *d* [46], with values interpreted as small (0.25), medium (0.55), or large (>0.93) [47]. Statistical analyses were conducted using R [48–53] with a statistical significance set a priori to $p < 0.05$.

Receiver operating characteristic (ROC) curve analyses were completed using a custom MATLAB script [54] to assess the sensitivity and specificity of % creak in distinguishing between the three groups for each stimuli type. ROC curves plot sensitivity versus 1—specificity at various thresholds. Sensitivity, or the true positive rate, is the proportion of actual positives correctly identified by the model, with higher values indicating better performance, whereas specificity is the proportion of actual negatives correctly identified (1—false positive rate), with higher specificity indicating better performance. A threshold was set to achieve 70% sensitivity, aiming to find a point on the curve that optimally balanced the tradeoff between the true positive rate and false positive rate across all group and stimuli combinations [55, 56]. ROC curves also provide the area under the curve (AUC), a measure of diagnostic accuracy. An AUC of 1 represents perfect classification, 0.5 no discriminative ability, and 0 perfectly incorrect classification [57].

3 | Results

The mean % creak and 95% CI are displayed in Figure 1 for each stimuli type and group. The ANCOVA revealed a statistically significant effect of group ($p<0.001$) with a medium effect size ($\eta_p^2=0.17$) and a small but significant effect of OS ($p=0.0187$; $\eta_p^2=0.02$) yet no significant effect of stimuli, interaction, nor presence of tremor (see Table 2). Post hoc tests revealed statistically significant differences between the AdLD group and controls ($p<0.001$; $d=1.02$ with a 95% CI [0.73–1.32]), indicating a large effect size, and between the AdLD and pMTD groups ($p<0.001$; $d=0.71$ with a 95% CI [0.42–0.99]), indicating a medium effect size, but not between the pMTD group and controls ($p=0.76$). Post hoc mediation analyses [58] were conducted to determine whether OS mediated the effect of group on % creak; OS was not found to be a mediator. Detailed results of the mediation analysis are presented in Data S1.

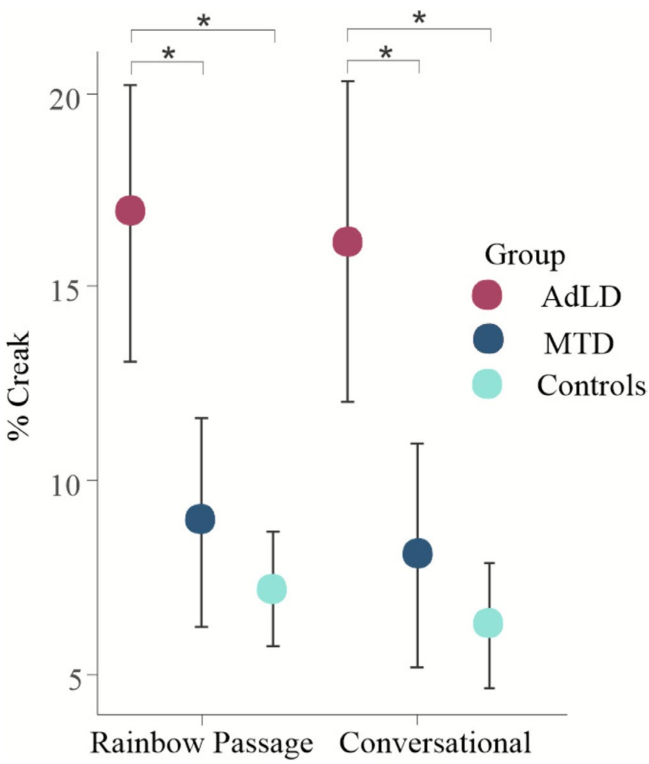

FIGURE 1 | Mean percent (%) creak and 95% confidence intervals are plotted for each stimuli type (Rainbow Passage and conversational speech) for each group (adductor laryngeal dystonia [AdLD], muscle tension dysphonia [PMTD], and controls). *Statistically significant difference at $p<0.05$.

