Machine Learning Model Predicts Listener Effort in ALS-related Dysarthria

Indu Navar Bingham[1], Esteban G. Roitberg[1,2], Julian Peller[1,3], Marcos A. Trevisan[4], Diego E. Shalom[4], Felipe Aguirre[1], Gastón Bujía[1], Iair Embon[1], Alan Taitz[5], Raquel Norel[6], Carla Agurto[6], Donna Harris[7,1], Amy Wright[1], Katie Seaver[8,1], Stacey Sullivan[9,1], Jordan R Green[8], Lyle W. Ostrow[7], Ernest Fraenkel[10], James D. Berry[9]

[1] Peter Cohen Foundation dba Everything ALS, Seattle, Washington and Los Altos, California, USA, [2] Escuela de Ciencia y Tecnología, Universidad Nacional de San Martín, Buenos Aires, Argentina., [3] Laboratorio de Neurociencias Cognitivas Computacionales, Humai, Buenos Aires, Argentina, [4] Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Departamento de Física - CONICET - Universidad de Buenos Aires, Instituto de Física Interdisciplinaria y Aplicada (INFINA). Buenos Aires, Argentina, [5] SRI International, Menlo Park, CA, USA. This work was performed working at Peter Cohen Foundation, [6] IBM Research Center, Yorktown Heights, NY, USA, [7] Lewis Katz School of Medicine at Temple University, Department of Neurology, Philadelphia, PA, USA, [8] MGH Institute of Health Professions, Charlestown Massachusetts, MA, USA, [9] Sean

Background

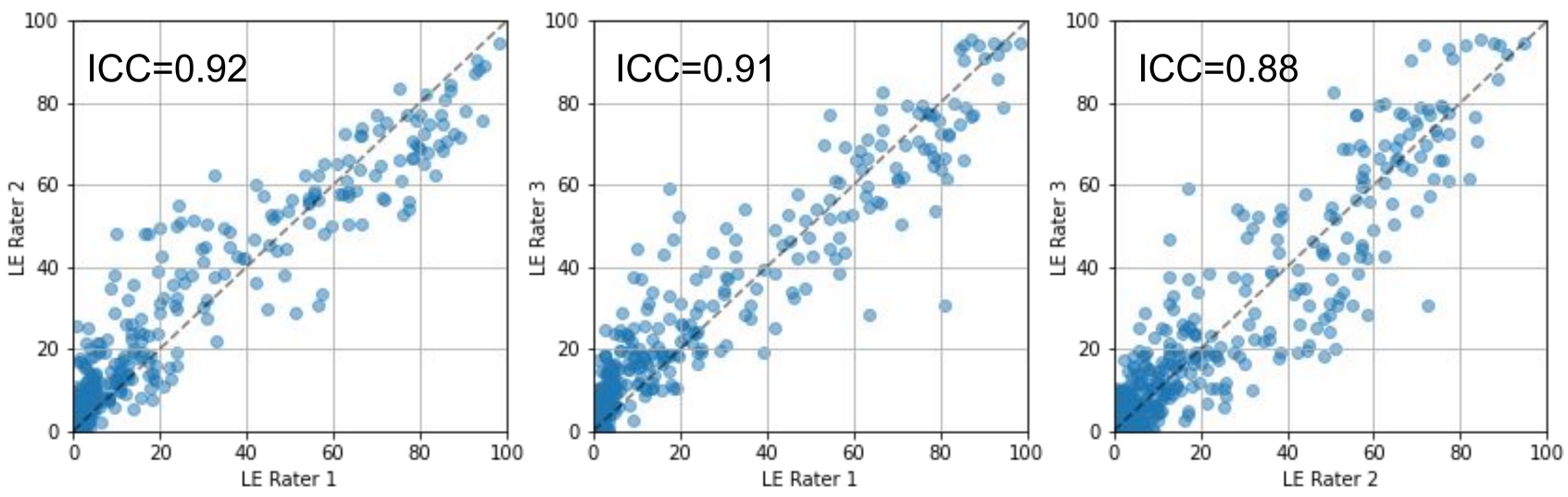
The dysarthria occurring in ALS involves various deteriorating speech subsystems, challenging accurate quantification of progression. Speech-Language Pathologists (SLPs) use intelligibility ratings and Listener Effort (LE) assessments to quantify the severity of dysarthria. Recently, several Machine Learning (ML) methods have been proposed to assess patients' speech impairment, generally focusing on predicting the ALSFRS-R speech-related question. We have used ML on speech recordings to predict LE, a quantitative, clinically relevant measure of dysarthric speech.

