INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Classification of bulbar ALS from kinematic features of the jaw and lips: Towards computer-mediated assessment

Andrea Bandini[1], Jordan R. Green[2], Lorne Zinman[3,4], Yana Yunusova[1,4,5]

[1]University Health Network: Toronto Rehabilitation Institute, Toronto, Canada
[2]MGH Institute of Health Professions, Boston, USA
[3]Neurology, Sunnybrook Health Sciences Centre, Toronto, Canada
[4]Brain Sciences, Sunnybrook Research Institute, Toronto, Canada
[5]Department of Speech-Language Pathology, University of Toronto, Canada

andrea.bandini@uhn.ca, jgreen2@mghihp.edu, Lorne.Zinman@sunnybrook.ca, yana.yunusova@utoronto.ca

Abstract

Recent studies demonstrated that lip and jaw movements during speech may provide important information for the diagnosis of amyotrophic lateral sclerosis (ALS) and for understanding its progression. A thorough investigation of these movements is essential for the development of intelligent video- or optically-based facial tracking systems that could assist with early diagnosis and progress monitoring. In this paper, we investigated the potential for a novel and expanded set of kinematic features obtained from lips and jaw to classify articulatory data into three stages of bulbar disease progression (i.e., pre-symptomatic, early symptomatic, and late symptomatic). Feature selection methods (Relief-F and mRMR) and classification algorithm (SVM) were used for this purpose. Results showed that even with a limited number of kinematic features it was possible to obtain good classification accuracy (nearly 80%). Given the recent development of video-based markerless methods for tracking speech movements, these results provide strong rationale for supporting the development of portable and cheap systems for monitoring the orofacial function in ALS.

Index Terms: amyotrophic lateral sclerosis, facial kinematics, lips, jaw, classification, feature selection

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that affects the upper and lower motor neurons in the brain, brainstem and spinal cord [1]. The progressive paralysis of the head and neck musculature during the disease, known as bulbar ALS, is associated with a more debilitating outcome and shorter survival than spinal (limb) signs and symptoms [2]. Bulbar dysfunction results in a considerable decline in communication abilities, with a negative impact on speech intelligibility and overall quality of life. Nearly all patients begin to experience bulbar signs and symptoms at some point during the course of the disease.

Currently, the diagnosis of ALS is based on patient report of symptoms and history, with high rate of misdiagnosis especially among patients with bulbar-onset [3]. Early detection of bulbar changes - particularly for those with spinal onset ALS and gradual bulbar changes - remains an important clinical challenge, since it may help predict the disease course and plan communication management with supportive interventions such as augmentative and alternative communication methods (AAC) [4-6].

The gold standard for the clinical evaluation of bulbar function in ALS is the measure of speaking rate (SR, defined as number of words per minute - WPM). SR has been reported to change earlier than speech intelligibility (SI) during the disease course and showed a linear decline over time [7-10]. Values of SR higher than 160 WPM during connected sentence reading have been considered normal or pre-symptomatic, values within the range of 120-160 WPM were associated with early changes of SI and values lower than 120 WPM denoted an abrupt decline of SI [7, 10-12]. Recent research, however, indicated that instrumental physiological measures of bulbar function may be even more sensitive to early changes in the bulbar ALS than SR [6, 10].

During the last two decades, kinematic analysis of speech movements has become an important means of bulbar ALS detection and understanding its progression. Several early studies [2, 13-15] investigated the link between orofacial movement changes and bulbar function decline, demonstrating that important information may be extracted by studying tongue, lip and jaw movements. One of the most evident features of disease progression in ALS is the increase in jaw and lip movement duration [2, 15]. Patients with ALS may also show reduced movement velocity, when compared to normal speakers, and increased jaw displacement possibly associated with a compensatory response to the decline in tongue function [12]. Our most recent studies indicated that among the speech subsystems (respiratory, phonatory, articulatory and resonatory), articulatory subsystem represented by movements of the jaw and lips, showed the highest association with the overall bulbar decline (measured by SR and SI) [6, 16]. These initial analyses suggested that a closer look at the jaw and lip movements might be diagnostically promising with a particular focus on the expansion of kinematic measures and comprehensive consideration of the semantics of the selected measures.

