Research Note

Video-Based Facial Movement Analysis in the Assessment of Bulbar Amyotrophic Lateral Sclerosis: Clinical Validation

Diego L. Guarin,[a] [iD] Babak Taati,[b,c,d] Agessandro Abrahao,[e,f,g] [iD] Lorne Zinman,[e,g,h] and Yana Yunusova[b,e,i] [iD]

[a] Department of Applied Physiology and Kinesiology, University of Florida, Gainesville [b] KITE–Toronto Rehabilitation Institute, University Health Network, Ontario, Canada [c] Department of Computer Science, University of Toronto, Ontario, Canada [d] Institute of Biomedical Engineering, University of Toronto, Ontario, Canada [e] Hurvitz Brain Sciences Research Program, Sunnybrook Research Institute, Toronto, Ontario, Canada [f] Harquail Centre for Neuromodulation, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada [g] Division of Neurology, Department of Medicine, University of Toronto, Ontario, Canada [h] L.C. Campbell Cognitive Neurology Research Unit, Cognitive Neurology, Sunnybrook Research Institute, University of Toronto, Ontario, Canada [i] Department of Speech-Language Pathology and Rehabilitation Sciences Institute, University of Toronto, Ontario, Canada

A B S T R A C T

Purpose: Facial movement analysis during facial gestures and speech provides clinically useful information for assessing bulbar amyotrophic lateral sclerosis (ALS). However, current kinematic methods have limited clinical application due to the equipment costs. Recent advancements in consumer-grade hardware and machine/deep learning made it possible to estimate facial movements from videos. This study aimed to establish the clinical validity of a video-based facial analysis for disease staging classification and estimation of clinical scores.
Method: Fifteen individuals with ALS and 11 controls participated in this study. Participants with ALS were stratified into early and late bulbar ALS groups based on their speaking rate. Participants were recorded with a three-dimensional (3D) camera (color + depth) while repeating a simple sentence 10 times. The lips and jaw movements were estimated, and features related to sentence duration and facial movements were used to train a machine learning model for multiclass classification and to predict the Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised (ALSFRS-R) bulbar subscore and speaking rate.
Results: The classification model successfully separated healthy controls, the early ALS group, and the late ALS group with an overall accuracy of 96.1%. Video-based features demonstrated a high ability to estimate the speaking rate (adjusted R^2 = .82) and a moderate ability to predict the ALSFRS-R bulbar subscore (adjusted R^2 = .55).
Conclusions: The proposed approach based on a 3D camera and machine learning algorithms represents an easy-to-use and inexpensive system that can be included as part of a clinical assessment of bulbar ALS to integrate facial movement analysis with other clinical data seamlessly

Amyotrophic lateral sclerosis (ALS) is a fast-progressing neurodegenerative disease affecting motor neurons in the brain, brainstem, and spinal cord. Over 70% of patients present with initial symptoms of weakness, loss of fine motor control, or spasticity in the arms and legs (also known as "spinal onset"). The remaining patients experience bulbar (i.e., speech and swallowing) changes at disease onset. For patients with spinal onset ALS, bulbar changes develop slowly and gradually. However, their detection is of utmost importance because the onset of bulbar disease often signifies a much more rapid and debilitating phase of disease progression and leads to reduced survival (Chiò et al., 2009). Early detection of bulbar disease and prediction of its progression in these

patients could lead to earlier intervention and more effective patient stratification and recruitment for clinical trials.

The bulbar subscore of the Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised (ALSFRS-R) is the most common measure currently used in the clinic to quantify the progression of bulbar disease (Plowman et al., 2017). ALSFRS-R, however, has many limitations, including a small number of items to measure specific dysfunctions (e.g., only three items measure bulbar dysfunction) and reduced sensitivity to disease progression, particularly in the early or late stages of the disease (Cudkowicz et al., 2013; Ilse et al., 2015; Rutkove, 2015). Because of this, speaking rate, defined as the number of words produced per minute (WPM) during a reading task, has been strongly suggested as a more suitable measure to track bulbar ALS, establish disease severity, and plan communication interventions (Ball et al., 2001; Barnett et al., 2020; Green et al., 2013; Yunusova et al., 2019). However, speaking rate can be affected by other premorbid factors such as reduced reading ability/disability or cognitive and respiratory impairments that can be present earlier in the disease than bulbar dysfunction (Hardiman, 2011; Strong et al., 1996; Yunusova et al., 2016). Emerging work suggests that movement-based (kinematic) articulatory measures, including range of motion, velocity, and acceleration, are highly specific to the neuromuscular deficits in ALS and might provide important information complementary to ALSFRS-R bulbar subscore and speaking rate (Hirose et al., 1982; Lee et al., 2018, 2020; Yunusova et al., 2008, 2010, 2012).

Previous studies suggested that articulatory kinematics could be affected before speech intelligibility and speaking rate in ALS, providing important information for disease characterization and prediction in the earliest stages of bulbar disease in ALS (Bandini, Green, Wang, et al., 2018; Rong et al., 2015; Shellikeri et al., 2016; Yunusova et al., 2010). Bandini et al. showed that kinematic features from the lips and jaw obtained from a sentence repetition task could classify individuals with ALS from healthy controls with an accuracy of 89% and distinguish presymptomatic (WPM > 160) from symptomatic (WPM < 160) patients with an overall accuracy of 87.4% (Bandini, Green, Wang, et al., 2018). Bandini et al. also showed that combining kinematic information from the lips and jaw with a measure of sentence duration could classify patients with presymptomatic (WPM > 160), early symptomatic (120 < WPM < 160), and late symptomatic (WPM < 120) ALS with an overall accuracy close to 80%. Sentence duration was algorithmically selected as the top classification feature overall; however, when used as the only feature, the classification algorithm showed an accuracy of around 50% (Bandini et al., 2017).

This work aims at building on the previous studies while improving on their clinical utility. The previous studies employed research-grade technologies such as optical tracking or electromagnetic articulography to obtain kinematic measures (e.g., Bandini et al., 2017, Bandini, Green, Taati, et al., 2018; Bandini, Green, Wang, et al., 2018; Wang et al., 2018). Novel consumer-grade technologies (e.g., three-dimensional [3D] cameras and webcams) paired with sophisticated machine learning models and algorithms for face alignment showed promise for digital clinical applications (Guarin et al., 2018; Jin & Tan, 2017). We recently demonstrated that a deep learning–based model for face alignment paired with a 3D camera could automatically localize the position of 68 facial landmarks in videos of patients with ALS with high accuracy; the average point-to-point Euclidean distance between the model's results and manual annotations was around 1.5% of the diagonal size of the face (Bandini et al., 2020). Promising results were also shown in a recent study where webcam video–based mouth and facial movements were combined with acoustic information to identify patients with ALS and predict the ALSFRS-R bulbar subscore (Neumann et al., 2021). However, it is not yet clear if it is possible to detect bulbar ALS across the course of the disease and predict its severity using video-based movement measures provided by a 3D camera and a face alignment algorithm.

