# Assessment of Speech and Fine Motor Coordination in Children With Autism Spectrum Disorder

**TANYA TALKAR** [1,2], **JAMES R. WILLIAMSON**[2], **DANIEL J. HANNON**[2], **HRISHIKESH M. RAO**[2], **SOPHIA YUDITSKAYA**[2], **KAJAL T. CLAYPOOL**[2], **DOUGLAS STURIM**[3], **LISA NOWINSKI**[4], **HANNAH SARO**[4], **CAROL STAMM**[4], **MARIA MODY**[5,6], **CHRISTOPHER J. MCDOUGLE**[4], **AND THOMAS F. QUATIERI** [1,2], **(Fellow, IEEE)**

[1]Speech and Hearing Bioscience and Technology, Division of Medical Sciences, Harvard Medical School, Boston, MA 02115, USA
[2]Human Health and Performance Systems, Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA 02142, USA
[3]Artificial Intelligence Technology and Systems, Massachusetts Institute of Technology Lincoln Laboratory, Lexington, MA 02142, USA
[4]Lurie Center for Autism, Massachusetts General Hospital, Lexington, MA 02421, USA
[5]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA 02129, USA
[6]Harvard Medical School, Boston, MA 02115, USA

Corresponding author: Tanya Talkar (ttalkar@g.harvard.edu)

**ABSTRACT** Autism Spectrum Disorder (ASD) is a developmental disorder characterized by difficulty in communication, which includes a high incidence of speech production errors. We hypothesize that these errors are partly due to underlying deficits in motor coordination and control, which are also manifested in degraded fine motor control of facial expressions and purposeful hand movements. In this pilot study, we computed correlations of acoustic, video, and handwriting time-series derived from five children with ASD and five children with neurotypical development during speech and handwriting tasks. These correlations and eigenvalues derived from the correlations act as a proxy for motor coordination across articulatory, laryngeal, and respiratory speech production systems and for fine motor skills. We utilized features derived from these correlations to discriminate between children with and without ASD. Eigenvalues derived from these correlations highlighted differences in complexity of coordination across speech subsystems and during handwriting, and helped discriminate between the two subject groups. These results suggest differences in coupling within speech production and fine motor skill systems in children with ASD. Our long-term goal is to create a platform assessing motor coordination in children with ASD in order to track progress from speech and motor interventions administered by clinicians.

**INDEX TERMS** Acoustic speech analysis, biomedical application, autism spectrum disorder, motor coordination, fine motor skills.

## I. INTRODUCTION

Children with Autism Spectrum Disorder (ASD) undergo a large number of speech interventions in the home, the clinic, and at school to address communication difficulties at the core of the disorder [1]. However, the effectiveness and tracking of these treatments have not been well documented [2]. Most progress is denoted by subjective observations from the

The associate editor coordinating the review of this manuscript and approving it for publication was Ganesh Naik.

clinician or caregivers. A goal of our research is to design a speech assessment protocol that can be used in the clinic and the home setting, from which we can derive objective speech biomarkers that can aid in tracking subtle but meaningful changes in speech production due to speech interventions and can inform intervention decisions.

Delay in the development of fine motor skills has been found to be common in ASD, and is typically a better predictor of diagnosis than gross motor skills [3]. In children, this can manifest in degraded control during handwriting

[4], [5] and drawing tasks [6]. Recent research has suggested that speech differences witnessed in ASD stem from the motor coordination and control deficits that are associated with facial and fine motor gestures in ASD [7]. Some children with ASD demonstrate difficulties with oral motor skills which are linked to speech production [8]. Correlations between deficits in fine motor skills and development of expressive and receptive language suggest, furthermore, that difficulties in speech production for children with ASD may contribute to communication difficulties [9], [10]. These findings of a co-occurrence of motor deficits across perceptuo-motor systems suggest that motor and language skills, including speech production, are interrelated in development [11].

There has been considerable research into acoustic biomarkers of ASD, such as fundamental frequency (F0) or intensity [12]. There has not yet emerged, however, a consistent set of objective acoustic biomarkers sensitive and specific to ASD for use in assessment, phenotyping, and tracking response to treatments, nor is there a complete understanding of interactions between acoustic biomarkers [12]. There is therefore a pressing need to develop quantifiable and sensitive biomarkers to understand speech production differences in individuals with ASD, and to fill the gap between advances in research into acoustic biomarkers and their use in the clinical setting.

Many acoustic analyses have focused on a single speech production subsystem, either articulatory (formants and facial features), laryngeal (fundamental frequency), or respiratory (intensity). One approach to addressing the interaction is through correlation structures across speech subsystems. These structures are formed from acoustic and facial features and have previously been used as a proxy measure of motor coordination. Features extracted from the structures have been used to predict clinical severity scores of individuals with major depressive disorder (MDD) [13], Parkinson's disease [14], and mild traumatic brain injury (mTBI) [15]. While these features have been successful in representing motor coordination deficits within speech subsystems, there are also relationships across speech subsystems. For example, lip, jaw, and laryngeal motion timings are constrained to facilitate their coordination [16], and pitch and loudness are positively correlated [17]. These relationships between subsystems are depicted in Fig. 1. The correlation features have also not been applied to fine motor skills, such as shape-drawing.

In this pilot study, we build upon acoustic analysis studies of ASD that have typically looked at individual subsystems [12]. We add new analyses to characterize motor coordination across speech production subsystems and for fine motor skills by looking at the patterns of the correlations across signals derived from speech and shape-drawing tasks. Furthermore, we create machine learning models with correlation structure features to discriminate between children with ASD and neurotypical controls.
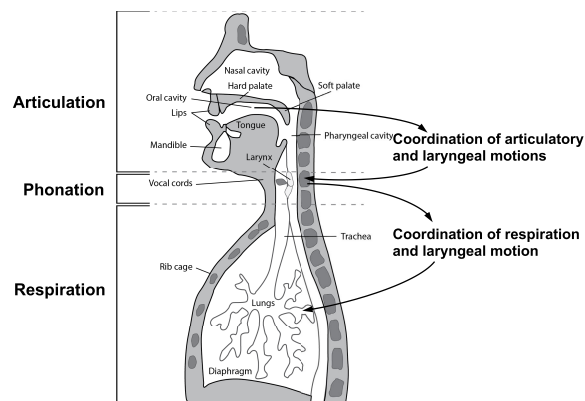


**FIGURE 1.** Major speech production subsystems of articulation, phonation, and respiration and their relationships. Speech production involves coordination within and between these subsystems as denoted in the figure.

## II. METHODS

### A. PARTICIPANTS

This study involved five children (5 males, ages 7.20 ± 0.40 years) with ASD and five neurotypical controls (4 males, ages 6.25 ± 0.43 years; 1 female, age 11 years), verbally and natively fluent in English and with age-appropriate reading skills. All subjects with ASD were evaluated by developmental specialists at the Massachusetts General Hospital (MGH) Lurie Center for Autism who verified a diagnosis of ASD based on a comprehensive diagnostic interview, direct observation, review of records, and formal testing, including the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2) [18], when necessary. The Primary Investigator then confirmed each ASD diagnosis with completion of a formal Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) checklist [19]. Children who had other neurodevelopmental or psychiatric disorders were excluded from enrollment through a thorough study screening process. All ASD subjects and controls were screened for global cognitive impairments by clinical interview and review of records, and subjects with intellectual disability were not included. For ASD subjects, IQ scores ranged from average to significantly above average (Full Scale IQ ranged from 105 to 135 with a mean of 124). All neurotypical controls had no reported or recorded history of speech, language, hearing, vision, or neurological disorders. The study was approved by the Institutional Review Boards (IRBs) at the Massachusetts Institute of Technology, MGH, and the US Army Human Research Protection Office (HRPO). Prior to participation, written, informed consent from one parent or guardian and verbal assent from the child were obtained.