Following the results reported in [6], we explored the potential of kinematic measures/features extracted from only facial articulators (lips and jaw) to classify data from patients with ALS into three classes: pre-symptomatic (SR $\geq$ 160 WPM), early symptomatic (120 WPM $\leq$ SR < 160 WPM) and late symptomatic (SR < 120 WPM) stages of bulbar decline. We also proposed an expanded set of kinematic features. In addition to the well-studied measures of duration, displacement and velocity of jaw/lips, we considered other

http://dx.doi.org/10.21437/Interspeech.2017-478

kinematic aspects such as acceleration (i.e. measure of neuromuscular control [17, 18]), jerk (i.e., measure of movement smoothness [19]) and surface geometry (e.g., measure of mouth symmetry). All of these measures have been associated in the literature with the presence of aberrant speech and limb motor control [17, 19]. In this paper, the classification power of the proposed feature set was tested by using multiclass Support Vector Machine (SVM). The most discriminative features as well as their response to the bulbar function decline were identified with feature selection algorithms. We hypothesize that class membership will be predicted by kinematic measures of lips and jaw with high accuracy.

2. Data collection

2.1. Participants and class definition

57 patients with ALS (31 male, 26 female) were selected from a larger sample (N=145) of patients that participated in a longitudinal study of bulbar dysfunction in ALS. All participants were diagnosed with possible or definite ALS as defined by the El Escorial Criteria for the World Federation of Neurology [20]. All participants were native speakers of English, at screening demonstrated normal hearing and showed no evidence of cognitive impairment as determined by a score of 26 or higher on the Montreal Cognitive Assessment [21]. 46 patients presented with spinal symptoms (e.g., limbs, trunk) at onset, whereas 11 patients reported bulbar symptoms at onset. Participants were recorded in multiple sessions (3.5 ± 2.1), spaced typically at 3 months. The average age of participants at session 1 was 58.2 ± 9.0 years. The ALS Functional Rating Scale - Revised (ALSFRS-R) questionnaire [22] was used to assess the impact of ALS on daily functions, and its bulbar subscore was calculated based on speech, swallowing and salivation symptom report. Bulbar clinical assessment was also performed through the measures of SI and SR, calculated using sentence intelligibility test (SIT) [23]. A total of 117 sessions were considered for this study. Sessions were divided into three balanced classes, according to [7, 10-12]: sessions with SR ≥ 160 WPM (class1, pre-symptomatic, 39 sessions), sessions with 120 WPM ≤ SR < 160 WPM (class2, early symptomatic, 39 sessions) and sessions with SR < 120 WPM (class3, late symptomatic, 39 sessions). 20 patients had data in class1 only, 13 in class2 only, 12 in class3 only, 4 in class1 and class2, and 8 in class2 and class3. Clinical information for the three groups is reported in Table 1. The study was approved by the Research Ethics Board at Sunnybrook Research Institute. All participants signed informed consent according to the requirements of the Declaration of Helsinki.

Table 1: *Clinical information of the three classes*

		SR (WPM)	SI (%)	ALSFRS-R	Bulbar subscore
class1	M	193.6	98.9	35.6	11.6
	SD	20.2	1.8	7.5	1.1
class2	M	140.8	96.5	34.9	9.5
	SD	10.0	4.3	5.0	1.7
class3	M	95.87	86.8	33.9	8.7
	SD	19.5	20.8	7.9	1.6

2.2. Speech task and data acquisition

Participants were instructed to repeat the sentence "Buy Bobby a puppy" 10 times at their normal comfortable speaking rate and loudness, briefly pausing between each repetition. Lip and jaw movements were recorded with a 3D optoelectronic motion capture system (Optotrak Certus, Northern Digital Inc., Waterloo, ON, Canada). Eleven active infrared markers (diameter 7 mm) were attached on the face of the participants, following the layout used in [14]. For this study, the following markers were considered: upper lip (UL), lower lip (LL), right and left mouth corners (RC, LC), and jaw right (JR). Trajectories of the markers were acquired with a sampling frequency (F_s) of 125 Hz. The position of each orofacial sensor was expressed relative to that of the 6 degrees of freedom (DoF) head marker. The origin of the 3D coordinate systems was defined anatomically as the tip of the midline between the incisors. This procedure was performed using a 4-marker digitizing probe, and the coordinate system transformation relative to the anatomic origin was completed during post-processing. Synchronized acoustic signals (F_s = 25 kHz, 16-bit resolution) were recorded with a high-quality earset microphone (Countryman E6, Countryman Associates, Inc., Menlo Park, CA, USA), which was positioned approximately 5 cm from the mouth.