In this study, we aimed to further understand the role of video-based kinematic measures and a simple timing measure (i.e., sentence duration) in disease detection and staging. We aimed to demonstrate that video-based articulatory kinematic features obtained with a 3D camera during the production of a short sentence can (a) distinguish healthy controls from individuals with ALS at different (early and late) stages of bulbar decline and (b) predict the severity of bulbar involvement as measured by speaking rate and ALSFRS-R bulbar subscore.

Based on previous results, we hypothesized that kinematic features would play a key role in separating patients in the early stage of bulbar ALS from controls, whereas sentence duration would be a key feature in distinguishing individuals in the later stage of bulbar ALS, where there is an evident decline in speaking rate. Moreover, we also hypothesized that both kinematic and durational measures would predict speaking rate and ALSFRS-R bulbar subscore across severity groups. The results of this study would establish the clinical validity of video-based face tracking in a sample of patients with ALS, supporting the use of this technology in the digital assessment of the disease (Goldsack et al., 2020).

Method

Participants

All participants signed informed consent according to the Declaration of Helsinki. The protocol was reviewed

and approved by the Sunnybrook Research Ethics Board. The data were collected from 15 participants with ALS (seven females) and 11 healthy controls (six females). One participant with ALS could not complete the task, and data from another participant were missing. Data from 13 individuals with ALS were analyzed. All participants were native speakers of English, had self-reported normal hearing, and scored within a normal range on a cognitive screen (i.e., Montréal Cognitive Assessment, score > 26; Nasreddine et al., 2005).

Participants in the ALS group were stratified into two subgroups based on their speaking rate: early and late bulbar ALS. All participants in the early group exhibited a Speech Intelligibility Test (SIT; Yorkston et al., 1995, 2007) speaking rate of < 155 WPM, which is below the 2 *SD*s of the mean speaking rate expected for the healthy controls as established in a large sample (Barnett et al., 2020). Participants in the late group demonstrated a speaking rate of < 120 WPM, chosen as a milestone after which a rapid bulbar deterioration and loss of speech intelligibility are expected (Rong et al., 2016). Table 1 presents a summary of the patients' demographic and clinical information.

Protocol/Procedure

Prior to kinematic data collection, overall and bulbar disease severity was assessed using the ALSFRS-R (Cedarbaum et al., 1999) and the SIT (Yorkston et al., 1995, 2007). During the SIT, participants read 11 randomly generated sentences containing five to 15 words. A single listener who is unfamiliar with the stimuli and the speakers transcribed the recordings orthographically (see details in Stipancic et al., 2018); we have used this method in our previous research in ALS (Barnett et al., 2020; Green et al., 2013; Rong et al., 2015, 2016). The speaking rate was measured as the average number of WPM. Speech intelligibility was calculated as the proportion of correctly transcribed words to the total number of words across sentences.

During the recording session, participants were seated in a quiet room in front of an Intel RealSense depth camera (SR300 or D400), recording depth and color videos at 50 frames per second at VGA resolution (640 × 480 pixels). A continuous light source was placed adjacent to the camera to provide uniform illumination. Participants repeated the sentence "Buy Bobby a puppy" 10 times at a comfortable speaking rate and loudness; the first repetition was not used for further analysis. The sentence was chosen as it elicits large articulatory motions of the lips and jaw.

We employed a previously validated deep learning–based model for facial alignment to track 68 facial landmarks in each video frame (Bandini et al., 2020); Figure 1 demonstrates the position of the 68 facial landmarks provided by the model. To be consistent with previous marker-based studies, only eight landmarks were used to estimate relevant facial movement features (Bandini et al., 2017; Bandini, Green, Wang, et al., 2018; Wang et al., 2016). These landmarks include the left and right medial canthus, the left and right oral commissures, the top and bottom vermilion borders at midline, the jaw, and the tip of the nose.
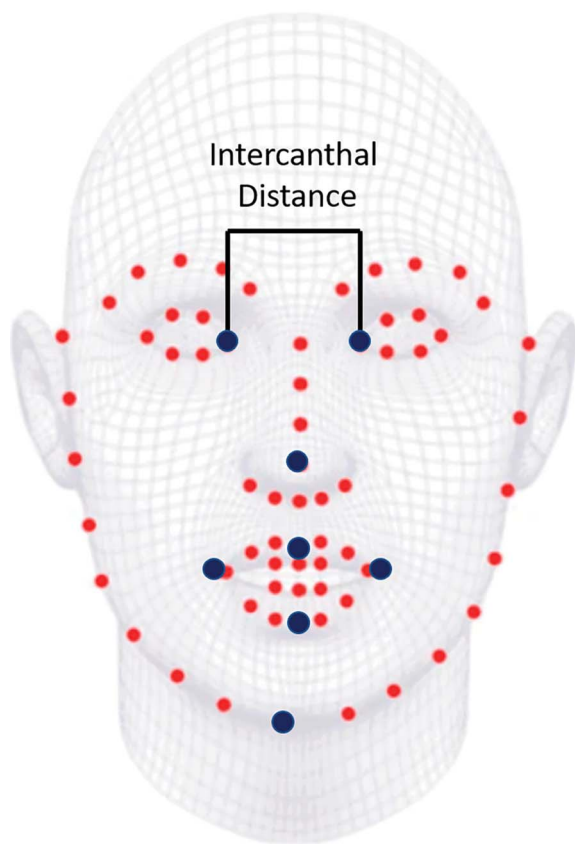
Depth and color videos were aligned using the camera's intrinsic information, and the real-world 3D coordinates of each landmark were computed following standard procedures (Guarin et al., 2020). Landmarks were normalized to have zero mean and the same scale in each video frame. Normalization was applied to make the measurements robust to changes in head pose and was achieved by dividing each landmark position by the average intercanthal distance (Sagonas et al., 2013). The time series

Table 1. Demographic and clinical information of the data set employed in this study.

Variable	Controls (*n* = 11)	Early ALS (*n* = 7)	Late ALS (*n* = 8)
Age (years)	61.2 ± 22.05 [22–78]	60.84 ± 7.31 [55–75]	63.75 ± 8.34 [45–73]
Sex (female/male)	6/5	2/5	5/3
Onset (spinal/bulbar)	—	6/1	4/4
Disease duration (months)	—	42 ± 48.06 [14–159]	59 ± 80.37 [10–264]
ALSFRS-R total	—	38.16 ± 0.89 [37–39]	35.33 ± 4.45 [26–40]
ALSFRS-R bulbar	—	10.33 ± 1.79 [7–12]	9.33 ± 1.88 [7–12]
Speech intelligibility (%)	98.2 ± 1.8 [94.5–100]	97.12 ± 3.49 [91.8–100]	68.54 ± 26.83 [28.2–92.7]
Speaking rate (words/min)	181.9 ± 20.7 [148.9–212.2]	141.79 ± 7.07 [130–150]	88.05 ± 14.94 [63.28–110]

Note. The means of each group are provided with standard deviations and ranges in brackets. ALS = amyotrophic lateral sclerosis; ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

Figure 1. Localization of 68 facial landmarks. The landmarks used in this study are highlighted in blue.

corresponding to the position of each landmark were low-pass filtered using a fifth-order polynomial filter to smooth the signal and facilitate the estimation of its derivatives (Pedregosa et al., 2011; Savitzky & Golay, 1964).

