Why Develop Language Resources for Autism?

Julia Parish-Morris, PhD
Assistant Professor
Center for Autism Research, CHOP
Perelman School of Medicine, UPenn

3rd International Symposium on Linguistic Diversity, Language Resources and Clinical Research

November 10, 2020
Today’s talk

Autism Spectrum Disorder (ASD)
• Overview
• Why we need large shared resources

A new kind of measurement

Example results
Today’s talk

- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources

- A new kind of measurement

- Example results
What is autism?

- Neurodevelopmental condition

- Behaviorally defined
  - Diagnosed using *behavior only*
  - No genetic test
  - No brain scan

- Symptom severity lies on a continuum - *Autism Spectrum Disorder*
Core features

- Impaired social communication
- Repetitive behaviors and restricted interests
  - Present since early childhood
  - Interferes with everyday functioning
Who has autism?

- Estimated prevalence approximately 1-1.5%
  - 1 in 54 U.S. school children (CDC, 2020)

- ~4:1 boy:girl ratio
  - Traditionally thought to vary with IQ
Today’s talk

- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources

- A new kind of measurement

- Example results
Why do we need large shared resources?
1. Diagnosis is expensive and difficult

- “Gold Standard” in the U.S. – expert clinician consensus
  - Autism Diagnostic Observation Schedule (ADOS)

- Based on observable behavior using human judgment

- Problem: imperfect agreement (kappa = .69)
Heterogeneity

• Symptoms vary between individuals
• Symptoms vary within an individual over the course of a lifetime
• ...and sometimes over the course of a day, or an hour!

Common co-occurring conditions - ASD rarely occurs alone
• Seizures, anxiety, ADHD, OCD, Tourette syndrome, language disorders, learning disorders, intellectual disability
Why do we need large shared resources?

3. Insufficiently granular measurement

- Current diagnosis and characterization methods are:
  - Expensive – small samples
  - Complicated, time-consuming
  - Rely on human judgment of behavior

- Mismatch
  - Rich genetic or imaging data maps to restricted yes/no dx category

- Need: highly quantifiable, fine-grained signal that is robust to practice effects
Why do we need large shared resources?

3. Insufficiently granular measurement

- Current diagnosis and characterization methods
  - Expensive – small samples
  - Complicated, time-consuming
  - Rely on human judgment of behavior

- Mismatch
  - Rich genetic or imaging data maps to restricted yes/no dx category

- Need: highly quantifiable, fine-grained signal that is robust to practice effects
Behavioral heterogeneity + small samples + poor measurement =

less-reproducible scientific results
suboptimal evidence base for interventions
worse outcomes
How to quantify autism?

- Autism manifests in the context of live social interaction (2 people)

- Need: High-dimensional, scalable method to capture time-synced human signals from interacting partners

  - Result: Precise behavioral characterization of two interacting systems
Quantifying social interaction

- Turn social interaction into *numbers*

- What you **do** (motor) and what you **say** (language)

  - **Motor**: computer vision
  - **Language**: computational linguistics
Analyzing the vocal signal: Challenges

- Language is highly multivariate (acoustics, words, grammar, conversational dynamics)

- “Normal” changes across development and sometimes across cultures

- Neurodevelopmental/psychiatric conditions have different profiles
  - Opportunity to create personalized profiles with different treatment indications
Today’s talk

- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources

- A new kind of measurement

- Example results
Dyadic biosensor

U.S. patent pending, J. Parish-Morris (Co-Inventor)

- Multi-channel directional microphones for automated analyses
- Video, audio, heart rate, skin conductance, accelerometers
- "Shovel ready" for machine learning

Fig. 1 The Biosensor captures everything participants say and do with perfect synchronization.
Today’s talk

- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources

- A new kind of measurement

- Example results
Example Study

**Goal:** Quantify restricted/repetitive behavior during naturalistic conversation using *computational linguistics* and *computer vision*

**Compare** behavioral diversity/entropy in adults with and without ASD in the domains of:

1. Language
2. Oral-motor movement
Participants

- Forty-four consenting adults, all native English speakers
  - Autism Spectrum Disorder (ASD): N=17
  - Typically developing (TD): N=27
- Diagnosed using according to DSM-5 criteria, informed by the Autism Diagnostic Observation Schedule – 2nd Edition

<table>
<thead>
<tr>
<th>Variable</th>
<th>ASD Mean (SD)</th>
<th>TD Mean (SD)</th>
<th>Statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>26.9 (7.3)</td>
<td>28.1 (8.4)</td>
<td>W = 234</td>
<td>0.923</td>
</tr>
<tr>
<td>Sex (Male, Female)</td>
<td>15, 2</td>
<td>23, 4</td>
<td>$\chi^2$: 0.08</td>
<td>0.774</td>
</tr>
<tr>
<td>Full-Scale IQ</td>
<td>102.1 (19.8)</td>
<td>111.7 (9.5)</td>
<td>W = 157</td>
<td>0.080</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>112.6 (22.1)</td>
<td>112.4 (11.2)</td>
<td>W = 215</td>
<td>0.736</td>
</tr>
<tr>
<td>ADOS Total</td>
<td>13.1 (3.0)</td>
<td>1.1 (0.9)</td>
<td>W = 442</td>
<td>&lt; 2e-8*</td>
</tr>
<tr>
<td>ADOS Social Affect</td>
<td>9.8 (2.3)</td>
<td>1.0 (0.9)</td>
<td>W = 442</td>
<td>&lt; 1e-8*</td>
</tr>
<tr>
<td>ADOS RRB</td>
<td>3.3 (1.5)</td>
<td>0.1 (0.3)</td>
<td>W = 441</td>
<td>&lt; 1e-9*</td>
</tr>
</tbody>
</table>

