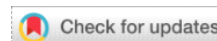




pISSN 2005-8063
eISSN 2586-5854
2026. 9. 30.
Vol.18 No.3
pp. 109-115



Vowel lengthening as a marker of dysarthria status in ALS: Evidence from read speech

Hayeun Jang*

Department of English Language and Literature, Sungkyunkwan University, Seoul, Korea

Abstract

Acoustic markers of bulbar involvement in amyotrophic lateral sclerosis (ALS) are typically validated against healthy controls. However, this cannot establish whether a measure tracks the disease or stage of bulbar involvement. Moreover, dysarthria status is correlated with the site of onset, since bulbar-onset disease is defined by early bulbar involvement. We analyzed a single standardized reading passage from 53 speakers (17 controls and 36 patients) using two contrasts: controls versus patients, and patients with versus without clinically judged dysarthria. Vowel duration was the only measure that was significant in both contrasts and survived correction for multiple comparisons. Patients with dysarthria produced vowels that were approximately 20% longer than those of controls and 23% longer than those of patients without dysarthria. Meanwhile, the latter did not differ from controls. This effect was unchanged after adjusting for the site of onset and after restricting the analysis to non-bulbar-onset patients. Moreover, it persisted in male-only reanalysis, leave-one-speaker-out resampling, and after adjusting for disease duration. The harmonics-to-noise ratio, fundamental frequency range, and pause duration distinguished patients from controls but did not differentiate patients by dysarthria status. Vowel space area contrasts were statistically indeterminate, in contrast to findings from the same cohort performing syllable repetition, indicating that measure selection and task selection interact.

Keywords: amyotrophic lateral sclerosis, dysarthria, vowel duration, site of onset, speech task

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder affecting upper and lower motor neurons. Bulbar involvement affects the cranial nerves controlling the lips, tongue, jaw, and larynx, producing a mixed spastic-flaccid dysarthria characterized by slowed speech, imprecise articulation, and altered voice quality.

Acoustic studies of dysarthria have reported reduced vowel space area (VSA) and vowel centralization (Chu, 2023; Kim et al., 2021; Youn et al., 2024), slower articulation rates with more frequent pauses (Kim, 2020; Lee et al., 2023), and reduced cepstral peak prominence (CPP) reflecting breathiness and strain (Seo & Seong, 2013).

1.1. The Choice of Speech Task

* hyjang@skku.edu, Corresponding author

Received 7 August 2026; Revised 1 September 2026; Accepted 3 September 2026

© Copyright 2026 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Standard assessment batteries for bulbar involvement in ALS combine three task types: sustained vowel phonation, diadochokinetic (DDK) syllable repetition, and passage reading (Green et al., 2013; Stegmann et al., 2024). The first two are attractive for automated analysis because they isolate laryngeal or articulatory subsystems under tightly controlled conditions, and much of the ALS acoustic literature has concentrated on them (Vashkevich & Rushkevich, 2021).

Studies comparing tasks directly, however, suggest that no single task is best for every question. For separating patients from healthy controls, non-linguistic tasks have generally performed well: Suhas et al. (2019) reported that DDK outperformed both sustained phonation and spontaneous speech for this classification, and sustained phonation has supported accurate ALS detection in several studies (Mallol-Ragolta et al., 2025; Vashkevich & Rushkevich, 2021).

For grading the degree of speech impairment, the ordering reverses. In a German-speaking ALS cohort, Gonzalez-Machorro et al. (2026) found that passage reading gave the closest agreement with the ALS Functional Rating Scale-Revised (ALSFRRS-R) speech subscore in both cross-sectional and within-speaker analyses, while sustained /a:/ performed worst; the authors suggest that read speech better reflects functional intelligibility, which is what the ALSFRRS-R speech item is designed to capture.

This distinction maps directly onto the two contrasts examined here. Our first contrast asks the detection question, on which non-linguistic tasks have an established advantage. Our second contrast is defined by the ALSFRRS-R speech subscore itself, and is therefore the question for which reading has been reported as the most informative task. The present study analyzes reading alone, and asks which acoustic measures extracted from a single standardized passage can support each contrast. Reading also has practical appeal for monitoring: it requires no instruction beyond the text, is short, and can be repeated identically across visits and remotely.