3. Methods

3.1. Feature extraction

3.1.1. Preprocessing

Marker trajectories were low-pass filtered (8-pole Butterworth, cut-off frequency 15 Hz) in order to remove high frequency noise [24]. Sentence repetitions within each audio recording were manually labeled with PRAAT [25] - temporal boundaries (start and end points) of each sentence were used to segment the kinematic data.

3.1.2. Kinematic features

For each sentence, 36 features were computed. The feature set included:

- Ten range of movement measures – Lips opening and width were calculated as the Euclidean distance between UL and LL, and RC and LC, respectively; maximum and average values were computed (Open_max, Open_avg, Width_max, Width_avg, measured in mm). LL and JR displacements (LL_path, JR_path) were calculated as the module of the vector between the origin and LL and JR, respectively; maximum, average and cumulative sum were computed (LL_path_max, LL_path_avg, LL_path_cml, JR_path_max, JR_path_avg, JR_path_cml, measured in mm).
- Six velocity measures – The velocity of LL and JR was calculated as the first derivative of LL_path and JR_path with time, respectively; maximum, minimum and average values were considered (vLL_max, vLL_min, vLL_avg, vJR_max, vJR_min, vJR_avg, measured in mm/s).
- Six acceleration measures - The acceleration of LL and JR was calculated as the second derivative of LL_path and JR_path with time, respectively; maximum, minimum and average values were considered (aLL_max, aLL_min, aLL_avg, aJR_max, aJR_min, aJR_avg, measured in mm/s^2).

- Six jerk measures – Jerk index of LL and JR was calculated as the third derivative of LL_path and JR_path with time, respectively; maximum, minimum and average values were considered (jLL_max, jLL_min, jLL_avg, jJR_max, jJR_min, jJR_avg, measured in mm/s^3).
- Seven surface measures – Two triangular surfaces were computed as the area of the triangles with vertices RC, UL, LL (S_R) and LC, UL, LL (S_L), then the total mouth surface (S) was calculated as the sum of S_R and S_L; maximum and average values (S_R_max, S_R_avg, S_L_max, S_L_avg, S_max, S_avg, measured in mm^2) as well as mean symmetry ratio between right and left parts (S_ratio_avg) was extracted.
- One duration measure – sentence duration (Sentence_dur) measured in seconds (s).

The above features were averaged across repetitions and a single feature vector was obtained for each session. The features were then expressed as Z-scores in order to standardize variables with different units of measurement.

3.2. Feature selection

Two feature selection algorithms were implemented in order to rank the features according to their discriminative power: the Relief-F algorithm [26] and the minimal-redundancy-maximal-Relevance (mRMR) criterion [27].

Relief-F [26] is an extension of the RELIEF algorithm for feature selection [28]. RELIEF evaluates the predictive power of features based on their ability to discriminate instances that are near each other. The evaluation is implemented by examining the two nearest neighbors of each instance - one from the same class (nearest hit) and one from the other class (nearest miss). If a feature has a high discriminative power, it is expected to have the same value for instances of the same class while distinguishing instances from different classes. RELIEF was extended in [26] to work with noisy/incomplete data and multi-class problems. In particular, Relief-F extends its search to *k* nearest hits and misses.

The mRMR criterion requires that the features fulfill two constraints: the maximal relevance and the minimum redundancy [27]. Through the first constraint, features with the highest relevance to target variable (Y) are selected, with relevance being estimated through the mutual information between the features and Y. The minimal redundancy constraint is employed to exclude features that have high dependency on each other, minimizing the cost function that depends on the mutual information among features.

Feature rankings were obtained by means of a 10-fold cross-validation (CV). As the feature ranking may change at each iteration of the CV, the final feature ranking was determined by averaging the ranks obtained during 10 iterations. These analyses were implemented using the Feature Selection Library (FSLib 2016) [29, 30], a free toolbox for Matlab that includes the aforementioned methods, among others.