Results of the ROC curve analyses are shown in Figure 2—Panel A for the Rainbow Passage and Panel B for conversational speech, with results detailed in Table 3. The AUC for % creak between AdLD and pMTD was 0.75 during conversational speech and 0.72 during the Rainbow Passage. The AUC for % creak between AdLD and controls was 0.77 during conversational speech and 0.75 during the Rainbow Passage. These results indicate acceptable discriminative ability [57]. The AUC for % creak between controls and pMTD was 0.49 during both conversational speech and the Rainbow Passage, indicating no discriminative ability [57]. With the threshold set at 70% sensitivity, for distinguishing AdLD from pMTD, we found a specificity of 72% for conversational speech and 66% for the Rainbow Passage. For distinguishing AdLD from controls, specificities were 68% for conversational speech and 66% for the Rainbow Passage.

4 | Discussion

The purpose of this study was twofold: to validate the discrimination of creak between speakers with AdLD, pMTD, and controls [30] in a larger sample size and to determine whether discrimination would remain during a conversational speaking task. Agathe and Claire [37] found increased creak in conversational speech compared to a reading passage in typical speakers, so we theorized that creak measured from a conversational speaking stimulus would have less discriminative ability than creak measured from the Rainbow Passage. However, we did not observe a statistically significant effect of stimuli on % creak in our study. Unlike Agathe and Claire [37], which was based on a sample of six college-aged, French-speaking females, our study used a larger, more generalizable sample of 50 control speakers, with both female and male speakers, spanning a broad age range. Moreover, the ability of automated creak to differentiate speakers with AdLD from pMTD and controls was the same regardless of stimulus type. These findings validate that creak can be used across speech contexts (reading passage, conversational speech).

Consistent with Marks et al. [30], there was a statistically significant effect of group on % creak for both conversational speech and the Rainbow Passage. Findings revealed a statistically significant difference in % creak between the AdLD group and controls with a large effect size and between the AdLD and pMTD groups with a medium effect size. As illustrated in Figure 1, the AdLD group exhibited a higher mean % creak with more variable % creak values, as reflected by the larger 95% CI. In

TABLE 2 | Results of the analysis of covariance.

Measure	Effect	df	F	p	η_p^2 [95% CI]	Effect size
% Creak	Group	2	29.232	<0.001*	0.17 [0.09–0.24]	Medium
	Stimuli	1	0.537	0.464	NS	—
	Overall severity of dysphonia	1	5.572	0.019	0.02 [0–0.06]	Small
	Presence of tremor	1	1.636	0.202	NS	—
	Group × stimuli	2	0.006	0.994	NS	—

Abbreviations: 95% CI = 95% confidence interval; — = not applicable for nonsignificant findings; % Creak = percentage of creak; NS = not significant.
*Significant at $p<0.05$.

