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

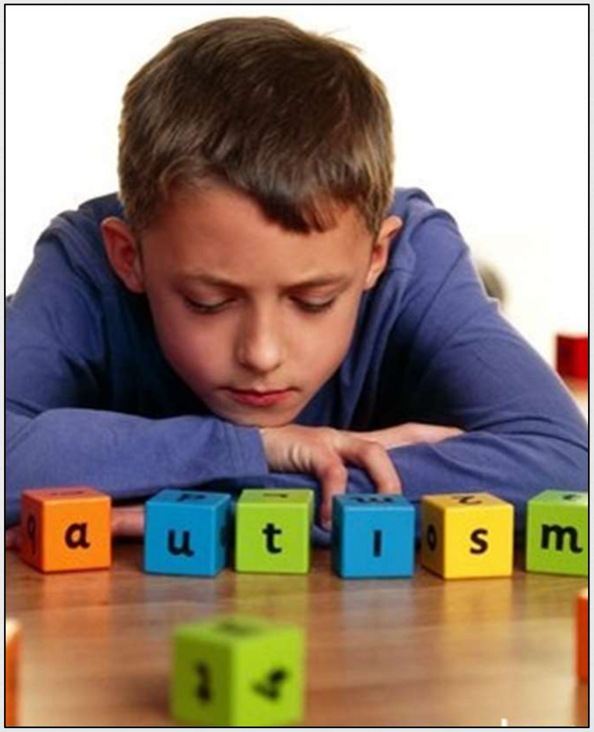
November 10, 2020

Penn Medicine

Children's Hospital of Philadelphia
RESEARCH INSTITUTE

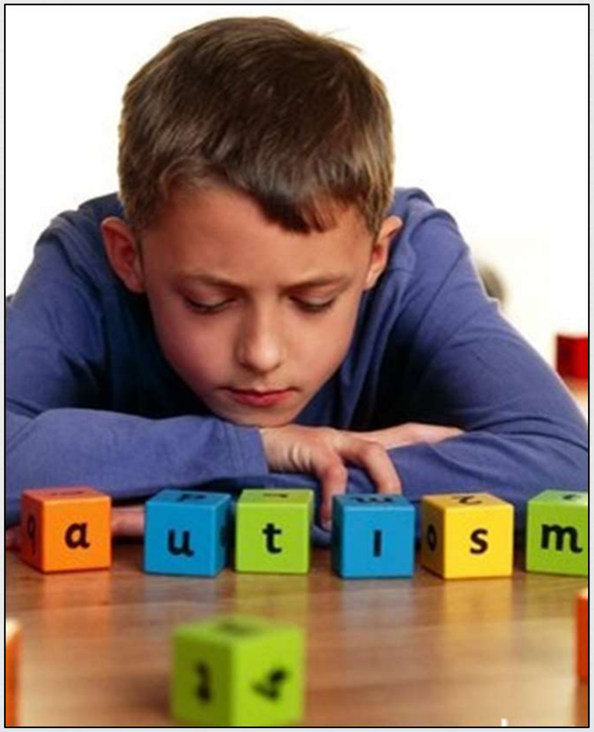
Today's talk

- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources
- A new kind of measurement
- Example results



Today's talk

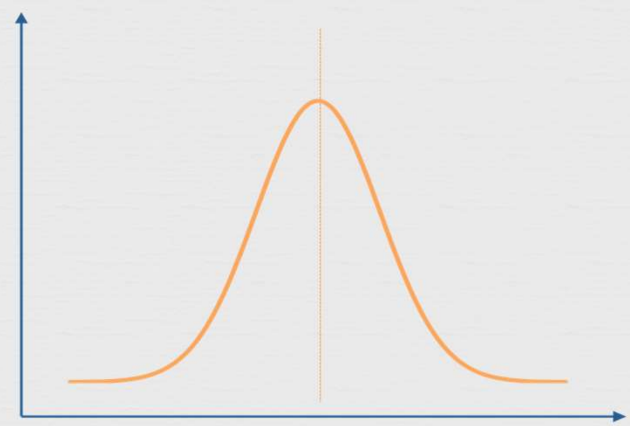
- Autism Spectrum Disorder (ASD)
  - Overview
  - Why we need large shared resources
- A new kind of measurement
- Example results



What is autism?