### B. PROTOCOL

Data was collected in a well-lit, electrically and acoustically shielded room at the Lurie Center. The protocol was displayed to the child on an iPad (Apple, Inc., Cupertino, CA)

through Powerpoint (Microsoft Corp., Redmond, WA) slides. A neuropsychologist administered the protocol through a separate, connected iPad and verbally instructed the child to follow prompts displayed on the slides. Audio was recorded at 48000 Hz with a wired DPA 4060 Microphone (DPA Microphones, Denmark) hanging from the ceiling above and slightly in front of the child, and video was recorded with a Canon XA20 Pro HD camera (Canon Inc., Japan) placed in front of the child. Each child first read aloud a simplified version of The Caterpillar passage [20], adapted to be age-appropriate and designed to induce phonetic contrasts and prosodic stress. One control subject had difficulty reading the passage and was therefore excluded from any analysis of The Caterpillar passage. Children were then instructed to repeat the diadochokinetic sequence ''pa-ta-ka'' as many times as possible in a single breath. Children then vocalized four different vowels ([a], [i], [u], [æ]) each in a single sustained breath for 5-7 seconds. Each child was led through two free speech tasks: 1) describing their favorite toy, sport, or hobby, and 2) describing the situations displayed in six different pictures.

Subsequently, the children proceeded through an adapted version of the Beery-Buktenica Developmental Test of Visual-Motor Integration [21] on a custom iPad application. The task involved simple drawings: vertical, horizontal, and diagonal lines, as well as squares, circles, triangles, plus signs and crosses. Handwriting measurements consisted of a time-series of spatial coordinates $\{x, y\}$ and pressures $\{z\}$, extracted using the jQuery Pressure.js module [22]. The time interval between successive data points was typically 16 ms. Variability in the time intervals was due to the module recording coordinates every time a pixel was displayed.

All children additionally underwent two clinical assessments of visual-motor skills. The neuropsychologist led them through the Beery-Buktenica Developmental Test of Visual-Motor Integration - 6th edition (VMI-6) [21], as well as the NEPSY-II Visuomotor Precision task [23]. In the latter, children were asked to quickly draw a line through a series of mazes without crossing the boundary of the maze.

### C. LOW-LEVEL FEATURE EXTRACTION
Speech data from the subjects was manually segmented using the VCode annotation tool [24]. Each segment represented a period of continuous speech from the subjects. Segments were further split to ensure that no segment had pauses longer than a second, as detected by an energy-based speech activity detection algorithm developed in MATLAB. This algorithm also allowed for elimination of segments without at least 3 seconds of continuous speech from further analysis. All recordings were manually checked to ensure the absence of any distortion in the audio and video recordings. A speech example from a neurotypical control reciting the diadochokinetic sequence 'pa-ta-ka' is provided in Fig. 2.

A set of low-level speech features was selected based on acoustic and facial parameters that have been used in studies to characterize individuals with ASD [12]. An example of
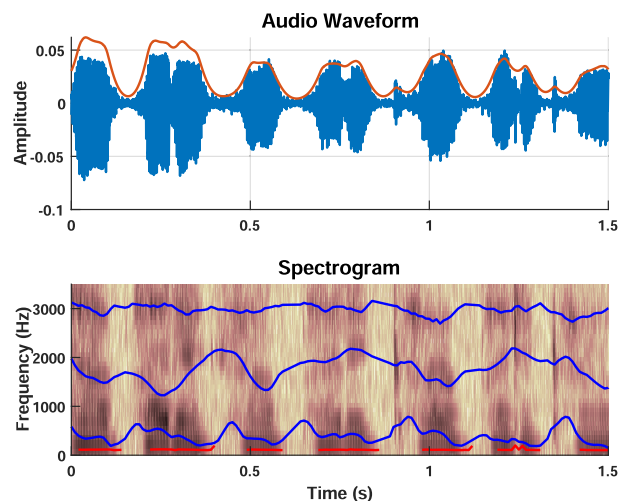


**FIGURE 2.** Waveform and spectrogram extracted from a control subject's recitation of the diadochokinetic sequence 'pa-ta-ka'. The extracted intensity time-series is overlaid on the waveform in orange. The extracted formants time-series are depicted on the spectrogram in blue, and the extracted F0 time-series is depicted in red.

these acoustic features extracted from a neurotypical control is provided in Fig. 2. The first three formant time-series (F1-F3) were estimated using the Kalman-based autoregressive moving average (KARMA) software tool with a 10 ms frame interval [25]. KARMA includes an energy-based speech activity detector that allows a Kalman smoother to extrapolate formants through silent gaps in the signal with no speech, thereby providing continuous time-series of formants.

Fundamental frequency (F0) was extracted using the $v_{fxrapt}$ tracker from the VOICEBOX speech processing toolbox [26], using frequency limits of 180-300 Hz, based on values of fundamental frequency for children of our target age range [27].

Intensity was extracted using a custom MATLAB script that provides a smooth contour of amplitude peaks based on an iterative time-domain signal envelope estimation [28], [29]. This technique captures both the contributions of the respiratory muscles and resonance-harmonics interaction to amplitude modulation of a speech envelope.

Facial action unit (FAU) features were extracted from video segments using the OpenFace toolbox [30], consistent with the Facial Action Coding System [31]. Each FAU represents a particular configuration of facial muscles, such as lip tightener or chin raiser, and has typically been used to recognize emotion [31]. The intensity of 9 FAUs, capturing the lips and jaw, were estimated with a sampling rate of 30 Hz. These FAUs are listed in the Appendix B. F0 and intensity were downsampled to a sampling rate of 100 Hz, while FAUs were interpolated to a sampling rate of 100 Hz using MATLAB's *interp1* function with spline interpolation.

Each handwritten shape was automatically parsed into different segments whenever the time interval between successive data points was greater than 40 ms or the euclidean distance was greater than 30 pixels. A segment was valid for

further analysis if it contained at least 10 data points. The positions and pressures of valid segments were then converted into time series of movement and pressure accelerations by convolving each channel with the discrete Laplacian filter, $[1, -2, 1]$, followed by smoothing with a 7-point Gaussian filter ($\sigma = 1$).

### D. HIGH-LEVEL CORRELATION FEATURE EXTRACTION

Multivariate auto- and cross-correlations of acoustic, facial, and handwriting low-level features were used to produce proxy measures of coordination within and across the underlying mechanisms of speech subsystems and fine motor skills [13]–[15].

Specifically, time-delay embedding was used to expand the dimensionality of the feature time series, resulting in correlation matrices with embedded auto- and cross-correlation patterns that represent coupling strengths across feature channels at multiple relative time delays. High complexity in a feature signal results in a narrow central band of high autocorrelation due to energy being distributed across a broad range of signal frequencies. Lower complexity in the feature signal corresponds to energy being concentrated in fewer frequencies, often with a wider central band of high autocorrelation. The correlation matrix eigenspectra quantify and summarize the frequency properties of the set of feature signals. Higher complexity across multiple signals is reflected in a more uniform distribution of eigenvalues, with lower complexity reflected in a larger proportion of the overall signal variability being concentrated in a small number of eigenvalues.

