

Digital markers of motor speech impairments in natural speech of patients with ALS-FTD spectrum disorders

Sanjana Shellikeri¹, Sunghye Cho², Sharon Ash¹, Carmen Gonzalez-Recober¹, Corey T. McMillan¹, Lauren Elman³, Colin Quinn³, Defne A. Amado³, Michael Baer³, David J Irwin¹, Lauren Massimo¹, Chris Olm¹, Mark Liberman^{2,4}, Murray Grossman¹, Naomi Nevler¹

1. Penn Frontotemporal Degeneration Center and Department of Neurology, University of Pennsylvania, Philadelphia, PA
2. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA
3. Penn ALS Clinic, University of Pennsylvania, PA
4. Department of Linguistics, University of Pennsylvania, Philadelphia, PA

Please address correspondence to:

Sanjana Shellikeri, PhD or Naomi Nevler, MD
Penn Frontotemporal Degeneration Center, Department of Neurology,
3620 Hamilton Walk,
Philadelphia, PA 19104
Email: sanjana.shellikeri@pennteam.upenn.edu or naomine@pennteam.upenn.edu

Funding Sources

This work is supported by funding from the American Academy of Neurology (2022-2784), the Alzheimer's Association (AARF-D-619473, AARF-D-619473-RAPID, AARF-21-851126), National Institute on Aging (AG073510-01, P01-AG066597, U19-AG062418, P30-AG072979, R01-AG054519 R01-NS109260), the Penn Institute on Aging, and Department of the Army (W81XWH-20-1-0531).

Data Sharing: Anonymized data will be shared by a reasonable request from any qualified investigator.

Author Contributions: SS, SC, SA, SS, CTM, DJI, ML, MG and NN had major contributions to the conception of the research question and overall design of the study. CGR, LE, CQ, DA, MB, LM and MG were involved in data acquisition and measurement. CTM and CO was involved in the neuroimaging analysis. SS performed all statistical analyses and comparisons in this study. SS, SC, SA, CGR, CTM, LE, CQ, DA, MB, DJI, LM, CO, ML, MG and NN were all involved in the manuscript drafting and have approved the final draft.

All authors declare no conflict of interest.

ABSTRACT

Background and objectives: Patients with ALS-FTD spectrum disorders (ALS-FTSD) have mixed motor and cognitive impairments and require valid and quantitative assessment tools to support diagnosis and tracking of bulbar motor disease. This study aimed to validate a novel automated digital speech tool that analyzes vowel acoustics from natural, connected speech as a marker for impaired articulation due to bulbar motor disease in ALS-FTSD.

Methods: We used an automatic algorithm called Forced Alignment Vowel Extraction (FAVE) to detect spoken vowels and extract vowel acoustics from 1 minute audio-recorded picture descriptions. Using automated acoustic analysis scripts, we derived two articulatory-acoustic measures: vowel space area (VSA, in Bark²) which represents tongue range-of-motion (size), and average second formant slope of vowel trajectories (F2 slope) which represents tongue movement speed. We compared vowel measures between ALS with and without clinically-evident bulbar motor disease (ALS+bulbar vs. ALS-bulbar), behavioral variant frontotemporal dementia (bvFTD) without a motor syndrome, and healthy controls (HC). We correlated impaired vowel measures with bulbar disease severity, estimated by clinical bulbar scores and perceived listener effort, and with MRI cortical thickness of the orobuccal part of the primary motor cortex innervating the tongue (oralPMC). We also tested correlations with respiratory capacity and cognitive impairment.

Results: Participants were 45 ALS+bulbar (30 males, mean age=61±11), 22 ALS-nonbulbar (11 males, age=62±10), 22 bvFTD (13 males, age=63±7), and 34 HC (14 males, age=69±8). ALS+bulbar had smaller VSA and shallower average F2 slopes than ALS-bulbar (VSA: $|d|=0.86$, $p=0.0088$; F2 slope: $|d|=0.98$, $p=0.0054$), bvFTD (VSA: $|d|=0.67$, $p=0.043$; F2 slope: $|d|=1.4$, $p<0.001$), and HC (VSA: $|d|=0.73$, $p=0.024$; F2 slope: $|d|=1.0$, $p<0.001$). Vowel measures declined with worsening bulbar clinical scores (VSA: $R=0.33$, $p=0.033$; F2 slope: $R=0.25$, $p=0.048$), and smaller VSA was associated with greater listener effort ($R=-0.43$, $p=0.041$). Shallower F2 slopes were related to cortical thinning in oralPMC ($R=0.50$, $p=0.03$). Neither vowel measure was associated with respiratory nor cognitive test scores.

Conclusions: Vowel measures extracted with automatic processing from natural speech are sensitive to bulbar motor disease in ALS-FTD and are robust to cognitive impairment.

1. INTRODUCTION

Amyotrophic lateral sclerosis – frontotemporal dementia spectrum disorders (ALS-FTSD) are a group of neurodegenerative diseases with mixed motor neuron and frontotemporal lobe degeneration that result in a range of symptoms, including progressive muscle weakness and atrophy, difficulty with speech and swallowing, and changes in behavior, cognition, and/or personality.¹ The variable and overlapping symptoms can make diagnosis and management challenging, and highlight the need for specific and reliable assessment tools to improve patient outcomes. Bulbar motor disease, impacting speech and swallowing functions, is a hallmark symptom of ALS. While only 30% of ALS patients begin with bulbar symptoms at onset (bulbar-onset ALS), over 90% develop bulbar disease over disease course.² The earliest and most frequent sequela of bulbar disease is dysarthria, a motor speech disorder from a loss of articulator muscle strength and control that could arise from upper and/or lower motor neuron degeneration. This typically results in poorly intelligible speech requiring increased listener effort to comprehend,³ which can lead to communication breakdown, social isolation, and deterioration in quality of life.⁴ Bulbar motor disease can also result in swallowing difficulties and greater aspiration risk, which are common causes of severe morbidity and mortality in this population.^{5,6} Timely diagnosis and monitoring of bulbar motor disease in ALS-FTSD is critical.

Across ALS clinics, patient-reported functional rating scales such as the bulbar subscale of the ALS Functional Rating Scale-revised (ALSFRS-r)⁷, are the most routinely used methods to assess bulbar motor disease. These questionnaires are subjective and capture bulbar impairment at a gross functional level; thus they are limited as markers for screening and tracking of disease progression. Speech is a real-world activity and complex motor skill; many studies suggest that it may be among the most informative modalities for assessing motor abnormalities.⁸ Emerging automatic digital speech analyses have shown promise in providing quantitative clinical measures of bulbar impairment in ALS using highly structured speech tasks.^{9,10} Yet, these structured tasks show poor face validity with natural real-life speech and thus can be less sensitive to mild bulbar motor impairment in early disease and subtle changes over time.¹¹⁻¹³ Furthermore, structured tasks can be difficult to collect in patients with comorbid cognitive impairments, which are up to 50% of patients with ALS,¹ due to issues with task comprehension or impaired reading skills. We expect that automatic analysis of speech can be

adapted to natural speech tasks, such as picture descriptions, and provide sensitive and informative measures for bulbar disease.

