



Detection of Subclinical Mild Traumatic Brain Injury (mTBI) Through Speech and Gait

Tanya Talkar^{1,2}, Sophia Yuditskaya², James R. Williamson², Adam Lammert³, Hrishikesh Rao², Daniel Hannon², Anne O'Brien⁴, Gloria Vergara-Diaz⁴, Richard DeLaura², Douglas Sturim², Gregory Ciccarelli², Ross Zafonte⁴, Jeffrey Palmer², Paolo Bonato⁴, Thomas F. Quatieri^{1,2}

¹Speech and Hearing Bioscience and Technology, Harvard University, Boston, MA, USA

²MIT Lincoln Laboratory, Lexington, MA, USA

³Worcester Polytechnic Institute, Worcester, MA, USA

⁴Spaulding Rehabilitation Hospital, Boston, MA, USA

ttalkar@g.harvard.edu, {sophia.yuditskaya, jrw}@ll.mit.edu, alammert@wpi.edu, {hrishikesh.rao, daniel.hannon}@ll.mit.edu, {aobrien1, gvergaradiaz}@partners.org, {richd, sturim, gregory.ciccarelli}@ll.mit.edu, rzafonte@mgh.harvard.edu, jspalmer@ll.mit.edu, pbonato@mgh.harvard.edu, quatieri@ll.mit.edu

Abstract

Between 15% to 40% of mild traumatic brain injury (mTBI) patients experience incomplete recoveries or provide subjective reports of decreased motor abilities, despite a clinically-determined complete recovery. This demonstrates a need for objective measures capable of detecting subclinical residual mTBI, particularly in return-to-duty decisions for warfighters and return-to-play decisions for athletes. In this paper, we utilize features from recordings of directed speech and gait tasks completed by ten healthy controls and eleven subjects with lingering subclinical impairments from an mTBI. We hypothesize that decreased coordination and precision during fine motor movements governing speech production (articulation, phonation, and respiration), as well as during gross motor movements governing gait, can be effective indicators of subclinical mTBI. Decreases in coordination are measured from correlations of vocal acoustic feature time series and torso acceleration time series. We apply eigenspectra derived from these correlations to machine learning models to discriminate between the two subject groups. The fusion of correlation features derived from acoustic and gait time series achieve an AUC of 0.98. This highlights the potential of using the combination of vocal acoustic features from speech tasks and torso acceleration during a simple gait task as a rapid screening tool for subclinical mTBI.¹

Index Terms: biomedical application, acoustic analysis, mild traumatic brain injury, motor coordination, machine learning

1. Introduction

Approximately 1.7 million people in the United States suffer a mild traumatic brain injury, primarily the result of events such as a sports injury, fall, or car accident [1, 2]. While most report full recoveries, between 15% to 40% of mTBI patients report incomplete recoveries and chronic subclinical impairments, typically related to decreased balance confidence, despite having clinical assessment scores that identify them as fully recovered

¹DISTRIBUTION 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.

[3]. Standard clinical assessments include clinical interviews, neuroimaging, physical examination of postural control, coordination, and vestibular deficits, as well as neuropsychological batteries to evaluate cognitive dysfunction [4, 5, 6]. In mission-critical military environments, such clinical assessments are often impracticable when making the decision to return warfighters to battle or military training [7]. In this paper, we address the need to develop sensitive and objective biomarkers utilizing fast, noninvasive recording capabilities to screen patients for incomplete mTBI recoveries.

Analysis of speech has shown promise as a noninvasive method of detecting neurological trauma. Features extracted from vowels recorded on a mobile device by athletes participating in a boxing match were able to predict the presence of a concussion with an accuracy of 98% [8]. Another study was able to achieve an AUC of 0.86 by combining temporal and spectral features from recordings of sentences and words collected from high school athletes [9]. In addition, proxy measures of motor coordination, derived from correlations of acoustic feature time series, have also shown promise in estimating cognitive task performance, which is degraded in patients with mTBI [10, 11]. Features representing articulatory coordination, derived from formant time series, were able to detect changes in cognitive performance for high school athletes with preclinical mTBI [10]. A combination of articulatory and facial expression feature correlations was applied toward Processing Speed Index scores of mTBI patients [11]. These studies highlight the promise of speech tasks as a non-invasive screening tool for mTBI and return-to-activity readiness.