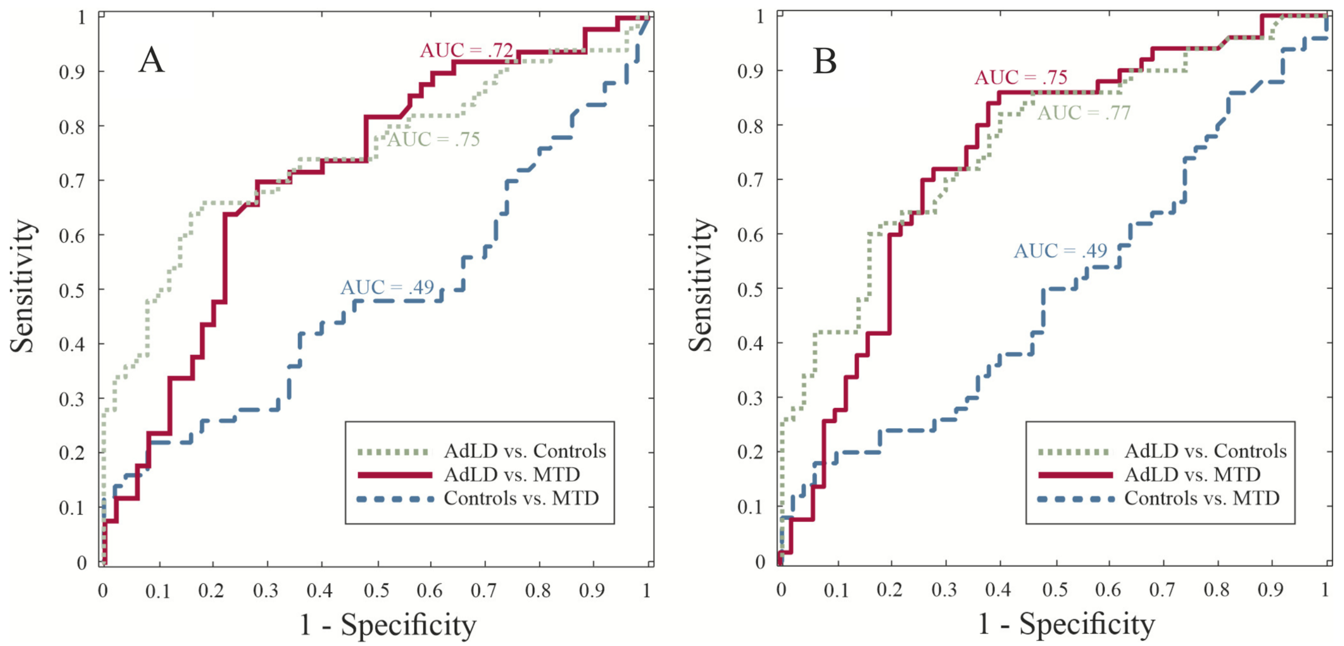

FIGURE 2 | Receiver–operator characteristic (ROC) curve plot illustrating the sensitivity and 1—specificity of percent (%) creak in differentiating speakers with adductor laryngeal dystonia (AdLD) from controls (green dotted lines), AdLD from muscle tension dysphonia (PMTD; red solid lines), and PMTD from controls (blue dashed lines). (A) ROC plot using the Rainbow Passage stimuli. (B) ROC plot using the conversational speech stimuli.

TABLE 3 | Receiver operating characteristic curve results comparing individuals with adductor laryngeal dystonia (AdLD), muscle tension dysphonia (PMTD), and controls.

Group	Stimuli type	AdLD vs. PMTD	PMTD vs. controls	AdLD vs. controls
AUC	Rainbow Passage	0.72	0.49	0.75
	Conversational	0.75	0.49	0.77
Sensitivity	Rainbow Passage	0.70	0.70	0.70
	Conversational	0.70	0.70	0.70
Specificity	Rainbow Passage	0.66	0.32	0.66
	Conversational	0.72	0.26	0.68
Threshold for 70% sensitivity (%)	Rainbow Passage	8.24	—	8.29
	Conversational	7.54	—	7.57

Abbreviations: % = percentage; — = not applicable due to AUC indicating no discriminative ability (<0.5); AUC = area under the curve.

contrast, the pMTD and control groups displayed lower mean % creak values, with their data points and 95% CI concentrated within a smaller range. This distribution suggests that % creak is more variable and higher in the AdLD group compared to the pMTD and control groups, confirming distinct creak patterns in the AdLD group.

Although there was a statistically significant difference in percent creak between individuals with AdLD and the two other groups at the group level, discrimination performance for both stimuli types was poorer than in previous work [30]. Whereas Marks et al. [30] found outstanding discriminative ability of % creak in differentiating AdLD from pMTD and from controls (AUCs of 0.86 and 0.94, respectively), in this larger sample, we found acceptable diagnostic accuracy across both types of stimuli (AUCs of 0.75 and 0.72 for AdLD vs. pMTD and 0.77 and 0.75 AdLD vs. controls) [57]. Lower AUC values in the present study compared to the previous may reflect differences in sample characteristics, such as larger sample size and differences in OS of dysphonia between the voice disorder groups. In the current study, OS was found to be a small yet significant covariate. These results suggest that OS of dysphonia had a small influence on % creak and may have contributed to distinguishing AdLD from the other two groups. Roy et al. [59] found that creak was only weakly correlated to listener-based severity ratings, with significance only found for speakers with AdLD. Together, these findings suggest that the relationship between OS and creak may be specific to the AdLD group, though the overall association remains weak. Future studies should include groups with higher OS that are perceptually matched, to determine whether observed differences can be attributed to factors beyond OS. The AUC values found in the present study had acceptable diagnostic

TABLE 4 | Sensitivity and specificity values for differentiating speakers with adductor laryngeal dystonia and muscle tension dysphonia using acoustic measures.

Measure	Sensitivity	Specificity
% Creak in the Rainbow Passage	0.70	0.66
% Creak in conversational speech	0.70	0.72
CSID [21]	0.67	0.64
LTAS [60]	0.61	0.68

Abbreviations: % Creak = percentage of creak; CSID = cepstral–spectral index of dysphonia; LTAS = long-term average spectrum.

accuracy [57]; however, some diagnostic uncertainty remains. Further research is needed to determine whether creak is suited as a screening or diagnostic tool before clinical application.