- Neurodevelopmental condition

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



- Symptom severity lies on a continuum - <u>A</u>utism <u>S</u>pectrum <u>D</u>isorder

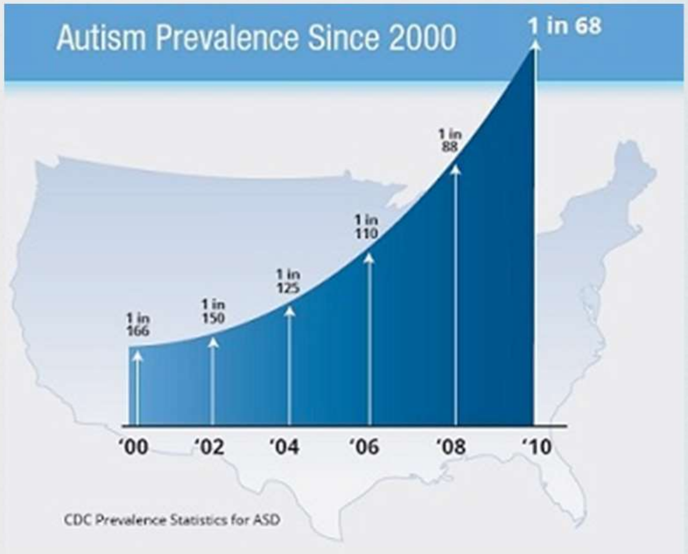
Core features

- **Impaired social communication**
- **Repetitive behaviors and restricted interests**
  - Present since early childhood
  - Interferes with everyday functioning

Who has autism?

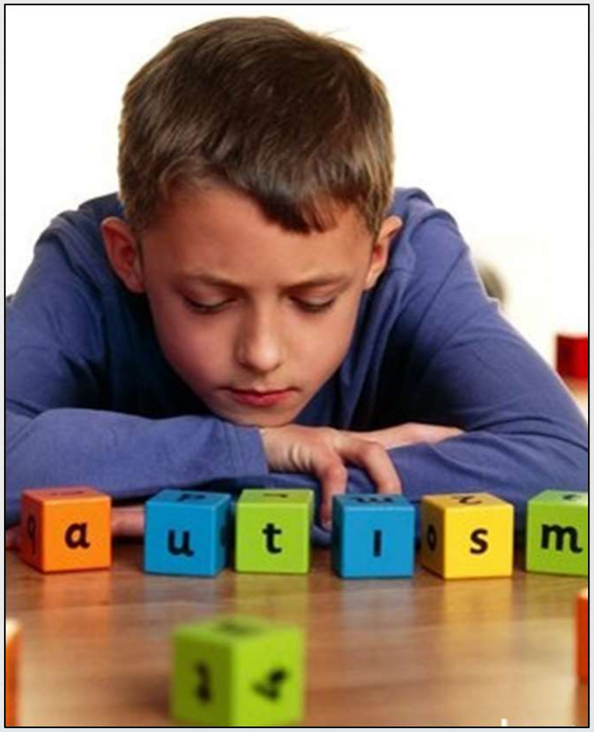
- Estimated prevalence approximately 1-1.5%
  - 1 in 54 U.S. school children (CDC, 2020)

- ~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

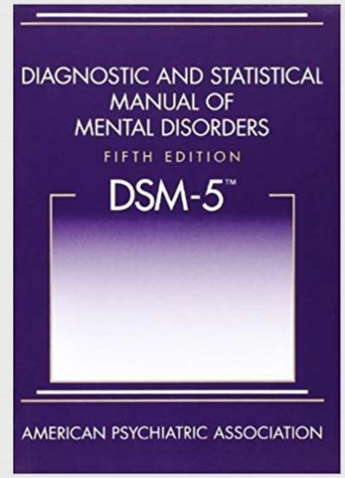


Why do we need large shared resources?
1. **Diagnosis is expensive and difficult**

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

- Based on <u>observable behavior</u> using <u>human judgment</u>

- <u>Problem</u>: imperfect agreement (kappa = .69)