Changes in gait have also been identified as potential markers of recovery from mTBI. Ready-to-play athletes who have clinically recovered from mTBI have demonstrated increased variability in their leg swing time, as well as increased time to reach stability during a walking task [12]. These findings suggest that increased variability in head-trunk coordination is present even in subclinical mTBI. Furthermore, coordination between the hip, the knee, and the ankle joints was found to be more variable in mTBI patients during a simple walking task [13]. Gait coordination, therefore, shows promise as an additional marker of mTBI recovery.

In this paper, we build upon these previous studies utilizing acoustic and gait features and apply these techniques to individ-

uals who have lingering subclinical impairments from an mTBI. We record a battery of speech tasks through a mobile device and collect torso acceleration during a gait task. We utilize correlation measures to quantify the coordination in three speech production subsystems - articulation, phonation, and respiration - as well as in gait movements. We use eigenvalues extracted from the correlation measures as inputs into two machine learning model architectures to discriminate between mTBI patients and healthy controls. To the best of our knowledge, this work represents the first joint use of speech and gait motor features for assessment of lingering, subclinical injury from an mTBI.

2. Data collection

2.1. Participants

Subjects were recruited as part of a larger, multi-modal collection effort at MIT Lincoln Laboratory to assess balance in mTBI patients. Ten healthy controls (four females; age 32.4 ± 14.0 years) and eleven subjects with a previous mTBI (five females; age 39.1 ± 17.18 years) participated in this study. Following the American Congress of Rehabilitation Medicine (ACRM) definition, a subject was grouped into the mTBI category if the subject had a Glasgow Coma Scale (GCS) of 13-15 or a post-traumatic amnesia not greater than 24 hours [14]. Subjects with neurological injury less than 1 week prior to the day of the experiment were excluded. Clinical assessments showed no deficits in areas of strength, range of motion, proprioception, and balance. However, all mTBI subjects complained of residual balance issues and decreased balance confidence from the previous injury. Healthy controls were recruited to match the sex, age, height, and weight distribution of the mTBI subject pool. Subjects that had other major neurological or psychiatric diseases were excluded. The experiment was approved by MIT's Committee on the Use of Humans of Experimental Subjects (COUHES) Institutional Review Board. All subjects signed an informed consent form prior to participation in the study.

2.2. Protocol

Speech, facial video, and walking accelerometry data was collected from subjects at the MIT Lincoln Laboratory Sensorimotor Technology Realization in Immersive Virtual Environments (STRIVE) Center. Speech was collected on an iPad (Apple Inc., Cupertino, CA) at 44100 Hz. All instructions for the speech tasks were presented using a custom-built application on the iPad. Subjects first read The Rainbow passage. They then were asked to repeat the diadochokinetic (DDK) sequence "pa-ta-ka" as many times as possible in a single breath. Finally, they were asked to describe the situation presented in two separate pictures. As part of a larger assessment of balance, subjects were asked to walk at a natural pace on a treadmill in the STRIVE center. The Vicon motion capture system (Vicon Motion Systems Ltd., Oxford, UK), with eight infrared cameras, was used to record the 3-D positional signals at 100 Hz of reflective markers which were placed on anatomical markers. Two sessions of gait were collected for each subject (Session 1 Duration: 320.1 ± 131.6 s, Session 2 Duration: 315.9 ± 127.1 s).