The discriminative accuracy of creak in the current study is consistent with previously reported acoustic metrics, such as the Cepstral Spectral Index of Dysphonia (CSID) and the long-term average spectrum (LTAS) in differentiating AdLD from pMTD [21, 51, 52]. However, unlike the other acoustic measures that require specific stimuli or complex calculations, creak does not rely upon task specificity; it can be automatically calculated from natural connected speech, either produced conversationally or read aloud, supporting its use in a more generalizable context. We used a 70% sensitivity threshold to prioritize sensitivity in identifying individuals with AdLD. At 70% sensitivity, % creak achieved specificity values of 66% for the Rainbow Passage and 72% for conversational speech for differentiating speakers with AdLD from those with pMTD, indicating its ability to identify true negatives and maintain effective diagnostic power. In comparison, the CSID has a sensitivity of 67% and specificity of 64% [21] and the LTAS offers a sensitivity of 61% and specificity of 68% (see Table 4) [60]. However, unlike metrics such as the CSID or LTAS which are dependent on voiced/voiceless contrasts or task-specific stimuli, creak demonstrated robust performance across mixed phonemes in connected speech, whether read or conversational, showing promise as a diagnostic tool in ecologically valid settings [39]. These findings suggest that creak is a versatile acoustic tool, supporting its utility in differentiating speakers with AdLD from those with pMTD across various speech contexts.

This study aimed to examine a large sample size of speakers with AdLD; however, there are inherent limitations associated with incorporating data from multiple study sites, including recording environment and diagnostic criteria. Although we did not find any difference in % creak among groups based on the recording environment, previous research has yielded mixed findings on the effect of recording environments on various acoustic measures [61, 62] and should be evaluated prospectively in the future. In the present study, control speakers had a range of OS from 0.81 to 17.75, reflecting speakers with no dysphonia to mild levels of dysphonia. The inclusion of control speakers with mild levels of dysphonia may have contributed to the lack of a significant difference in % creak between speakers with pMTD and controls; controls. The study did not account for signs of bowing other than the multidisciplinary diagnosis of pMTD.

In the current study, participants with AdLD were diagnosed by board-certified otolaryngologists across the United States based on Ludlow's consensus criteria, including task-specificity [24]. However, an inherent limitation is the lack of agreement among laryngologists, SLPs, and neurologists across the United States [24]. Further, the vocal tremor screening tool [43] used here was originally designed for telephone diagnosis, but in the present study, expert listeners applied it to the same acoustic recordings analyzed in this study. The screening tool is not a definitive diagnostic tool and it is possible that some speakers with a co-occurrence of tremor may have gone undetected. Thus, it is possible that diagnostic heterogeneity may have confounded our results. It is also possible that ongoing effects of Botox contributed to laryngeal and/or oscillatory function in the speakers with AdLD who had previously received Botox. Further investigation is warranted regarding the discriminative validity of creak between Botox-naïve speakers with AdLD and speakers with pMTD. Finally, phonemic context of conversational speech was not controlled, which may influence symptom expression. Further research on the effects of phonemic context in creak is warranted.

5 | Conclusion

Automated creak estimates differentiated speakers with AdLD from pMTD and controls with similar performance across conversational and read speaking tasks. These results support discriminative and ecological validity for creak as a discriminative measure for AdLD.

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders (DC015570 to C.E.S. and DC020349 to K.L.M.), Boston University Sargent College Student Research Grants (K.L.M. and D.A.D.), an American Speech Language Hearing Foundation Speech Science Grant (K.L.M.), and a Graduate Fellow Award from the Rafik B. Hariri Institute for Computing and Computational Science and Engineering (D.A.D.).

Conflicts of Interest

Cara E. Stepp has received consulting fees from Altec Inc./Delsys Inc., companies 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 competing financial or nonfinancial interests existed at the time of publication.

References

1. A. Blitzer, M. F. Brin, K. Simonyan, L. J. Ozelius, and S. J. Frucht, "Phenomenology, Genetics, and CNS Network Abnormalities in Laryngeal Dystonia: A 30-Year Experience," *Laryngoscope* 128, no. Suppl 1 (2018): S1–S9, https://doi.org/10.1002/lary.27003.