1.2. Within-Patient Contrasts and the Site-of-Onset Problem

Most of this evidence rests on a single comparison: patients versus healthy controls. That comparison establishes that a measure is sensitive to the presence of disease, but not what the measure indexes. For clinical monitoring the relevant question is different: does the measure change as bulbar involvement develops? Answering it requires contrasts within the patient group.

We therefore evaluate six acoustic measures across two contrasts. The first is controls versus patients with ALS, which asks whether the measure is sensitive to the disease at all. The second is patients with versus without clinically judged dysarthria, which asks whether the measure tracks the stage of bulbar involvement. A measure useful for monitoring should be significant in the second contrast, not only the first.

The second contrast carries a confound that is intrinsic to ALS rather than incidental to any one cohort. Bulbar-onset disease is defined by early bulbar involvement, so bulbar-onset patients will be over-represented among those with dysarthria in any cross-sectional sample. A comparison of dysarthric with non-dysarthric patients is therefore partly a comparison of bulbar-onset with limb-onset disease, and any measure that differs between onset phenotypes for reasons unrelated to bulbar staging will appear to track dysarthria. Site of onset is rarely modelled in

acoustic studies of ALS. We record it here and use it in three ways: as a covariate, as a stratification variable, and as an exclusion criterion, retaining only non-bulbar-onset patients in the strictest analysis.

Because ALS cohorts are necessarily small and clinically heterogeneous, we also apply procedures uncommon in this literature: correction for multiplicity, sex-stratified reanalysis, leave-one-speaker-out (LOO) resampling, equivalence testing for null results, and explicit examination of clinical confounders including site of onset, disease duration, and respiratory function. As will be shown, the last of these proves decisive for interpreting voice quality measures.

2. Method

2.1. Participants

Speech samples were collected prospectively from patients with ALS or motor neuron disease and healthy controls at a tertiary referral hospital between December 2019 and October 2020. The present study is a retrospective acoustic-phonetic analysis of this previously collected dataset. The analyzed cohort comprised the 53 speakers in Table 1 (17 controls and 36 patients with ALS) for whom both a usable recording of the reading passage and complete clinical ratings were available.

Patients met revised El Escorial criteria and were assessed with the ALSFRRS-R. Dysarthria status was determined by the ALSFRRS-R speech subscore: a score of 4 indicated absence of dysarthria, and a score below 4 indicated its presence. All 12 patients classified as non-dysarthric scored 4; the 24 dysarthric patients scored 3 (n=18) or 2 (n=6). Site of onset was recorded as bulbar, cervical, or lumbosacral; one patient with concurrent bulbar and lumbosacral onset was grouped with bulbar onset for analysis.

Speech samples from this cohort performing a different task have been reported previously. Kim et al. (2023) analyzed corner vowel parameters from repeated consonant-vowel syllables and consonant-vowel-consonant monosyllabic words recorded in the same sessions, using the same dysarthria classification, in 43 patients and 20 controls. The present analysis draws on the subset of those sessions for which the reading passage and complete clinical ratings were both available, and examines the connected reading passage, which was not analyzed in that report, together with measures of pause structure, spectral tilt, and periodicity that connected speech makes available. Where relevant we compare the two sets of findings, since the substantial overlap

in speakers allows task effects to be considered with clinical heterogeneity held largely constant.

2.2. Speech Materials

Recordings were made in a quiet room using a digital voice recorder (ICD-PX370, Sony, Tokyo, Japan). Participants read the "Travel" passage at a comfortable rate. Standardized Korean paragraphs of this kind are established materials in Korean clinical practice and research: they have been used to evaluate treatment outcomes in multiple system atrophy (Park, 2018), to measure speaking rate in spastic dysarthria (Nam & Park, 2023), to compare prosody across utterance tasks (Kim, 2020), and as the basis for automatic detection of inappropriate pauses in dysarthric speech (Kim et al., 2024; Lee et al., 2023). All measures reported here

Table 1. Participant characteristics

Feature	Control	ALS without dysarthria	ALS with dysarthria
n	17	12	24
Female / Male	9 / 8	1 / 11	14 / 10
Age (years)	49.7 (13.2)	58.3 (9.7)	57.8 (13.2)
ALSFRS-R total	-	37.0 (4.0)	36.1 (7.3)
ALSFRS-R bulbar	-	11.3 (0.8)	9.1 (1.3)
ALSFRS-R speech	-	4.0 (0.0)	2.8 (0.4)
FVC (% predicted) a	-	78.9 (17.5)	75.0 (20.2)
Years since onset	-	4.4 (5.4)	2.7 (1.7)
Site of onset: Bulbar	-	0	10
Site of onset: Cervical	-	9	7
Site of onset: Lumbosacral	-	3	7

Values are mean $\pm$ SD.