3. Feature extraction & classification

3.1. Low-level feature extraction

A speech activity detection algorithm developed in MATLAB eliminated beginning and ending pauses from all speech task

recordings. Four different acoustic feature time series and their delta-time series were extracted from each recording. Fundamental frequency (F0) [15] and Mel-frequency Cepstral Coefficient (MFCC) time series were extracted using the Praat software package at 1000 Hz and 200 Hz respectively [16]. The first three formant time series (F1-F3) were extracted using the Kalman-based autoregressive moving average (KARMA) software tool at 100 Hz [17]. KARMA provides a continuous time series of formants using an energy-based voice detector that allows a Kalman smoother to estimate formants through silent gaps in the signal. The envelope of the signal was extracted at 100 Hz using a custom MATLAB script that provides a smooth contour of amplitude peaks based on an iterative time-domain signal envelope estimation [18]. This technique captures both the effect of resonance-harmonic interactions and of the respiratory muscles on amplitude modulation of a speech envelope.

Torso accelerations were extracted using the Vicon system using 3-D positional signals collected at 100 Hz from a reflective marker positioned on the sternum. A Laplacian filter, $[-1, 2, -1]\delta_t$, where $\delta_t = 0.01s$, was convolved with the time series of position values in each 3-D axis. To reduce noise in the acceleration values, Accelerations with absolute value greater than $2g/\delta_t$ were removed and replaced with a linearly interpolated value. Next, the acceleration time series were smoothed in the time domain by convolving each time series with a Gaussian filter with $\sigma = 3\delta_t$. High level features were extracted independently from 20 s frames of the 3-D smoothed acceleration signals, with 10 s overlap between successive frames.

3.2. High-level feature extraction

Multivariate auto- and cross-correlations of acoustic and gait low-level features were used as proxy measures of the motor coordination within and across the underlying speech subsystems and gross motor systems [10, 11, 19]. Specifically, time-delay embedding was used to expand the dimensionality of the acoustic and walking time series, resulting in correlation matrices that represent coupling strengths across feature channels at multiple relative time delays. The eigenspectra of these channel-delay correlation matrices quantify and summarize the frequency properties of the set of feature signals [19].

Channel-delay correlation matrices for speech tasks were calculated for 8 combinations of features: F0, envelope, MFCC, formants, $F0 \times$ envelope, $F0 \times$ formants, envelope \times formants, $F0 \times$ envelope \times formants. Correlation matrices were also calculated using their delta-time series equivalents, leading to 16 different feature combinations for each speech task. When correlated with F0, the envelope and formant time series were interpolated to 1000 Hz using spline interpolation. An automatic masking technique was used to include only voiced segments, relying on a Voice Activity Detection time series from Praat. Correlation matrices for the gait task were calculated across the torso acceleration time series.

Each correlation matrix contains the correlation coefficients between the low-level feature time series at defined time delays, creating the embedding space (Figure 1). For each feature or combination of features, four matrices were constructed at four delay scales with delay spacings of 1, 3, 7, and 15 data samples for the acoustic feature time series, and of 3, 7, 15, and 31 data samples for the torso acceleration time series. These delay scales allow the feature coupling patterns to be characterized at multiple time scales. Each matrix comparing n signals had dimensionality $n*15 \times n*15$. For each speech task, there were 16 different sets of matrices, representing each combination of fea-

tures. The gait task had one set of matrices. Eigenvalues from all resulting matrices were extracted in rank-order. Eigenvalues from each delay scale for a single task and feature combination were concatenated to form a final feature vector with $n * 15 * 4$ elements, which was used for classification. The eigenvalues from individual delay scales were also used for characterization through analysis of Cohen’s d effect size patterns.

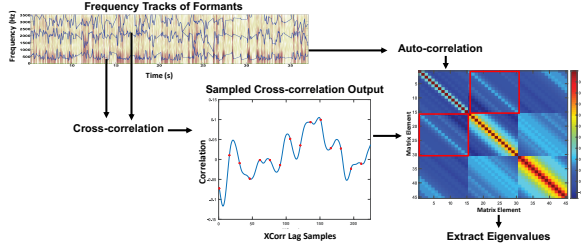


Figure 1: Example of the high-level feature extraction using auto- and cross-correlations across F1, F2, and F3 time series.