Hypothesis: Based on remote speech recordings, Listener Effort can be reliably predicted by Machine Learning models.

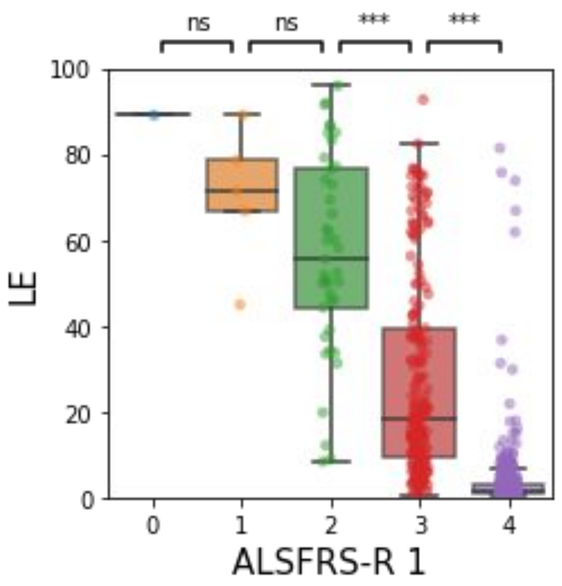
Methods

- The EverythingALS Speech Study obtained longitudinal clinical information and speech recordings from 292 participants.
- A subset of 125 participants was selected (105 pALS, 20 controls), encompassing 2124 recordings from 708 unique participant assessments.
- LE was manually rated in this subset by trained SLPs, demonstrating excellent inter-rater reliability.
- This process ensured a highly reliable dataset, on top of which various ML strategies were applied to predict LE.

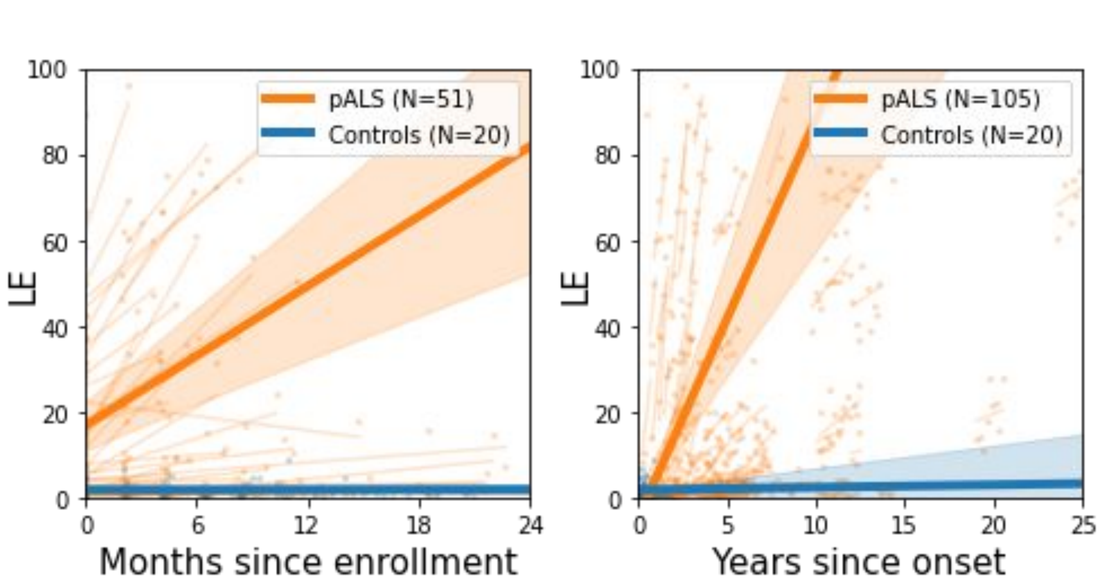
Results: Listener Effort



Very good/excellent pairwise LE inter-rater reliability.