ALS, amyotrophic lateral sclerosis; ALSFRS-R, ALS functional rating scale-revised.

derive from this connected read speech; no sustained phonation or DDK task is analyzed.

2.3. Acoustic Measures

Speech was force-aligned using the Montreal Forced Aligner; parameters were extracted frame-wise using PraatSauce. Vowel duration was obtained from aligned TextGrids. Pause duration was computed as the gap between the offset of one word and the onset of the next in the aligned word tier, so that boundaries produced without an intervening silence yielded no pause. Gaps of 10 ms or shorter were excluded as alignment artifacts, although the shortest interval observed was 30 ms, which is the aligner's minimum assignable duration. The preceding syntactic boundary was labeled using the Kiwi morphological analyzer. CPP, harmonics-to-noise ratio (HNR), and the first-to-second harmonic amplitude difference (H1-H2) indexed periodicity, additive noise, and spectral tilt. Fundamental frequency was converted to a semitone range. VSA was computed from the first two formants of the corner vowels /a/ (transcribed /æ/ in the aligner's output), /i/, and /u/, using each speaker's mean formant values across all tokens of each vowel: $.5 \times [F2a(F1i-F1u) + F2i(F1u-F1a) + F2u(F1a-F1i)]$.

2.4. Statistical Analysis

Token-level measures were analyzed with linear mixed-effects models (lme4 in R) with random intercepts for speaker and vowel identity. Speaker-level measures (VSA, F0 range) were analyzed with multiple linear regression. Sex and scaled age were covariates.

Duration measures were log-transformed; raw-scale residuals

were severely right-skewed (maximum standardized residual >18). Log coefficients are reported as proportional differences.

All contrasts were computed directly using emmeans with Tukey adjustment rather than inferred from separate comparisons against controls, and were additionally corrected across the six measures within each contrast (Holm). Pause contrasts were corrected across the six boundary types.

To separate dysarthria status from site of onset, the patient-subgroup contrast was estimated under three specifications: (a) unadjusted; (b) adjusted for site of onset as a three-level covariate; and (c) restricted to non-bulbar-onset patients, comprising 9 non-dysarthric and 7 dysarthric cervical-onset patients and 3 non-dysarthric and 7 dysarthric lumbosacral-onset patients. Specification (c) is the strictest, since it removes every patient whose diagnosis entails early bulbar involvement. A parallel three-group comparison adding controls to the non-bulbar-onset patients was also computed.

The following procedures were applied uniformly to all measures rather than selectively: sex-stratified reanalysis in male participants only; LOO resampling; equivalence testing using 90% confidence intervals against bounds of 5,000 Hz² (VSA) and 2 dB (H1-H2); adjustment for disease duration; and direct testing of measure by group interactions at the speaker level. Because this is a reanalysis of an existing dataset, these procedures were specified during analysis rather than registered in advance.

3. Results

Table 2. Acoustic measures across two contrasts

Measure	Contrast 1: Control - ALS	Contrast 2: Non-dysarthric - Dysarthric
Vowel duration (log)	-0.126 [-0.239, -0.013] $p=.029$, $p\text{-adj}=.117$	-0.208 [-0.361, -0.055] $p=.005$, $p\text{-adj}=.033$
HNR (dB)	2.14 [0.52, 3.77] $p=.011$, $p\text{-adj}=.066$	-1.58 [-3.96, 0.81] $p=.255$
F0 range (st)	-4.06 [-7.35, -0.76] $p=.017$, $p\text{-adj}=.085$	2.80 [-2.06, 7.66] $p=.352$
CPP (dB)	1.06 [-0.45, 2.57] $p=.163$	-2.07 [-4.22, 0.08] $p=.061$
H1-H2 (dB)	1.40 [-2.25, 5.05] $p=.443$	2.83 [-2.57, 8.22] $p=.421$
VSA (Hz ²)	9,442 [-13,838, 32,723] $p=.419$	22,435 [-11,663, 56,534] $p=.259$

Estimates with 95% CI; $p\text{-adj}$ is Holm-corrected across the six measures within each contrast.