The mathematical details of this method are provided in [13] and a visual explanation of the method is in Fig.3. For each speech segment, channel-delay correlation matrices were calculated from various combinations of formants (F1, F2, F3), F0, intensity, and FAU time-series. Each matrix contained the correlation coefficients between the time-series at defined time delays to create the embedding space. Four matrices were computed at four delay scales (10, 30, 70, 150 ms) with 15 time-delays used per scale. These delay scales allow for characterization of coupling of signals at the frame rate and at longer time scales. These scales, in particular, have been used for characterization of psychomotor slowing in MDD [13]. Each matrix comparing $n$ signals

has a dimension of ($n*15 \times n*15$). For all correlations, an automatic masking technique was used to include only segments that contained speech, using a custom speech activity detection algorithm. Eigenvalues of all resulting matrices were extracted by rank-order, from largest to smallest. For each task in the protocol, the eigenvalues across individual segments that made up the task were averaged to obtain the mean eigenspectrum for the task, which were used as feature vectors for comparison and classification analysis. For a comparison of $n$ signals, this will yield $n * 15$ eigenvalues.

For the handwriting segments, matrices were calculated from the smoothed derivatives of the $\{x, y, z\}$ values in each segment. Because the segments could be as short as 10 data points, only a single delay scale, with 7 time-delays and delay spacing of one data point (16 ms), was used. Eigenvalues were extracted in the same way as they were for the speech segments. For a comparison of the 3 signals, this yielded 21 eigenvalues. The eigenspectrum from each segment was treated independently in comparison and classification analysis.

### E. CHARACTERIZATION AND CLASSIFICATION

Cohen's d effect sizes of the eigenvalues for ASD vs control subjects were computed to characterize complexity of cross-correlations of signals [13]–[15]. The eigenspectra were also used as features in constructing Gaussian Mixture Models (GMMs) to classify subjects as ASD or control. This process used leave-one-subject-out (LOSO) cross-validation. For each of the speech feature sets and tasks, the first six principal component analysis (PCA) features were extracted from the eigenspectra which were normalized (z-scored) across the subjects, capturing at least 90% of the variability in the eigenspectra. For each cross-validation fold, an ensemble of 10 GMMs was created from these PCA features in the training data. Supervised adaptation of the GMM means was used to create ASD and control GMMs, a technique typically used in speaker recognition [32]. For each speech feature set and task combination, the ASD prediction score for the test subject was the log-likelihood ratio of the sum of the likelihoods of ASD GMMs and control GMMs in the ensemble.

Classification of ASD subjects for handwriting followed the same GMM architecture, using a single PCA feature extracted from the eigenspectra of each handwriting segment. The cumulative prediction per subject was obtained from the mean of the prediction scores across all of the subject's handwriting segments. Receiver operating characteristic (ROC) curves were computed from all prediction scores from speech and handwriting tasks and features, and the area under the curve (AUC) of the ROCs was computed. For speech tasks, this resulted in a single AUC per feature combination and task. For handwriting tasks, independent AUCs were calculated using either the mean of the prediction scores across all segments for a subject or by treating each prediction score independently.
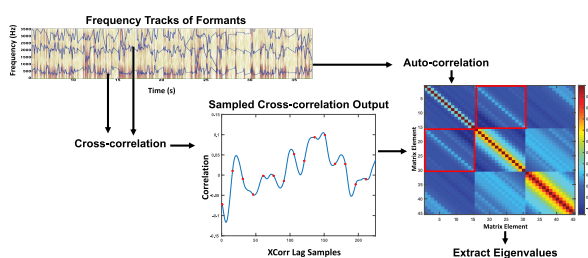


**FIGURE 3. Example correlation structure pipeline for eigenspectra feature calculation using formants (F1-F3).**

## III. RESULTS

### A. CLINICAL ASSESSMENTS

All children demonstrated intact visual-motor integrations skills with scores on the Beery-Buktenica Developmental Test of Visual-Motor Integration - 6th Edition (VMI-6) [21] falling in the average to high average range. The VMI-6 task produces a standard score with a mean of 100 with a standard deviation of 15. Children with ASD had an average VMI-6 standard score of $97.8 \pm 5.84$ and neurotypical controls had an average score of $106.2 \pm 5.16$, showing no significant difference with a two-sided t-test (p = 0.05). In contrast to solidly intact visual-motor integration skills, children with ASD trended towards slightly reduced visual-motor precision on the NEPSY-II Visuomotor Precision task [23]. The NEPSY-II Visuomotor Precision task produces a scaled score with a mean of 10 with a standard deviation of 3. Children with ASD had an average score of $6.8 \pm 1.79$ and neurotypical controls had an average score of $8.6 \pm 1.52$, showing no significant difference between the two groups as well (p = 0.12).

### B. LOW-LEVEL FEATURES

As a baseline using low-level features, Tables 1 and 2 lists the Cohen's d effect sizes of the mean and variance of the extracted acoustic, facial, and handwriting acceleration features between the ASD and control groups. Values above 0 indicate that the univariate statistic was larger in the ASD group as compared to the control group. For each task, the mean and variance of the feature was calculated for each segment within the task, and then averaged across all segments to provide a univariate statistic for the subject. Out of the features, only the variance of F2 during the 'pa-ta-ka' task and free speech and the variance of F3 during sustained vowels showed both high effect sizes (Cohen's $d \geq 1.0$) and a p-value less than 0.05 with a two-sided t-test between the ASD and control groups, possibly indicating greater variability and less control over articulators. The majority of the features did not show a large separation between the two groups, consistent with many previous studies [12]. Exact values of the univariate statistics are provided in Appendix A.

**TABLE 1.** Cohen's d effect sizes comparing means and variances of features extracted from speech tasks for control and ASD subjects. p < 0.05*.

| | | Caterpillar | Pa-Ta-Ka | Sustained Vowels | Free Speech |
|---|---|---|---|---|---|
| **F0** | Mean | 0.95 | 0.07 | 0.12 | 0.09 |
| | Var. | 0.04 | 1.04 | 0.70 | -0.45 |
| **F1** | Mean | -0.53 | 0.41 | 0.97 | 0.05 |
| | Var. | -0.69 | 1.15 | -0.49 | -0.29 |
| **F2** | Mean | -0.45 | -1.40 | 0.39 | -0.70 |
| | Var. | -0.10 | 1.63* | -0.74 | 2.04* |
| **F3** | Mean | 1.11 | -0.72 | -0.41 | -0.31 |
| | Var. | -0.32 | 0.05 | 1.52* | 0.57 |
| **Intensity** | Mean | -0.48 | 0.39 | 0.50 | -0.21 |
| | Var. | -0.27 | 0.75 | 0.29 | 1.03 |

**TABLE 2.** Cohen's d effect sizes comparing distributions of means and standard deviations of handwriting acceleration features that are extracted from position (x, y) and pressure (z) measurements from the parsed segments across all handwriting tasks for control and ASD subjects. p < 0.05*.

| | x | y | z |
|---|---|---|---|
| Mean | 0.85 | 0.53 | 0.34 |
| St. Dev. | 0.24 | 0.16 | 0.44 |

### C. HIGH-LEVEL FEATURES

Two major patterns were observed when looking at the correlation structures and the eigenspectra across the ASD and control groups. The first was higher complexity in the correlation of F0 and intensity during the 'pa-ta-ka' task in the ASD group (Fig. 4), and the second was lower complexity in the correlation of formants and FAUs during the reading task, and in the correlation of handwriting features during the shape drawing tasks (Fig. 5). This was seen using the correlation structures and the Cohen's d effect size patterns of the eigenspectra.
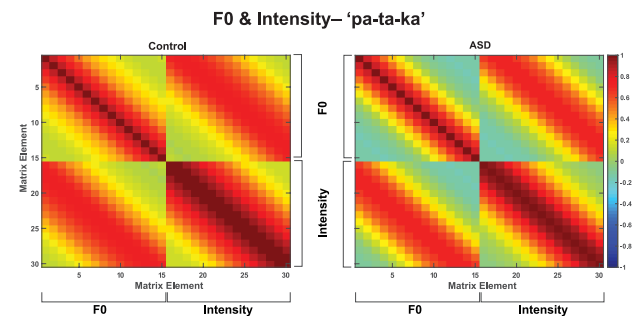


**FIGURE 4.** Average correlation matrix of F0 and intensity from control subjects (left) and ASD subjects (right) during the 'pa-ta-ka' task.