Kinematic Features

For each sentence, 64 features were extracted. The feature set included the following:

- One feature corresponds to sentence duration.
- Five features of the lip opening—estimated as the Euclidean distance between the lower and upper lip—include the mean; interquartile range; and 95th, 50th, and 5th percentiles of the signal.
- Twenty-one features of the lower lip movement— estimated as the Euclidean distance between the lower lip and the nose—include the cumulative sum of the lower lip movement signal and the mean; interquartile range; 95th, 50th, and 5th percentiles of the velocity, speed, acceleration, and jerk signals. In addition, 95th and 5th percentiles were used to represent near maximums of the signal to avoid the effect

of outliers. The velocity estimates were directional in nature; 95th percentile indicated near maximum during opening, whereas 5th percentile indicated near maximum during closing movements (Bandini et al., 2017).

- Twenty-one features of the jaw movement— estimated as the Euclidean distance between the jaw and the nose—include the cumulative sum of the jaw movement signal and the mean; interquartile range; 95th, 50th, and 5th percentiles of the velocity, speed, acceleration, and jerk signals.
- Five features of the left lip area—estimated as the area enclosed by the top lip, bottom lip, and left side of the mouth—include the mean; interquartile range; and 95th, 50th, and 5th percentiles.
- Five features of the right lip area—estimated as the area enclosed by the top lip, bottom lip, and right side of the mouth—include the mean; interquartile distance; and 95th, 50th, and 5th percentiles.
- Five features of the overall mouth area—estimated as the sum of the left and right mouth areas— include the mean; interquartile distance; and 95th, 50th, and 5th percentiles.
- One feature related to mouth symmetry—estimated as the mean ratio between the left and right mouth areas.

These features were selected based on previous research indicating that measures related to the range of motion, velocity, acceleration, jerk of the lips and jaw, and surface/ symmetry of the mouth might be sensitive to bulbar disease in ALS (Bandini et al., 2017; Wang et al., 2016).

Statistical Analyses

Classification

To differentiate the patients at early and late stages of bulbar ALS and control speakers, we used a gradient boosted (XGBoost) decision trees algorithm. The XGBoost is an ensemble-based machine learning algorithm comprising a collection of decision trees that perform nonlinear classification/ regression by recursively partitioning the data into disjoint branches (Hastie et al., 2009). XGBoost incorporates significant advantages of tree-based methods, handles different types of predictor variables, accommodates missing data, has no need for data normalization or elimination of outliers, can fit complex nonlinear relationships, and automatically handles interaction effects/correlation between predictors (Elith et al., 2008).

Three classification tests were performed using different feature groups: (a) sentence duration as the only feature, (b) kinematic features as the only features, and (c) sentence duration and kinematic features combined. The output of the XGBoost model was the probability (a value

between 0 and 1) that a sentence repetition was performed by a control (Class 0), a patient in the early ALS group (Class 1), or one in the late ALS group (Class 2). Probability scores for each repetition were averaged to obtain a final score for each subject.

Classification performance was evaluated using leave-one-subject-out cross-validation (LOSO-CV). All repetitions belonging to one participant were used as the test set for each repetition of the LOSO-CV, and the classifier was trained with repetitions from other participants. Model performance was evaluated by computing the overall accuracy and precision (i.e., the ratio between the true positives and the sum of true positives and false positives), recall (i.e., the ratio between the true positives and the sum of true positives and false negatives), and F1 scores (i.e., the harmonic mean of precision and recall) for each class (Artstein & Poesio, 2008). The XGBoost algorithm provided a ranked list of the most important features for classification. The classification procedure was repeated by removing features with a lower ranking, one at a time, until the best classification performance for the validation set was achieved. Statistical comparison among groups between features providing the best model was performed using a Welch's t test. Statistical significance was considered at $\alpha = .05$.

Prediction of Clinical Measures

Using the XGBoost algorithm for regression, we estimated two clinical measures: speaking rate and the ALSFRS-R bulbar subscore. Three regression tests were performed using three different feature groups: (a) sentence duration as the only feature, (b) kinematic features as the only features, and (c) sentence duration and kinematic features combined. The output of the XGBoost model was the predicted clinical measure of a participant based on a single trial. The predicted clinical measure for all repetitions was averaged to obtain a final value for each subject.

Model prediction performance was evaluated using LOSO-CV as well. Model performance was evaluated by computing the root-mean-square error (RMSE) and the adjusted coefficient of determination (adjusted R^2) between the actual and predicted values. Adjusted R^2 is a modification of R^2 that considers the number of regressors in the model. When a new regressor is added to the model, the adjusted R^2 increases only if the prediction error improves by more than expected by chance (Nakagawa et al., 2017). The XGBoost algorithm provided a ranked list of the most important features for prediction. The regression procedure was repeated by removing features with a lower ranking one at a time until the best regression performance for the validation set was obtained based on the adjusted R^2.

Results

Classification

Table 2 presents the classification results obtained in the three tests performed in this study. The average classification accuracy was 53.8% when the sentence duration was the only feature used, 80.7% when kinematic features alone were used, and 96.1% when sentence duration and kinematic features were used together. When only sentence duration was used for classification, the algorithm perfectly separated the late group from the other two groups. However, it could not distinguish between patients in the early bulbar group and controls. When only kinematic features were used, classification improved, but the accuracy only reached 80%. Finally, the algorithm