ALS, amyotrophic lateral sclerosis; CPP, cepstral peak prominence.

3.1. Overview

Table 2 presents all six measures across the two contrasts. Only vowel duration reached significance in both contrasts. HNR and F0 range separated patients from controls but showed no difference between patient subgroups. CPP, H1-H2, and VSA reached significance in neither.

3.2. Vowel Duration

Vowel duration was the only measure to survive Holm correction, doing so in the patient-subgroup contrast ($p\text{-adj}=0.033$). Back-transformed, patients with dysarthria produced vowels approximately 20% longer than controls (95% CI: 5%–36%) and approximately 23% longer than patients without dysarthria (95% CI: 6%–43%). Patients without dysarthria did not differ from controls (+0.027 log units, $p=.922$).

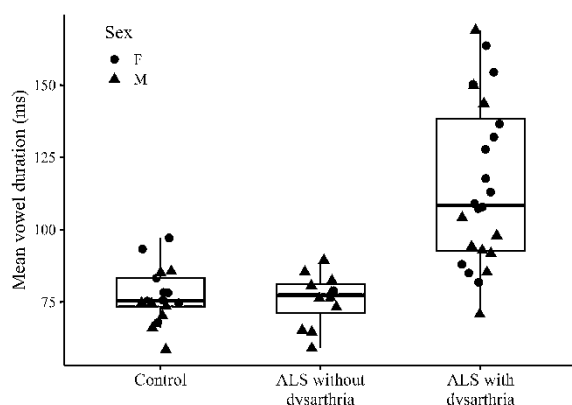


Figure 1. Mean vowel duration by group.

The shape of the effect is as informative as its significance: the non-dysarthric group matched controls rather than falling between controls and the dysarthric group, as shown in Figure 1. Speech timing on a reading task appears unaffected until dysarthria is clinically apparent. The effect held under every robustness procedure. In male participants only ($n=8/11/10$), it was if anything stronger: controls versus dysarthric patients -0.250 ($p=.005$), non-dysarthric versus dysarthric -0.247 ($p=.004$), controls versus non-dysarthric -0.003 ($p=.999$). LOO resampling retained significance in 100% of refits (p range .0002–.0042). Results were unchanged on the raw scale (36.7 ms, $p<.001$) and after trimming residuals beyond 3 SD (0.153, $p=.012$).

Adjusting for disease duration did not attenuate the patient-subgroup difference (unadjusted $b=0.232$, $p=.005$; adjusted $b=0.230$, $p=.007$), and disease duration itself was not predictive ($b=-0.003$, $p=.917$).

3.3. Separating Dysarthria Status from Site of Onset

All 10 bulbar-onset patients were dysarthric, as expected. Cervical and lumbosacral onset were not separated, however, dividing 9 to 7 and 3 to 7 between non-dysarthric and dysarthric patients (Table 1), which allows the group effect to be estimated within non-bulbar-onset strata.

Vowel duration was essentially unchanged across the three specifications (0.208, 0.192, 0.198), and site of onset itself had

no effect (cervical $b=-0.101$, $p=.281$; lumbosacral $b=-0.078$, $p=.390$). Adding controls to the non-bulbar-onset patients (17 controls, 12 non-dysarthric, 14 dysarthric) reproduced the full-cohort pattern: dysarthric patients differed from controls (-0.147 , $p=.016$) from non-dysarthric patients (-0.169 , $p=.018$), while non-dysarthric patients did not differ from controls (0.022, $p=.927$).

HNR behaved differently. The patient-subgroup contrast reached significance only with site of onset as a covariate (-2.87 , $p=.041$), and not unadjusted ($p=.255$) or with bulbar-onset patients excluded (-1.48 , $p=.132$). Restricted to non-bulbar onset, both patient subgroups differed from controls (2.76, $p=.002$; 1.63, $p=.039$) but not from each other (-1.12 , $p=.298$). CPP showed the same pattern more weakly, with only the controls versus non-dysarthric contrast approaching significance (2.48, $p=.052$). The apparent voice quality advantage of dysarthric patients in the full cohort thus reflected the bulbar-onset patients, whose HNR values were comparatively high.