Fig. 4 depicts the average of correlation matrices of control subjects and ASD subjects from the 'pa-ta-ka' task using auto- and cross-correlations of F0 and intensity. There are narrower bands of high auto-correlation in the diagonal blocks for F0 and intensity for the ASD subjects as compared to the control subjects. The signals in the ASD group, therefore, have a shorter time period where their signals are highly auto-correlated. This indicates that energy is being distributed across a broad range of signal frequencies, which means that these signals have higher complexity in the ASD group. This higher complexity indicates decreased precision of hitting F0 and intensity targets in the ASD group as compared to neurotypical controls.

Fig. 5 shows the Cohen's d effect sizes of the eigenvalue features derived from the correlation matrices for the combination of formants and FAUs captured during The Caterpillar passage, and for the acceleration of $x$, $y$, and pressure (z) time-series from the segments of all handwriting tasks, features which aided in high discriminability when used in a GMM. Effect sizes greater than 0 mean that the eigenvalue was greater for ASD subjects as compared to control subjects. The pattern in both of these plots indicates that ASD subjects
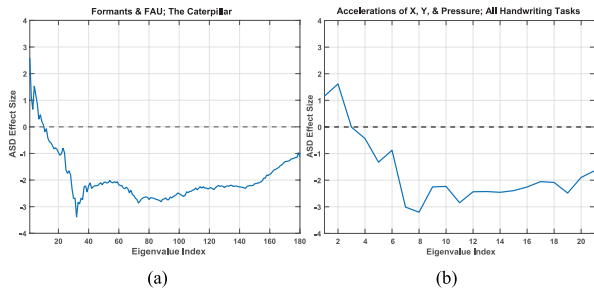
**FIGURE 5.** (a) Cohen's d effect sizes of eigenvalues derived from the correlations of formants and FAUs during The Caterpillar passage. There are 180 eigenvalues for 12 total time series (9 FAUs + 3 formants) and 15 time delays. (b) Cohen's d effect sizes of eigenvalues derived from the correlations of displacement (X,Y) and pressure (Z) accelerations during handwriting tasks. There are 21 eigenvalues for 3 signals and 7 time delays. Both (a) and (b) are displayed as a function of eigenvalue rank (from largest to smallest). Effect sizes that are greater in magnitude than 1.41 have a p-value < 0.05.

have lower complexity in the underlying signals as compared to control subjects during these tasks. ASD subjects, therefore, perhaps have more coupling in their movements to hit formant, FAU, and handwriting targets, as opposed to more independence of movements in control subjects.

### D. CLASSIFICATION

Fig. 6 displays the AUC values generated from ROC curves of GMM log-likelihood values. Eigenvalues derived from F0 during free speech, and the correlation of F0 and formants from The Caterpillar passage were able to perfectly discriminate between the two groups (AUC = 1.00, p = 0.00). FAUs from The Caterpillar passage both individually and correlated with formants yielded an AUC of 1.00. There were also many other tasks that resulted in an AUC greater than 0.77, which corresponded to p-value less than 0.05. This cutoff is denoted as a red dashed line in the plots in Fig. 6. Analysis across some subsystems, such as F0 and intensity in the 'pa-ta-ka'
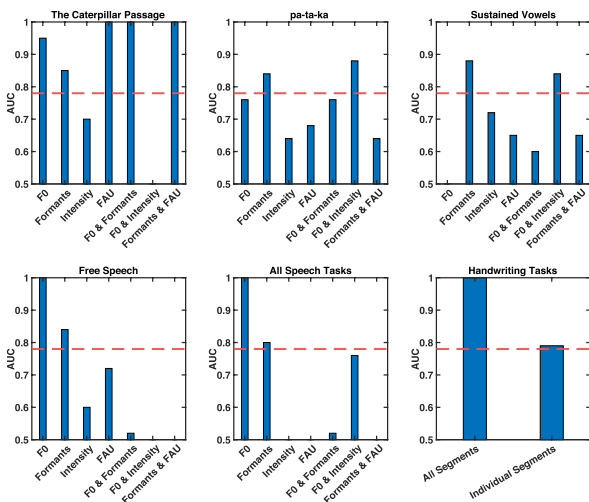


**FIGURE 6.** AUC of ROC curves derived from GMM log-likelihood classification for combinations of speech tasks and features and the handwriting tasks. The red dotted line depicts an AUC of 0.77, representing a p-value < 0.05.

task (AUC = 0.88, p = 0.0006) and in the vowel task (AUC = 0.84, p = 0.006), improved performance as compared to analyzing them separately. Furthermore, the results from the handwriting data (AUC = 0.79, p = 0.03 per individual segment, AUC = 1.00, p = 0.00 by averaging prediction scores across segments) highlights that these features can be used to assess fine motor coordination.

### IV. DISCUSSION

This paper presents a pilot study assessing motor coordination in five children with ASD and five neurotypical controls using surrogate measures derived from multiple modalities, focusing on correlations across and within feature representations of speech subsystems and handwriting movements. We found that eigenvalues derived from correlation structures of time-series of the features, can be used to discriminate between the ASD and control subject groups with an AUC of 1.00 for certain speech tasks and features. Furthermore, our results suggest a characterization of the motor skills of the ASD subjects. We found that correlations of F0 and intensity features derived from the 'pa-ta-ka' speech signal showed higher complexity in the ASD subjects than in the neurotypical controls, indicating decreased precision of movement. On the other hand, ASD subjects showed lower complexity of signals with correlations of features such as formants, FAUs, and handwriting position and pressure accelerations. That features from both speech and handwriting tasks were successful at discriminating between ASD and control groups may reflect similar underlying mechanisms linking fine motor skills to the development of speech production [9], [10].

Overall, the differences in discriminative ability of features across different tasks, as well as the differences in the patterns witnessed for different features, suggests that the effect of limited motor coordination on speech production is highly nuanced and subsystem or task dependent. This agrees with studies that indicate that coordination during speech is task dependent, such as differences in coupling between upper- and lower-lip movement in 2-year-olds during bilabial and "nonlabial" speech tasks [33]. The differences seen across features and tasks may also be influenced by the emotional content present in the task. For example, the discriminative ability of FAUs in The Caterpillar passage may be a reflection of differences in processing and production of prosody and emotion that are witnessed in children with ASD [34], which would be important to separate out from measures of motor skills. This may also help to explain discrepancies seen between studies analyzing acoustic features in ASD [12], such as a wider F0 range witnessed in spontaneous production [35] as opposed to a non-significant difference between the F0 range for children with ASD and neurotypical controls during a sentence imitation task [36].