Table 2. Classification results using the three categories of features used for classification in this study.
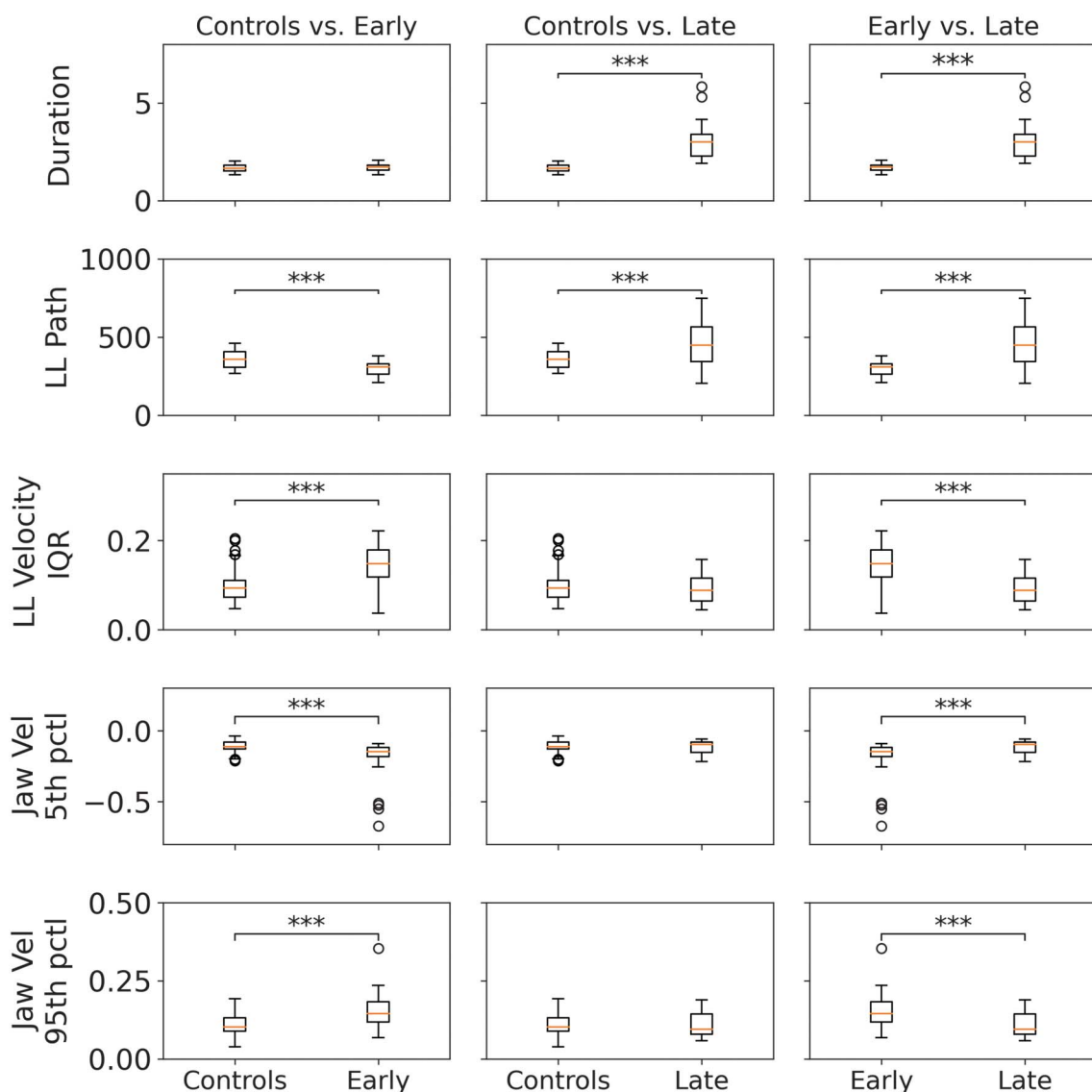
Duration				
Accuracy (%)	Class	Precision (%)	Recall (%)	F1 score (%)
53.8	Control	46.2	54.5	50.0
	Early	0.0	0.0	0.0
	Late	100.0	100.0	100.0
Kinematics				
Accuracy (%)	Class	Precision (%)	Recall (%)	F1 score (%)
80.7	Control	84.6	100.0	91.7
	Early	71.4	71.4	71.4
	Late	83.3	62.5	71.4
Duration + kinematics				
Accuracy (%)	Class	Precision (%)	Recall (%)	F1 score (%)
96.2	Control	91.7	100.0	95.7
	Early	100.0	85.7	92.3
	Late	100.0	100.0	100.0

perfectly separated the late bulbar group from the other two groups when sentence duration and kinematic features were combined. It demonstrated an excellent ability to separate controls from patients in the early group, with only one misclassified individual in each group.

The XGBoost algorithm selected five features as relevant for classification. Selected features included sentence duration and four kinematic features, including the cumulative sum of the lower lip movement signal, interquartile range of the lower lip velocity, and jaw velocity at 5% and 95th percentile, indicating near maximum jaw velocity during mouth opening and closing gestures, respectively.

Figure 2 presents box-and-whiskers plots comparing relevant features among the different groups. Sentence duration was not significantly different between controls and patients in the early bulbar ALS group ($t = 1.12$, $p = .26$); in contrast, the late bulbar ALS group demonstrated a sentence duration that was significantly longer than that observed for controls ($t = 14.57$, $p < .01$) and the early bulbar ALS group ($t = 14.11$, $p < .01$). The cumulative sum of the lower lip movement signal was significantly higher for controls than the early bulbar ALS group ($t = 7.67$, $p < .01$) and significantly higher for the late bulbar ALS group than controls ($t = 6.26$, $p < .01$) and the early

Figure 2. Differences in features extracted from the mouth area, including sentence duration, the cumulative sum of the lower lip movement signal, interquartile range (IQR) of the lower lip velocity, maximum jaw velocity when opening the mouth (95th percentile), maximum jaw velocity when closing the mouth (5th percentiles), and median overall mouth area between controls and the early bulbar group (left), controls and the late bulbar group (middle), and the early and late bulbar groups (right). A Welch's *t* test was performed to evaluate the difference in features between groups (*$p < .05$, **$p < .01$, and ***$p < .001$). LL = lower lip; Vel = velocity; pctl = percentile.

bulbar ALS group ($t = 9.51$, $p < .01$). The interquartile range of the lower lip velocity was significantly higher for the early bulbar ALS group than for controls ($t = 7.47$, $p < .01$) and the late bulbar ALS group ($t = 8.32$, $p < .01$). However, it was not significantly different between controls and the late bulbar ALS group ($t = 1.33$, $p = .18$). The 95% of jaw velocity (opening gesture) was significantly higher for the early bulbar ALS group than for controls ($t = 4.64$, $p < .01$) and the late bulbar ALS group ($t = 4.37$, $p < .01$). However, it was not significantly different between controls and the late bulbar ALS group ($t = 0.25$, $p = .67$). The 5% jaw velocity (closing gesture) was significantly higher (a larger negative number) for the early bulbar ALS group than for controls ($t = 4.44$, $p < .01$) and the late bulbar ALS group ($t = 4.12$, $p < .01$). However, it was not significantly different between controls and the late bulbar ALS ($t = 0.41$, $p = .67$).

Estimation of Clinical Measures

Predicting Speaking Rate

Table 3 presents the regression results for estimating the speaking rate. Results showed that a model considering sentence duration and kinematic features demonstrated superior performance in estimating speaking rate as reflected in the highest adjusted R^2 and the lowest RMSE. Figure 3 presents the measured and estimated speaking rate obtained using the relevant features selected by the XGBoost algorithm; these features included sentence duration and three kinematic features (cumulative sum of the lower lip path, 95th percentile of jaw jerk, and the symmetry between the left and right mouth areas). Results demonstrated a close relationship between the measured and estimated values (supported by an adjusted R^2 of .82).
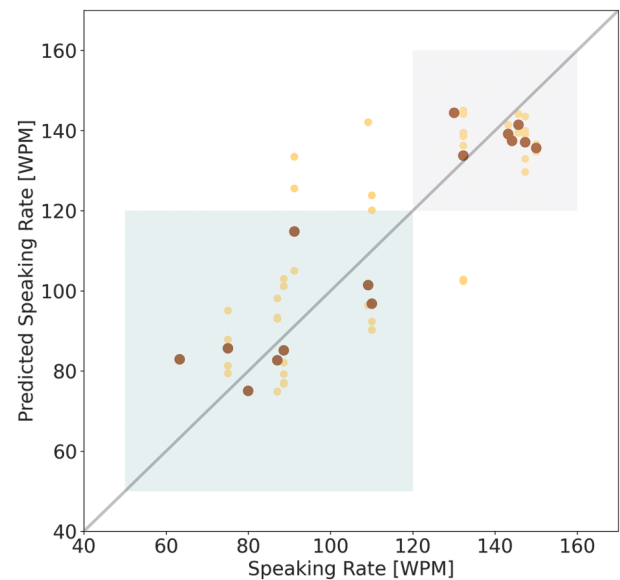
Estimating ALSFRS-R Bulbar Subscore