The sex imbalance (non-dysarthric group 11 of 12 male; Table 1) follows partly from the same source, since bulbar-onset patients were predominantly female and cervical-onset patients predominantly male. Restricting to male participants (8 controls, 11 non-dysarthric, 10 dysarthric) gave the same picture: vowel duration held and was if anything stronger (controls versus dysarthric -0.250 , $p=.005$; non-dysarthric versus dysarthric -0.247 , $p=.004$; controls versus non-dysarthric -0.003 , $p=.999$), whereas for HNR only the controls versus non-dysarthric contrast reached significance (2.61, $p=.030$) and CPP reached it nowhere (all $p>.20$).

3.4. Measures that Index Diagnosis but not Dysarthria Status

HNR and F0 range distinguished patients from controls ($p=.011$ and .017) but not patient subgroups ($p=.255$ and .352), and neither survived correction across measures even in contrast 1 ($p\text{-adj}=.066$ and .085).

Pause duration showed a similar pattern. The group by boundary type interaction was not significant (contrast 1: $\chi^2(5)=4.90$, $p=.428$; contrast 2: $\chi^2(10)=13.94$, $p=.176$), so we report pause duration averaged across boundary types rather than decomposing it by boundary: both patient subgroups paused longer than controls (-0.460 , $p=.033$; -0.434 , $p=.006$) but did not differ from each other (0.026, $p=.986$). The two control contrasts differ in robustness, however. Because the shortest interval in our data is 30 ms, set by the temporal resolution of the aligner (Section 2.3), we repeated the analysis excluding gaps below 50, 100, 150, 200, and 250 ms. The difference between controls and dysarthric patients held at every threshold ($p=.004$ to .028), whereas the difference between controls and non-dysarthric patients was significant only at 30 ms and declined monotonically thereafter (-0.460 to -0.180 ; $p\geq.088$ at all higher thresholds), indicating that it rests on very short gaps that are not clearly separable from stop closures.

3.5. Secondary Analyses: Null Results, Measure Interactions, and Clinical Correlates

Equivalence testing was applied to H1-H2, using half the

control-group standard deviation (3.37 dB) as the bound. This criterion is determined by the control distribution alone and is therefore independent of the group comparisons being evaluated. Every contrast was classified as indeterminate. Because the choice of bound is consequential, we also report the smallest bound at which each contrast could be declared equivalent: for H1-H2 these range from 4.5 to 7.5 dB, the largest being comparable in magnitude to the effect of sex on that measure (−7.61 dB). Equivalence testing was not applied to VSA. The control-group standard deviation (44,371 Hz²) reached 71% of the group mean (62,761 Hz²), so no bound could be specified that would be both defensible as a negligible difference and smaller than the measure's own dispersion in this sample. The appropriate conclusion for VSA and H1-H2 is that this sample lacked the precision to detect or exclude group differences.

Modeling standardized vowel duration and CPP jointly at the speaker level, the interaction was not significant for contrast 1 [$F(1,49)=0.13$, $p=.716$] but was for contrast 2 [$F(2,48)=5.09$, $p=.010$], driven by the patient-subgroup difference ($p=.005$). It is carried by vowel duration rather than CPP, and therefore confirms that the temporal measure discriminates where the voice quality measure does not.

Subtracting the speech item from the ALSFRS-R bulbar domain yields a non-speech bulbar score (0–8) that was strongly skewed toward intact function, with 23 of 36 patients scoring 7 or 8. It predicted none of the acoustic measures (vowel duration $b=-0.027$, $p=.492$; HNR $b=-0.315$, $p=.574$; CPP $b=-0.449$, $p=.348$), though the limited variance makes these tests uninformative. Continuous speech scores were also weak predictors of vowel duration ($b=-0.075$, $p=.068$). However, documented bulbar symptom status classified 33 of 36 patients identically to the dysarthria criterion. Under this grouping the vowel duration result was reproduced (controls versus symptomatic $p=.011$; symptomatic versus asymptomatic $p=.030$). Given the 92% overlap between criteria, we treat this as supporting rather than independent evidence.

4. Discussion

4.1. Vowel Lengthening as a Correlate of Dysarthria Status

Of six acoustic measures, only vowel duration distinguished patients by dysarthria status, and it was the only measure to survive correction for multiple comparisons. The effect persisted after adjustment for site of onset, after excluding all bulbar-onset patients, in male-only reanalysis, in 100% of LOO refits, on the raw measurement scale, after outlier trimming, and after adjustment for disease duration.