Utilization of these correlation measures have typically been restricted to formants or FAUs [13]–[15]. Analysis of correlations revealed similar results across speech subsystems and handwriting in support of our hypothesis that

these measures may serve to detect motor coordination differences across speech subsystems and during fine motor tasks. However, using correlations across speech subsystems did not always improve discrimination between the groups, which could be attributed to the difficulty of obtaining additional discriminable information in a small number of subjects. Further exploration is warranted with a larger group of subjects. It would also be important to ensure age matching of ASD and control subjects, as development of speech typically leads to increased independence of articulators [37]. We also want to ensure sex matching and inclusion of more females to ensure that there is no sex bias in the measures being developed.

As we extend the use of these correlation features to function as objective measures of progress related to speech interventions, it will be important to quantify the relationship between the eigenvalue patterns and changes in the underlying signals. For example, we aim to quantify how changes in the frequency spectrum of a signal are directly reflected in changes in the eigenspectra for that signal. We also aim to quantify the contribution of auto-correlation of a signal vs. cross-correlation across signals to the eigenspectra representing the time-delay embedding space of multiple signals. In addition, it will be important to understand how these measures explicitly map to the articulators and muscles forming the signals. For acoustic features, we plan to apply acoustic to articulatory inversion techniques, using the correlation measures on a more direct representation of the articulators, which has shown promising results in studies of major depressive disorder [38].

Our next phase of this study aims to create a multimodal platform to assess motor control and coordination in children with ASD as they undergo speech interventions. With this expansion, we aim to validate our pilot results with a larger number of subjects. We also plan to further quantify the relationship between fine motor skills and speech production, and to analyze longitudinal changes in our correlation measures for an individual to assess their potential for use in clinical settings. In particular, we plan to incorporate eyetracking as another measure of fine motor coordination, given observations of impaired responses to object tracking and saccades in ASD [39]. We plan to work with clinicians to map any insights about lower or higher complexity in speech and handwriting features to existing interventions. This mapping may help clinicians make decisions about treatment trajectories based on the objective features. Furthermore, it could inform any changes they make if they find that a specific intervention is not helping, thereby providing a quantitative way for clinicians to make intervention decisions.

## APPENDIX A
## UNIVARIATE STATISTICS FOR ALL TASKS
This appendix contains tables that describe the statistics of the raw values of the acoustic features extracted for each of the speech tasks.

Table 3 contains the univariate statistics for ASD and control subjects during a reading of The Caterpillar passage.

**TABLE 3.** Univariate summary statistics and p-values of two-sample t-tests comparing values of features extracted from The Caterpillar passage from ASD and control subjects.

| The Caterpillar | | ASD | Control | P-Value |
|---|---|---|---|---|
| **F0** | Mean | 236.23 ± 28.01 Hz | 239.34 ± 12.13 Hz | 0.84 |
| | Var. | 240.23 ± 134.25 Hz | 209.05 ± 83.70 Hz | 0.14 |
| **F1** | Mean | 439.57 ± 54.87 Hz | 465.20 ± 39.10 Hz | 0.46 |
| | Var. | 24.13 ± 5.97 kHz | 27.51 ± 2.88 kHz | 0.34 |
| **F2** | Mean | 1986.70 ± 106.90 Hz | 2033.65 ± 99.28 Hz | 0.52 |
| | Var. | 86.79 ± 38.45 kHz | 90.28 ± 24.55 kHz | 0.88 |
| **F3** | Mean | 3226.81 ± 52.45 Hz | 3145.34 ± 93.99 Hz | 0.14 |
| | Var. | 37.99 ± 22.49 kHz | 46.07 ± 27.89 kHz | 0.64 |
| **Intensity** | Mean | 0.02 ± 0.02 | 0.03 ± 0.01 | 0.49 |
| | Var. | 0.05 ± 0.06 | 0.04 ± 0.01 | 0.69 |

**TABLE 4.** Univariate summary statistics and p-values of two-sample t-tests comparing values of features extracted from the diadochokinetic task from ASD and control subjects.

| Diadochokinetic | | ASD | Control | P-Value |
|---|---|---|---|---|
| **F0** | Mean | 235.89 ± 19.66 Hz | 239.34 ± 12.13 Hz | 0.91 |
| | Var. | 295.67 ± 370.28 Hz | 195.14 ± 88.40 Hz | 0.30 |
| **F1** | Mean | 427.56 ± 76.15 Hz | 404.76 ± 65.78 Hz | 0.54 |
| | Var. | 22.01 ± 12.56 kHz | 29.47 ± 14.08 kHz | 0.11 |
| **F2** | Mean | 1949.64 ± 158.86 Hz | 1906.37 ± 95.85 Hz | 0.06 |
| | Var. | 50.68 ± 30.26 kHz | 35.39 ± 9.78 kHz | 0.03 |
| **F3** | Mean | 3231.04 ± 195.75 Hz | 3133.96 ± 127.69 Hz | 0.29 |
| | Var. | 40.34 ± 16.59 kHz | 26.44 ± 12.95 kHz | 0.93 |
| **Intensity** | Mean | 0.04 ± 0.03 | 0.03 ± 0.02 | 0.55 |
| | Var. | 0.15 ± 0.24 | 0.10 ± 0.10 | 0.26 |

Table 4 contains the univariate statistics for ASD and control subjects during a repetition of the sequence 'pa-ta-ka'.

Table 5 contains the univariate statistics for ASD and control subjects across all sustained vowel tasks.

Table 6 contains the univariate statistics for ASD and control subjects during the free speech tasks.

## APPENDIX B
## FACIAL ACTION UNITS
Table 7 contains the list of the Facial Action Units (FAUs) that were used in analysis.

## ACKNOWLEDGMENT

**TABLE 5.** Univariate summary statistics and p-values of two-sample t-tests comparing values of features extracted from the sustained vowel task from ASD and control subjects.

| Vowels | | ASD | Control | P-Value |
|---|---|---|---|---|
| **F0** | Mean | 237.67 ± 19.91 Hz | 235.49 ± 29.99 Hz | 0.85 |
| | Var. | 46.08 ± 41.05 Hz | 36.99 ± 52.84 Hz | 0.50 |
| **F1** | Mean | 404.62 ± 92.73 Hz | 308.95 ± 41.86 Hz | 0.16 |
| | Var. | 14.64 ± 13.53 kHz | 12.69 ± 14.33 kHz | 0.46 |
| **F2** | Mean | 1815.56 ± 202.87 Hz | 1776.00 ± 148.60 Hz | 0.56 |
| | Var. | 104.32 ± 45.32 kHz | 50.88 ± 47.18 kHz | 0.27 |
| **F3** | Mean | 3265.02 ± 232.57 Hz | 3245.08 ± 96.85 Hz | 0.53 |
| | Var. | 64.91 ± 76.66 kHz | 35.37 ± 26.59 kHz | 0.04 |
| **Intensity** | Mean | 0.07 ± 0.05 | 0.05 ± 0.02 | 0.46 |
| | Var. | 0.05 ± 0.01 | 0.01 ± 0.01 | 0.66 |

**TABLE 6.** Univariate summary statistics and p-values of two-sample t-tests comparing values of features extracted from the free speech task from ASD and control subjects.