3.3. Model architecture

Two different machine learning models were used to classify subjects into the healthy control or mTBI groups. Both models used 10-fold cross validation, with a control and mTBI subject left out in each fold. In both models, the model output score for free speech was the sum of model output scores of a subject over both free speech tasks. The model output score for the gait task was the average score across all 20s segments. Model output scores were used to compare model performance using the area under the receiver operating characteristic curve (AUC).

3.3.1. Gaussian Mixture Model

For each task and feature, the top two principal components from principal component analysis (PCA) were extracted from the concatenated eigenspectra feature vector (Figure 2). For each cross-validation fold, an ensemble of ten Gaussian Mixture Models (GMMs) was created using the PCA features from the training set. Supervised adaptation of each GMM was used to create separate control and mTBI GMMs [20]. The likelihood of a training subject belonging to each group was the log of the sum of the likelihoods of the subject for the control GMMs or mTBI GMMs in the ensemble. The final model output score was the ratio of those log-likelihoods.

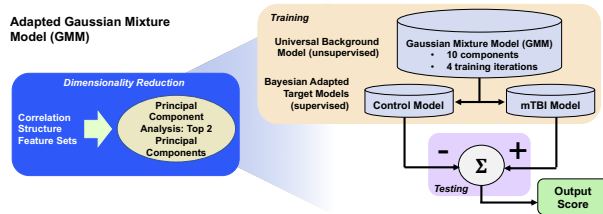


Figure 2: Model architecture of the Gaussian Mixture Models.

3.3.2. Convolutional Neural Network

The Convolutional Neural Network (CNN) consisted of two convolutional layers, a max-pooling layer, and a final fully-connected linear layer (Figure 3). The concatenated eigenspectra feature vectors for each task and feature were normalized

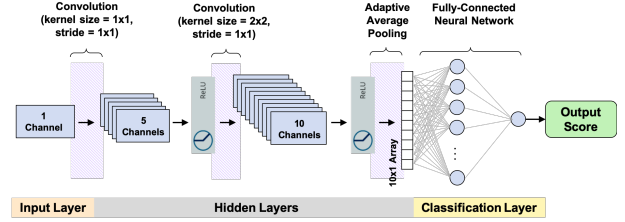


Figure 3: Model architecture of the Convolutional Neural Network.

within each training fold. The model was trained for 100 epochs with a batch size of 4 subjects, balanced to include two control and two mTBI subjects. The model used binary cross entropy loss and the Adam optimizer to update gradients. The output from the final layer was passed through a softmax layer to get a model output score for each subject. Ten models were created for each cross-validation fold of a task and feature, and the final model output score was the average score over all of the models for each subject.

3.3.3. Fusion

Features were fused if they produced sufficient cross-validation accuracy within each training fold. Specifically, model output scores and training set AUCs were saved for each subject, feature, and task combination. Fusion was first calculated for a given task across all features. For a given subject, a mask was created across all features. The mask value for a held out test subject and feature was 1 if the associated training set AUC was greater than 0.7. Model output scores for a subject were summed across all features in that task that met the criterion, and then divided by the number of included features. This is described in eq. (1), where s refers to a specific subject, t refers to a specific task, and f refers to a specific feature.

$$score_{s,t} = \frac{\sum_{f \in Features} mask_{s,t,f} * score_{s,t,f}}{\sum_{f \in Features} mask_{s,t,f}} \quad (1)$$

These scores were used to determine the AUC of the fusion of models for a single task. The sum of these scores across all tasks for a subject was used as the fusion score for a fusion across all tasks and features, and subsequently used to determine the AUC. Fusions were performed independently for GMM and CNN models.

4. Results

4.1. Model performance

Table 1 lists the AUC values for fusions of features within a task as well as across speech and walking tasks. AUCs above 0.8 are in bold and the standard error is denoted. Fusing across all speech tasks led to an AUC of 0.84 with the GMM and 0.96 with the CNN. A fusion of all tasks led to an AUC of 0.98 from the GMM and 0.94 from the CNN. Performance on individual speech tasks was higher with the CNN, however, performance on gait and fused tasks was comparable between the CNN and GMM approaches.