2. M. P. Cannito, M. Doiuchi, T. Murry, and G. E. Woodson, "Perceptual Structure of Adductor Spasmodic Dysphonia and Its Acoustic Correlates," *Journal of Voice* 26, no. 6 (2012): 818.e5–818.13, https://doi.org/10.1016/j.jvoice.2012.05.005.

3. B. Haslinger, P. Erhard, C. Dresel, F. Castrop, M. Roettinger, and A. O. Ceballos-Baumann, ""Silent Event-Related" fMRI Reveals Reduced Sensorimotor Activation in Laryngeal Dystonia," *Neurology* 65, no. 10 (2005): 1562–1569.

4. C. M. Sapienza, T. Murry, and W. S. Brown, "Variations in Adductor Spasmodic Dysphonia: Acoustic Evidence," *Journal of Voice* 12, no. 2 (1998): 214–222, https://doi.org/10.1016/S0892-1997(98)80041-6.

5. C. M. Sapienza, S. Walton, and T. Murry, "Acoustic Variations in Adductor Spasmodic Dysphonia as a Function of Speech Task," *Journal of Speech, Language, and Hearing Research* 42, no. 1 (1999): 127–140, https://doi.org/10.1044/jslhr.4201.127.

6. C. M. Sapienza, S. Walton, and T. Murry, "Adductor Spasmodic Dysphonia and Muscular Tension Dysphonia: Acoustic Analysis of Sustained Phonation and Reading," *Journal of Voice* 14, no. 4 (2000): 502–520, https://doi.org/10.1016/S0892-1997(00)80008-9.

7. C. M. Sapienza, M. P. Cannito, T. Murry, R. Branski, and G. Woodson, "Acoustic Variations in Reading Produced by Speakers With Spasmodic Dysphonia Pre-Botox Injection and Within Early Stages of Post-Botox Injection," *Journal of Speech, Language, and Hearing Research* 45, no. 5 (2002): 830–843, https://doi.org/10.1044/1092-4388(2002/067).

8. K. L. Marks, T. F. Feaster, S. Baker, D. CME, P. C. Doyle, and C. E. Stepp, "Spectral Aggregate of the High-Passed Fundamental Frequency and Its Relationship to the Primary Acoustic Features of Adductor Laryngeal Dystonia," *Journal of Speech, Language, and Hearing Research* 65, no. 11 (2022): 4085–4095, https://doi.org/10.1044/2022_JSLHR-22-00157.

9. P. Keating, M. Garellek, and J. Kreiman, "Acoustic Properties of Different Kinds of Creaky Voice," *Journal of the Acoustical Society of America* 153 (2015): A297.

10. A. B. Patel, S. F. Bansberg, C. H. Adler, D. G. Lott, and L. Crujido, "The Mayo Clinic Arizona Spasmodic Dysphonia Experience: A Demographic Analysis of 718 Patients," *Annals of Otology, Rhinology, and Laryngology* 124, no. 11 (2015): 859–863, https://doi.org/10.1177/0003489415588557.

11. D. Wolraich, C. Vasile Marchis-Crisan, N. Redding, S. L. Khella, and N. Mirza, "Laryngeal Tremor: Co-Occurrence With Other Movement Disorders," *ORL* 72, no. 5 (2010): 291–294, https://doi.org/10.1159/000317032.

12. C. R. Baylor, K. M. Yorkston, and T. L. Eadie, "The Consequences of Spasmodic Dysphonia on Communication-Related Quality of Life: A Qualitative Study of the Insider's Experiences," *Journal of Communication Disorders* 38, no. 5 (2005): 395–419.

13. R. Epstein, J. Stygall, and S. Newman, "The Short-Term Impact of Botox Injections on Speech Disability in Adductor Spasmodic Dysphonia," *Disability and Rehabilitation* 19, no. 1 (1997): 20–25, https://doi.org/10.3109/09638289709166441.

14. E. Smith, M. Taylor, M. Mendoza, J. Barkmeier, J. Lemke, and H. Hoffman, "Spasmodic Dysphonia and Vocal Fold Paralysis: Outcomes of Voice Problems on Work-Related Functioning," *Journal of Voice* 12, no. 2 (1998): 223–232, https://doi.org/10.1016/S0892-1997(98)80042-8.