Though all speech production subsystems may be affected,¹⁴ the articulatory subsystem which is controlled mostly by the tongue, is especially vulnerable to motor neuron disease.^{15,16} Available automatic speech platforms in ALS most commonly measure speaking rate in words per minute (WPM) but this measure can be confounded by respiratory or cognitive impairments.^{17,18} Three dimensional (3D) movement tracking studies link impaired articulation in ALS to slower and smaller tongue movements.¹⁵ Here we developed a novel digital speech analysis tool that assesses the articulatory subsystem by analyzing vowel acoustics that are directly related to tongue movements (speed and size) as a marker of bulbar motor disease in ALS. Speech features at the phoneme-level (i.e., at the level of the speech sound, such as vowels) may better reflect articulatory motor control processes than utterance-level features like speaking rate.¹⁹ This study assessed vowel measures of natural speech as potential markers of bulbar disease in ALS-FTSD. We hypothesized that vowel articulation measures of natural speech will be selectively impaired in ALS-FTSD patients with bulbar motor disease, and would be related to clinical bulbar motor signs and symptoms, speech intelligibility, and degeneration in oral regions of the primary motor cortex.

2. METHODS

2.1. Participants

Participants were selected from the University of Pennsylvania Integrated Neurodegenerative Disease Database (INDD²⁰) as of December 31st, 2022, based on a clinical diagnosis of ALS with or without cognitive/behavioural impairments, i.e., ALS-FTSD¹, or bvFTD,²¹ English as a primary language, and availability of at least one digitized picture description speech sample. The first available sample was used for analysis. bvFTD patients with evidence of a motor syndrome across lifespan clinical data were excluded. We also included speech data from historical healthy controls (HC) without a known cognitive or motor disorder.

For ALS, inclusion criteria included availability of clinical motor data within 3 months of speech sampling (average time interval between motor exam and speech sampling = 0.66 + 1.18 months), which was used to stratify patients into those with and without bulbar motor disease, herein referred to as ALS+bulbar and ALS-nonbulbar. Specifically, a score < 12 on the

ALSFERS-r bulbar score²²; and/or a score > 0 on the Penn Upper Motor Neuron scale (PUMNs) bulbar score²³ indicated presence of bulbar disease. We also collected forced vital capacity in seated position (%FVC) in ALS patients, which is closely linked to respiratory capacity. Other clinical tests included the Edinburgh Cognitive Assessment Scale – North American version (ECAS-NA)²⁴, available in $n=61$ ALS patients (85%) within 9 months of speech sampling (average time between ECAS and speech sampling = $0.85 + 2.22$ months), and the Mini Mental State Exam (MMSE), available in $n=42$ ALS (58%) and all bvFTD and HC within 9 months of recording (average time between MMSE and speech sampling = $2.80 + 5.67$ months).

The final cohort included 72 ALS patients (of which $n=16$ with cognitive impairments based on published cut-offs on ECAS-NA or MMSE); 25 bvFTD patients; and 32 HC. Seven ALS patients had bulbar onset with initial symptoms in speech or swallowing, while the remaining patients had limb onset (35 arm-onset, 27 leg-onset) or social/ behavioural onset ($n=3$). Irrespective of symptom onset site, $n=49$ ALS patients were classified as ALS+bulbar and $n=23$ as ALS-nonbulbar. We verified that ALS-nonbulbar patients did not present with lower motor neuron bulbar signs on their neurological exam. Demographic and clinical characteristics of the final groups are summarized in Table 1.

2.3. Digital speech data collection and processing

All participants provided a natural speech sample by describing the Cookie Theft picture scene from the Boston Diagnostic Aphasia Examination.²⁵ A subset of the ALS patients ($n = 24$, 34%) provided a second picture description on the same day, elicited by the picnic scene from the Western Aphasia Battery.²⁶ Participants were asked to describe the picture(s) in as much detail as they could. The speech samples were recorded on a single channel without compression at a sampling rate of 16 KHz and bit depth of 16 in a quiet room with minimal background noise. Groups did not differ in average sample duration (79 ± 20 seconds).

Recordings were transcribed by trained transcribers and processed with automated algorithms to derive acoustic measures of tongue movement size and speed (described below).

We also calculated speaking rate (WPM) as total words divided by total time (including pauses), as it is a commonly used marker of bulbar motor disease in ALS.

2.4 Automatic vowel acoustic analysis

Acoustics pertain to the physical properties of sounds. Specifically, we analyze formant frequencies which are the natural resonant frequencies of the vocal tract based on the position of

the tongue²⁷, Vowel acoustics, specifically formant frequencies, are strongly related to tongue movement during articulation:²⁷ the first formant, F1, is associated with the height of the tongue in the mouth during speech; the second formant, F2, is associated with the anterior and posterior dimension of tongue position. Vowels were automatically processed and analyzed using the Forced Alignment & Vowel Extraction (FAVE) system.²⁸ First, FAVE aligned the acoustic signal to the text transcript, creating a Praat TextGrid file²⁹ labeled at the phoneme (i.e., consonant and vowel) and word levels. FAVE's automatic phoneme-level segmentation shows high congruency with gold-standard manually defined boundaries.³⁰ Next, FAVE tagged all vowels with the CMU word pronunciation dictionary³¹ and extracted duration and formant values for each vowel token. F1 and F2 were extracted at the “steady-state” of each vowel token (approximately mid-point of the vowel when it is most stable), and at periodic increments through the vowel trajectory (at 20%, 35%, 50%, 65%, and 80% of vowel duration). We only analyzed stressed vowels, which were assigned by the pronunciation dictionary, and analyzed all available words except fillers, i.e., “um” and “uh”. bvFTD group produced fewer vowels on average than ALS groups (Table 1) which we adjusted for in the models.

Formant values in the logarithmic Hz scale were transformed to the linear Bark scale using $[(26.81 * F) / (1960 + F)] - 0.53$, where F= frequency in Hz.³² The Bark scale is perceptually uniform and is defined so that the critical bands of human hearing each have a width of one Bark. Bark scaling is beneficial in studies of dysarthria as it more closely corresponds to the perceptual organization of speech sounds in the auditory system, facilitating the interpretation of results in terms of speech perception and intelligibility. Bark transformation has shown to reduce the absolute differences between healthy females' and males' VSAs and preserve strong relationships with intelligibility in studies of dysarthria.³³

2.4.1. Articulatory-acoustic vowel measures of size and speed

Articulatory-acoustic measures that previously demonstrated sensitivity to bulbar dysfunction in ALS from structured speech tasks were adapted for natural speech^{34,35}. When both picture descriptions were available for a participant, vowels were pooled across the two pictures as the average vowel durations and formant values did not differ between the two picture stimuli ($p > 0.96$, repeated-samples t-tests). We calculated two articulatory-acoustic measures per participant:

