Assessment of Speech Subsystem Interactions in Children with Autism Spectrum Disorder

[1,2]Tanya Talkar, [2]James R. Williamson, PhD, [2]Daniel Hannon, PhD, [2]Hrishikesh Rao, PhD, [2]Sophia Yuditskaya, MS, [2]Douglas Sturim, PhD, [2]Kajal Claypool, PhD, [3]Lisa Nowinski, PhD, [3]Hannah Saro, MS, [3]Carol Stamm, PhD, [4]Maria Mody, PhD, [3]Christopher McDougle, MD, [2]Thomas F. Quatieri, ScD

[1]Speech and Hearing Bioscience and Technology, Harvard University, Boston, MA
[2]MIT Lincoln Laboratory, Lexington, MA
[3]Massachusetts General Hospital Lurie Center for Autism, Boston, MA
[4]Massachusetts General Hospital Martinos Brain Imaging Center, Charlestown, MA

Background

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that affects individuals throughout their lifespan. There is evidence to suggest that highly verbal children with ASD have a greater number of phonological errors during conversation as compared to their neurotypical equivalents, which is hypothesized to contribute to communication difficulties[1]. We hypothesize that some speech production difficulties stem from motor coordination and control deficits that are also associated with facial and dexterity fine motor impairments seen in ASD[2,3]. We aimed to measure motor coordination of four underlying speech production subsystems: vocal tract, larynx, respiratory system, and face, by calculating correlations across features extracted from audio and video recordings. Our initial findings support our hypothesis, showing differences in measures of motor coordination across speech subsystems as well as across different linguistic contexts, highlighting difficulties in reaching acoustic and articulatory targets for children with ASD.

Methods

Speech acoustic features from audio and facial features from video were collected as part of a larger multimodal data collection at the Massachusetts General Hospital Lurie Center for Autism. We obtained MIT/MGH/US Army HRPO IRB approvals prior to the initial data collection. The protocol included a reading of an age-appropriate adaptation of The Caterpillar passage, a diadochokinetic sequence 'Pa-Ta-Ka', four vowels (/a/, /i/, /u/, /ae/) sustained for 5-7 seconds each, and free speech samples, including a Q&A session and picture description tasks. 5 controls (4 males; ages 6,6,6,7; 1 female; age 11) and 5 highly verbal ASD subjects (5 males; ages 7,7,7,7,8 years), were involved in the collection.

Correlations of acoustic and facial features work as a proxy measure of coordination within and across the underlying mechanisms of speech subsystems[4]. This was assessed using eigenvalues from a multi-scale correlation structure derived using the first three formants, F0, and intensity extracted from the audio, and facial action units (FAUs) extracted from the video. This technique has been shown to be successful in assessing speech motor coordination in depression and Parkinson's disease[4]. We computed Cohen's d effect sizes for the average eigenvalues of the ASD class as compared to the controls. Additionally, the eigenvalues were used to discriminate between the ASD and control groups, using a leave-one-subject-out cross validation procedure to create Gaussian Mixture Models (GMMs). Performance was assessed using Receiver Operating Characteristic (ROC) curves by extracting out the area under the curve (AUC). GMMs were calculated independently for correlations of individual and combinations of acoustic and facial features.

Results & Discussion

There was no significant difference between the groups when comparing means and standard deviations of the raw acoustic and facial features. However, eigenvalues derived from the correlation structures and values in the correlation structures indicated higher variability in F0 and intensity time-series in children with ASD, particularly in the Pa-Ta-Ka task. Correlation analysis also indicated higher variability in F1 and F2 signals during the sustained vowel task in the ASD group. In contrast, patterns seen in correlations of formants and FAUs indicated higher variability in formant and FAU time-series for The Caterpillar passage in control subjects. These patterns,

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. This material is based upon work supported by the Under Secretary of Defense for Research and Engineering under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Under Secretary of Defense for Research and Engineering.

together, may indicate that children with ASD have a harder time precisely moving muscle groups in speech subsystems to reach the acoustic and articulatory trajectories achieved by control subjects in these tasks, as the signals are reflections of the underlying movements.

Eigenvalues derived from correlations across formants and FAUs from The Caterpillar passage, free speech, and on the Pa-Ta-Ka task, furthermore, yielded positive effect sizes (+3) at large eigenvalues, and negative effect sizes (-5) at smaller eigenvalues, reflecting a more isotropic time-embedded scatter distribution of these time-series in control subjects. This implies that formant and FAU time-series are less coupled in control subjects as compared to in ASD subjects, potentially indicating higher independence of articulator movement in control subjects.