Table 4 presents the regression results for estimating the ALSFRS-R bulbar subscore. Results showed that a model with only kinematic features demonstrated better performance than the models that included duration alone or duration and kinematic features combined. Figure 4 presents the measured and predicted ALSFRS-R bulbar subscore with two kinematic features selected as relevant by the XGBoost algorithm (95th percentile of lower lip velocity and the

Figure 3. Relation between measured (x-axis) and predicted (y-axis) speaking rate in patients with amyotrophic lateral sclerosis (ALS) using a model with duration and kinematic features combined. Light dots indicate the prediction for each repetition of "Buy Bobby a puppy." Darker dots represent the repetitions' average. Shaded areas separate the early (≥ 120 WPM) from the late (< 120 WPM) bulbar ALS group data. WPM = words per minute.

symmetry between the left and right mouth areas). Results demonstrated a relatively close fit between the measured and predicted values supported by an adjusted R^2 of .55 and an RMSE of 0.91. Results also showed that predicted scores for patients with the lowest ALSFRS-R were highly variable, indicating a large repetition variability in performance in patients with a more advanced bulbar disease.
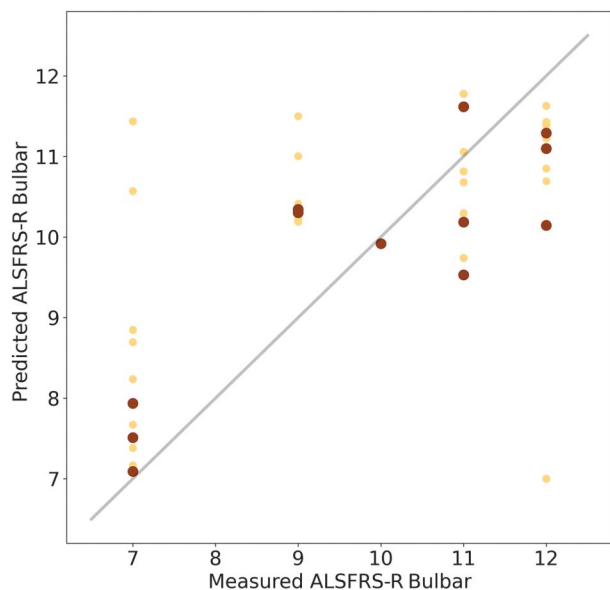
Discussion

This study demonstrated that video-based facial kinematic features provided essential information, complementary to that provided by a timing measure of sentence duration, to support the identification of patients with ALS, particularly in the early stage of the disease. Moreover, our results indicated that it was possible to predict the bulbar subscore of ALSFRS-R by using kinematic features alone

Table 3. Adjusted R^2 score and root-mean-square error (RMSE) were obtained for this study's three feature groups used for predicting speaking rate.

Group	Adjusted R^2	RMSE (words/min)
Duration	.65	15.07
Kinematics	.60	14.66
Duration + kinematics	**.82**	**11.42**

Note. Bold data indicate goodness of fit.

Table 4. Adjusted R^2 score and root-mean-square error (RMSE) were obtained for this study's three feature groups used for predicting bulbar Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

Group	Adjusted R^2	RMSE (score)
Duration	−.95	2.11
Kinematics	**.55**	**0.88**
Duration + kinematics	.52	0.91

Note. Bold data indicate goodness of fit.

Figure 4. Relation between measured (*x*-axis) and predicted (*y*-axis) bulbar subscore of Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised (ALSFRS-R) in patients with amyotrophic lateral sclerosis (ALS) using kinematic features. Light dots indicate the prediction for each repetition of "Buy Bobby a puppy." Darker dots represent the repetitions' average.

and the speaking rate by combining the duration of a short sentence with a few kinematic features. These results suggested that combining kinematic and timing/durational measures in a multimodal approach might improve bulbar disease detection and tracking specificity.

A substantial effort has been employed to develop and validate interpretable, automatic, and remote bulbar disease assessment methods. Automatic acoustic analysis has emerged as the leading candidate for objectively measuring of speech decline associated with neurodegenerative diseases (Orozco-Arroyave et al., 2020; Singh & Xu, 2020; Stegmann et al., 2020). However, many acoustic features are sensitive to the recording conditions, the encoding used for data transmission, and noise levels (Deliyski et al., 2005; Weerathunge et al., 2021). Certain acoustic measures may also show reduced test–retest reliability (Carding et al., 2004; Caverlé & Vogel, 2020; Lopes et al., 2017; Maryn et al., 2009). This work supports the ongoing efforts to develop multimodal assessment approaches that combine acoustic and kinematics analyses. Such multimodal assessments might provide a more robust characterization of bulbar disease throughout its course (Neumann et al., 2021; Sun et al., 2018; Walsh & Smith, 2012) and particularly, as our results suggested, in the earlier stages of bulbar disease. Our results agreed with a recent study by Neumann et al. (2021), in which acoustic and video-based kinematic features were combined to differentiate patients with bulbar symptomatic ALS (ALSFRS-R < 12)

from controls and patients with bulbar presymptomatic ALS (ALSFRS-R = 12). Neumann et al. showed that a multimodal approach improved disease detection and staging, with kinematic features playing an important role in predicting the ALSFRS-R score in the bulbar symptomatic and asymptomatic groups. In our study, we did not evaluate bulbar presymptomatic patients or employed acoustic features (except for a timing measure of the sentence duration, which can be estimated based on the acoustic signal); a full array of acoustic features will be the focus of our future studies. Regardless of the differences, this study and the study of Neumann et al. are critical first steps toward clinically validating video-based systems for remote assessment of ALS.

Feature interpretability is a critical aspect of any automatic disease diagnosis and motoring system. Our approach provided clinically relevant, easy-to-interpret features that might lead to a better understanding and characterization of bulbar involvement in ALS. In agreement with previous studies performed with research-grade equipment, our results showed that the lower lip and jaw overall movement and velocity were important features to differentiate between healthy controls and patients in the early and late bulbar ALS groups. Specifically, we observed that the early ALS group demonstrated significantly larger movements and higher jaw velocity in opening and closing mouth movements. This result is consistent with previous studies comparing patients with ALS at different stages of the disease (Bandini et al., 2017; Bandini, Green, Wang, et al., 2018; Yunusova et al., 2010) and with studies comparing healthy controls with patients with ALS in the early stages of the disease (Mefferd et al., 2012; Perry et al., 2018). Such behavior had been suggested to compensate for the decline in tongue function, which starts early in the disease (Rong & Green, 2019). This compensatory mechanism does not seem to be present in the late bulbar group, likely due to muscle weakness. We further observed that the late bulbar ALS group demonstrated a significantly larger overall movement of the lower lips—as measured by the cumulative sum of the lower lip movement signal than the other groups. The increase in the cumulative movement of the lips is consistent with the increased sentence duration observed in the late bulbar ALS group (Bandini et al., 2017).

