Linguistic Markers of Autism Spectrum Disorder:

Classification Sensitivity and Specificity of Language Produced During Clinical Evaluations

Julia Parish-Morris¹, Christopher Cieri², Mark Liberman², Neville Ryant², & Robert T. Schultz¹

¹Center for Autism Research, CHOP
²Linguistic Data Consortium, UPenn
• Why we are interested

• The CAR/LDC ADOS Project

• Four Features and Clinical Correlates within ASD

• Future Directions
• Why we are interested

• The CAR/LDC ADOS Project

• Four Features and Clinical Correlates within ASD

• Future Directions
• Natural language
  • Highly nuanced outward signal of internal brain activity
  • Fundamentally social

• Most children with ASD acquire language; nearly all vocalize

• Can applying HLT and Big Data methods help us reliably identify and understand ASD?

Why we are interested
• **Variable vocalization** throughout development:
  • Differences evident in infancy
  • Language delay as toddlers/preschoolers
  • Difficulty being understood/trouble understanding humor and sarcasm
  • Conversational quirks (unusual word use, turn-taking, synchrony, accommodation)

• **Real-life Effects** of pragmatic language problems:
  • Difficulty forming/maintaining friendships
  • Increased risk of being bullied
  • Difficulty with romantic relationships
  • Difficulty maintaining employment
• Many small variations accumulate to create an odd impression.
• It’s hard to “put your finger on” what exactly differs, so it’s tricky to treat!

Language in ASD
• Natural language:
  • Nuanced signal (marriage of cognitive and motoric systems)
  • No practice effects

• Can identify and extract features ("linguistic markers")

• Specific linguistic features associated with:
  • Depression
  • Dementia
  • PTSD
  • Schizophrenia

• …Autism

Clinical computational linguistics
On average, individuals with ASD:

- Produce idiosyncratic or unusual words more often than typically developing peers (Ghaziuddin & Gerstein, 1996; Prud’hommeaux, Roark, Black, & Van Santen, 2011; Rouhizadeh, Prud’Hommeaux, Santen, & Sproat, 2015; Rouhizadeh, Prud’hommeaux, Roark, & van Santen, 2013; Volden & Lord, 1991)

- Repeat words or phrases more often than usual (echolalia; van Santen, Sproat, & Hill, 2013)

- Use filler words “um” and “uh” differently than matched peers (Irvine, Eigsti, & Fein, 2016)

- Wait longer before responding in the course of conversation (Heeman, Lunsford, Selfridge, Black, & Van Santen, 2010)

- Produce speech that differs on pitch variables; these can be used to classify samples as coming from children with ASD or not (Asgari, Bayestehtashk, & Shafran, 2013; Kiss, van Santen, Prud’hommeaux, & Black, 2012; Schuller et al., 2013)
• Why we are interested

• The CAR/LDC ADOS Project

• Four Features and Clinical Correlates within ASD

• Future Directions
• Process and analyze recorded language samples from Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012)
  • Conversation and play-based assessment of autism symptoms
  • Recorded for reliability and clinical supervision, coded on a scale, then filed away

• 600+ at CAR alone, thousands more across the U.S. and in Europe; never compiled

• Associated with rich metadata that includes family history, social, cognitive, and behavioral phenotype, genes, and neuroimaging

ADOS Project
Goals of pilot effort:

• Assess feasibility

• Identify and extract linguistic features

• Machine learning classification

• Correlate features with clinical phenotype

ADOS Project
• Time aligned, verbatim, orthographic transcripts (~20 minutes of conversation)

• New transcription specification developed by LDC resembles those used for conversational speech

• 4 transcribers and 2 adjudicators from LDC and CAR produced a “gold standard” transcript for analysis and for evaluation/training of future transcriptionists

• Simple comparison of word level identity between CAR’s adjudicated transcripts and LDC’s transcripts: 93.22% overlap on average, before a third adjudication resolved differences between the two

• Transcripts force-aligned to audio

Transcription
• Pilot sample
• N=100
• Mean age=10-11 years
• Primarily male
• 65 ASD, 18 TD, 17 Non-ASD mixed clinical
• Average full scale IQ, verbal IQ, nonverbal IQ
• Why we are interested

• The CAR/LDC ADOS Project

• Four Features and Clinical Correlates within ASD

• Future Directions

Today's Talk
• Which words did participants use most frequently?

• 20 most “ASD-like” words:
  • \{nsv\}, know, he, a, now, no, uh, well, is, actually, mhm, w-, years, eh, right, first, year, once, saw, was
  • \{nsv\} stands for “non-speech vocalization”, meaning sounds that with no lexical counterpart, such as imitative or expressive noise
  • “uh” appears in this list, as does “w-”, a stuttering-like disfluency.

• 20 least “ASD-like” words:
  • like, um, and, hundred, so, basketball, something, dishes, go, york, or, if, them, \{laugh\}, wrong, be, pay, when, friends.
  • “um” appears, as does the word “friends, and laughter

Word Choice
• Decreased use of “friend” words correlates with increased social impairment as rated by clinicians, Pearson’s $r = -.35$, $p = .03$. 