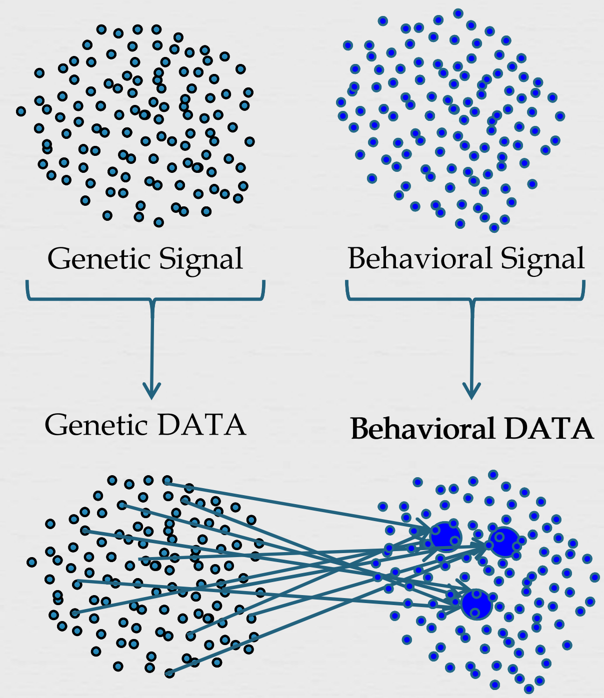
Why do we need large shared resources?
2. **Extreme phenotypic heterogeneity**

- **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

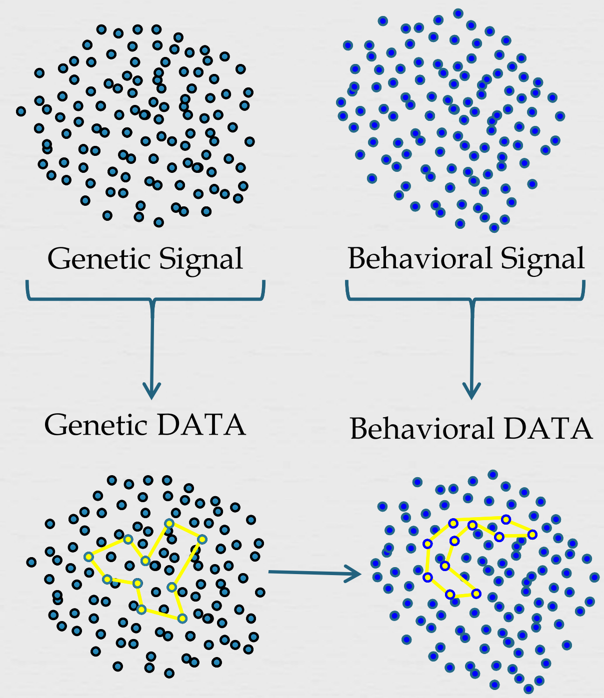
- Need: highly quantifiable, fine-grained signal that is robust to practice effects

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)
  - <u>Motor</u>: computer vision
  - <u>Language</u>: computational linguistics

Analyzing the vocal signal: Challenges

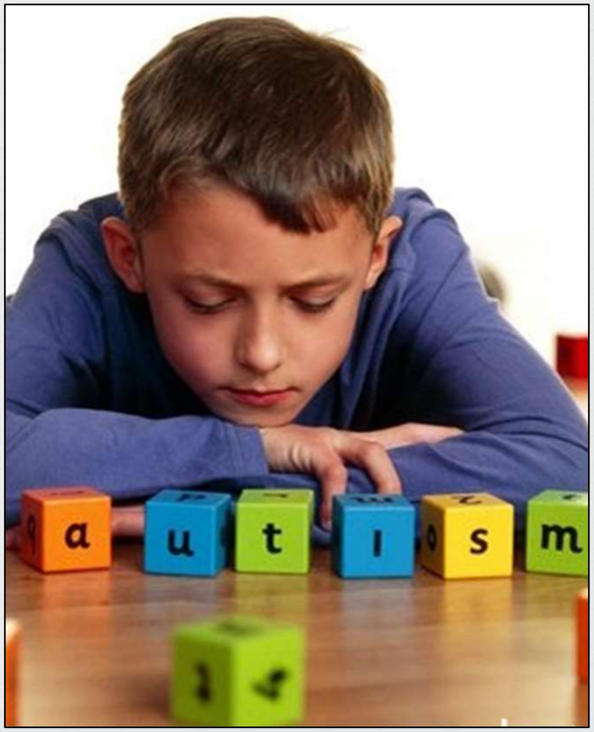
- Language is highly <u>multivariate</u> (acoustics, words, grammar, conversational dynamics)