3.3. Classification

A multi-class SVM was chosen to classify data into the three classes defined in sect. 2.1. SVM seeks to build the optimal hyperplane that divides the feature space into two classes (i.e., the hyperplane that maximize the margin between classes). SVMs are widely used in literature, including speech literature [31-33], particularly in cases when data are not linearly separable. In this case, features can be mapped to a higher-dimensional space where they are linearly separable by means of kernel functions, such as polynomial or Gaussian functions. A Gaussian radial basis function kernel was chosen for this analysis. The multi-class problem was addressed through a one-vs-one strategy: a binary classifier was trained for each different pair of class labels, for a total of 3 binary SVMs. The classification accuracy was assessed with a 10-fold CV repeated 10 times.

3.4. Statistical analysis

A Kruskal-Wallis test was performed to determine whether the kinematic variables selected by the feature selection algorithms changed across the three stages of bulbar function decline. Post-hoc Wilcoxon rank-sum tests were performed in order to identify statistically significant differences between the classes. The statistical analysis and the classification were performed in Matlab, v. 2016b (Statistics and Machine Learning Toolbox).

4. Results

Using the whole feature set, the overall achieved accuracy of the multiclass SVM was 74.5 ± 2.4 %. To test if the accuracy would increase with a lower number of features, we computed the classification accuracy with feature sets of increasing size (from 1 to 36) and chose the number of features to retain as a trade-off between the classification accuracy and the dimension of the feature set, following the principle of parsimony [31]. Figure 1 shows that the accuracy of 75.9 % was obtained with 16 features, and with only 5 features the accuracy is ~73 %. With 21 features, a global maximum of 78.6 % was achieved (Figure 1). These results were produced with the feature ranking provided by the Relief-F algorithm. Other measures of the classification performance (i.e., precision, recall and F-measure) using the feature set provided by Relief-F, are reported in Table 2.

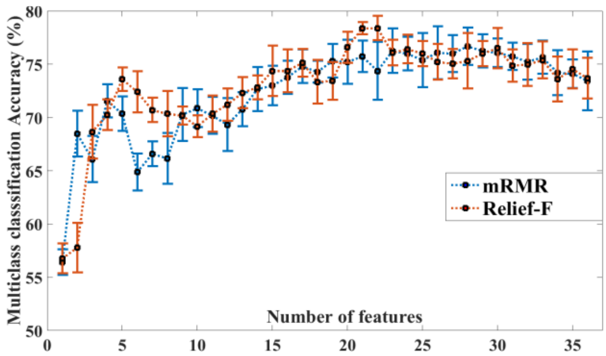

Figure 1: *Classification accuracy of the multi-class SVM with increasing dimension of the feature set.*

Table 2: *Precision (TP/TP+FP), recall (TP/TP+FN) and F-Measure (harmonic mean of precision and recall) obtained with the feature set provided by Relief-F.*

	class1	class2	class3
Precision	0.69	0.76	0.98
Recall	0.91	0.69	0.75
F-Measure	0.79	0.72	0.85

The first 21 features ranked by Relief-F, listed in decreasing relevance, were: Sentence_dur, JRpath_cml, aJR_min,

Table 3: *Statistically significant features among groups. Post-hoc Wilcoxon rank-sum tests were performed in order to compare features of class 2 and 3 with features from class 1 (* p < .05, ** p < .01, *** p < .001).*

