USC University of Southern California

USC Viterbi School of Engineering

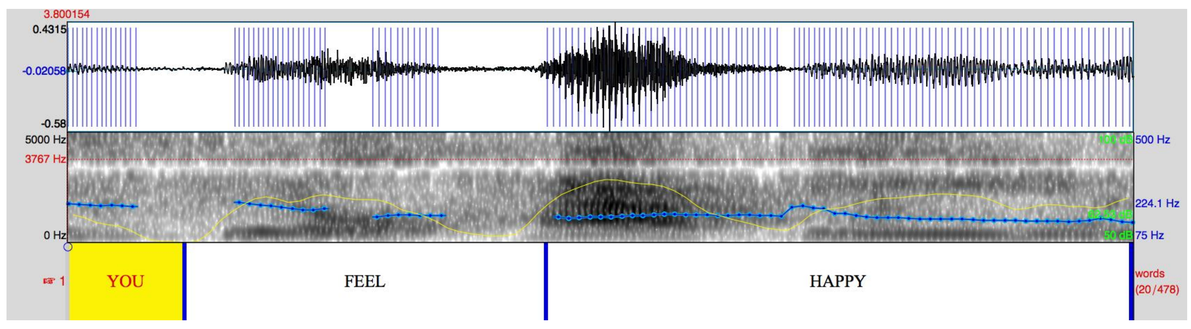
Ming Hsieh Department of Electrical Engineering

Spontaneous-Speech Acoustic-Prosodic Features Of Children with Autism and the Interacting Psychologist

Daniel Bone[1], Matthew P. Black[1], Chi-Chun Lee[1], Marian E. Williams[2], Pat Levitt[3], Sungbok Lee[1], Shrikanth Narayanan[1]
USC: SAIL[1], University Center for Excellence in Developmental Disabilities[2], Zilka Neurogenic Institute[3]

Motivation & Introduction

Prosody- Intonation, volume, duration, and rate of speech
suprasegmental aspects have communicative function
segmental aspects relate to voice quality



Atypical prosody is often reported in children with **Autism Spectrum Disorders (ASD)**, but heterogeneous and qualitatively described

"slow, rapid, jerky and irregular in rhythm, odd intonation or inappropriate pitch and stress, markedly flat and toneless, or consistently abnormal volume"

We investigate various word- and phonetic- level spontaneous speech features to quantify the *qualitatively described* atypical prosody

Additionally, we find the *psychologist's* acoustic-prosodic features inform their perception of the child's behavior

The USC CARE Corpus

Category	Count/Statistic
Age (years)	mean: 9.8, std. dev.: 2.5, range: 5.8-14.7
Gender	male: 22, female: 6
Native language	Spanish: 8, English: 9, Sp.&Eng.: 4, unk: 7
Ethnicity	Hispanic/Latino: 20, White/White+Other: 8
ADOS module	#3: 28
ADOS diagnosis	autism: 17, ASD: 5, below ADOS cutoffs: 6

Acoustic-Prosodic Features

25 features (functionals on feature contours) per person, per session

Intonation and Volume (turn-end prosody) (12 functionals):
2nd-order polynomial (intercept, slope, and curvature) of pitch and intensity

Rate (9 functionals):
Syllabic speaking rate, vowel and consonant duration

Voice Quality (4 functionals):
Jitter and Shimmer- peak-to-peak variations in pitch period and amplitude

Analysis of Acoustic-Prosodic Features

Child's Acoustic-Prosodic Features

Spearman's rank correlation coefficients ($p<0.05$)

	Code Label: Atyp. Pros.	Comm. Total	Soc.Int. Total	C&SI Total
Child's Acoustic-Prosodic Features				
f0_slope μ	-0.45	$\mathbf{-0.57}$	$\mathbf{-0.50}$	$\mathbf{-0.56}$
f0_curve μ		-0.46	-0.41	-0.45
Int_intercept σ		$+0.39$		
Jitter *median*	$+0.42$	$+0.39$	$+0.41$	$+0.41$
Jitter *iqr*	$\mathbf{+0.55}$	$+0.47$	$\mathbf{+0.48}$	$\mathbf{+0.50}$
syl_SR-nonBoundary $q_{0.9}$		-0.41		

Both participant's features correlate with the child's rated atypicality

Child features- suggest 'monotonic' speech, variable volume, atypical voice quality, and slower rate of speech.

Psychologist features- suggest psychologist's speech behavior changes depending on her perception of the child (e.g., higher jitter and generally slower speech rate).

Psychologist's Acoustic-Prosodic Features

Spearman's rank correlation coefficients ($p<0.05$)

	Code Label: Atyp. Pros.	Comm. Total	Soc.Int. Total	C&SI Total
Psychologist's Acoustic-Prosodic Features				
f0_slope μ		$+0.38$		
f0_intercept σ	$+0.44$	$\mathbf{+0.62}$	$+0.40$	$+0.47$
f0_slope σ		$+0.47$		
f0_curve σ	$+0.42$	$\mathbf{+0.58}$		$+0.39$
Jitter *median*	$\mathbf{+0.53}$	$\mathbf{+0.77}$	$\mathbf{+0.58}$	$\mathbf{+0.69}$
Jitter *iqr*	$+0.46$	$\mathbf{+0.57}$	$+0.39$	$+0.47$
syl_SR-Boundary $q_{0.9}$		$+0.46$		
syl_SR-nonBoundary $q_{0.9}$			$\mathbf{-0.48}$	-0.43
vowel_dur σ		$\mathbf{+0.59}$		

Predictive Tasks

Psychologist's acoustic-prosody is **more informative** of the child's rated ASD severity than the child's features based on multiple linear regression prediction.

We can potentially leverage this info to model interaction strategies.

Correlations of prosodic feature sets' predictions with ADOS code labels. [,**,***]≡ α=[0.10,0.05,0.01]*

Child's Acoustic-Prosodic Feature	Code Label: Atyp. Pros.	Comm. Total	Soc.Int. Total	C&SI Total
Child		$\mathbf{0.36^{*}}$		$\mathbf{0.37^{*}}$
Psychologist		$\mathbf{0.61^{***}}$	$\mathbf{0.61^{***}}$	$\mathbf{0.45^{**}}$
Both		$\mathbf{0.63^{***}}$		$\mathbf{0.50^{***}}$

Discussion & Future Work

The results suggest the **psychologist is attuning** to the child's behavioral cues, deliberately or spontaneously.

Future work

Model the temporal patterning of interaction

Is atypical prosody global (thin-slices) or local (bouts)?

Model strategies of the psychologist

Collect normative data from typically developing children to model non-linear variability in speech prosody

`http://sail.usc.edu`

This work was supported in part by NSF and NIH.

Ming Hsieh Institute
Ming Hsieh Department of Electrical Engineering