- (1) Vowel space area (VSA): the 2-dimensional area of the triangle in F1xF2 space, in Bark², formed by the three “corner” vowels, which are vowels produced at the extremes of tongue displacements. VSA serves as an acoustic proxy for a speaker’s articulatory working space, or tongue movement area.³⁵ The area is calculated between the following vowels at steady-state: /i/ as in “cookie”, /a/ as in “hot”, /u/ as in “stool,” and /ʊ/ as in “cookie”, using $F1_{/i/} * (F2_{/a/} - F2_{/u+ʊ/}) + F1_{/a/} * (F2_{/u+ʊ/} - F2_{/i/}) + F1_{/u+ʊ/} * (F2_{/i/} - F2_{/a/}) / 2$. As denoted, we combined /u/ and /ʊ/ to form a single vowel class representing the highest, back-most vowel, as these two vowels largely overlapped in acoustic space, and were both produced half as many times as the other two vowels, /i/ and /a/. A median of 6 vowel tokens were produced per vowel class per recording. We calculated the triangular VSA created between the centroids of these three vowel classes, i.e., mean F1 and F2 across vowel tokens within a vowel class (see graphical display of VSA in Supplementary Materials A).
- (2) Mean F2 slope: the rate of change of F2, in Bark/ms, averaged across all vowel trajectories, i.e., every stressed consonant-vowel-consonant transition. Slopes were calculated for each vowel token using $F2_{\max} - F2_{\min} / \text{vowel duration}$, across 20-80% of the vowel trajectory, then averaged per person. F2 slope is an acoustic measure of the rate of change by the tongue in the front-back axis, the primary articulatory gesture associated with consonant-vowel-consonant transitions.³⁶ A shallower slope indicates longer durations and less movement excursion.

2.5. Listening task and perceived listener effort

A random subset of Cookie Theft recordings from patients with ALS (n = 40, 56% of the total group) were rated perceptually for “listener effort”, a construct of speech intelligibility that is often described as the perceived cognitive resource necessary for speech recognition.³⁷ The subset with listener effort data did not differ from the whole ALS group in demographics or ALSFRS-r bulbar score. Listener effort is more sensitive to early speech changes than intelligibility, which remains intact at the expense of increased effort of listeners.³⁸

Listener effort was rated via REDCap, an online survey platform. Two independent listeners (S.S. and S.A.) who were blinded to the bulbar group designation were asked about their perceived effort of listening to 30-second clips of the Cookie Theft recordings (mean duration = 29.07 ± 0.16 seconds). S.S. selected the audio clips manually by arbitrarily selecting

the largest section with the most voicing as seen on the waveform, which is blinded to articulatory information. Listeners were asked to answer the following question: “How effortful was it for you to understand the speaker? Remember, we are asking how hard you worked, not how well you did.” Answers were recorded on a visual analog scale, which consisted of a slider on an unmarked scale that the rater could position freely, with end points labeled “very” (converted to a score of 100) and “not at all” (score of 0).³⁹

Inter-rater agreement for listener effort was good (ICC = 0.72, $p < 0.001$) between the two raters, established by the intraclass correlation coefficient (ICC) on a mean-rating, consistency, 2-way random-effects model. A subset of recordings which was rated twice by each rater (24 samples, 20%). The intra-rater reliability (i.e., consistency within each rater) showed moderate to good consistency (ICC = 0.84 (S.S.) and 0.72 (S.A.), both $p < 0.001$) on a single-rating agreement 2-way mixed-effects model. The average listener effort rating between the two raters was used for analysis.

2.6. Neuroimaging methods

High resolution T1 volumetric brain MRI scans were available for 18 ALS patients (9 ALS+bulbar, 11 ALS-nonbulbar) collected on a Siemens 3T Trio scanner within 1 year of speech sampling (10 ± 6.56 months). Due to the small sample size, we tested associations across both ALS groups. We statistically adjusted for the ALS group and time between MRI and speech in our models. The MRI subset did not differ from the whole ALS group in demographics or disease severity (ALSFRS-r total scores). We also confirmed that the bulbar group stratification at speech sampling matched the condition at MRI, i.e., none of the ALS-nonbulbar patients had developed bulbar signs or symptoms at the time of MRI scan, and all ALS+bulbar patients had evidence of bulbar disease at the time of MRI. The clinical motor scores remained the same for all but two patients, for whom there was a 2-point progression on the ALSFRS-r bulbar scale from speech sampling to their next closest MRI dates.

Briefly, T1-weighted images were preprocessed using `antsCorticalThickness.sh` (<https://github.com/ftdc-picsl/antsct-aging>⁴⁰). The Schaefer 400 parcellation⁴¹ was warped to the individual T1, which divides each hemisphere into 200 roughly equal-sized areas. Mean cortical thickness values were extracted from each label. Average cortical thickness for the bilateral orobuccal homunculus (oral primary motor cortex (PMC)) and hand PMC regions were calculated by combining values across anatomically pre-defined Schaeffer labels relating to

tongue motion and finger tapping motion in healthy controls on task-based fMRI⁴². The oral PMC region was composed of label numbers 43, 45, and 47 (left hemisphere), and labels 245, 246, and 248 (right hemisphere). The hand PMC region was the average of labels 53, 57, and 60 (left hemisphere) and labels 255, 256, and 259 (right hemisphere).⁴²

2.7. Statistical Considerations

Vowel measures were compared between ALS+bulbar, ALS-nonbulbar, bvFTD, and HC groups using analyses of covariance (ANCOVAs) adjusting for group differences in age and total vowels (Table 1). We conducted pairwise comparison of means using the Tukey method for the Group factor, which adjusts for pairwise comparisons and takes into account the effects of the covariables. We tested group differences between ALS+bulbar and ALS-nonbulbar after adjusting for %FVC and ECAS ALS-Specific scores.

We examined associations with bulbar motor severity in ALS+bulbar patients using a sum of ALSFRS-r bulbar and PUMNs bulbar scores (summed bulbar score). Lower values on ALSFRS-r but higher values on PUMNs indicate more bulbar disease; therefore, we reversed the PUMNs so a maximum summed bulbar score of 16 indicated absence of all bulbar signs and symptoms. Four ALS+bulbar patients without PUMNs scores were omitted from this analysis. Linear regression models adjusted for ECAS ALS-Specific score and %FVC. As these not significant covariables and were only available in a subset of patients, we report Pearson's correlations in the manuscript and full model results in Supplementary Materials B.

We also examined relations with perceived listener effort, serving as a parameter of communication limitation largely related to the speaker's intelligibility.³⁷ We compared 17 ALS speakers requiring zero listener effort vs. 23 speakers rated >0 using t-tests, and tested associations with listener effort in the 23 ALS speakers rated >0 using linear regressions, adjusting for ALS-Specific score and %FVC. We report Pearson's correlations here, and full model results in Supplementary Materials B.

To test our hypothesized independence from cognitive and respiratory deficits, we examined Pearson's correlations with ECAS ALS-Specific scores and %FVC in the ALS patients. We conducted identical analyses as above with speaking rate as a routinely-used comparative measure of bulbar motor disease. Supplementary Materials B reports all tested clinical correlations.