Two features of the pattern are worth emphasis. First, patients without dysarthria matched controls almost exactly (−0.003 log units in the male-only analysis), rather than showing an intermediate value. Vowel duration therefore appears to index established rather than incipient bulbar involvement and is better suited to confirming and quantifying change than to anticipating it. Second, the magnitude (roughly 20% longer vowels) is substantial relative to the measurement, though whether it supports classification of individual patients was not tested here and would require a separate discriminability analysis.

4.2. Dysarthria Status and Disease Markers

Because all bulbar-onset patients were dysarthric, a naive comparison of dysarthric with non-dysarthric patients is partly a comparison of onset phenotypes. However, cervical- and lumbosacral-onset patients occurred in both subgroups, making the separation testable. For vowel duration, the separation succeeded: vowel lengthening tracks the presence of dysarthria rather than the associated onset phenotype. Conversely, for HNR, the separation revealed a confound; the difference was driven by the comparatively high HNR of the bulbar-onset patients.

HNR, F0 range, and pause duration separated patients from controls but showed no difference between patient subgroups. For pause duration the picture is partly resolved: dysarthric patients paused longer than controls robustly across pause-detection thresholds, whereas the corresponding difference for non-dysarthric patients depended on very short intervals and is left open (Section 3.4). What is clear in both cases is the absence of any difference between the patient subgroups (0.026, $p=.986$), which is the comparison that bears on staging. For HNR the same conclusion held after removing bulbar-onset patients, with both subgroups differing from controls and not from each other.

This dissociation has a practical implication. A measure validated only against healthy controls has been shown to be sensitive to the presence of ALS, which is not the clinical question, since patients are already diagnosed. Monitoring requires sensitivity to change in bulbar status, and only the within-patient contrast tests for it. Our results show that the two questions can give different answers for the same measure, and that a measure can be robustly sensitive to the first while being uninformative about the second.

We cannot determine what these measures are responding to. Respiratory involvement affects subglottal pressure and thus both phonatory stability and pause structure, and among patients with available spirometry, lower forced vital capacity predicted lower HNR (standardized $b=1.39$, $p=.029$, $n=24$).

A further consideration concerns what the grouping variable itself measures. The ALSFRS-R speech item rates functional intelligibility on a single ordinal scale and does not separate articulatory from phonatory contributions; a patient whose voice quality has degraded while intelligibility is preserved through compensation is still rated 4. That non-dysarthric patients differed from controls on HNR while the two patient subgroups did not differ from each other is therefore consistent with phonatory change that the perceptual item does not register. For vowel duration the position is the opposite: non-dysarthric patients matched controls almost exactly, so we have no evidence of timing change concealed by the scale. Acoustic measurement may complement perceptual rating where the scale is insensitive, but on our data this applies to the phonatory measures rather than to the temporal one.

4.3. Task Dependence and the Role of Reading

Comparing our results with a syllable repetition study in the same cohort (Kim et al., 2023) isolates task effects. Two findings converge: vowel duration was prolonged in dysarthric patients relative to both controls and non-dysarthric patients in both tasks, confirming it does not serve as an early indicator. Two findings diverge: VSA was significantly reduced in the syllable task but not in reading, whereas HNR separated patients from controls in reading

but not in syllables. Therefore, measure and task selection are interdependent, and comparisons across studies should be made task-wise.

Our findings align with the literature reporting that reading is the most informative task for grading speech impairment. Reading offers ecological and practical value by sampling connected speech with the timing and coarticulatory demands that make speech intelligible, which is highly advantageous for longitudinal monitoring.

Regarding the divergence in VSA, previous reports consistently found reduced VSA in dysarthric speakers (Chu, 2023; Kim et al., 2021; Youn et al., 2024), including in the syllable-based analysis of this cohort (Kim et al., 2023). We did not replicate it in reading, but our result does not contradict those reports: equivalence testing classified every VSA contrast as indeterminate, with confidence intervals wide enough to accommodate substantial differences in either direction.

Beyond the measurement noise that connected speech introduces, the temporal findings suggest a substantive possibility. Turner et al. (1995) showed that speaking rate and vowel space are coupled in ALS, with slower rate associated with a more expanded vowel space, and Han et al. (2013) reported a comparable effect of experimentally reduced rate in cerebral palsy. If our dysarthric patients were slowing sufficiently to reach articulatory targets in reading, a reduction in VSA would not be expected at this stage, whereas the shorter, more rapid productions of a syllable repetition task would leave less room for such compensation.