15. N. D. Hogikyan, W. P. Wodchis, C. Spak, and P. R. Kileny, "Longitudinal Effects of Botulinum Toxin Injections on Voice-Related Quality of Life (V-RQOL) for Patients With Adductory Spasmodic Dysphonia," *Journal of Voice* 15, no. 4 (2001): 576–586, https://doi.org/10.1016/S0892-1997(01)00060-1.

16. S. Morzaria and E. J. Damrose, "A Comparison of the VHI, VHI-10, and V-RQOL for Measuring the Effect of Botox Therapy in Adductor Spasmodic Dysphonia," *Journal of Voice* 26, no. 3 (2012): 378–380, https://doi.org/10.1016/j.jvoice.2010.07.011.

17. E. P. Silverman, C. Garvan, R. Shrivastav, and C. M. Sapienza, "Combined Modality Treatment of Adductor Spasmodic Dysphonia," *Journal of Voice* 26, no. 1 (2012): 77–86, https://doi.org/10.1016/j.jvoice.2010.08.004.

18. D. Isetti and T. Meyer, "Workplace Productivity and Voice Disorders: A Cognitive Interviewing Study on Presenteeism in Individuals With Spasmodic Dysphonia," *Journal of Voice* 28, no. 6 (2014): 700–710, https://doi.org/10.1016/j.jvoice.2014.03.017.

19. T. K. Meyer, A. Hu, and A. D. Hillel, "Voice Disorders in the Workplace: Productivity in Spasmodic Dysphonia and the Impact of Botulinum Toxin," *Laryngoscope* 123, no. Suppl 6 (2013): S1–S14, https://doi.org/10.1002/lary.24292.

20. T. K. Meyer, C. Spiekerman, R. Kaye, et al., "Association of Laryngeal Botulinum Neurotoxin Injection With Work Productivity for Patients With Spasmodic Dysphonia," *JAMA Otolaryngology. Head & Neck Surgery* 147, no. 9 (2021): 804–810, https://doi.org/10.1001/jamaoto.2021.1745.

21. N. Roy, A. Mazin, and S. N. Awan, "Automated Acoustic Analysis of Task Dependency in Adductor Spasmodic Dysphonia Versus Muscle Tension Dysphonia," *Laryngoscope* 124, no. 3 (2014): 718–724, https://doi.org/10.1002/lary.24362.

22. M. P. Cannito, L. B. Chorna, J. C. Kahane, and J. P. Dworkin, "Influence of Consonant Voicing Characteristics on Sentence Production in Abductor Versus Adductor Spasmodic Dysphonia," *Journal of Voice* 28, no. 3 (2014): 394.e13–394.e22, https://doi.org/10.1016/j.jvoice.2013.10.010.

23. M. L. Erickson, "Effects of Voicing and Syntactic Complexity on Sign Expression in Adductor Spasmodic Dysphonia," *American Journal of Speech-Language Pathology* 12, no. 4 (2003): 416–424, https://doi.org/10.1044/1058-0360(2003/087).

24. C. L. Ludlow, R. Domangue, D. Sharma, et al., "Consensus-Based Attributes for Identifying Patients With Spasmodic Dysphonia and Other Voice Disorders," *JAMA Otolaryngology. Head & Neck Surgery* 144, no. 8 (2018): 657–665, https://doi.org/10.1001/jamaoto.2018.0644.

25. N. Roy, S. C. Mauszycki, R. M. Merrill, M. Gouse, and M. E. Smith, "Toward Improved Differential Diagnosis of Adductor Spasmodic Dysphonia and Muscle Tension Dysphonia," *Folia Phoniatrica et Logopaedica* 59, no. 2 (2007): 83–90.

26. C. Ludlow, "Management of the Spasmodic Dysphonias," *Voice Disorders: Diagnosis and Treatment* 115, no. 4 (1995): 436–454.

27. N. Roy, M. Gouse, S. C. Mauszycki, R. M. Merrill, and M. E. Smith, "Task Specificity in Adductor Spasmodic Dysphonia Versus Muscle Tension Dysphonia," *Laryngoscope* 115, no. 2 (2005): 311–316, https://doi.org/10.1097/01.mlg.0000154739.48314.ee.

28. F. X. Creighton, E. Hapner, A. Klein, A. Rosen, H. A. Jinnah, and M. M. Johns, "Diagnostic Delays in Spasmodic Dysphonia: A Call for Clinician Education," *Journal of Voice* 29, no. 5 (2015): 592–594, https://doi.org/10.1016/j.jvoice.2013.10.022.