We examined associations with cortical atrophy in oral PMC and hand PMC in ALS using linear regressions, adjusting for ALS subgroup and time between MRI and speech recording. All statistical tests were two-tailed with a significance threshold of $\alpha = 0.05$. Statistical analyses were performed with R version 4.1.0 and RStudio version 1.4.1717.

2.8. Standard Protocol Approvals, Registrations, and Patient Consents

The Institutional Review Board of the Hospital of the University of Pennsylvania approved the study of human subjects, and all participants agreed to participate in the study by informed written consent.

3. RESULTS

3.1. VSA and F2 slope are exclusively reduced in ALS+bulbar

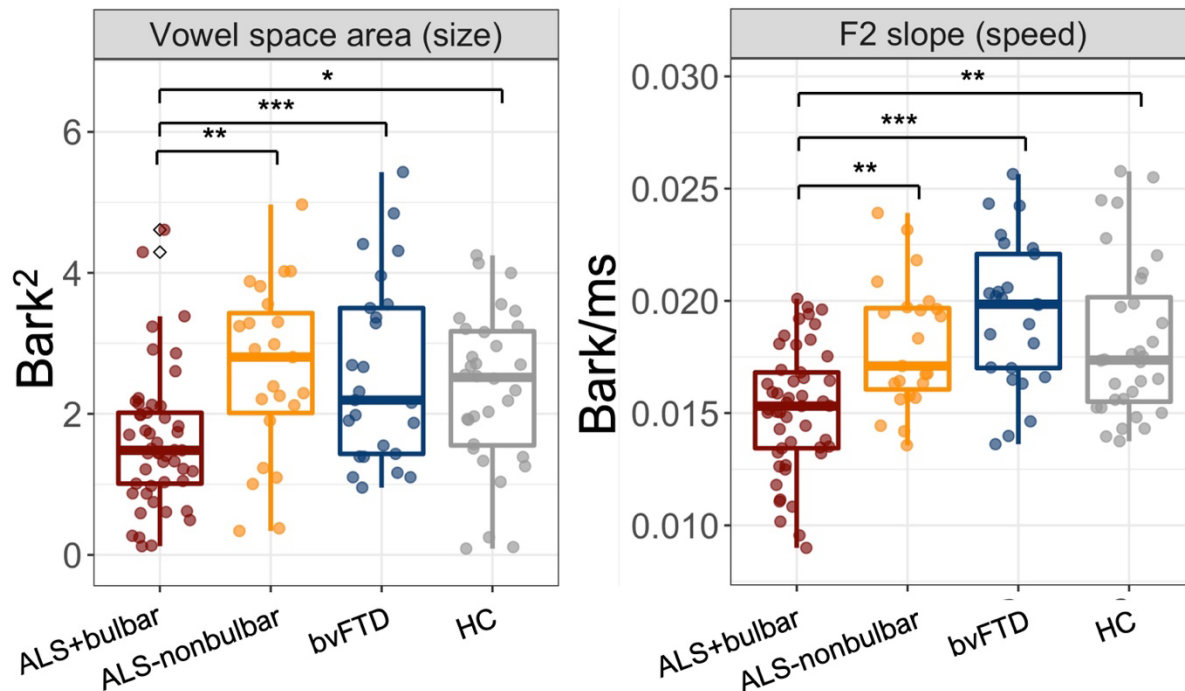


Figure 1. Vowel space area (measure of tongue movement size) and mean F2 slope (measure of speed) are significantly reduced in ALS patients with bulbar motor disease (ALS+bulbar) compared to ALS patients without bulbar disease (ALS-nonbulbar), behavioural variant frontotemporal dementia (bvFTD), and healthy controls (HC). All other contrasts were non-significant, $p > 0.05$. $*$ = $p < 0.05$, $**$ = $p < 0.01$, $***$ = $p < 0.001$; Tukey's adjusted HSD.

Figure 1 shows the vowel measures for the four groups. Patients with ALS+bulbar had significantly smaller VSA than ALS-nonbulbar (Tukey's adjusted $p = 0.0089$, $|d| = 0.94$), bvFTD

($p < 0.001$, $|d| = 0.89$), and HC ($p = 0.014$, $|d| = 0.73$). ALS+bulbar had significantly shallower F2 slopes than ALS-nonbulbar ($p = 0.0026$, $|d| = 1.04$), bvFTD ($p < 0.001$, $|d| = 1.47$), and HC ($p = 0.0015$, $|d| = 0.99$). The two vowel measures were not reduced in ALS-nonbulbar nor in bvFTD compared to HC, $p > 0.53$. Age and total vowels were not significant covariables for either measure ALS+bulbar vs. ALS-nonbulbar group differences remained significant after adjusting for %FVC and ALS-Specific scores, and neither were significant covariables (see Supplementary Materials B for post-hoc linear regression model results).

By contrast, speaking rate was reduced in both ALS+bulbar and bvFTD groups, compared to ALS-nonbulbar and HC (Table 1).