4.4. Limitations

Several limitations apply to this study. The cross-sectional design cannot establish the order in which acoustic changes emerge. Because all bulbar-onset patients were dysarthric, the two variables are partly confounded, and the strictest analyses rest on only 26 patients. Group sizes in this study were unbalanced (17 healthy controls, 12 non-dysarthric ALS, and 24 dysarthric ALS patients). While unbalanced designs do not bias fixed-effect estimates in linear mixed-effects models, they do distribute precision unevenly. Consequently, confidence intervals for pairwise comparisons involving the control group were wider due to its relatively smaller sample size. However, the principal finding of this study (i.e., the significant difference between the two ALS subgroups) does not involve the control group and thus remains robust against this specific imbalance. Nevertheless, the limited precision in estimating the control baseline, combined with the fact that the control group was on average eight years younger than the patient groups, warrants cautious interpretation of any direct comparisons made against the healthy baseline. The small sample size means null results for VSA and H1-H2 are indeterminate. Grouping was based on the ALSFRS-R speech subscore, which overlaps conceptually with temporal acoustic outcomes and which rates intelligibility without separating articulatory from phonatory change (Section 4.2), and disease duration was based on patient recall. Pause measurement depends on the temporal resolution of the aligner: the shortest interval available was 30 ms, and one of our two control contrasts was sensitive to whether intervals near that limit were retained, so pause-based measures require an explicit detection criterion before they can be compared across studies. Finally, speech was elicited using a single reading passage, which may not generalize to spontaneous speech, and residual normality was

rejected for all token-level models.

4.5. Toward Automated Monitoring

Vowel duration in read speech has properties that suit automated monitoring: the passage requires no clinician judgment or calibration of effort, extraction needs only forced alignment and segment measurement, and the effect is large relative to the measurement. A plausible implementation would have patients read the same passage on a smartphone at fixed intervals, tracking mean vowel duration against their own baseline, since within-speaker change is the relevant quantity and using each patient as their own reference removes the between-speaker variability that dominates cross-sectional comparison. Two caveats bear on this. Our design establishes a group-level difference rather than individual-level discriminability, and because the measure indexes established rather than incipient involvement, it is suited to quantifying change once it has begun rather than to anticipating onset. And automated measures inherit the properties of the tools that produce them, as our pause results illustrate: whether intervals near the aligner's resolution limit were retained changed one conclusion. Longitudinal validation with fixed and reported processing settings would be the next step.

5. Conclusion

Evaluating six acoustic measures from a standardized reading passage in 53 speakers, vowel duration was the only measure that distinguished patients by dysarthria status rather than merely patients from controls. Patients with dysarthria produced vowels approximately 20% longer than controls and 23% longer than patients without dysarthria, who did not differ from controls. The effect was unchanged after adjusting for site of onset and after excluding all bulbar-onset patients, and it survived correction for multiplicity, sex-stratified reanalysis, leave-one-speaker-out resampling, and adjustment for disease duration.

HNR, fundamental frequency range, and pause duration separated patients from controls but not by dysarthria status. For HNR the apparent patient-subgroup difference in the full cohort was attributable to bulbar-onset patients and did not survive their exclusion, illustrating that covariate adjustment can create rather than remove such an appearance.

VSA contrasts were statistically indeterminate in reading, although the same cohort showed significant reductions in a syllable repetition task, indicating that conclusions about a measure are conditional on the task from which it is extracted. We propose vowel duration in connected reading as the measure in this set best suited to monitoring bulbar involvement in ALS, and recommend that studies of dysarthria staging in ALS record and model site of onset.

Acknowledgement

We thank the clinical team at Seoul National University Hospital, including the authors of Kim et al., for patient recruitment and clinical assessment, and all participants for their contribution to this study.

