Symptom Severity and Acoustically-Identified Groups of Speakers with and without Autism Spectrum Disorder

Ethan Weed[1], Riccardo Fusaroli[1], Jessica Mayo[2], Inge-Marie Eigsti[3]

[1] Aarhus University, [2] Yale School of Medicine, [3] University of Connecticut

Background: The speech of people with Autism Spectrum Disorder (ASD) has often been described as atypical, but there is little consensus on what acoustic features constitute the atypicality (Fusaroli et al., 2017). This raises the possibility that acoustically-different profiles of atypical speech may exist within ASD speech, as well as the possibility that these profiles may correspond to measures of symptom severity.

Objectives: Study goals were (1) to identify potential sub-groups of typical and atypical speakers, using acoustic features of prosody and voice, and (2) to investigate how acoustically-identified groups of speakers with and without ASD relate to symptom severity scores on the Social Responsiveness Scale (Constantino & Gruber, 2012).

Methods: We analyzed speech recordings (8 scripted sentences per participant) from 15 adolescents diagnosed with ASD (mean(SD) age = 14.4 (1.48) years) with IQ scores in the typical range, and 15 adolescents with typical development (TD; mean(SD) age = 14.1(1.91) years); groups did not differ on chronological age or full-scale IQ. Participants in both the ASD and the NT groups demonstrated average to high average performance on standardized language measures (see Mayo, 2015, for details). From these speech recordings, we extracted acoustic measures of prosody and voice quality, and used a community-detection algorithm (Csardi, M. G.,2013; Reichardt & Bornholdt, 2006) to find groups of speakers who shared a common acoustic profile. We then compared these groups on all subscales of the SRS.

Results: The community-detection algorithm identified four groups of speakers whose acoustic profiles were more similar to each other than to members of the other groups (Figure 1 panel A). Groups 2 and 4 consisted primarily of TD participants (81%), and were distinguished acoustically by differences in speech rate and utterance duration. Groups 1 and 3 consisted primarily of ASD participants (79%), and were distinguished from each other primarily by diverging patterns of H1H2 (associated with breathiness, (Hillebrand et al, 1994)), with group 3 showing higher mean H1H2 and variation in H1H2, and group 1 showing lower mean H1H2, than any other group. These four groups were significantly related to performance on all the SRS subscales and on total SRS score ($F(3, 25) = 6.31$, $p = 0.00245$), with groups 1 and 3 rating higher on all SRS subscales than groups 2 and 4. Group 3 in particular was distinguished by a slower speech rate, longer utterances, and a higher variation of breathiness.

Conclusions: We identified four acoustic profiles among the speakers, and these profiles related significantly to the SRS subscales; the highest SRS scores were seen in the group with the most breathiness and greatest variation in breathiness. It is likely that these acoustic measures contribute to community perceptions of atypicality (Sasson et al, 2017) and ASD "frankness" (de Marchena & Miller, 2016). Results provide a foundation for exploration of how (and whether) to intervene with prosodic and other speech qualities.
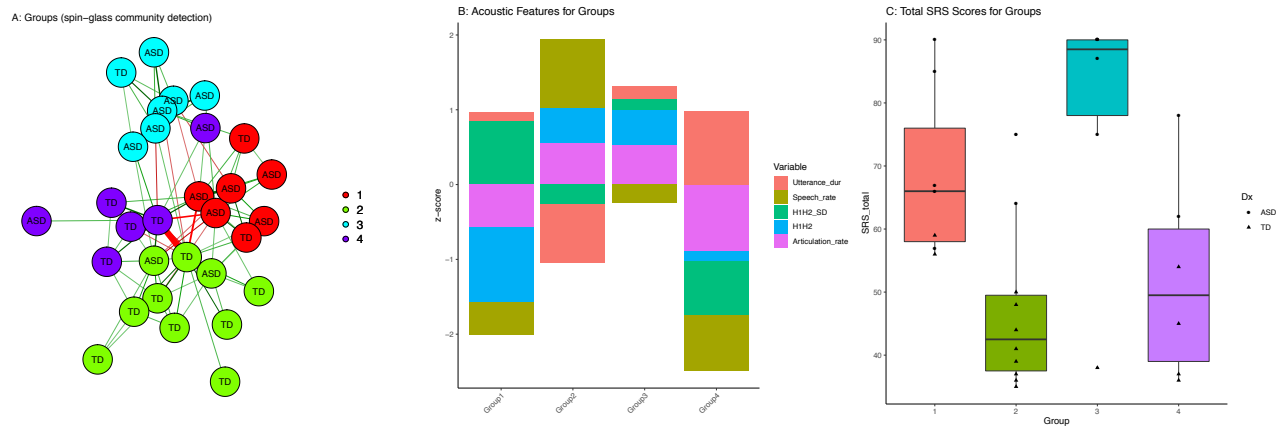
Figure 1: Network of participants, grouped by acoustic profile (A), acoustic features defining profiles (B), and SRS scores grouped by acoustic profile (C)