3.3. VSA and F2 slope correlate with bulbar scores but not cognitive or respiratory scores

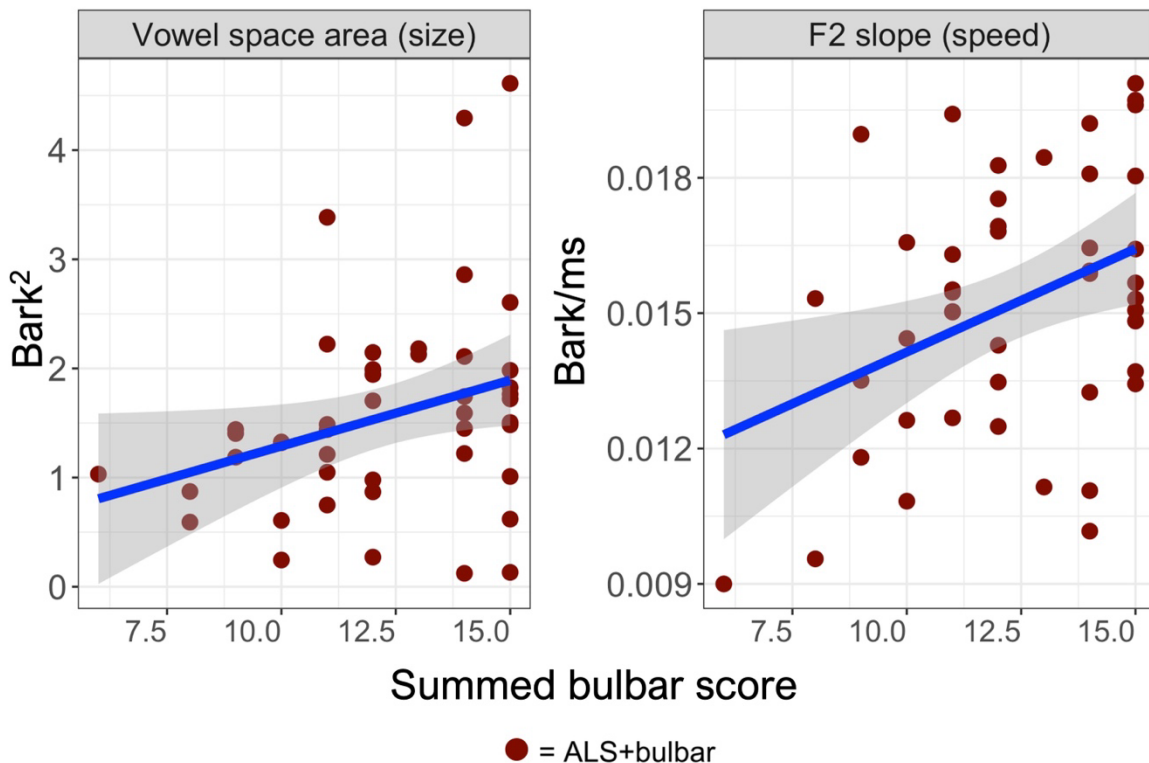


Figure 2. Worse bulbar disease (lower summed bulbar scores composed of ALSFRS-r bulbar and PUMNs bulbar scores) is significantly associated with smaller vowel space area and shallower F2 slopes in ALS+bulbar patients.

Figure 2 shows the significant relationship between vowel measures and summed bulbar scores in ALS+bulbar. A lower summed bulbar score, which indicates greater bulbar motor disease, was significantly correlated with a smaller VSA, $R = 0.33$, $p = 0.043$, and shallower F2

slopes, $R=0.38$, $p=0.011$. ECAS ALS-Specific scores were not significant covariables in either model (Supplementary Materials B).

Neither vowel measure correlated with %FVC ($p>0.20$) nor ALS-Specific scores ($p>0.80$). Speaking rate did not correlate with summed bulbar scores ($p=0.10$) but rather was significantly associated with %FVC, $R=0.31$, $p=0.009$, and ECAS ALS-Specific scores, $R=0.42$, $p=0.0007$ (Supplementary Materials C).

3.4. VSA correlates with listener effort

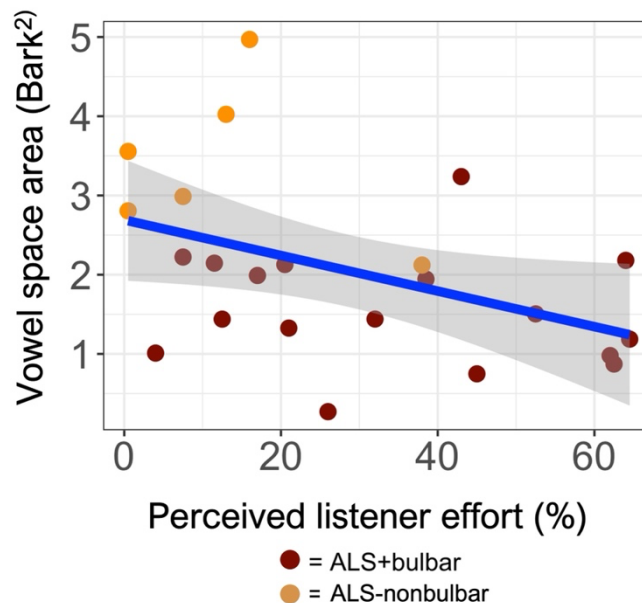


Figure 3. Greater perceived listener effort is significantly associated with smaller vowel space area in ALS.

VSA was significantly reduced in ALS patients with perceived listener effort >0 compared to patients requiring 0 listener effort ($p=0.041$, $|d|=0.45$, plot not shown). Greater perceived listener effort was associated with smaller VSA, $R= -0.43$, $p=0.041$ (Figure 3). ECAS ALS-Specific score was not a significant covariable in the model (Supplementary Materials B).

F2 slope did not differ between listener effort groups nor relate to listener effort, $p>0.17$. Speaking rate was not associated with listener effort, $p=0.30$ (Supplementary Materials C).

3.5. F2 slope correlates with cortical thickness of the orobuccal homunculus of the primary motor cortex but not hand motor area

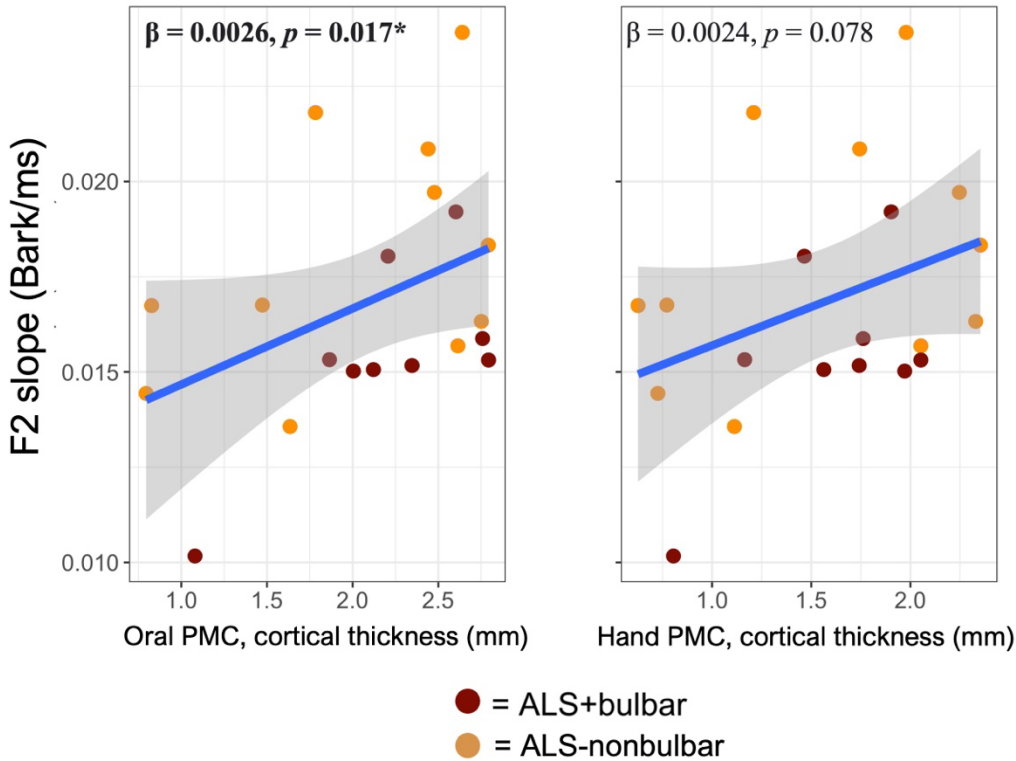


Figure 4. Shallower mean F2 slope was significantly related to greater cortical thinning in orobuccal homunculus, i.e., oral primary motor cortex (oralPMC) (left panel). Oral PMC consisted of average thickness across 3 left and 3 right hemispheric regions previously found to correspond with tongue motion in healthy controls. F2 slope was not significantly related to atrophy in hand PMC (right panel). Beta coefficients and p values for each ROI are noted.