- "Normal" <u>changes across development</u> and sometimes <u>across cultures</u>

- Neurodevelopmental/psychiatric conditions have <u>different profiles</u>
  - *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



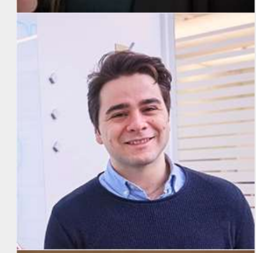
Dyadic biosensor

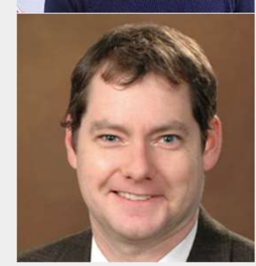
U.S. patent pending, J. Parish-Morris (Co-Inventor)

Keith Bartley

Bob Schultz

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





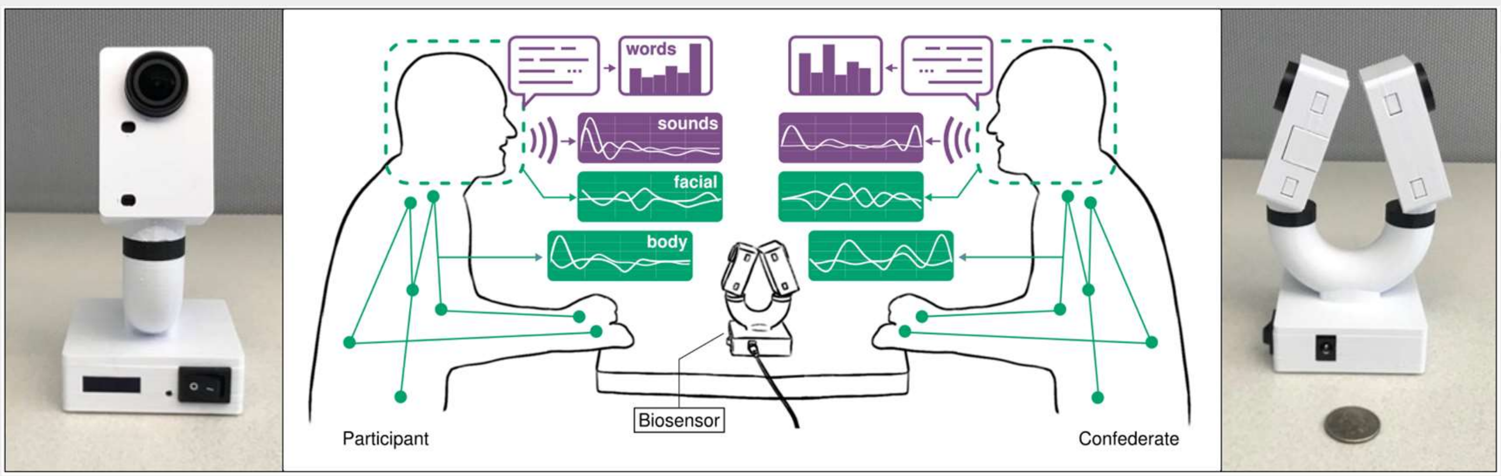
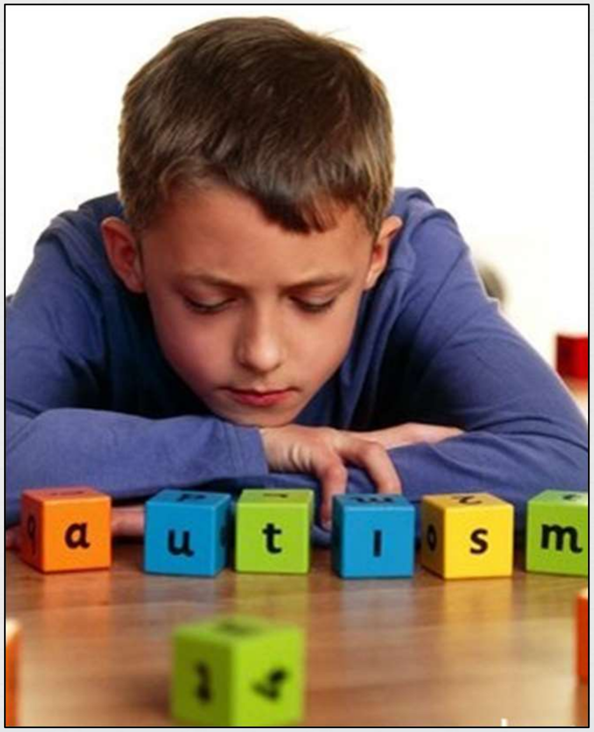