	class1	class2	class3	Kruskal-Wallis test
LLpath_cml (mm)	3682.34±819.41	4441.92±1188.70**	5846.57±1977.53***	$H(2)=30.67$, $p<.001$
JRpath_cml (mm)	10085.63±1400.27	12247.01±2307.84***	16824.41±5210.84***	$H(2)=52.19$, $p<.001$
vLL_min (mm/s)	-168.37±63.33	-139.74±61.31*	-115.03±41.70***	$H(2)=18.57$, $p<.001$
vJR_max (mm/s)	148.82±48.92	120.88±50.81**	114.04±35.12***	$H(2)=13.08$, $p=.001$
vJR_min (mm/s)	-148.39±48.84	-120.64±51.21**	-102.62±31.89***	$H(2)=19.18$, $p<.001$
aLL_max (mm/s^2)	5468.48±2025.47	4134.08±1820.99***	3404.94±1330.59***	$H(2)=25.55$, $p<.001$
aLL_min (mm/s^2)	-6178.79±2266.87	-4354.11±2588.85***	-3269.82±1879.63***	$H(2)=40.29$, $p<.001$
aJR_max (mm/s^2)	3878.60±1324.46	2800.83±1306.47***	2325.28±1106.21***	$H(2)=30.43$, $p<.001$
aJR_min (mm/s^2)	-5189.68±1885.78	-3422.17±2053.22***	-2632.3±1500.94***	$H(2)=41.53$, $p<.001$
jLL_max (mm/s^3)	367415.05±140919.66	256109.97±157083.76***	180636.21±92832.32***	$H(2)=40.68$, $p<.001$
jLL_min (mm/s^3)	-272736.33±100713.84	-213474.79±125506.61***	-187298.63±92772.20***	$H(2)=20.88$, $p<.001$
jJR_max (mm/s^3)	241971.07±177143.62	149836.42±100881.88***	135707.65±132811.80***	$H(2)=29.59$, $p<.001$
jJR_min (mm/s^3)	-199462.39±93598.57	-138057.89±100417.50***	-140985.27±115257.89***	$H(2)=24.32$, $p<.001$
Sentence_dur(s)	1.56±0.19	1.85±0.30***	2.59±0.71***	$H(2)=56.17$, $p<.001$

aLL_min, LLpath_cml, jLL_max, aJR_max, aLL_max, jLL_min, vJR_min, jJR_min, jJR_max, vLL_min, vJR_max, jJR_avg, Width_avg, aJR_avg, jLL_avg, Width_max, vJR_avg, and aLL_avg. Results of the Kruskal-Wallis test indicated that 14 out of 21 selected features reported statistically significant differences across the three groups (Table 3). In particular, range of movement measures (LLpath_cml and JRpat_cml) and Sentence_dur increased with the decline of bulbar function, whereas the other measures (peaks of velocity, acceleration and jerk) showed a decrease across the three classes.

5. Discussion and conclusion

This study reported results of classification of kinematic data from jaw and lips during speech into different stages of bulbar function decline in ALS. The main finding was that with a selected number of clinically interpretable kinematic features it was possible to obtain relatively good classification accuracy (nearly 80%), discriminating between pre-symptomatic, early symptomatic, and late symptomatic stages of bulbar decline. It seems, however, that the classifier was better able to distinguish between late symptomatic and pre- and early symptomatic groups than between the latter two groups (Table 2), suggesting that larger kinematic changes occurred when SR became lower than 120 WPM.

Since the features were clinically interpretable, we were able to comprehensively describe the facial kinematic during bulbar function decline. Consistent with previous reports, maximum velocity of lips and jaw decreased with disease progression, whereas sentence duration and cumulative paths travelled by lips and jaw markers increased. The reported changes in acceleration and jerk were novel but consistent with the observed velocity results, suggesting that movements of lips and jaw became slower (decrease of velocity and acceleration) and smoother (decrease of jerk) with the decline in bulbar function. The increase of the total distance travelled by lips and jaw was consistent with previous findings [12], linking this behaviour to a mechanism for compensating for the decline in tongue function. This behaviour can also be due to the longer sentence duration.

A number of proposed measures (acceleration and jerk above all) changed considerably from early symptomatic to late symptomatic stages of bulbar decline, suggesting that these measures can be used to provide useful information about disease progression through jaw and lip musculature. Surface features were not retained for classification, and did not show significant differences. Thus, no changes in the mouth geometry were visible during speech movements.

To the best of our knowledge, this is the first study where measures of lip and jaw movements were used to classify different stages of bulbar severity in ALS. Wang et al., 2016 [32] used Support Vector Regression (SVR) to predict the intelligible speaking rate from acoustic and articulatory data with a smaller set of patients. The authors demonstrated that adding articulatory measures of lips and tongue to the acoustic feature set extracted from sentences led to an improvement of the prediction performance, confirming that important indicators of bulbar function decline can be identified in the articulatory subsystem [32].

Further work will include a larger dataset, in order to replicate current results. Given the recent developments of markerless methods for tracking orofacial movements during speech [17, 34], our results provide the rationale for implementation of facial tracking technologies to the assessment of articulatory speech function. The use of facial features for assessing motor speech disorders is highly feasible and, based on our data, clearly indicated in the case of bulbar dysfunction in ALS. The use of objective kinematic features in conjunction with machine learning techniques may be helpful for the development of automatic systems that support clinicians in the detection of early signs of bulbar disease and tracking disease course with contactless methods.