**Word Choice and Clinical Phenotype in ASD**
• Word choice correctly classified 68% of ASD participants and 100% of typical participants

• Naïve Bayes, leave-one-out cross validation and weighted log-odds-ratios calculated using the “informative Dirichlet prior" algorithm (Monroe et al., 2008)

• Receiver Operating Characteristic (ROC) analysis revealed good sensitivity and specificity; AUC=85%

Classification: Word Choice
• Rates of um production across the ASD and TD groups (um/(um+uh))

• ASD group produced UM during 61% of their filled pauses (CI: 54%-68%)

• TD group produced UM during 82% of their filled pauses (CI: 75%-88%)

• Minimum value for the TD group was 58.1%, and 23 of 65 participants in the ASD group fell below that value.

Fluency Differences
• Increased “um” use in ASD participants correlated with decreased ADOS comparison scores
  • Spearman’s rho = -.25, \( p = .05 \)

• This was due to significant sex differences in um/uuh use; controlling for sex eliminated the effect

• Typical sex differences in um/uuh use (Wieling et al., 2016) may be preserved in ASD

**Fluency and Clinical Phenotype in ASD**
• Mean word duration as a function of phrase length

• TD participants spoke the fastest
  (overall mean word duration of 376 ms, CI 369-382, calculated from 6891 phrases)

• Followed by the non-ASD mixed clinical group (mean=395 ms; CI 388-401, calculated from 6640 phrases)

• Followed by the ASD group with the slowest speaking rate (mean=402 ms; CI: 398-405, calculated from 24276 phrases)
• Faster speech associated with higher verbal IQ
  • Spearman’s rho = -.26, $p = .04$

Rate and Clinical Phenotype in ASD
• Gap between speaker turns
• Too short = interrupting or speaking over a conversational partner
• Too long (awkward silences) interrupt smooth social exchanges
• ASD slower than TD

Differences in Latency to Respond
Latency to Respond and Clinical Phenotype in ASD

- Longer latency to respond associated with more social impairment according to clinical ratings (ADOS social affect score)
  - Spearman’s rho = .33, \( p = .007 \)
• Median absolute deviation from the median (MAD)
  • Outlier-robust measure of dispersion in F0 distribution
  • Calculated in semitones relative to speaker’s 5th percentile

• MAD values are both higher and more variable within the ASD and non-ASD mixed clinical group than the TD group
  • ASD: median: 1.99, IQR: 0.95
  • Non-ASD: median: 1.95, IQR: 0.80
  • TD: median: 1.47, IQR: 0.26

**Fundamental Frequency**
• MAD associated positively with clinician ratings of social impairment, Pearson’s $r = .27$, $p = .03$

• …and negatively with parent reported adaptive functioning in the communication domain, Pearson’s $r = -.29$, $p = .02$
• ASD and TDC differ on a variety of linguistic features

• Features correlate with clinician ratings of social impairment, as well as with parent report of adaptive functioning

• Emerging collaborations include more ADOS recordings with phenotypic data, neuroimaging, and genetics from heterogeneous samples (including mixed clinical and more females with ASD)

• BUT these particular language samples are still expensive and time-consuming to obtain

• **NEED**: Scalable, inexpensive methods to collect natural language from large, diverse samples

**Discussion**
• Why we are interested

• The CAR/LDC ADOS Project

• Four Features and Clinical Correlates within ASD

• Future Directions

Today's Talk
• **Phone bank**
  • Inexpensive student worker asks ADOS questions
  • Child and parent language samples, questionnaires, online IQ
  • Nationally representative cohort

• **Computerized Social Affective Language Task (C-SALT)**
  • Completely self-contained app (UNITY)
  • Records language and social affect in schools, clinics and homes
  • Controlled recording is conducive to automated approaches (reduces need for transcription)

• **Combine data sources to improve predictive power**
  • Motor, language, medical records, parent/teacher report, clinical judgment, performance tasks, imaging, genetics

**Future Directions**
• **Support clinical decision-making and improve access**
  - Low-cost, remote screening
  - Direct behavioral observation: record in clinics, integrate into EHR
  - Inform identification efforts and assist in differential diagnosis

• **Identify behavioral markers of underlying (treatable) pathobiology**
  - Profiles of individual strengths and weaknesses ➔ link to biology = personalized treatment planning and improved outcomes
  - Granular assessment of response to intervention – dense sampling

• **Give participants and families more information about themselves**
  - Online feedback
  - Monitor growth trajectories

**Applications**
• Participants and families!
• CAR and LDC clinicians and staff
  • Special thanks to Leila, Emily & Cait
• Key collaborators
  • Bob Schultz
  • Mark Liberman
  • Chris Cieri
  • Neville Ryant (I still don’t have a pic)
• Funding sources
  • Autism Science Foundation
  • McMorris Autism Program
  • NIH K12

Acknowledgements
Questions?