The global measure of speaking rate is the most common estimate of bulbar disease severity in ALS (Barnett et al., 2020; Yorkston et al., 1995, 2007). Our results indicated that it is possible to predict speaking rate by using a model that combined the duration of a short sentence with kinematic features from the lips and jaw, specifically using the features of cumulative distance, jerk, and mouth symmetry. These results suggest that sentence duration and kinematic measures were complementary to each other. Moreover, our results demonstrated a

moderate relationship between articulatory kinematics of the lips and the ALSFRS-R bulbar subscore, suggesting that the ALSFRS-R bulbar subscore might be influenced by the movement characteristics (i.e., lip velocity and mouth symmetry). Notably, lips symmetry played prominently in predicting disease severity in both severity models (i.e., speaking rate and ALSFRS-R), but not for patient classification. These results suggest continuous asymmetric facial musculature weakens with disease progression in ALS, which should be evaluated in detail in future studies (see also Bandini, Green, Taati, et al., 2018). In contrast, weighting in the classification models was more prominently placed on the sentence duration measure, so the classification algorithm discarded information provided by many kinematic features, including symmetry.

Limitations

The main limitation of this study was the small sample size in both patient and control groups, affected by difficulty in in-person data collection due to the COVID-19 pandemic. Based on this limitation, future studies are required to validate our results with a larger population.

Considering that the bulbar decline is triphasic, with many patients with spinal onset disease exhibiting a presymptomatic phase (speaking rate > 155 WPM; ALSFRS-R bulbar subscore = 12; Rong et al., 2015), our future work will focus on the detection of neuromuscular abnormalities in the bulbar musculature prior to the decline in speaking rate and ALSFRS-R bulbar subscore. These results would lead to better management planning for these patients as early as possible and improved stratification for clinical trials.

In this study, only a single speech task was analyzed. A more comprehensive set of results might be obtained by analyzing additional speech and nonspeech tasks, for example, items of the cranial nerve exam (e.g., lip spread and eyebrows elevation). Moreover, we averaged all the results obtained for repetitions from a single patient, ignoring repetition variability. Repetition variability might be essential in characterizing bulbar decline in ALS (Kuruvilla-Dugdale & Mefferd, 2017). Future studies should include intertrial variability as a potential indicator of bulbar decline.

Conclusions and Clinical Implications

Overall, these preliminary results supported the validity of using a 3D video camera paired with a deep learning algorithm to evaluate bulbar dysfunction in ALS. Our results highlighted the potential utility of video-based articulatory kinematic features in identifying and staging bulbar ALS, validating the current video-based method. They

largely agreed with previous classification results obtained using high-precision research-grade instruments (Bandini et al., 2017; Bandini, Green, Taati, et al., 2018; Bandini, Green, Wang, et al., 2018; Rowe et al., 2020; Wang et al., 2018). They expanded existing results by showing that predicting speaking rate and ALSFRS-R scores from a video recording of a short sentence is possible. A task of reading/repeating a short sentence is recommended in an evaluation protocol for those with cognitive impairments, respiratory diseases, or limited reading ability who might not be able to successfully complete a more length reading task (Yunusova et al., 2016) and for remote assessment, as automatic analysis might be easier to perform in short recordings (Neumann et al., 2021). Overall, a simple and inexpensive setup consisting of a 3D camera and machine learning algorithms allowed us to make meaningful observations about the neuromuscular changes due to bulbar disease in ALS. This type of consumer-grade technology has a high chance of being adopted by clinicians; it requires a minimum initial investment and is easy to use. Video data could be acquired passively as part of other examinations without altering the current workflow for clinical evaluation. The next step of this work is to establish how to use consumer-grade 2D video cameras and machine learning methods to achieve the same results.

Data Availability Statement

Data from 11 participants with amyotrophic lateral sclerosis (out of 15) and from 11 healthy controls (out of 11) are available for research purposes as part of the Toronto NeuroFace data set (Bandini et al., 2020). More information can be found at: https://slp.utoronto.ca/faculty/yana-yunusova/speech-production-lab/datasets/.

Acknowledgments

References

Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics, 34*(4), 555–596. https://doi.org/10.1162/coli.07-034-R2

Ball, L. J., Willis, A., Beukelman, D. R., & Pattee, G. L. (2001). A protocol for identification of early bulbar signs in amyotrophic lateral sclerosis. *Journal of the Neurological Sciences, 191*(1–2), 43–53. https://doi.org/10.1016/S0022-510X(01)00623-2

Bandini, A., Green, J. R., Taati, B., Orlandi, S., Zinman, L., & Yunusova, Y. (2018). Automatic detection of amyotrophic lateral sclerosis (ALS) from video-based analysis of facial movements: Speech and non-speech tasks. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (pp. 150–157). https://doi.org/10.1109/FG.2018.00031

Bandini, A., Green, J. R., Wang, J., Campbell, T. F., Zinman, L., & Yunusova, Y. (2018). Kinematic features of jaw and lips distinguish symptomatic from presymptomatic stages of bulbar decline in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 61*(5), 1118–1129. https://doi.org/10.1044/2018_JSLHR-S-17-0262

Bandini, A., Green, J. R., Zinman, L., & Yunusova, Y. (2017). Classification of bulbar ALS from kinematic features of the jaw and lips: Towards computer-mediated assessment. *Interspeech, 2017,* 1819–1823. https://doi.org/10.21437/Interspeech.2017-478

Bandini, A., Rezaei, S., Guarin, D., Kulkarni, M., Lim, D., Boulos, M. I., Zinman, L., Yunusova, Y., & Taati, B. (2020). A new dataset for facial motion analysis in individuals with neurological disorders. *IEEE Journal of Biomedical and Health Informatics, 25*(4), 1111–1119. https://doi.org/10.1109/JBHI.2020.3019242

Barnett, C., Green, J. R., Marzouqah, R., Stipancic, K. L., Berry, J. D., Korngut, L., Genge, A., Shoesmith, C., Briemberg, H., Abrahao, A., Kalra, S., Zinman, L., & Yunusova, Y. (2020). Reliability and validity of speech & pause measures during passage reading in ALS. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 21*(1–2), 42–50. https://doi.org/10.1080/21678421.2019.1697888

Carding, P. N., Steen, I. N., Webb, A., MacKenzie, K., Deary, I. J., & Wilson, J. A. (2004). The reliability and sensitivity to change of acoustic measures of voice quality. *Clinical Otolaryngology and Allied Sciences, 29*(5), 538–544. https://doi.org/10.1111/j.1365-2273.2004.00846.x

Caverlé, M. W. J., & Vogel, A. P. (2020). Stability, reliability, and sensitivity of acoustic measures of vowel space: A comparison of vowel space area, formant centralization ratio, and vowel articulation index. *The Journal of the Acoustical Society of America, 148*(3), 1436–1444. https://doi.org/10.1121/10.0001931