References

- Chu, H. (2023). *Vowel space characteristics when producing meaningless and meaningful words in speakers with spastic dysarthria* (Master's thesis). Catholic University of Pusan, Busan, Korea.
- Gonzalez-Machorro, M., von Heynitz, R., Hanslmeier, J., Grimm, F., Deac, A. I., Gründel, A., Cordts, I., & Schuller, B. (2026). Towards speech impairment prediction in German-speaking individuals with amyotrophic lateral sclerosis. *arXiv*. <https://arxiv.org/abs/2606.17616>.
- Green, J. R., Yunusova, Y., Kuruvilla, M. S., Wang, J., Pattee, G. L., Synhorst, L., Zinman, L., & Berry, J. D. (2013). Bulbar and speech motor assessment in ALS: Challenges and future directions. *Amyotrophic lateral sclerosis and frontotemporal degeneration*, 14(7-8), 494-500.
- Han, J., Sung, J. E., Sim, H. S., & Lee, Y. M. (2013). Effects of speaking rate manipulation and the severity of dysarthria on speech intelligibility and acoustic parameters in persons with cerebral palsy. *Journal of Speech-Language & Hearing Disorders*, 22(1), 35-54.
- Kim, G., Lee, Y., Ryu, J., & Kwon, S. (2021). A comparative study on vowel space characteristics in normal and motor speech disorder groups. *Journal of Speech-Language & Hearing Disorders*, 30(2), 51-59.
- Kim, J., Kim, S., Kim, M., Jang, J., & Koo, M. (2024, June). Summarized analysis and rehabilitation suggestion for dysarthric speech using large language models. *Proceedings of the 2024 Korea Computer Congress*. Jeju, Korea.
- Kim, J. A., Jang, H., Choi, Y., Min, Y. G., Hong, Y. H., Sung, J. J., & Choi, S. J. (2023). Subclinical articulatory changes of vowel parameters in Korean amyotrophic lateral sclerosis patients with perceptually normal voices. *PLOS ONE*, 18(10), e0292460.
- Kim, M. (2020). *A comparative study on the characteristics of prosody through utterance tasks in adults with dysarthria and normal adults* (Master's thesis). Catholic University of Pusan, Busan, Korea.
- Lee, J., Choi, Y., & Koo, M. (2023, June). Detection of break index in dysarthric speech using large-scale speech recognition. *Proceedings of the 2023 Korea Computer Congress* (pp. 376-378). Jeju, Korea.
- Mallol-Ragolta, A., Gonzalez-Machorro, M., von Heynitz, R., Scherzer, K., Cordts, I., & Schuller, B. W. (2025, June). Early detection of ALS in absence of speech impairments with computer audition. *Proceedings of International Conference on Artificial Intelligence in Medicine* (pp. 273-282). Pavia, Italy.
- Nam, H., & Park, H. (2023). The characteristics of flaps sound duration in spastic dysarthria. *Journal of Speech-Language & Hearing Disorders*, 32(1), 9-15.
- Park, Y. (2018). Efficacy of intensive treatment of dysarthria for people with multiple system atrophy. *Phonetics and Speech Sciences*, 10(4), 163-171.
- R Core Team. (2024). R: A language and environment for statistical computing (version 4.4.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Seo, I., & Seong, C. (2013). Voice quality of dysarthric speakers in connected speech. *Phonetics and Speech Sciences*, 5(4), 33-41.
- Stegmann, G., Krantsevich, C., Liss, J., Charles, S., Bartlett, M., Shefner, J., Rutkove, S., ... Berisha, V. (2024). Automated speech analytics in ALS: Higher sensitivity of digital articulatory precision over the ALSFRS-R. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 25(7-8), 767-775.
- Suhas, B. N., Patel, D., Rao, N., Belur, Y., Reddy, P., Atchayaram, N., Yadav, R., ... Ghosh, P. K. (2019, September). Comparison of speech tasks and recording devices for voice based automatic classification of healthy subjects and patients with amyotrophic lateral sclerosis. *Proceedings of Interspeech 2019* (pp. 4564-4568). Graz, Austria.
- Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 38(5), 1001-1013.
- Vashkevich, M., & Rushkevich, Y. (2021). Classification of ALS patients based on acoustic analysis of sustained vowel phonations. *Biomedical Signal Processing and Control*, 65, 102350.
- Youn, D., Lee, H., An, Y., Lee, Y., & Lee, Y. (2024, November). Characteristics of vowel space parameter for patients with dysarthria. *Proceedings of the 23rd Korean Speech-Language & Hearing Association Academic Conference* (p. 161). Daegu, Korea.

• **Hayeun Jang**, Corresponding author

Associate Professor, Dept. of English Language and Literature,
Sungkyunkwan University
25-2, Seonggyungwan-ro, Jongno-gu, Seoul 03063, Korea
Tel: +82-2-780-0254
Email: hyjang@skku.edu
Areas of interest: Phonology, Phonetics