4.2. Characterization of speech and gait

Figure 4 depicts the Cohen’s d effect sizes for the eigenspectra extracted from correlations of formants (delay scale: 150 ms; 15

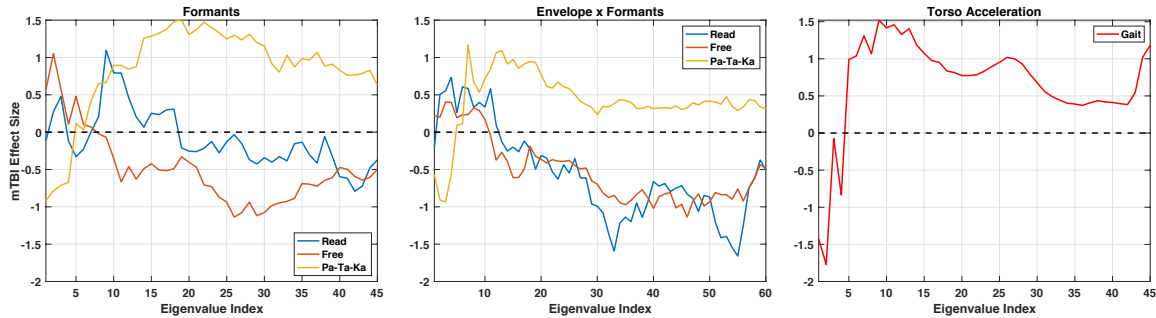


Figure 4: Cohen’s d effect sizes of eigenspectra extracted from correlations of formants (left), speech envelope and formants (middle), and torso accelerations (right).

Table 1: AUC performance on fused tasks with standard error. Any AUC ≥ 0.8 is **bolded**.

Task Fusion	GMM	CNN
Free Speech (FS)	0.76 \pm 0.11	0.89 \pm 0.07
Read Speech (RS)	0.76 \pm 0.11	0.90 \pm 0.07
’pa-ta-ka’ (DDK)	0.67 \pm 0.12	0.89 \pm 0.07
Gait (G)	0.85 \pm 0.09	0.87 \pm 0.08
FS and RS and DDK	0.84 \pm 0.09	0.96 \pm 0.05
FS and RS and DDK and G	0.98 \pm 0.03	0.94 \pm 0.05

samples), speech envelope and formants (delay scale: 70 ms; 7 samples), and torso acceleration features (delay scale: 70 ms; 7 samples). These features yielded AUCs ≥ 0.80 in both the CNN and the GMM models for individual tasks. Effect sizes > 0 indicate that the eigenvalue was larger for the mTBI group. The eigenvalues are ranked from largest to smallest. A pattern of positive effect sizes for larger eigenvalues and negative effect sizes for smaller eigenvalues indicates higher complexity of signals for the control group, and the opposite pattern indicates higher complexity of signals for the mTBI group [19].

In read and free speech, we see lower complexity of formants and the interaction between the speech envelope and formants for mTBI subjects in free speech and during the Rainbow passage. This suggests that mTBI subjects have more coupling between their articulators during reading and free speech, as opposed to the complexity that control subjects have to reach acoustic targets. However, we see higher complexity in these features from the mTBI subjects during the recitation of ’pa-ta-ka’. This suggests less precision in articulator movements for mTBI subjects during the ’pa-ta-ka’ task as compared to control subjects. This also extends to the gait task, where there is higher complexity in torso accelerations for mTBI subjects as compared to control subjects, suggesting that movements for the mTBI group are not as controlled nor precise.

5. Discussion

In this paper, we described a simple protocol and analytical approach to detect subclinical mTBI, which can be used for return-to-duty decisions for warfighters or return-to-play decisions for sports players. Utilizing correlation features across vocal acoustic and torso acceleration time series yielded an AUC of 0.98 with a GMM and an AUC of 0.94 using a CNN. The high performance of both of these models further validates the utility of using features from simple speech and gait tasks as markers in

detecting mTBI, agreeing with the evidence provided in previous studies [8, 9, 10, 11, 12, 13]. It also highlights that lingering motor coordination issues from an mTBI may not be detectable by existing clinical tests used to determine recovery.