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., & Nakanishi, A. (1999). The ALSFRS-R: A revised ALS Functional Rating Scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *Journal of the Neurological Sciences, 169*(1–2), 13–21. https://doi.org/10.1016/s0022-510x(99)00210-5

Chiò, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., Traynor, B. G., & Eurals Consortium. (2009). Prognostic factors in ALS: A critical review. *Amyotrophic Lateral Sclerosis, 10*(5–6), 310–323. https://doi.org/10.3109/17482960802566824

Cudkowicz, M. E., van den Berg, L. H., Shefner, J. M., Mitsumoto, H., Mora, J. S., Ludolph, A., Hardiman, O., Bozik, M. E., Ingersoll, E. W., Archibald, D., Meyers, A. L., Dong, Y., Farwell, W. R., Kerr, D. A., & EMPOWER Investigators. (2013). Dexpramipexole versus placebo for patients with amyotrophic lateral sclerosis (EMPOWER): A randomised, double-blind, Phase 3 trial. *The Lancet. Neurology, 12*(11), 1059–1067. https://doi.org/10.1016/S1474-4422(13)70221-7

Deliyski, D. D., Shaw, H. S., & Evans, M. K. (2005). Adverse effects of environmental noise on acoustic voice quality measurements. *Journal of Voice, 19*(1), 15–28. https://doi.org/10.1016/j.jvoice.2004.07.003

Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology, 77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Goldsack, J. C., Coravos, A., Bakker, J. P., Bent, B., Dowling, A. V., Fitzer-Attas, C., Godfrey, A., Godino, J. G., Gujar, N., Izmailova, E., Manta, C., Peterson, B., Vandendriessche, B., Wood, W. A., Wang, K. W., & Dunn, J. (2020). Verification, analytical validation, and clinical validation (V3): The foundation of determining fit-for-purpose for biometric monitoring technologies (BioMeTs). *npj Digital Medicine, 3*(1), 1–15. https://doi.org/10.1038/s41746-020-0260-4

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. https://doi.org/10.3109/21678421.2013.817585

Guarin, D. L., Dempster, A., Bandini, A., Yunusova, Y., & Taati, B. (2020). Estimation of orofacial kinematics in Parkinson's disease: Comparison of 2D and 3D markerless systems for motion tracking. 540–543. https://doi.org/10.1109/FG47880.2020.00112

Guarin, D. L., Dusseldorp, J., Hadlock, T. A., & Jowett, N. (2018). A machine learning approach for automated facial measurements in facial palsy. *JAMA Facial Plastic Surgery, 20*(4), 335–337. https://doi.org/10.1001/jamafacial.2018.0030

Hardiman, O. (2011). Management of respiratory symptoms in ALS. *Journal of Neurology, 258*(3), 359–365. https://doi.org/10.1007/s00415-010-5830-y

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hirose, H., Kiritani, S., & Sawashima, M. (1982). Patterns of dysarthric movement in patients with amyotrophic lateral sclerosis and pseudobulbar palsy. *Folia Phoniatrica et Logopaedica, 34*(2), 106–112. https://doi.org/10.1159/000265636

Ilse, B., Prell, T., Walther, M., Hartung, V., Penzlin, S., Tietz, F., Witte, O.-W., Strauss, B., & Grosskreutz, J. (2015). Relationships between disease severity, social support and health-related quality of life in patients with amyotrophic lateral sclerosis. *Social Indicators Research, 120*(3), 871–882. https://doi.org/10.1007/s11205-014-0621-y

Jin, X., & Tan, X. (2017). Face alignment in-the-wild: A survey. *Computer Vision and Image Understanding, 162,* 1–22. https://doi.org/10.1016/j.cviu.2017.08.008

Kuruvilla-Dugdale, M., & Mefferd, A. (2017). Spatiotemporal movement variability in ALS: Speaking rate effects on tongue, lower lip, and jaw motor control. *Journal of Communication Disorders, 67,* 22–34. https://doi.org/10.1016/j.jcomdis.2017.05.002

Lee, J., Bell, M., & Simmons, Z. (2018). Articulatory kinematic characteristics across the dysarthria severity spectrum in individuals with amyotrophic lateral sclerosis. *American Journal of Speech-Language Pathology, 27*(1), 258–269. https://doi.org/10.1044/2017_AJSLP-16-0230

Lee, J., Rodriguez, E., & Mefferd, A. (2020). Direction-specific jaw dysfunction and its impact on tongue movement in individuals with dysarthria secondary to amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 63*(2), 499–508. https://doi.org/10.1044/2019_JSLHR-19-00174

Lopes, L. W., Batista Simões, L., Delfino da Silva, J., da Silva Evangelista, D., da Nóbrega E Ugulino, A. C., Oliveira Costa

Silva, P., & Jefferson Dias Vieira, V. (2017). Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses. *Journal of Voice, 31*(3), 382.e15–382.e26. https://doi.org/10.1016/j.jvoice.2016.08.015

Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America, 126*(5), 2619–2634. https://doi.org/10.1121/1.3224706

Mefferd, A. S., Green, J. R., & Pattee, G. (2012). A novel fixed-target task to determine articulatory speed constraints in persons with amyotrophic lateral sclerosis. *Journal of Communication Disorders, 45*(1), 35–45. https://doi.org/10.1016/j.jcomdis.2011.09.002

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface, 14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montréal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Neumann, M., Roesler, O., Liscombe, J., Kothare, H., Suendermann-Oeft, D., Pautler, D., Navar, I., Anvar, A., Kumm, J., Norel, R., Fraenkel, E., Sherman, A. V., Berry, J. D., Pattee, G. L., Wang, J., Green, J. R., & Ramanarayanan, V. (2021). *Investigating the utility of multimodal conversational technology and audiovisual analytic measures for the assessment and monitoring of amyotrophic lateral sclerosis at scale*. ArXiv. http://arxiv.org/abs/2104.07310

Orozco-Arroyave, J. R., Vásquez-Correa, J. C., Klumpp, P., Pérez-Toro, P. A., Escobar-Grisales, D., Roth, N., Ríos-Urrego, C. D., Strauss, M., Carvajal-Castaño, H. A., Bayerl, S., Castrillón-Osorio, L. R., Arias-Vergara, T., Künderle, A., López-Pabón, F. O., Parra-Gallego, L. F., Eskofier, B., Gómez-Gómez, L. F., Schuster, M., & Nöth, E. (2020). Apkinson: The smartphone application for telemonitoring Parkinson's patients through speech, gait and hands movement. *Neurodegenerative Disease Management, 10*(3), 137–157. https://doi.org/10.2217/nmt-2019-0037

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*(85), 2825–2830.