Reduced F2 slope was significantly related to cortical thinning in oral PMC in ALS after adjusting for time interval between MRI and speech sampling and ALS group, but not with hand PMC (Figure 4). Full model results are in Supplementary Materials B. VSA was not associated with cortical thickness in either motor region, $p > 0.62$.

4. DISCUSSION

Given the devastating impact of bulbar motor disease on vital speech and swallowing functions, developing objective clinical markers that can reliably support diagnosis and tracking of bulbar motor disease is crucial in patients with ALS-FTSD, who present with mixed motor, respiratory, and cognitive impairments. In this study, we provide clinical validation data for vowel measures of natural speech that reflect tongue movement size (vowel space area) and

speed (F2 slope), which are sensitive to bulbar motor disease in ALS based on multiple findings. Firstly, both measures were impaired only in ALS patients with evidence of bulbar motor signs or symptoms, compared to ALS patients without bulbar motor impairment, bvFTD, and healthy controls. Secondly, vowel measures declined with worse bulbar motor disease. Lastly, vowel space area was related to perceived listener effort and F2 slope was related to atrophy in orobuccal homunculus areas. Importantly, neither vowel measure was related to cognitive or respiratory impairment.

Our study found that vowel-based measures are sensitive indicators of neurodegenerative bulbar changes in ALS. Our findings are consistent with previous studies of structured speech tasks reporting smaller VSA and shallower F2 slopes in ALS patients with bulbar disease, which was related to reduced speech intelligibility and slower and smaller tongue movements.⁴³⁻⁴⁵ But traditional methods of measuring vowels are time-consuming and require expert training limiting their use as clinical markers, as vowel detection and formant extraction require hand-labelling the speech signal. Fortunately, recent advances in automatic algorithms such as FAVE employed in this study, have made it possible to analyze large samples of natural speech more efficiently. Natural speech may be a better measure of articulatory impairment than structured speech tasks, which are commonly used in existing digital speech assessment tools but may be less sensitive to subtle neurodegenerative changes. A study comparing natural speech (monologues) to structured speech tasks (sentence repetitions and passage readings) in Parkinson's disease found that the former captured more articulatory deficits, likely due to its high face validity with real-life communication.⁴⁶ Our findings and previous studies highlight the sensitivity of vowel-based measures to ALS motor disease processes⁴⁷, and future comparisons of natural vs. structured speech in ALS patients could further inform the task specificity of motor speech control and validate our approach. Our results represent a significant step toward the development of a fully automated tool for the large-scale evaluation of articulatory impairment during natural speech due to neurodegeneration.

Our results highlight the clinical specificity of vowel measures in detecting motor impairment in ALS, in comparison to speaking rate which is commonly used to assess bulbar motor disease. We found that vowel measures were selectively impaired in ALS patients with bulbar motor disease, whereas speaking rate (WPM) was also affected in non-motor patients with cognitive impairment (bvFTD). Correlational analyses indicated that vowel measures declined

due to bulbar motor disease and resulted in greater listener effort, but were not related to respiratory or cognitive impairment. In contrast, speaking rate declined due to multiple motor and cognitive sources of impairment in ALS, and was not related to listener effort. Thus, speaking rate may not be a suitable motor-specific marker in patients with concomitant cognitive impairments, while vowel measures may offer greater diagnostic and tracking precision as a clinical tool. Future investigations should test the responsiveness of vowel measures in early or presymptomatic stages of bulbar motor disease.

We found that smaller VSA was related to greater perceived listener effort, suggesting that a smaller tongue working space results in vowels being produced more centrally that are harder to distinguish from each other. On the other hand, tongue movement speed, as measured by mean F2 slope, did not show a similar relationship to listener effort. This can be explained by previous studies in ALS that found a reduction in articulation rate prior to a decline in speech intelligibility, suggesting that slow speech may be a compensatory strategy used by patients to maintain intelligibility.⁴⁸ VSA may also be particularly sensitive to swallowing impairment, specifically in the oral stage where tongue movements play a crucial role in bolus formation and manipulation. Studies in stroke patients have shown a strong link between VSA and swallowing efficiency.⁴⁹ Given the importance of tongue working space to both speech and swallowing, further research should examine the relationship between VSA and swallowing parameters in ALS.

Our imaging analyses revealed a moderately strong relationship between tongue articulation speed (F2 slope) and cortical thinning in the average bilateral oral primary motor areas that innervate tongue muscles. However, we did not find a similar relationship with tongue movement size (VSA). Slow articulation is a cardinal feature of spastic dysarthria caused by upper motor neuron loss,⁵⁰ while reduced tongue movement size caused by muscle weakness is a feature of flaccid dysarthria related to lower motor neuron loss in the brainstem. This may partially explain the lack of association with VSA in our exploratory MRI analysis, which was limited to the motor cortex. Since the availability of MRI data was limited, we analyzed both ALS+bulbar and ALS-nonbulbar patients together, despite their differing clinical presentations. Surprisingly, these two patient groups largely overlapped in the range of cortical thinning in oralPMC. However, some ALS-nonbulbar cases may have been misclassified due to missed subtle bulbar UMN signs, such as a brisk jaw jerk, which could explain why they have a

relatively thin oralPMC and shallow F2 slopes. Indeed, chart review revealed that the four ALS-nonbulbar cases with the thinnest oralPMC developed bulbar UMN signs within 6 months from MRI, while the remaining ALS-nonbulbar cases did not develop bulbar disease for 9-36 months after MRI. Conversely, the ALS+bulbar cases may have more brainstem lower motor neuron involvement and a spared motor cortex, as lower motor neuron involvement in bulbar muscles has an overt presentation of atrophy, fasciculations, and “slurred” speech, which is likelier to be classified as ALS+bulbar. Future studies should aim to examine associations between articulatory-acoustic metrics and brainstem volumetrics to better understand the underlying neural mechanisms of speech production in ALS patients.

Analyzing natural speech tasks can be challenging due to the uncontrolled nature of the speech produced, which can result in diverse and unpredictable lexical content and varying amounts of speech data compared to structured speech. By using picture description as the task, we mitigate some of this variability and elicit topic-controlled speech with specific words that are produced by all speakers (e.g., “cookie”, “stool”, and “water”). Considering coarticulation effects of surrounding consonants on vowel formants, we maximized vowel data by pooling data from multiple picture descriptions when available. While these strategies were effective, additional speech tasks that elicit longer samples of natural speech, such as descriptions of a story elicited by a wordless picture book or conversations on controlled topics, could be valuable. These naturalistic tasks have the advantage of being easy to administer and executed by patients with varying degrees of cognitive impairment. The present study is also limited by small sample sizes, primarily due to the low availability of MRI and speech data in ALS patients with severe bulbar impairment. These patients often experience respiratory dysfunction and have difficulty tolerating the supine position necessary for MRI. To improve upon these limitations, future studies and collaborative efforts should be encouraged to collect speech data from the widest possible range of patients. Although this study demonstrated strong clinical validity of vowel articulation measures to bulbar motor disease in ALS, the repeatability (test-retest reproducibility) of the novel metrics should be established in future.