6. Acknowledgements

The authors would like to acknowledge the funding support from NIH grant R01DC013547 and from the University Health Network – Toronto Rehabilitation Institute. We are thankful to the participants and their families for taking part in this project. We also thank Madhura Kulkarni for her assistance with this project.

7. References

[1] M. C. Kiernan, S. Vucic, B. C. Cheah, M. R. Turner, A. Eisen, O. Hardiman, J. R. Burrell, and M. C. Zoing, "Amyotrophic lateral sclerosis," *The Lancet*, vol. 377, pp. 942–955, 2011.

[2] S. E. Langmore and M. Lehman, "The orofacial deficit and dysarthria in ALS," *Journal of Speech and Hearing Research*, vol. 37, pp. 28–37, 1994.

[3] M. R. Turner, J. Scaber, J. A. Goodfellow, M. E. Lord, R. Marsden, K. Talbot, "The diagnostic pathway and prognosis in bulbar-onset amyotrophic lateral sclerosis," Journal of the Neurological Sciences, vol. 294, no. 1, pp. 81-85, 2010.

[4] A. Chen, C. G. Garrett, "Otolaryngologic presentations of amyotrophic lateral sclerosis," *Otoloaryngology—Head and Neck Surgery*, vol 123, no. 3, pp. 500-504, 2005.

[5] E. K. Hanson, K. M.Yorkston, D. Britton, "Dysarthria in amyotrophic lateral sclerosis: a systematic review of characteristics, speech treatment, and augmentative and alternative options," *Journal of Medical Speech-Language Pathology*, vol. 19, no. 3, pp. 12-30, 2011.

[6] P. Rong, Y. Yunusova, J. Wang, J. R. Green, "Predicting early bulbar decline in amyotrophic lateral sclerosis: a speech subsystem approach," *Behavioural Neurology*, no.183021, pp. 1-11, 2015.

[7] K. M Yorkston, E. A. Strand, R. Miller, A. Hillel, K. Smith, "Speech deterioration in amyotrophic lateral sclerosis: implications for the timing of intervention," *Journal of Medical Speech-Language Pathology*, vol. 1, no. 1, pp. 35-46, 1993.

[8] M Nishio, S Niimi, "Changes over time in dysarthric patients with amyotrophic lateral sclerosis (ALS): a study of changes in speaking rate and maximum repetition rate (MRR)," *Clinical Linguistics & Phonetics*, vol. 14, no. 4, pp. 485-497, 2000.

[9] L. J. Ball, A. Willis, D. R. Beukelman, G. L. Pattee, "A protocol for identification of early bulbar signs in amyotrophic lateral sclerosis," *Journal of the Neurological Sciences*, vol. 191, pp. 43-53,2001.

[10] J. R. Green, Y. Yunusova, M. S. Kuruvilla, J. Wang, G. L. Pattee, L. Synhorst, L. Zinman, J. D. Berry, "Bulbar and speech motor assessment in ALS: challenges and future directions," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 14, pp. 494-500, 2013.

[11] G. S. Turner, K. Tjaden, G. Weismer, "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis," *Journal of Speech and Hearing Research*, vol. 38, pp. 1001-1013, 1995.

[12] S. Shellikeri, J. R. Green, M. Kulkarni, P. Rong, R. Martino, L. Zinman, Y. Yunusova, " Speech movement measures as markers of bulbar disease in amyotrophic lateral sclerosis," *Journal of Speech, Language, and Hearing Research*, vol. 59, no. 5, pp. 887-899, 2016.

[13] R. DePaul, B. R. Brooks, "Multiple orofacial indices in amyotrophic lateral sclerosis," *Journal of Speech and Hearing Research*, vol. 36, pp. 1158-1167, 1993.

[14] Y. Yunusova, J. R. Green, M. J. Lindstrom, L. J. Ball, G. L. Pattee, L. Zinman, "Kinematic of disease progression in bulbar ALS," *Journal of Communication Disorders*, vol.43, no. 1, pp. 6-20, 2010.

[15] Y. Yunusova, G. Weismer, J. R. Westbury, M. J. Lindstrom, "Articulatory movements during vowels in speakers with dysarthria and healthy controls," *Journal of Speech, Language, and Hearing Research*, vol. 51, pp. 596-611, 2008.