Perry, B. J., Martino, R., Yunusova, Y., Plowman, E. K., & Green, J. R. (2018). Lingual and jaw kinematic abnormalities precede speech and swallowing impairments in ALS. *Dysphagia, 33*(6), 840–847. https://doi.org/10.1007/s00455-018-9909-4

Plowman, E. K., Tabor, L. C., Wymer, J., & Pattee, G. (2017). The evaluation of bulbar dysfunction in amyotrophic lateral sclerosis: Survey of clinical practice patterns in the United States. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 18*(5–6), 351–357. https://doi.org/10.1080/21678421.2017.1313868

Rong, P., & Green, J. R. (2019). Predicting speech intelligibility based on spatial tongue–jaw coupling in persons with amyotrophic lateral sclerosis: The impact of tongue weakness and jaw adaptation. *Journal of Speech, Language, and Hearing Research, 62*(8S), 3085–3103. https://doi.org/10.1044/2018_JSLHR-S-CSMC7-18-0116

Rong, P., Yunusova, Y., Wang, J., & Green, J. R. (2015). Predicting early bulbar decline in amyotrophic lateral sclerosis: A speech subsystem approach. *Behavioural Neurology, 2015,* e183027. https://doi.org/10.1155/2015/183027

Rong, P., Yunusova, Y., Wang, J., Zinman, L., Pattee, G. L., Berry, J. D., Perry, B., & Green, J. R. (2016). Predicting speech intelligibility decline in amyotrophic lateral sclerosis based on the deterioration of individual speech subsystems. *PLOS ONE, 11*(5), Article e0154971. https://doi.org/10.1371/journal.pone.0154971

Rowe, H. P., Gutz, S. E., Maffei, M. F., & Green, J. R. (2020). Acoustic-based articulatory phenotypes of amyotrophic lateral sclerosis and Parkinson's disease: Towards an interpretable, hypothesis-driven framework of motor control. *Interspeech, 2020,* 4816–4820. https://doi.org/10.21437/Interspeech.2020-1459

Rutkove, S. B. (2015). Clinical measures of disease progression in amyotrophic lateral sclerosis. *Neurotherapeutics, 12*(2), 384–393. https://doi.org/10.1007/s13311-014-0331-9

Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2013). 300 Faces in-the-Wild Challenge: The first facial landmark localization challenge. In *2013 IEEE International Conference on Computer Vision Workshops* (pp. 397–403). https://doi.org/10.1109/ICCVW.2013.59

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Shellikeri, S., Green, J. R., Kulkarni, M., Rong, P., Martino, R., Zinman, L., & Yunusova, Y. (2016). Speech movement measures as markers of bulbar disease in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 59*(5), 887–899. https://doi.org/10.1044/2016_JSLHR-S-15-0238

Singh, S., & Xu, W. (2020). Robust detection of Parkinson's disease using harvested smartphone voice data: A telemedicine approach. *Telemedicine Journal and E-Health, 26*(3), 327–334. https://doi.org/10.1089/tmj.2018.0271

Stegmann, G. M., Hahn, S., Liss, J., Shefner, J., Rutkove, S., Shelton, K., Duncan, C. J., & Berisha, V. (2020). Early detection and tracking of bulbar changes in ALS via frequent and remote speech analysis. *npj Digital Medicine, 3*(1), 132–135. https://doi.org/10.1038/s41746-020-00335-x

Stipancic, K. L., Yunusova, Y., Berry, J. D., & Green, J. R. (2018). Minimally detectable change and minimal clinically important difference of a decline in sentence intelligibility and speaking rate for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 61*(11), 2757–2771. https://doi.org/10.1044/2018_JSLHR-S-17-0366

Strong, M. J., Grace, G. M., Orange, J. B., & Leeper, H. A. (1996). Cognition, language, and speech in amyotrophic lateral sclerosis: A review. *Journal of Clinical and Experimental Neuropsychology, 18*(2), 291–303. https://doi.org/10.1080/01688639608408283

Sun, Y., Ng, M. L., Lian, C., Wang, L., Yang, F., & Yan, N. (2018). Acoustic and kinematic examination of dysarthria in Cantonese patients of Parkinson's disease. *2018 11th International Symposium on Chinese Spoken Language Processing (ISCSLP)* (pp. 354–358). https://doi.org/10.1109/ISCSLP.2018.8706615

Walsh, B., & Smith, A. (2012). Basic parameters of articulatory movements and acoustics in individuals with Parkinson's disease. *Movement Disorders, 27*(7), 843–850. https://doi.org/10.1002/mds.24888

Wang, J., Kothalkar, P. V., Kim, M., Bandini, A., Cao, B., Yunusova, Y., Campbell, T. F., Heitzman, D., & Green, J. R. (2018). Automatic prediction of intelligible speaking rate for individuals with ALS from speech acoustic and articulatory

samples. *International Journal of Speech-Language Pathology, 20*(6), 669–679. https://doi.org/10.1080/17549507.2018.1508499

Wang, J., Kothalkar, P. V., Kim, M., Yunusova, Y., Campbell, T. F., Heitzman, D., & Green, J. R. (2016). Predicting intelligible speaking rate in individuals with amyotrophic lateral sclerosis from a small number of speech acoustic and articulatory samples. *Workshop on Speech and Language Processing for Assistive Technologies, 2016,* 91–97. https://doi.org/10.21437/SLPAT.2016-16

Weerathunge, H. R., Segina, R. K., Tracy, L., & Stepp, C. E. (2021). Accuracy of acoustic measures of voice via telepractice videoconferencing platforms. *Journal of Speech, Language, and Hearing Research, 64*(7), 2586–2599. https://doi.org/10.1044/2021_JSLHR-20-00625

Yorkston, K. M., Beukelman, D. R., Hakel, M., & Dorsey, M. (2007). *Speech intelligibility test for windows*. Institute for Rehabilitation Science and Engineering at Madonna Rehabilitation Hospital.

Yorkston, K. M., Miller, R. M., & Strand, E. A. (1995). *Management of speech and swallowing in degenerative diseases*. Pro-Ed.

Yunusova, Y., Graham, N. L., Shellikeri, S., Phuong, K., Kulkarni, M., Rochon, E., Tang-Wai, D. F., Chow, T. W., Black, S. E., Zinman, L. H., & Green, J. R. (2016). Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). *PLOS ONE, 11*(1), Article e0147573. https://doi.org/10.1371/journal.pone.0147573

Yunusova, Y., Green, J. R., Greenwood, L., Wang, J., Pattee, G. L., & Zinman, L. (2012). Tongue movements and their acoustic consequences in amyotrophic lateral sclerosis. *Folia Phoniatrica et Logopaedica, 64*(2), 94–102. https://doi.org/10.1159/000336890

Yunusova, Y., Green, J. R., Lindstrom, M., Ball, L., Pattee, G., & Zinman, L. (2010). Kinematics of disease progression in bulbar ALS. *Journal of Communication Disorders, 43*(1), 6–20. https://doi.org/10.1016/j.jcomdis.2009.07.003

Yunusova, Y., Plowman, E. K., Green, J. R., Barnett, C., & Bede, P. (2019). Clinical measures of bulbar dysfunction in ALS. *Frontiers in Neurology, 106*. https://doi.org/10.3389/fneur.2019.00106

Yunusova, Y., Weismer, G., Westbury, J. R., & Lindstrom, M. J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research, 51*(3), 596–611. https://doi.org/10.1044/1092-4388(2008/043)