Fig.1 The Biosensor captures everything participants say and do with perfect synchronization.

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[1]

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

1. Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule–Second edition (ADOS-2). *Los Angeles: Western Psychological Services*.

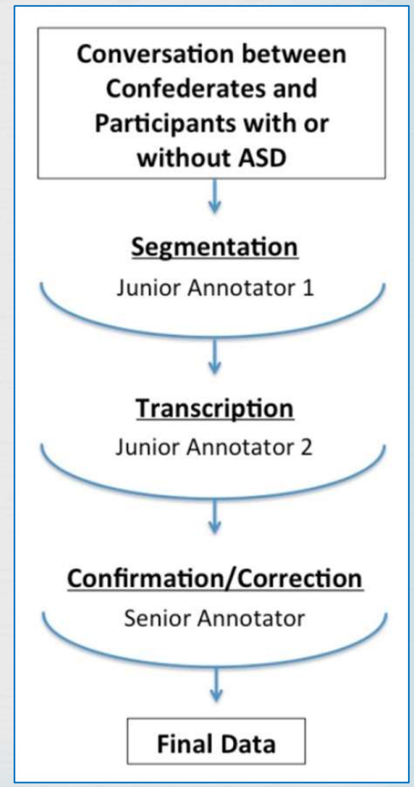
Paradigm

- 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

1. Ratto, A. B., Turner-Brown, L., Rupp, B. M., Mesibov, G. B., & Penn, D. L. (2011). Development of the contextual assessment of social skills (CASS): A role play measure of social skill for individuals with high-functioning autism. *Journal of autism and developmental disorders*, *41*(9), 1277-1286.

Lexical pipeline

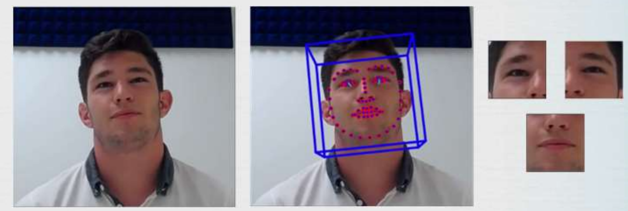
- Verbatim, orthographic, time-aligned transcription of utterances by participant and confederate
- Reliable, blinded annotators using Xtrans[1]
- All spoken words included, no stemming, stopwords remain
  - Diversity includes morphological differences like "want" and "wanted"
- Total word count and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

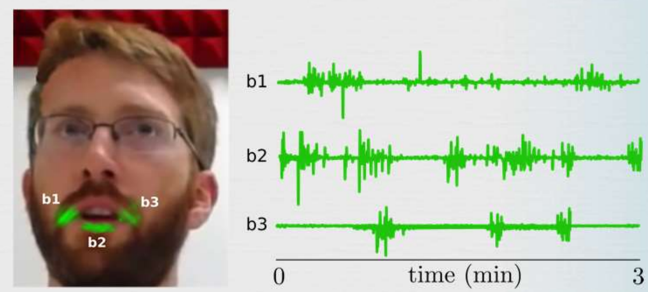


1. Glenn, M. L., Strassel, S. M., & Lee, H. (2009). XTrans: A speech annotation and transcription tool. In *Tenth Annual Conference of the International Speech Communication Association*.
2. Rinker, T. W. (2017). qdap: Quantitative Discourse Analysis Package. 2.3.0. Buffalo, New York. http://github.com/trinker/qdap