[16] P. Rong, Y. Yunusova, J. Wang, L. Zinman, G. L. Pattee, J. D. Perry, B. Perry, J. R. Green, "Predicting speech intelligibility decline in amyotrophic lateral sclerosis based on the deterioration of individual speech subsystems," *PLoS ONE*, vol. 11, no. 5, p. e0154971, 2016.

[17] A. Bandini, S. Orlandi, F. Giovannelli, A. Felici, M. Cincotta, D. Clemente, P. Vanni, G. Zaccara, C. Manfredi, "Markerless analysis of articulatory movements in patients with Parkinson's disease," *Journal of Voice*, vol. 30, no. 6, pp. 766.e1-766.e11, 2016.

[18] R. Netsell, B. Daniel, C. G. Celesia, "Acceleration and weakness in parkinsonian dysarthria," *Journal of Speech and Hearing Disorders*, vol. 40, no. 2, pp. 170-178, 1975.

[19] D. Goldvasser, C. A. McGibbon, D. E. Krebs, "High curvature and jerk analyses of arm ataxia," *Biological Cybernetics*, vol. 84, no. 2, pp. 85-90, 2001.

[20] B. R. Brooks, R. G. Miller, M. Swash, T. L. Munsat, "El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis," *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, vol. 1, no. 5, pp. 293-299, 2000.

[21] Z. S., A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, H. Chertkow, "The Montreal cognitive assessment (MoCA): a brief screening tool for mild cognitive impairment," *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695-699, 2005.

[22] J. M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, A. Nakanishi, "The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function," *Journal of the Neurological Sciences*, vol. 169, no. 1-2, pp. 13-21, 1999.

[23] K. Yorkston, D. Beukelman, M Hakel, M. Dorsey, "Speech Intelligibility Test [Computer Software]," Lincoln, NE: Madonna Rehabilitation Hospital, Institute for Rehabilitation Science and Engineering, 1996.

[24] J. R. Green, J. Wang, D. L. Wilson, "SMASH: a tool for articulatory data processing and analysis," in *INTERSPEECH 2013 – 14th Annual Conference of the International Speech Communication Association, August 25-29, Lyon, France, Proceedings*, 2013, pp. 1331-1335.

[25] P. P. G. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9-10, pp. 341-345, 2001.

[26] M. Robnik-Šikonja, I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, no. 1-2, pp. 23-69, 2003.

[27] H. Peng, F. Long, C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005.

[28] K. Kira L. A. Rendell, "A practical approach to feature selection," In D. Sleeman and P.Edwards, editors, *Proceedings of International Conference on Machine Learning*, 1992, pp. 249-256.

[29] G. Roffo, "Feature Selection Library (Matlab Toolbox)," *arXiv Preprint arXiv*, 1607.01327, 2016.

[30] G. Roffo, S. Melzi, M. Cristani, "Infinite feature selection," In *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 4202-4210.

[31] A. Tsanas, M. Little, P. McSharry, J. Spielman, L. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5 pp. 1264-1271, 2012.

[32] J Wang, P. V. Kothalkar, M. Kim, Y. Yunusova, T. F. Campbell, D. Heitzman, J. R. Green, "Predicting intelligible speaking rate in individuals with amyotrophic lateral sclerosis from a small number of speech acoustic and articulatory samples," In *SLPAT 2016 – 7th Workshop in Speech and Language Processing for Assistive Technologies, San Francisco, USA, Proceedings*, 2017, pp. 91-97.

[33] A. Bandini, S. Orlandi, H. J. Escalante, F. Giovannelli, M. Cincotta, C. A. Reyes-Garcia, P. Vanni, G. Zaccara, C. Manfredi, "Analysis of facial expressions in Parkinson's disease through video-based automatic methods," *Journal of Neuroscience Methods*, vol. 281, pp. 7-20, 2017.

[34] A. Bandini, S. Ouni, P. Cosi, S. Orlandi, C. Manfredi, "Accuracy of a markerless acquisition technique for studying speech articulators," in *INTERSPEECH 2015 – 16th Annual Conference of the International Speech Communication Association, September 6-10, Dresden, Germany, Proceedings*, 2015, pp. 2162-2166.