With these caveats in mind, we conclude that automatic vowel analysis of natural speech can provide highly objective and motor-specific measures of bulbar disease in patients with ALS spectrum disorders. The proposed measures are robust to cognitive and respiratory dysfunction, are clinically interpretable, and can be derived from a brief natural speech sample.

REFERENCES

1. Strong MJ, Abrahams S, Goldstein LH, et al. Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria. *Amyotroph Lateral Scler Frontotemporal Degener.* 2017;18(3-4):153-174. doi:10.1080/21678421.2016.1267768
2. Carpenter RJ, McDonald TJ, Howard FM. The otolaryngologic presentation of amyotrophic lateral sclerosis. *Otolaryngology.* 1978;86(3 Pt 1):ORL-479-ORL-484. doi:10.1177/019459987808600319
3. Chen A, Garrett CG. Otolaryngologic presentations of amyotrophic lateralsclerosis. *Otolaryngol Head Neck Surg (1979).* 2005;132(3):500-504. doi:S0194599804018753 [pii] 10.1016/j.otohns.2004.09.092
4. Hecht M, Hillemacher T, Gräsel E, et al. Subjective experience and coping in ALS. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders.* Published online 2002. doi:10.1080/146608202760839009
5. Perry BJ, Nelson J, Wong JB, Kent DM. Predicting dysphagia onset in patients with ALS: the ALS dysphagia risk score. *Amyotroph Lateral Scler Frontotemporal Degener.* Published online 2021. doi:10.1080/21678421.2021.1961805
6. Paillisse C, Lacomblez L, Dib M, Bensimon G, Garcia-Acosta S, Meininger V. Prognostic factors for survival in amyotrophic lateral sclerosis patients treated with riluzole. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders.* Published online 2005. doi:10.1080/14660820510027035
7. Plowman EK, Tabor LC, Wymer J, Pattee G. The evaluation of bulbar dysfunction in amyotrophic lateral sclerosis: survey of clinical practice patterns in the United States. *Amyotroph Lateral Scler Frontotemporal Degener.* Published online 2017. doi:10.1080/21678421.2017.1313868
8. Robin J, Harrison JE, Kaufman LD, Rudzicz F, Simpson W, Yancheva M. Evaluation of speech-based digital biomarkers: review and recommendations. *Digit Biomark.* 2020;4(3):99-108.
9. Stegmann GM, Hahn S, Liss J, et al. Early detection and tracking of bulbar changes in ALS via frequent and remote speech analysis. *NPJ Digit Med.* Published online 2020. doi:10.1038/s41746-020-00335-x
10. Stegmann GM, Hahn S, Liss J, et al. Repeatability of Commonly Used Speech and Language Features for Clinical Applications. *Digit Biomark.* Published online 2020. doi:10.1159/000511671
11. Bunton K. Speech versus nonspeech: Different tasks, different neural organization. *Semin Speech Lang.* Published online 2008. doi:10.1055/s-0028-1103390
12. Liberman AM. Why is speech so much easier than reading? In: *Reading and Spelling: Development and Disorders.* ; 1998.
13. Tasko SM, McClean MD. Variations in articulatory movement with changes in speech task. Published online 2004.
14. Green JR, Yunusova Y, Kuruvilla MS, et al. Bulbar and speech motor assessment in ALS: Challenges and future directions. *Amyotroph Lateral Scler Frontotemporal Degener.* Published online 2013. doi:10.3109/21678421.2013.817585

15. Shellikeri S, Green JR, Kulkarni M, et al. Speech movement measures as markers of bulbar disease in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*. Published online 2016. doi:10.1044/2016_JSLHR-S-15-0238
16. Waito AA, Wehbe F, Marzouqah R, et al. Validation of Articulatory Rate and Imprecision Judgments in Speech of Individuals With Amyotrophic Lateral Sclerosis. *Am J Speech Lang Pathol*. Published online December 8, 2021:1-13. doi:10.1044/2020_AJSLP-20-00199
17. Yunusova Y, Graham NL, Shellikeri S, et al. Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Kassubek J, ed. *PLoS One*. 2016;11(1):e0147573. doi:10.1371/journal.pone.0147573
18. Mitchell HL, Hoit JD, Watson PJ. Cognitive-linguistic demands and speech breathing. *Journal of Speech, Language, and Hearing Research*. 1996;39(1):93-104.
19. Tjaden K. Segmental articulation in motor speech disorders. *Motor speech disorders: essays for Ray Kent*. Published online 2007:151-186.
20. Toledo JB, van Deerlin VM, Lee EB, et al. A platform for discovery: The University of Pennsylvania Integrated Neurodegenerative Disease Biobank. *Alzheimer's and Dementia*. Published online 2014. doi:10.1016/j.jalz.2013.06.003
21. Rascovsky K, Hodges JR, Kipps CM, et al. Diagnostic criteria for the behavioral variant of frontotemporal dementia (bvFTD): current limitations and future directions. *Alzheimer Dis Assoc Disord*. 2007;21(4):S14-S18.
22. Cedarbaum JM, Stambler N, Fuller C, et al. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci*. 1999;169(1-2):13-21.
23. Quinn C, Edmundson C, Dahodwala N, Elman L. Reliable and efficient scale to assess upper motor neuron disease burden in amyotrophic lateral sclerosis. *Muscle Nerve*. 2020;61(4):508-511.
24. Niven E, Newton J, Foley J, et al. Validation of the Edinburgh Cognitive and Behavioural Amyotrophic Lateral Sclerosis Screen (ECAS): A cognitive tool for motor disorders. *Amyotroph Lateral Scler Frontotemporal Degener*. Published online 2015. doi:10.3109/21678421.2015.1030430
25. Helm-Estabrooks N. Boston Diagnostic Aphasia Examination. In: *Encyclopedia of Clinical Neuropsychology*. ; 2018. doi:10.1007/978-3-319-57111-9_868
26. Kertesz A. *Aphasia and Associated Disorders: Taxonomy, Localization, and Recovery*. Holt Rinehart & Winston; 1979.
27. Yunusova Y, Weismer G, Westbury JR, Lindstrom MJ. Articulatory Movements During Vowels in Speakers With Dysarthria and Healthy Controls. *Journal of Speech, Language, and Hearing Research*. 2008;51(3):596-611. doi:10.1044/1092-4388(2008/043)
28. Rosenfelder I, Fruehwald J, Evanini K, et al. FAVE (forced alignment and vowel extraction) suite version 1.1. 3. Published online 2014.
29. Boersma P, van Heuven V. Speak and unSpeak with PRAAT. *Glott International*. 2001;5(9/10):341-347.
30. Evanini K, Isard S, Liberman M. Automatic formant extraction for sociolinguistic analysis of large corpora. Published online 2009.
31. Kominek J, Black AW. The CMU Arctic speech databases. In: *Fifth ISCA Workshop on Speech Synthesis*. ; 2004.