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)[1]
- Registration
  - Part-based registration[2]
  - Video stabilization to eliminate jitter
- Quantification
  - Facial Bases method[3]
  - 60 mouth bases
  - Normalized the total activation count of each basis by the maximum count observed for the same basis of confederates





1. Baltrušaitis, T., Zadeh, A., Lim, Y.C. & Morency, L.P.(2018). OpenFace 2.0: Facial Behavior Analysis Toolkit, *IEEE International Conference on Automatic Face and Gesture Recognition.*
2. Sariyanidi, E., Gunes, H., & Cavallaro, A. (2015). Automatic analysis of facial affect: A survey of registration, representation, and recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(6), 1113-1133.
3. Sariyanidi, E., Gunes, H., & Cavallaro, A. (2017). Learning bases of activity for facial expression recognition. *IEEE Transactions on Image Processing*, *26*(4), 1965-1978.

Statistical approach

- 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)

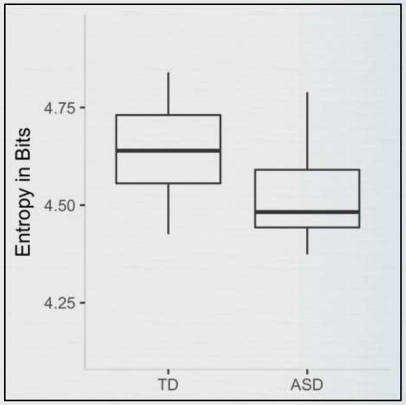
$$H = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

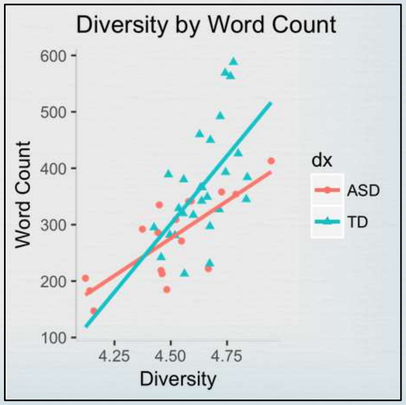
  - **High entropy** is expected when participants make a rich set of facial expressions and produce a variety of words while speaking
  - **Low entropy** is expected when participants generate a restricted set of mouth movements and produce repetitive speech patterns
- Tests:
  - Wilcoxon rank sum tests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

1. Cover, T. M., & Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

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 $t=-3.51$, $p=0.001$*

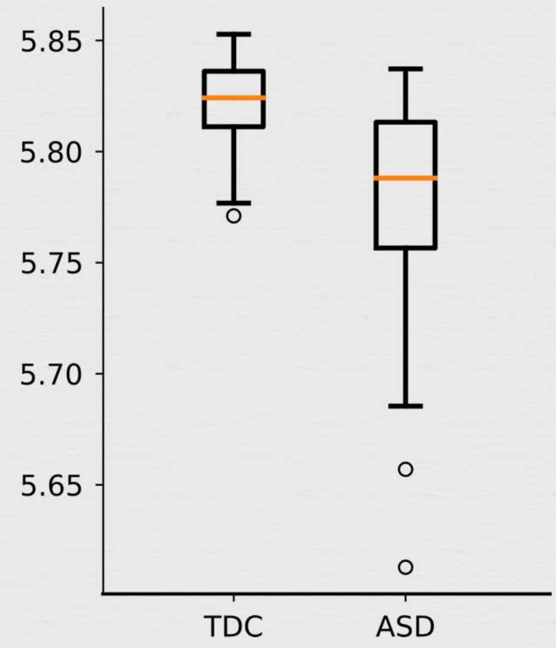




1. Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell Labs Technical Journal*, *30*(1), 50-64.
2. Witten, I. H., & Bell, T. C. (1990). Source models for natural language text. *International Journal of Man-Machine Studies*, *32*(5), 545-579.

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


THANK
YOU

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