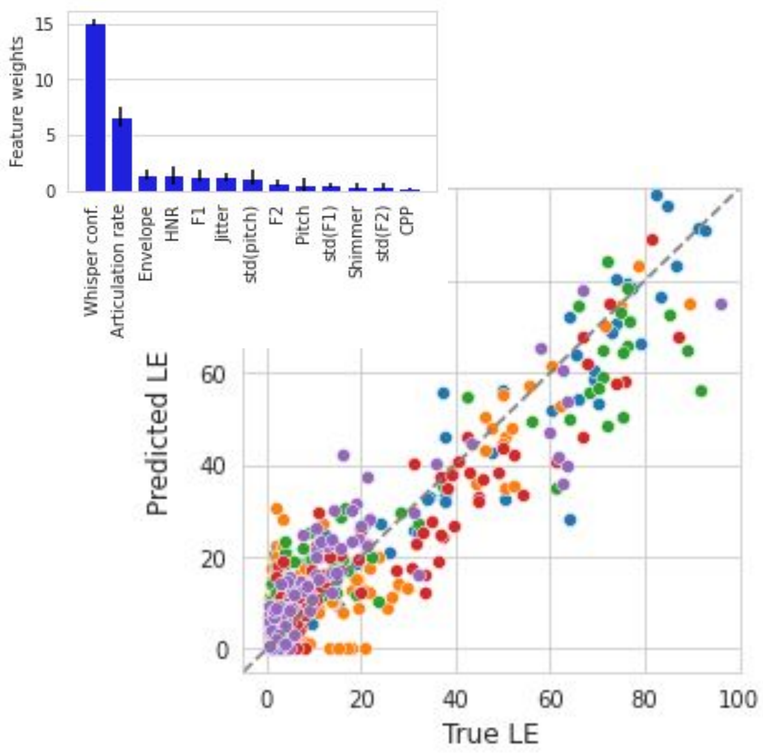
Lower ALSFRS-RSE scores and higher LE denote lower speech function.



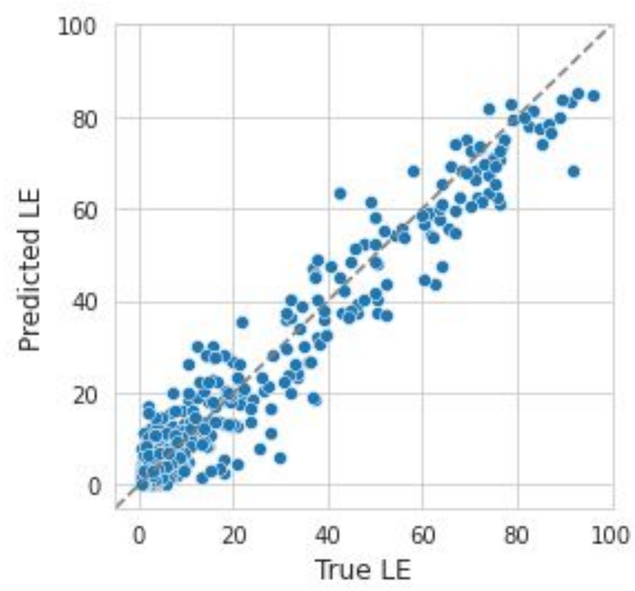
Higher slope of decline of LE for PALS than controls in both all participants and those with onset <3yrs prior to enrollment.

Results: Model

- A simple Lasso regression model yielded an R^2 of 0.83 ± 0.07, with only two features proving relevant for making the prediction: **Speaking Rate** and **Whisper Confidence**.
- The latter feature represents the confidence of the automatic speech recognition system (Whisper) in its own transcription accuracy.



- More complex approaches, such as using an ensemble including traditional and deep learning models, achieved an R^2 of 0.94 ± 0.03.



Conclusions

Listener Effort, as rated by Speech Language Pathologists, can be robustly estimated using a simple Machine Learning strategy, with only two speech features proving relevant for the prediction: **Speaking Rate** and **Whisper Confidence**.