Paradigm

Contextual Assessment of Social Skills (CASS)\(^1\)

3-minute semi-structured assessment of conversational ability designed to mimic real-life first-time encounters
- Framed as “getting to know each other”; no specific prompts provided

CASS confederates:
- 10 undergraduate students or BA-level research assistants
- Trained to speak for no more than 50% of the time
- Wait 10s to initiate the conversation; wait 5s before re-initiating conversation after pauses

Lexical pipeline

- Verbatim, orthographic, time-aligned transcription of utterances by participant and confederate
- Reliable, blinded annotators using Xtrans\textsuperscript{1}
- All spoken words included, no stemming, stopwords remain
  - Diversity includes morphological differences like “want” and “wanted”
- Total word count and entropy calculated per speaker using qdap\textsuperscript{2}

Oral-motor pipeline

- 3 steps: Face detection, face registration, and movement quantification

- Detection and localization of landmarks (eyes, lip corners, nose etc.)
  - Publicly available tool (OpenFace)

- Registration
  - Part-based registration
  - Video stabilization to eliminate jitter

- Quantification
  - Facial Bases method
  - 60 mouth bases
  - Normalized the total activation count of each basis by the maximum count observed for the same basis of confederates

---

Entropy: the amount of information a data modality carries

Shannon entropy\(^{1}\) where \(b\) is the base of the logarithm \((b=2;\) our measure of entropy is in bits\)

\[
H = - \sum_{i=1}^{n} p(x_i) \log_b p(x_i)
\]

- **High entropy** is expected when participants make a rich set of facial expressions and produce a variety of words while speaking
- **Low entropy** is expected when participants generate a restricted set of mouth movements and produce repetitive speech patterns

Tests:
- Wilcoxon rank sum tests with continuity correction; exploratory correlation analyses
- Linear mixed models (lme4) or simple linear models in R
  - Random effects of confederate identity and fixed effects of sex, age, and IQ checked for significance; excluded when non-significant
  - Facial analyses included speaking length and head motion as covariates

Results: Lexical entropy

- Reduced entropy in participants with ASD as compared to TD participants, $t(42)=2.85$, $p=0.007$, Cohen’s $d=0.82$
  - The effect of diagnosis on entropy was significant after accounting for age, IQ, and gender, $t(39)=3.25$, $p=0.002$
  - Diversity of confederate language did not differ by participant diagnosis, $t(35.26)=0.17$, $p=0.86$
- There was a (neurotypical) association between word count and entropy$^{1,2}$
  - A second model tested the interactive effect of word count and diagnosis on participant lexical diversity
- The slope of the relationship between word count and diversity was greater in the TD group than the ASD group, interaction $t=-3.51$, $p=0.001^*$

Results: Oral-motor entropy

- Reduced mouth movement diversity in the ASD group as compared to the TD group (Cohen's $d=1.0$, $t=-2.73$, $p=0.009$)
  - Model included head movement and speech length as covariates
  - Difference remained significant when age, sex, and IQ were included as covariates (Cohen's $d=1.0$, $t=-2.52$, $p=0.016$)
  - No covariates contributed significantly to the model
Discussion

- **Take home**: Reduced behavioral diversity, across domains, captures an underlying dimension of restriction and repetition that distinguishes autistic adults from typical controls.

- Restriction in mouth movement (motor) not driven by restriction in words produced (cognitive) – uncorrelated – contributing unique variance.
Future Research

- **Build** a large, shared resource of ASD conversations at LDC to accelerate the pace of discovery
- **Test** real-world effects of subtle linguistic differences in ASD (e.g., likelihood of referral, peer friendships)
- **Explore** linguistic markers of complex phenotypes (e.g., depression, ASD + anxiety, ASD + depression or ADHD)
- **Develop** targeted social communication interventions that are *personalized* for the unique challenges faced by a given individual
Acknowledgements

- Participants and families
- CAR clinicians, students, and staff

- Key collaborators:
  - Bob Schultz, Director of CAR
  - Mark Liberman & Chris Cieri & Sunghye Cho, LDC @ Upenn
  - Clare Harrop, UNC Psychiatry
  - Joe Donaher, CHOP Center for Childhood Communication
  - Ani Nenkova, UPenn Computer & Information Science
  - Ted Brodkin, UPenn Psychiatry
  - Jami Young, CHOP PolicyLab

- Funding sources
  - Autism Science Foundation
  - McMorris Family Foundation
  - NIMH R01, NIDCD R03, Roche Ltd, CHOP RI, EAC