29. N. Roy, "Differential Diagnosis of Muscle Tension Dysphonia and Spasmodic Dysphonia," *Current Opinion in Otolaryngology & Head and Neck Surgery* 18, no. 3 (2010): 165–170, https://doi.org/10.1097/MOO.0b013e328339376c.

30. K. L. Marks, M. E. Díaz Cádiz, L. E. Toles, et al., "Automated Creak Differentiates Adductor Laryngeal Dystonia and Muscle Tension Dysphonia," *Laryngoscope* 133, no. 10 (2023): 2687–2694, https://doi.org/10.1002/lary.30588.

31. S. A. Borrie and C. R. Delfino, "Conversational Entrainment of Vocal Fry in Young Adult Female American English Speakers," *Journal of Voice* 31, no. 4 (2017): 513.e25–513.e32, https://doi.org/10.1016/j.jvoice.2016.12.005.

32. K. Dallaston and G. Docherty, "The Quantitative Prevalence of Creaky Voice (Vocal Fry) in Varieties of English: A Systematic Review of the Literature," *PLoS One* 15, no. 3 (2020): e0229960, https://doi.org/10.1371/journal.pone.0229960.

33. L. Wolk, N. B. Abdelli-Beruh, and D. Slavin, "Habitual Use of Vocal Fry in Young Adult Female Speakers," *Journal of Voice* 26, no. 3 (2012): e111–e116, https://doi.org/10.1016/j.jvoice.2011.04.007.

34. I. Yuasa, "Creaky Voice: A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women?," *American Speech* 85 (2010): 315–337, https://doi.org/10.1215/00031283-2010-018.

35. L. Davidson, "The Versatility of Creaky Phonation: Segmental, Prosodic, and Sociolinguistic Uses in the World's Languages," *WIREs Cognitive Science* 12 (2021): e1547, https://doi.org/10.1002/wcs.1547.

36. R. J. Podesva, "Phonation Type as a Stylistic Variable: The Use of Falsetto in Constructing a Personal," *Journal of SocioLinguistics* 11, no. 4 (2007): 478–504, https://doi.org/10.1111/j.1467-9841.2007.00334.x.

37. B. L. Agathe and P. Claire, "The Influence of Language and Speech Task Upon Creaky Voice Use Among Six Young American Women Learning French," in *Interspeech 2013* (International Speech Communication Association, 2013).

38. G. Mehta and A. Cutler, "Detection of Target Phonemes in Spontaneous and Read Speech," *Language and Speech* 31, no. 2 (1988): 135–156, https://doi.org/10.1177/002383098803100203.

39. M. Nakamura, K. Iwano, and S. Furui, "Differences Between Acoustic Characteristics of Spontaneous and Read Speech and Their Effects on Speech Recognition Performance," *Computer Speech & Language* 22, no. 2 (2008): 171–184, https://doi.org/10.1016/j.csl.2007.07.003.

40. G. B. Kempster, B. R. Gerratt, K. Verdolini Abbott, J. Barkmeier-Kraemer, and R. E. Hillman, "Consensus Auditory–Perceptual Evaluation of Voice: Development of a Standardized Clinical Protocol," *American Journal of Speech-Language Pathology* 18, no. 2 (2009): 124–132, https://doi.org/10.1044/1058-0360(2008/08-0017).

41. Grant Fairbanks, *Voice and Articulation Drillbook* (Grant Fairbanks, 1960).

42. M. Gamer, *irr: Various Coefficients of Interrater Reliability and Agreement* (Httpcranr-Proj, 2010), https://cir.nii.ac.jp/crid/1572824499179920768.

43. D. M. Johnson, E. R. Hapner, A. M. Klein, M. Pethan, and M. M. Johns, "Validation of a Telephone Screening Tool for Spasmodic Dysphonia and Vocal Fold Tremor," *Journal of Voice* 28, no. 6 (2014): 711–715, https://doi.org/10.1016/j.jvoice.2014.03.009.

44. T. Drugman, J. Kane, and C. Gobl, "Data-Driven Detection and Analysis of the Patterns of Creaky Voice," *Computer Speech & Language* 28, no. 5 (2014): 1233–1253, https://doi.org/10.1016/j.csl.2014.03.002.