Analysis of individual features provided insight into the differences of coordination between the mTBI and control subject groups. The mTBI group had lower complexity of formant time series and of speech envelope and formant time series as compared to the control group during the free speech and read speech tasks, suggesting more coupling between these time series. On the other hand, mTBI subjects had less precision and more independence of these features in the ’pa-ta-ka’ task as well as of torso accelerations during the gait task. The difference seen between these tasks suggests different task-dependent demands of motor coordination [21]. The difference in these demands, therefore, potentially manifest in different ways in mTBI subjects as compared to control subjects. Further validation is required to see whether these patterns are witnessed in a larger group of subjects. Future work will also focus on longitudinal analysis of these patterns from acute mTBI through the recovery stage, which may aid clinicians in guiding therapy and recovery strategies [22].

We utilized complexity of feature time series as a proxy for motor coordination, which has been used in previous studies [10, 11, 19]. However, to further characterize articulator movement and better understand speech motor coordination differences between the mTBI patients and healthy individuals, we hope to use acoustic-to-articulatory methods to approximate the movements of the underlying articulators. We can use the articulator time series to assess articulator coordination through the correlation approach, which has already shown promise in characterization of Major Depressive Disorder (MDD) [23].

Our speech protocol can be recorded on a mobile device, which allows for the protocol to be conducted in low-resource settings with relative ease. Mobile devices also contain accelerometers, which offer easy collection and analysis of torso acceleration data in the field [24] and can be used to characterize and identify an individual’s gait [25]. We plan to assess whether accelerometry data collected by strapping the device to the torso can provide the same insights into motor coordination as the data collected in this study using the Vicon motion capture system. In this way, the protocol could be administered completely using a mobile device for use in any setting.

6. Acknowledgements

TT was supported by NIH-NIDCD under Grant T32DC000038.