32. Traunmüller H. Analytical expressions for the tonotopic sensory scale. *J Acoust Soc Am*. 1990;88(1):97-100.
33. Kim H, Hasegawa-Johnson M, Perlman A. Vowel contrast and speech intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica*. 2011;63(4):187-194.
34. Yunusova Y, Green JR, Greenwood L, Wang J, Pattee GL, Zinman L. Tongue Movements and Their Acoustic Consequences in Amyotrophic Lateral Sclerosis. *Folia Phoniatrica et Logopaedica*. 2012;64(2):94-102. doi:10.1159/000336890
35. Lee J, Littlejohn MA, Simmons Z. Acoustic and tongue kinematic vowel space in speakers with and without dysarthria. *Int J Speech Lang Pathol*. 2017;19(2):195-204.
36. Yunusova Y, Weismer G, Westbury JR, Lindstrom MJ. Articulatory movements during vowels in speakers with dysarthria and healthy controls. Published online 2008.
37. Whitehill TL, Wong CCY. Contributing factors to listener effort for dysarthric speech. *J Med Speech Lang Pathol*. 2006;14(4):335-342.
38. Fletcher AR, Wisler AA, Gruver ER, Borrie SA. Beyond Speech Intelligibility: Quantifying Behavioral and Perceived Listening Effort in Response to Dysarthric Speech. *Journal of Speech, Language, and Hearing Research*. Published online 2022:1-11.
39. Picou EM, Moore TM, Ricketts TA. The effects of directional processing on objective and subjective listening effort. *Journal of Speech, Language, and Hearing Research*. 2017;60(1):199-211.
40. Tustison NJ, Cook PA, Klein A, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *Neuroimage*. Published online 2014. doi:10.1016/j.neuroimage.2014.05.044
41. Schaefer A, Kong R, Gordon EM, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral cortex*. 2018;28(9):3095-3114.
42. Schaefer A, Kong R, Gordon EM, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral cortex*. 2018;28(9):3095-3114.
43. Weismer G, Kent RD, Hodge M, Martin R. The acoustic signature for intelligibility test words. *J Acoust Soc Am*. 1988;84(4):1281-1291.
44. Yunusova Y, Green JR, Lindstrom MJ, Ball LJ, Pattee GL, Zinman L. Kinematics of disease progression in bulbar ALS. *J Commun Disord*. Published online 2010. doi:10.1016/j.jcomdis.2009.07.003
45. Turner GS, Tjaden K, Weismer G. The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*. 1995;38(5):1001-1013.
46. Rusz J, Cmejla R, Tykalova T, et al. Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task. *J Acoust Soc Am*. 2013;134(3):2171-2181.
47. Rowe HP, Shellikeri S, Yunusova Y, Chenausky K v, Green JR. Quantifying articulatory impairments in neurodegenerative motor diseases: A scoping review and meta-analysis of interpretable acoustic features. *Int J Speech Lang Pathol*. Published online 2022:1-14.
48. Rong P, Yunusova Y, Green JR. Speech intelligibility decline in individuals with fast and slow rates of ALS progression. In: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. ; 2015.

49. Park EJ, Yoo SD, Kim HS, et al. Correlations between swallowing function and acoustic vowel space in stroke patients with dysarthria. *NeuroRehabilitation*. 2019;45(4):463-469.
50. Darley FL, Aronson AE, Brown JR. Differential diagnostic patterns of dysarthria. *J Speech Hear Res*. Published online 1969. doi:10.1044/jshr.1202.246

Table 1. Demographics and clinical characteristics of participants. Normally distributed variables are summarized in mean (sd) and non-normal variables in median [IQR].

Variable	Group				p-value
	ALS+bulbar	ALS- nonbulbar	bvFTD	HC	
n	49	23	25	32	
Sex = n Male (%)	33 (67.3)	11 (47.8)	14 (56.0)	16 (50.0)	0.31
Age, years	61.2 (10.9) ^{HC}	61.0 (9.6) ^{HC}	61.1 (6.6) HC	69.3 (7.9) ALSb, ALSnonb, HC	0.001**
Education, years	14.0 [12.0, 16.0]	16.0 [12.5, 16.0]	16.0 [16.0, 16.0]	16.0 [13.8, 18.0]	0.12
Disease duration at test, months	26.0 [17.0, 41.0]	30.0 [15.5, 52.5]	44.0 [31.0, 77.0]	-	0.08
Bulbar disease duration at test, months	19.2 (28.3)	-	-	-	-
ALSFRS-r total (max=48)	32.3 (7.7)	37.9 (4.9)	-	-	-
ALSFRS-r bulbar score (max=12)	9.8 (1.7)	12.0 (0.0)	-	-	-

PUMNs bulbar score (max=4)	1.5 (1.2)	0.0 (0.0)	-	-	-
MMSE total score (max=30)	28.0 [26.2, 29.0]	29.0 [26.8, 30.0] ^{bvFTD}	25.0 [22.8, 27.0] ALSnonb, HC	29.0 [28.0, 30.0] ^{bvFTD}	<0.001* **
ECAS-NA ALS-Specific score (max=100)	84.0 [78.0, 89.0]	86.0 [79.0, 89.2]	-	-	0.46
Cognitively impaired, n (%) [^]	12 (24.5)	4 (17.4)	17 (70.8)	-	-
Forced Vital Capacity (%)	71.7 (24.4)	85.0 (20.9)	-	-	0.03*
Speaking rate (WPM)	117.9 [97.4, 141.4] ^{ALSnonb,} HC	151.7 [144.1, 169.0] ^{ALSb,} bvFTD	112.2 [79.4, 152.8] ALSnonb, HC	145.9 [125.5, 169.2] ^{ALSb,} bvFTD	<0.001* **
Total number of vowels	88.9 (46.9) bvFTD	108.2 (47.8) bvFTD	56.7 (25.3) ALSb, ALSnonb	81.4 (39.2)	<0.001* **
<p>Significant pairwise comparisons reflecting Tukey's adjusted <i>p</i>-values are shown as superscripts: ^{ALSnonb} = significantly different from ALS-nonbulbar; ^{ALSb} = significantly different from ALS-bulbar; ^{bvFTD} = significantly different from bvFTD; ^{HC} = significantly different from HC.</p>					

^ = cognitively impaired status was determined by published North American age- and education-specific thresholds on the ECAS-NA or MMSE cut-off of 26 based on availability of data. Disease duration was calculated as interval from disease onset to speech collection date. Bulbar disease duration was calculated as interval from initial report of bulbar signs and/or symptoms on ALSFRS-r or PUMNs to speech sampling date. Abbreviations: ALSFRS-r = ALS Functional Rating Scale- revised, lower values indicate worse symptoms; ALS+bulbar = ALS patients with evidence of bulbar motor disease based on ALSFRS-r and PUMNs bulbar scales. ALS-nonbulbar = ALS patients without any evidence of bulbar motor disease. bvFTD = behavioural variant frontotemporal dementia without motor syndrome. HC= healthy controls, PUMNs = Penn Upper Motor Neuron scale, higher values indicate more neurological signs. ECAS-NA = Edinburgh Cognitive Assessment Scale- North American version, lower values indicate more cognitive impairment. MMSE = Mini Mental State Examination, lower values indicate more cognitive impairment. WPM = words per minute of speech.