45. R. S. Witte and J. S. Witte, *Statistics* (John Wiley & Sons, 2017).

46. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Routledge, 2013), https://doi.org/10.4324/9780203771587.

47. L. Gaeta and C. R. Brydges, "An Examination of Effect Sizes and Statistical Power in Speech, Language, and Hearing Research," *Journal of Speech, Language, and Hearing Research* 63, no. 5 (2020): 1572–1580, https://doi.org/10.1044/2020_JSLHR-19-00299.

48. D. Bates, M. Maechler, B. Bolker, et al., "lme4: Linear Mixed-Effects Models Using 'Eigen' and S4," 2024, accessed June 25, https://cran.r-project.org/web/packages/lme4/index.html.

49. A. Kassambara, "rstatix: Pipe-Friendly Framework for Basic Statistical Tests," 2023, accessed June 25, 2024, https://cran.r-project.org/web/packages/rstatix/index.html.

50. J. Kloke and J. McKean, "Rfit: Rank-Based Estimation for Linear Models," 2024, accessed June 25, 2024, https://cran.r-project.org/web/packages/Rfit/index.html.

51. R Core Team, *R: A Language and Environment for Statistical* (R Core Team, 2023), https://www.R-project.org/.

52. W. Revelle, "psych: Procedures for Psychological, Psychometric, and Personality Research," 2024 published online March 18, 2024, accessed June 25, 2024, https://cran.r-project.org/web/packages/psych/index.html.

53. H. Wickham, R. François, L. Henry, et al., "dplyr: A Grammar of Data Manipulation," 2023, accessed June 25, 2024, https://cran.r-project.org/web/packages/dplyr/index.html.

54. Mathworks, *MATLAB: The Language of Technical Computing: Computation, Visualization, Programming.* (Mathworks, 2023).

55. T. R. Vetter, P. Schober, and E. J. Mascha, "Diagnostic Testing and Decision-Making: Beauty Is Not Just in the Eye of the Beholder," *Anesthesia and Analgesia* 127, no. 4 (2018): 1085–1091, https://doi.org/10.1213/ANE.0000000000003698.

56. E. J. Mascha, "Identifying the Best Cut-Point for a Biomarker, or Not," *Anesthesia and Analgesia* 127, no. 4 (2018): 820–822, https://doi.org/10.1213/ANE.0000000000003680.

57. J. N. Mandrekar, "Receiver Operating Characteristic Curve in Diagnostic Test Assessment," *Journal of Thoracic Oncology* 5, no. 9 (2010): 1315–1316, https://doi.org/10.1097/JTO.0b013e3181ec173d.

58. D. Tingley, T. Yamamoto, K. Hirose, L. Keele, and K. Imai, "Mediation: R Package for Causal Mediation Analysis," *Journal of Statistical Software* 59 (2014): 1–38.

59. N. Roy, S. N. Awan, S. Jennings, J. Jensen, and R. M. Merrill, "Adductor Laryngeal Dystonia Versus Muscle Tension Dysphonia: Examining the Utility of Automated Acoustic Analysis to Detect Task Dependency as a Distinguishing Feature," *Journal of Speech, Language, and Hearing Research* 67, no. 10 (2024): 3612–3630, https://doi.org/10.1044/2024_JSLHR-24-00104.

60. D. R. Houtz, N. Roy, R. M. Merrill, and M. E. Smith, "Differential Diagnosis of Muscle Tension Dysphonia and Adductor Spasmodic Dysphonia Using Spectral Moments of the Long-Term Average Spectrum," *Laryngoscope* 120, no. 4 (2010): 749–757, https://doi.org/10.1002/lary.20741.

61. P. Bottalico, J. Codino, L. C. Cantor-Cutiva, et al., "Reproducibility of Voice Parameters: The Effect of Room Acoustics and Microphones," *Journal of Voice* 34, no. 3 (2020): 320–334, https://doi.org/10.1016/j.jvoice.2018.10.016.

62. B. van der Woed, M. Wu, V. Parsa, P. C. Doyle, and K. Fung, "Evaluation of Acoustic Analyses of Voice in Nonoptimized Conditions," *Journal of Speech, Language, and Hearing Research* 63, no. 12 (2020): 3991–3999, https://doi.org/10.1044/2020_JSLHR-20-00212.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** lary70332-sup-0001-supinfo.docx.