| Free Speech | | ASD | Control | P-Value |
|---|---|---|---|---|
| **F0** | Mean | 235.66 ± 8.17 Hz | 235.04 ± 8.25 Hz | 0.90 |
| | Var. | 339.92 ± 125.83 Hz | 334.71 ± 106.79 Hz | 0.69 |
| **F1** | Mean | 452.63 ± 37.17 Hz | 452.63 ± 55.47 Hz | 0.94 |
| | Var. | 31.06 ± 3.59 kHz | 27.83 ± 3.71 kHz | 0.66 |
| **F2** | Mean | 1949.94 ± 36.05 Hz | 1991.72 ± 56.92 Hz | 0.30 |
| | Var. | 98.70 ± 15.36 kHz | 80.79 ± 21.23 kHz | 0.01 |
| **F3** | Mean | 3175.55 ± 21.25 Hz | 3176.23 ± 39.97 Hz | 0.64 |
| | Var. | 41.83 ± 8.16 kHz | 35.39 ± 9.78 kHz | 0.39 |
| **Intensity** | Mean | 0.03 ± 0.01 | 0.03 ± 0.01 | 0.75 |
| | Var. | 0.07 ± 0.04 | 0.06 ± 0.05 | 0.14 |

**TABLE 7.** Table of Facial Action Units (FAUs) derived from OpenFace [30] that were used in analysis.

| FAU Table | |
|---|---|
| FAU # | Description |
| 10 | Upper Lip Raiser |
| 12 | Lip Corner Puller |
| 14 | Dimpler |
| 15 | Lip Corner Depressor |
| 17 | Chin Raiser |
| 20 | Lip Stretcher |
| 23 | Lip Tightener |
| 25 | Lips Part |
| 26 | Jaw Drop |

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.©2020 Massachusetts Institute of Technology. Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (February 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work.

## REFERENCES

[1] V. A. Green, K. A. Pituch, J. Itchon, A. Choi, M. O'Reilly, and J. Sigafoos, "Internet survey of treatments used by parents of children with autism," *Res. Develop. Disabilities*, vol. 27, no. 1, pp. 70–84, Jan. 2006.

[2] R. Paul, "Interventions to improve communication in autism," *Child Adolescent Psychiatric Clinics North Amer.*, vol. 17, no. 4, pp. 835–856, Oct. 2008.

[3] A. N. Bhat, R. J. Landa, and J. C. Galloway, "Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders," *Phys. Therapy*, vol. 91, no. 7, pp. 1116–1129, Jul. 2011.

[4] N. Grace, B. P. Johnson, N. J. Rinehart, and P. G. Enticott, "Are motor control and regulation problems part of the ASD motor profile? A handwriting study," *Develop. Neuropsychol.*, vol. 43, no. 7, pp. 581–594, Oct. 2018.

[5] C. T. Fuentes, S. H. Mostofsky, and A. J. Bastian, "Children with autism show specific handwriting impairments," *Neurology*, vol. 73, no. 19, pp. 1532–1537, Nov. 2009.

[6] R. Booth, R. Charlton, C. Hughes, and F. Happé, "Disentangling weak coherence and executive dysfunction: Planning drawing in autism and attention–deficit/hyperactivity disorder," *Phil. Trans. Roy. Soc. London. B, Biol. Sci.*, vol. 358, no. 1430, pp. 387–392, Feb. 2003.

[7] J. P. McCleery, N. A. Elliott, D. S. Sampanis, and C. A. Stefanidou, "Motor development and motor resonance difficulties in autism: Relevance to early intervention for language and communication skills," *Frontiers Integrative Neurosci.*, vol. 7, p. 30, Apr. 2013.

[8] J. Amato, Jr., and D. Slavin, "A preliminary investigation of oromotor function in young verbal and nonverbal children with Autism," *Infant-Toddler Intervent., Transdisciplinary J.*, vol. 8, no. 2, pp. 84–175, 1998.

[9] M. A. Gernsbacher, E. A. Sauer, H. M. Geye, E. K. Schweigert, and H. H. Goldsmith, "Infant and toddler oral- and manual-motor skills predict later speech fluency in autism," *J. Child Psychol. Psychiatry*, vol. 49, no. 1, pp. 43–50, Jan. 2008.

[10] B. Choi, K. A. Leech, H. Tager-Flusberg, and C. A. Nelson, "Development of fine motor skills is associated with expressive language outcomes in infants at high and low risk for autism spectrum disorder," *J. Neurodevelopmental Disorders*, vol. 10, no. 1, p. 14, Dec. 2018.

[11] M. Mody, A. M. Shui, L. A. Nowinski, S. B. Golas, C. Ferrone, J. A. O'Rourke, and C. J. McDougle, "Communication deficits and the motor system: Exploring patterns of associations in autism spectrum disorder (ASD)," *J. Autism Develop. Disorders*, vol. 47, no. 1, pp. 155–162, Oct. 2016.

[12] R. Fusaroli, A. Lambrechts, D. Bang, D. M. Bowler, and S. B. Gaigg, "'Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis," *Autism Res.*, vol. 10, no. 3, pp. 384–407, Aug. 2016.

[13] J. R. Williamson, D. Young, A. A. Nierenberg, J. Niemi, B. S. Helfer, and T. F. Quatieri, "Tracking depression severity from audio and video based on speech articulatory coordination," *Comput. Speech Lang.*, vol. 55, pp. 40–56, May 2019.

[14] K. M. Smith, J. R. Williamson, and T. F. Quatieri, "Vocal markers of motor, cognitive, and depressive symptoms in Parkinson's disease," in *Proc. 7th Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Oct. 2017, pp. 71–78.

[15] B. S. Helfer, T. F. Quatieri, J. R. Williamson, L. Keyes, B. Evans, W. N. Green, T. Vian, J. Lacirignola, T. Shenk, T. Talavage, and J. Palmer, "Articulatory dynamics and coordination in classifying cognitive change with preclinical mTBI," presented at the 15th Annu. Conf. Int. Speech Commun. Assoc., Sep. 2014.

[16] V. Gracco and A. Lofqvist, "Speech motor coordination and control: Evidence from lip, jaw, and laryngeal movements," *J. Neurosci.*, vol. 14, no. 11, pp. 6585–6597, Nov. 1994.

[17] P. Gramming, J. Sundberg, S. Ternström, R. Leanderson, and W. H. Perkins, "Relationship between changes in voice pitch and loudness," *J. Voice*, vol. 2, no. 2, pp. 118–126, Jan. 1988.

[18] C. Lord, M. Rutter, P. C. DiLavore, S. Risi, K. Gotham, and S. L. Bishop, *Autism Diagnostic Observation Schedule*, 2nd ed. Torrance, CA, USA: Western Psychological Corporation, 2012.

[19] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, 5th ed. Arlington, TX, USA: American Psychiatric Publishing, 2013.

[20] R. Patel, K. Connaghan, D. Franco, E. Edsall, D. Forgit, L. Olsen, L. Ramage, E. Tyler, and S. Russell, "'The caterpillar': A novel reading passage for assessment of motor speech disorders," *Amer. J. Speech-Lang. Pathol.*, vol. 22, no. 1, pp. 1–9, Feb. 2013.

[21] K. E. Beery, *Beery VMI: The Beery-Buktenica Developmental Test of Visual-Motor Integration*. Minneapolis, MN, USA: Pearson, 2004.

[22] S. Yamartino. (2020). *Pressure.js*. Accessed: Nov. 11, 2019. [Online]. Available: https://pressurejs.com/index.html

[23] B. L. Brooks, E. M. Sherman, and E. Strauss, "NEPSY-II: A developmental neuropsychological assessment," *Child Neuropsychol.*, vol. 16, no. 1, pp. 80–101, 2009.

[24] J. Hagedorn, J. Hailpern, and K. G. Karahalios, "VCode and VData: Illustrating a new framework for supporting the video annotation workflow," in *Proc. Work. Conf. Adv. Vis. Interfaces (AVI)*. New York, NY, USA: ACM, 2008, pp. 317–321.