7. References

- [1] M. Faul, L. Xu, M. Wald, and V. Coronado, "CDC-TBI in the US Report-Traumatic Brain Injury-Injury Center," *Centers for Disease Control and Prevention*, 2010.
- [2] E. Finkelstein, P. S. Corso, and T. R. Miller, *The incidence and economic burden of injuries in the United States*. Oxford University Press, USA, 2006.
- [3] M. Peterson, B. D. Greenwald *et al.*, "Balance problems after traumatic brain injury," *Archives of physical medicine and rehabilitation*, vol. 96, no. 2, pp. 379–380, 2015.
- [4] C. on the Review of the Department of Veterans Affairs Examinations for Traumatic Brain Injury, "Diagnosis and Assessment of Traumatic Brain Injury," *Evaluation of the Disability Determination Process for Traumatic Brain Injury in Veterans*, 2019.
- [5] C. Prince and M. E. Bruhns, "Evaluation and Treatment of Mild Traumatic Brain Injury: The Role of Neuropsychology," *Brain Sciences*, vol. 7, no. 8, 2017.
- [6] K. I. Ustinova, C. L. A. A. Dull, and J. Perkins, "Physical therapy for correcting postural and coordination deficits in patients with mild-to-moderate traumatic brain injury," *Physiotherapy Theory and Practice*, vol. 31, no. 1, 2015.
- [7] R. E. Furger, L. D. Nelson, E. B. Lerner, and M. A. McCrea, "Frequency of factors that complicate the identification of mild traumatic brain injury in level I trauma center patients," *Concussion*, vol. 1, no. 2, 2016.
- [8] M. Falcone, N. Yadav, C. Poellabauer, and P. Flynn, "Using isolated vowel sounds for classification of mild traumatic brain injury," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 7577–7581.
- [9] L. Daudet, N. Yadav, M. Perez, C. Poellabauer, S. Schneider, and A. Huebner, "Portable mTBI assessment using temporal and frequency analysis of speech," *IEEE journal of biomedical and health informatics*, vol. 21, no. 2, pp. 496–506, 2016.
- [10] B. S. Helfer, T. F. Quatieri, J. R. Williamson, L. Keyes, B. Evans, W. N. Greene, T. Vian, J. Lacirignola, T. Shenk, T. Talavage *et al.*, "Articulatory dynamics and coordination in classifying cognitive change with preclinical mTBI," in *Fifteenth Annual Conference of the International Speech Communication Association*, 2014.
- [11] A. C. Lammert, J. R. Williamson, A. Hess, T. Patel, T. F. Quatieri, H. J. Liao, A. Lin, and K. J. Heaton, "Noninvasive estimation of cognitive status in mild traumatic brain injury using speech production and facial expression," in *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2017, pp. 105–110.
- [12] K. C. Powers, J. M. Kalmar, and M. E. Cinelli, "Dynamic stability and steering control following a sport-induced concussion," *Gait & posture*, vol. 39, no. 2, pp. 728–732, 2014.
- [13] H.-L. Chen, T.-W. Lu, and L.-S. Chou, "Effect of concussion on inter-joint coordination during divided-attention gait," *Journal of medical and biological engineering*, vol. 35, no. 1, pp. 28–33, 2015.
- [14] T. Kay, D. E. Harrington, R. Adams, T. Anderson, S. Berrol, K. Cicerone, C. Dahlberg, D. Gerber, R. Goka, P. Harley *et al.*, "Definition of mild traumatic brain injury," *Journal of Head Trauma Rehabilitation*, vol. 8, no. 3, pp. 86–87, 1993.
- [15] P. Boersma, "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," in *Institute of Phonetic Sciences*, vol. 17. University of Amsterdam, 1993, pp. 97–110.
- [16] P. Boersma and D. Weenink, "Praat: doing phonetics by computer, version 6.0.37." 2018. [Online]. Available: <http://www.praat.org/>
- [17] D. D. Mehta, D. Rudoy, and P. J. Wolfe, "Kalman-based autoregressive moving average modeling and inference for formant and antiformant tracking," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1732–1746, 2012.
- [18] R. L. Horwitz-Martin, T. F. Quatieri, E. Godoy, and J. R. Williamson, "A vocal modulation model with application to predicting depression severity," in *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 2016, pp. 247–253.
- [19] 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," *Computer Speech & Language*, vol. 55, pp. 40–56, 2019.
- [20] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digital signal processing*, vol. 10, no. 1-3, pp. 19–41, 2000.
- [21] W. Ziegler, "Task-related factors in oral motor control: Speech and oral diadochokinesis in dysarthria and apraxia of speech," *Brain and language*, vol. 80, no. 3, pp. 556–575, 2002.
- [22] C. Noufi, A. C. Lammert, D. D. Mehta, J. R. Williamson, G. Ciccarelli, D. Sturim, J. R. Green, T. F. Campbell, and T. F. Quatieri, "Vocal Biomarker Assessment Following Pediatric Traumatic Brain Injury: A Retrospective Cohort Study," *Proc. Interspeech 2019*, pp. 3895–3899, 2019.
- [23] C. Espy-Wilson, A. C. Lammert, N. Seneviratne, and T. F. Quatieri, "Assessing Neuromotor Coordination in Depression Using Inverted Vocal Tract Variables," *Proc. Interspeech 2019*, pp. 1448–1452, 2019.
- [24] J. R. Williamson, A. Dumas, G. Ciccarelli, A. R. Hess, B. A. Telfer, and M. J. Buller, "Estimating load carriage from a body-worn accelerometer," in *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. IEEE, 2015, pp. 1–6.
- [25] F. Juefei-Xu, C. Bhagavatula, A. Jaech, U. Prasad, and M. Savvides, "Gait-id on the move: Pace independent human identification using cell phone accelerometer dynamics," in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2012, pp. 8–15.