ROC curves formed from the GMMs yielded AUCs (**Figure 1**) that were above chance for many of the tasks and correlations of features. Moreover, many were significant with a p-value of 0.05 for any AUC value above 0.78. Specifically, an AUC of 1.0 was found when utilizing F0 during free speech tasks, and when using either a combination of formants and FAUs or a combination of F0 and formants during The Caterpillar passage. Combinations of features across speech subsystems in some cases improved performance, such as with F0 and intensity during 'Pa-Ta-Ka' (AUC = 0.88), highlighting differences in coordination across speech subsystems in ASD subjects.

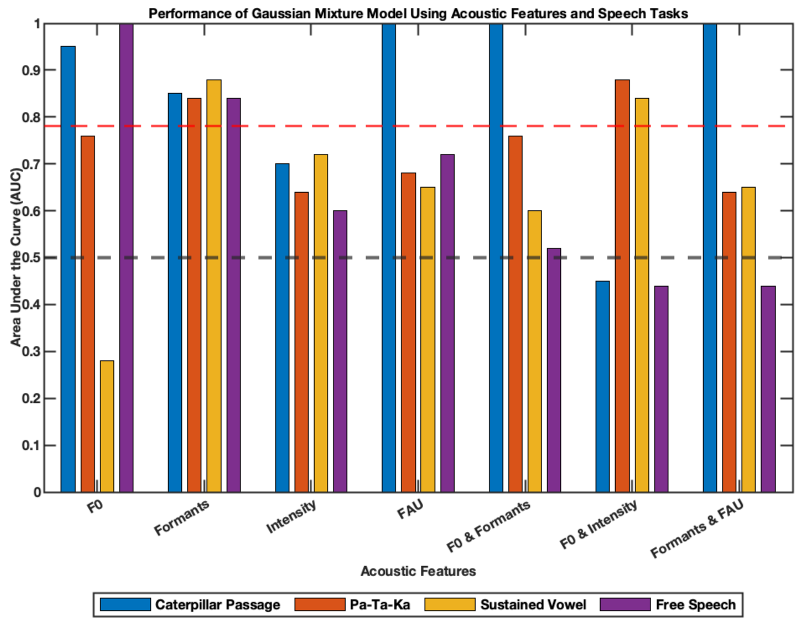

Figure 1: *AUC of ROC Curves Derived from GMM log-likelihood classification for combinations of speech tasks and features extracted. Red dotted line indicates an AUC of 0.78, above which, values were significant with a p-value of 0.05.*

Conclusions

While further validation will need to be conducted on a larger sample size, the initial results from this pilot analysis suggest that correlations between acoustic and facial measures of speech subsystems (vocal tract, larynx, face, and respiratory systems) can provide insight into the speech motor coordination differences present in children with ASD. In addition, there were differences seen in motor coordination measures across different speech tasks. We aim to use these speech motor coordination features to form an assessment of children with ASD, which can be tracked during the course of speech interventions. Future work will also incorporate speech control models to analyze the coordination of individual articulators across speech subsystems using inversion techniques. The aim will be to develop a neurocomputational model to focus on the interaction of the neurological feedforward and feedback pathways controlling the coordination of these subsystems.

References

1. Cleland, J., Gibbon, F. E., Peppé, S. J., O'Hare, A., & Rutherford, M. (2010). Phonetic and phonological errors in children with high functioning autism and Asperger syndrome. International Journal of Speech-Language Pathology, 12(1), 69-76.
2. Choi, B., Leech, KA., Tager-Flusberg, H., Nelson, C.A. (2018). Development of fine motor skills is associated with expressive language outcomes in infants at high and low risk for autism spectrum disorder, Journal of Neurodevelopmental Disorders 2018 10:14.
3. Mody, M., Shui, A.M, Nowinski, L.A., Golas, S.B., Ferrone, C., O'Rourke, J.A., McDougle, C.J. (2017). Communication Deficits and the Motor System: Exploring Patterns of Associations in Autism Spectrum Disorder (ASD). J Autism Dev Disord 2017;47:155-62.
4. Quatieri, T. F., Williamson, J. R., Smalt, C. J., Perricone, J., Patel, T., Brattain, L., Eddy, M. (2017). Multimodal biomarkers to discriminate cognitive state. The Role of Technology in Clinical Neuropsychology, 409.