[25] D. D. Mehta, D. Rudoy, and P. J. Wolfe, "Kalman-based autoregressive moving average modeling and inference for formant and antiformant tracking," *J. Acoust. Soc. Amer.*, vol. 132, no. 3, pp. 1732–1746, Sep. 2012.

[26] M. Brookes. (2020). *VOICEBOX: A Speech Processing Toolbox for MATLAB*. Accessed: Nov. 11, 2019. [Online]. Available: https://github.com/ImperialCollegeLondon/sap-voicebox

[27] D. N. Sorenson, "A fundamental frequency investigation of children ages 6–10 years old," *J. Commun. Disorders*, vol. 22, no. 2, pp. 115–123, Apr. 1989.

[28] R. L. Horwitz-Martin, T. F. Quatieri, E. Godoy, and J. R. Williamson, "A vocal modulation model with application to predicting depression severity," in *Proc. IEEE 13th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jun. 2016, pp. 247–253.

[29] A. Röbel and X. Rodet, "Efficient spectral envelope estimation and its application to pitch shifting and envelope preservation," in *Proc. 8th Int. Conf. Digit. Audio Effects*, Madrid, Spain, 2005, pp. DAFX1–DAFX6.

[30] T. Baltrusaitis, A. Zadeh, Y. C. Lim, and L.-P. Morency, "OpenFace 2.0: Facial behavior analysis toolkit," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 59–66.

[31] R. Ekman, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. New York, NY, USA: Oxford Univ. Press, 1997.

[32] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digit. Signal Process.*, vol. 10, nos. 1–3, pp. 19–41, Jan. 2000.

[33] J. L. Ruark and C. A. Moore, "Coordination of lip muscle activity by 2-year-old children during speech and nonspeech tasks," *J. Speech, Lang., Hearing Res.*, vol. 40, no. 6, pp. 1373–1385, Dec. 1997.

[34] R. B. Grossman and H. Tager-Flusberg, "'Who said that?' matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD," *J. Autism Develop. Disorders*, vol. 42, no. 12, pp. 2546–2557, Mar. 2012.

[35] J. J. Diehl and R. Paul, "Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders," *Appl. Psycholinguistics*, vol. 34, no. 1, pp. 135–161, Jan. 2013.

[36] J. J. Diehl and R. Paul, "Acoustic differences in the imitation of prosodic patterns in children with autism spectrum disorders," *Res. Autism Spectr. Disorders*, vol. 6, no. 1, pp. 123–134, Jan. 2012.

[37] J. R. Green, C. A. Moore, M. Higashikawa, and R. W. Steeve, "The physiologic development of speech motor control," *J. Speech, Lang., Hearing Res.*, vol. 43, no. 1, pp. 239–255, Feb. 2000.

[38] C. Espy-Wilson, A. C. Lammert, N. Seneviratne, and T. F. Quatieri, "Assessing neuromotor coordination in depression using inverted vocal tract variables," in *Proc. Interspeech*, Sep. 2019, pp. 1448–1452.

[39] B. P. Johnson, J. A. G. Lum, N. J. Rinehart, and J. Fielding, "Ocular motor disturbances in autism spectrum disorders: Systematic review and comprehensive meta-analysis," *Neurosci. Biobehav. Rev.*, vol. 69, pp. 260–279, Oct. 2016.

**TANYA TALKAR** received the B.S. and M.Eng. degrees in computer science and electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2016 and 2017, respectively. She is currently pursuing the Ph.D. degree in the speech and hearing bioscience and technology program with Harvard University, Cambridge.

She is also conducting Research with the Human Health and Performance Systems Group, MIT Lincoln Laboratory, Lexington, MA, USA. She is under an NIH T32 Training Grant. Her research interests include assessment of motor coordination, and development of neurocomputational models for detection, and monitoring of neurological and neurotrauma conditions, such as autism spectrum disorder and mild traumatic brain injury.

Ms. Talkar is a member of the Tau Beta Pi, the Eta Kappa Nu, ASA, and NSSLHA. She received the MIT EECS-Finn Undergraduate Research and Innovation Scholar Award and the National Merit Scholarship.

**JAMES R. WILLIAMSON** received the B.S. degree in psychology from the University of Massachusetts, Amherst, MA, USA, in 1987, and the Ph.D. degree in cognitive and neural systems from Boston University, Boston, MA, USA, in 1995.

From 1996 to 2000, he was a Postdoctoral Fellow and a Research Assistant Professor in cognitive and neural systems, investigating self-organizing neural network models of vision and pattern recognition. Since 2001, he has been with the MIT Lincoln Laboratory, where he is currently a Technical Staff Member with the Human Health and Performance Systems Group. His research interests include detecting and modeling changes in neurocognitive and physical state through analysis of speech, facial expressions, gait, eye movements, EEG, fMRI, and other body sensing modalities.

Dr. Williamson has been part of teams with the MIT Lincoln Laboratory that took first place in the AVEC Depression Sub-Challenges, in 2013 and 2014, and was part of a team that received the 2015 Lincoln Laboratory Team Award for Research and Development on Vocal and Facial Biomarkers. He also received the Best Reviewer Award from the IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, in 2014.

**DANIEL J. HANNON** received the B.A. degree in psychology from the Nazareth College, Rochester, NY, USA, in 1985, and the M.Sc. and Ph.D. degrees in experimental psychology from Brown University, Providence, RI, USA, in 1987 and 1991, respectively.

He is currently a Research Psychologist with the Human Health and Performance Systems Group, MIT Lincoln Laboratory. He is also a Licensed Mental Health Practitioner. He has worked extensively with patients on the Autism spectrum and their families. His research interests include mental health-related research and development topics.

**HRISHIKESH M. RAO** received the B.E. degree *(summa cum laude)* in biomedical engineering from the University of Minnesota, Twin Cities, MN, USA, and the Ph.D. degree in biomedical engineering from Duke University, Durham, NC, USA, in 2016.

He is currently a Technical Staff Member with the MIT Lincoln Laboratory, Lexington, MA, USA. His research interests include analysis of physiological signals, gathered using wearable technology, for estimating physiological and cognitive state, measurement and mathematic modeling of eye movements to assess neurological motor dysfunction or injury, and disease states.
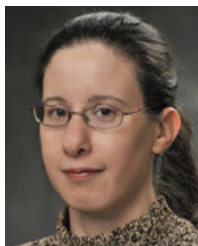
**SOPHIA YUDITSKAYA** received the B.S., M.Eng., and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2002, 2005, and 2010, respectively.

At MIT, she conducted thesis research applying machine learning models to detect a preverbal child's emotional state using vocal attributes in a dense, naturalistic, longitudinal audio dataset, in order to facilitate the study of emotional factors in early child development. She is currently an Associate Member of technical staff with the MIT Lincoln Laboratory working with diverse biosignals, including vocal, facial, gait, oculomotor, and EEG. She is developing machine learning analytics that aim to facilitate diagnosis, biomarker identification, severity assessment, and decision support for targeted therapies. Her primary research interests include neurocognitive modeling to understand the processes, underlying sensorimotor coordination, and neuroplasticity for neurological disorder characterization and rehabilitation.

**KAJAL T. CLAYPOOL** received the B.E. degree in computer engineering from the Manipal Institute of Technology, Manipal, India, the Ph.D. degree in computer science from the Worcester Polytechnic Institute, Worcester, MA, USA, in 2002, and the postdoctoral master's degree in biomedical informatics from the Harvard Medical School, Boston, MA, USA, in 2017.

She was a Technical Staff Member with Oracle Corporation and an Assistant Professor with the Department of Computer Science, University of Massachusetts Lowell. She was active with the Database Research Community for more than 20 years. She is currently a Senior Staff Member with the Human Health and Performance Systems Group. She is particularly interested in using statistics and machine learning to investigate the pseudo-intervention of physical activity on obesity, aging, and Alzheimer's disease. She is also actively engaged in research and technology development in big data management. She has authored numerous publications in these and related areas. Her current research interests include translational informatics to understand the associations between environment, omics, disease, information integration, information retrieval, data analytics, and secure architectures.

**DOUGLAS STURIM** received the B.S.E.E. degree from Washington University in St. Louis, St. Louis, MO, USA, in 1985, the M.S. degree in electrical engineering from the Rensselaer Polytechnic Institute, Troy, NY, USA, in 1992, and the Ph.D. degree in electrical engineering from Brown University, Providence, RI, USA, in 1999 (Navy Fellowship).

In 1999, he joined the Artificial Intelligence Technology and Systems Group, as a Technical Staff Member. He was also a Technical Staff Member with TRW, from 1985 to 1990. He performed research in sonar tracking and estimation with the Naval Undersea Warfare Center, from 1990 to 1999. He has contributed numerous publications to journals and conferences. His research interests include robust speaker and language identification.

Dr. Sturim is a member of the Sigma Xi. He received the Outstanding Student Award from the Rensselaer Polytechnic Institute.

**LISA NOWINSKI** received the bachelor's degree in psychology and spanish from the University of the Pacific, Stockton, CA, USA, in 2004, and the Ph.D. degree in counseling, clinical, and social psychology from the University of California Santa Barbara, Santa Barbara, CA, USA, in 2009.

She held a Predoctoral Internship and a Postdoctoral Fellowship in pediatric neuropsychology with the Massachusetts General Hospital Learning and Emotional Assessment Program (LEAP). As the Director of Clinical Psychology with the Lurie Center for Autism, she is responsible for all clinical psychology services and training. She directs a robust assessment and therapy program. She runs a competitive Postdoctoral Fellowship-offering specialized training to clinical psychologists who seek to build expertise in the assessment and care of individuals with ASD. She specializes in diagnostic and neuropsychological assessment of individuals with autism spectrum disorder across the lifespan. Her work focuses on collaborative models of assessment that involve parents and caregivers as active partners throughout the assessment process.

Dr. Nowinski is also a member of the Autism Treatment Network's Behavioral Scientist Committee, where she has worked with professionals from around the country on several initiatives, including the development of a Puberty Toolkit for teens and families with ASD.

**HANNAH SARO** was born in Boston, MA, USA, in 1991. She received the B.S. degree in social work from Bridgewater State University, Bridgewater, MA, USA, in 2014. She is currently pursuing the M.S. degree in child and developmental psychology with Southern New Hampshire University, Manchester, NH, USA.

From 2014 to 2017, she was a Research and a Staff Assistant. Since 2017, she has been a Clinical Research Coordinator with the MGH Lurie Center for Autism, Lexington, MA, USA. Her research interests include autism spectrum disorders, Angelman syndrome, and other developmental disorders.

Ms. Saro holds her Professional Licensure from the Board of Social Work, as a Licensed Social Worker (LSW).

**CAROL STAMM** received the B.S. degree in biological sciences from SUNY Binghamton, Binghamton, NY, USA.

She was a Clinical Research Coordinator for many years, having spent the majority of them at the University of Rochester Medical Center, Rochester, NY, USA. She is currently a Clinical Research Coordinator with the Lurie Center. She has clinical research trials experience in oncology, dementia, spinal cord injury, endocrinology, and neurodevelopmental paediatrics, including autism which provides her with a solid foundation to continue her passion for working with and supporting the children and the families living with autism.

**MARIA MODY** received the Ph.D. degree in speech and hearing sciences from The City University of New York, New York, NY, USA, in 1993, and the Ph.D. degree (Training) in speech language pathology from the Albert Einstein College of Medicine, The Bronx, NY, USA.

With early funding from the Human Frontier Science Program, she focused on speech development in infants with ENST, Paris, France. She is currently a Cognitive Neuroscientist and a Faculty Member with the Martinos Center for Biomedical Imaging, Massachusetts General Hospital, and the Harvard Medical School. She is an Editor of the book *Neural Mechanism of Language* (Springer, 2017). Her research interests include developing neuroimaging paradigms to examine spoken and written language processes in normal and atypical development.

**CHRISTOPHER J. MCDOUGLE** received the B.A. degree in chemistry from Valparaiso University, Valparaiso, IN, USA, in 1981, and the M.D. degree in medicine from the Indiana University School of Medicine, Indianapolis, IN, USA, in 1986.

From 1986 to 1990, he was a Resident in psychiatry with the Yale University School of Medicine, New Haven, CT, USA. From 1992 to 1995, he was a Child and Adolescent Psychiatry Fellow with the Yale Child Study Center, New Haven. From 1990 to 1997, he was an Assistant Professor and an Associate Professor of psychiatry with the Yale University School of Medicine. From 1997 to 2011, he was an Endowed and a tenured Professor of psychiatry with the Indiana University School of Medicine and the Chairman of the Department of Psychiatry, from 2000 to 2011. Since 2011, he has been the Director of the Lurie Center for Autism, Massachusetts General Hospital, and the Nancy Lurie Marks Professor in autism with the Harvard Medical School, Boston, MA, USA. He is an Editor or Co-Editor of three books and more than 300 peer-reviewed articles. He has received significant and ongoing research funding from the National Institutes of Health, Bethesda, MD, USA. His research interests include neuropsychopharmacology of autism spectrum disorder, other neurodevelopmental disorders, and immune-nervous system interaction during development as it pertains to autism and related disorders.

Dr. McDougle was a recipient of the Frank J. Menolascino Award for Psychiatric Services for Persons with Mental Retardation/Developmental Disabilities, American Psychiatric Association, in 2009, and the George Tarjan Award for Contributions in Developmental Disabilities, American Academy of Child and Adolescent Psychiatry, in 2012. He serves as an Associate Editor for *Journal of Autism and Developmental Disorders*.

**THOMAS F. QUATIERI** (Fellow, IEEE) received the B.S. degree *(summa cum laude)* from Tufts University, Medford, MA, USA, in 1973, and the S.M., E.E., and Sc.D. degrees from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1975, 1977, and 1979, respectively.

He is currently a Technical Senior Member with the MIT Lincoln Laboratory, Lexington, focused on speech and auditory signal processing and neuro-biophysical modeling with application to detection and monitoring of neurological, neurotraumatic, and stress conditions. He holds a Faculty Appointment with the Harvard-MIT Speech and Hearing Bioscience and Technology Program. He is the author of more than 200 publications, holds 12 patents, and has authored the textbook *Discrete-Time Speech Signal Processing: Principles and Practice*.

Dr. Quatieri is a member of the Tau Beta Pi, the Eta Kappa Nu, the Sigma Xi, ICSA, ARO, and ASA. He has served on many the IEEE Signal Processing and the Speech Technical Committees. He was a recipient of the four IEEE TRANSACTIONS Best Paper Awards and the 2010 MIT Lincoln Laboratory Best Paper Award. He led the Lincoln Laboratory Team that took first place in the 2013 and 2014 AVEC Depression Challenges and that received the 2015 MIT Lincoln Laboratory Team Award for their work on vocal and facial biomarkers. He serves as an Associate Editor for *Computer Speech and Language*.

● ● ●