

Automated Identification of Stress and Focus Assignment

Emily Tucker Prud'hommeaux, Jan van Santen, Lois Black
Department of Biomedical Computer Science, Center for Spoken Language Understanding
Oregon Health & Science University



OREGON
HEALTH & SCIENCE
UNIVERSITY

Background

- Evaluation of expressive prosodic ability important for diagnosis of neurodevelopmental disorders such as ASD.
- Existing methods for assessment of prosody require that judgments be made at the time of examination.
- Accurate automated analysis of prosody could increase both efficiency and accuracy in clinical evaluations of prosodic ability.

Objectives

1. Establish reliability of real-time judgments of stress and focus assignment.
2. Determine whether complex automated measures of acoustic features are comparable to naive listener judgments and real-time clinical assessments.

Method

Data

Speakers

- 15 ASD, 13 TD, 15 meeting some but not all criteria for ASD.
- Age 4-8; performance IQ > 70.

Prosodic Tasks

1. *Lexical Stress*: Repeat disyllabic nonsense word with initial or final stress.
2. *Emphatic Stress*: Repeat four-word sentence with emphasis on one word [1].
3. *Focus*: Correct inaccurate description of a picture by emphasizing the correct word [2].

Scoring

The following scores for each of the three tasks were correlated with one another on a per-utterance and per-speaker basis:

Real-time examiner scores

During examination, one of 4 clinicians immediately assessed the correctness of each response, yielding *real-time scores*.

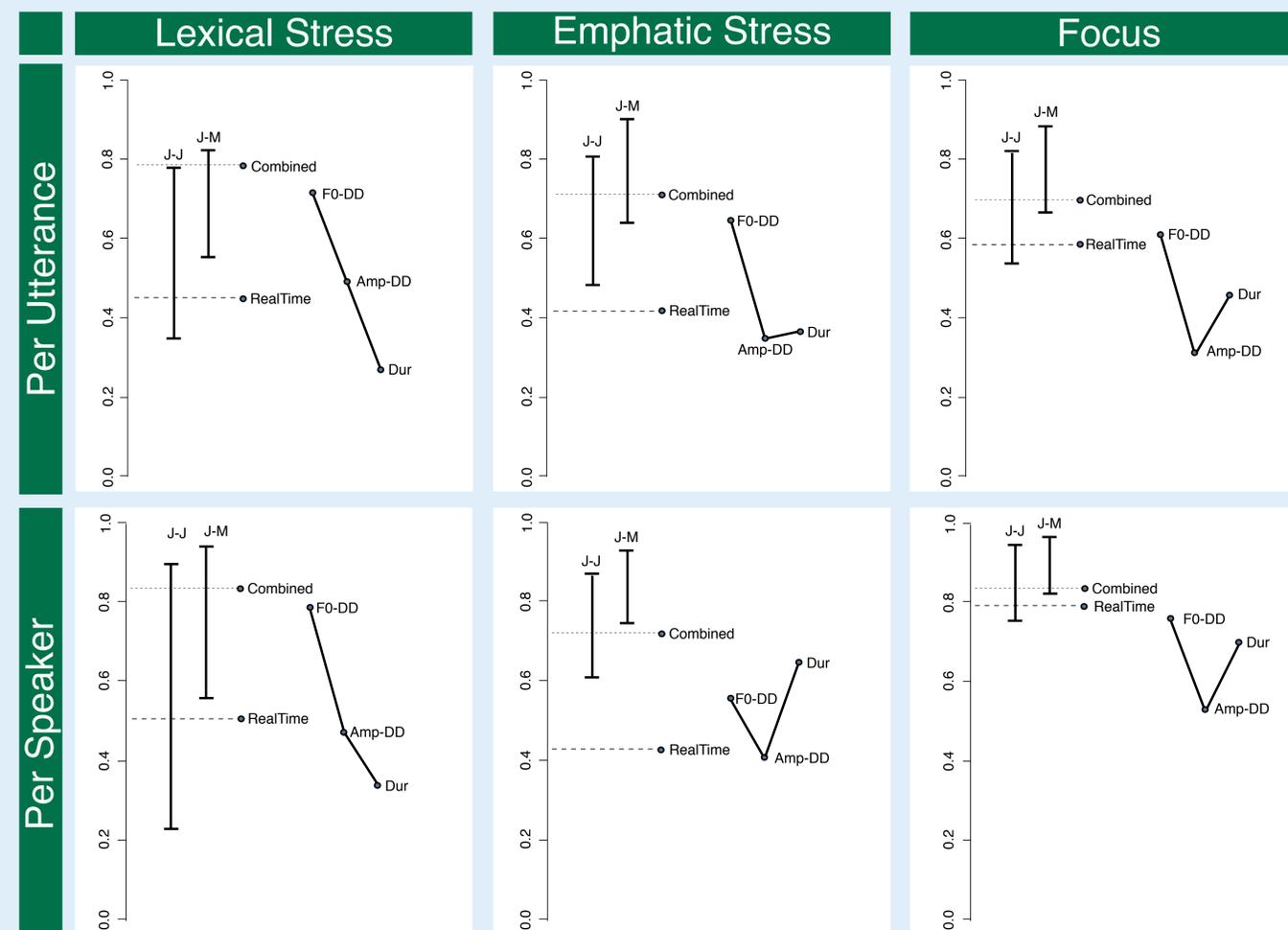
Web-based perceptual experiment

- Six naive *judges* listened to recordings of minimal pairs of responses for each tasks.
- Each minimal pair from a single speaker with same content but different target prosody.
- Judges identified the intended meaning of the two utterances (e.g., of two recordings, which one was meant to be “BLUE cow” rather than “blue COW”).

Automated acoustic analysis

- *Pitch and energy trajectories* and *phoneme duration* information extracted from recordings of the children's responses.
- Data analyzed using an innovative *dynamic difference (DD)* measure that captures the difference in the pitch and amplitude dynamics of the two recordings of a minimal pair.
- Measures of melody, timing, and intensity *combined using multiple linear regression* to create a single complex score for each utterance pair.

Results: Correlations



Conclusions

- Automated digital measures comparable in reliability to judges' scores and superior to real-time clinical judgments on both a per-utterance and a per-speaker basis.
- Including automated objective measures of prosody alongside traditional real-time judgments could enhance both accuracy and reliability in clinical assessments of prosodic ability.

References & Sponsors

- [1] Shriberg, L., Allen, C., McSweeney, J., Wilson, D., 2001. PEPPER: Programs to examine phonetic and phonologic evaluation records (Computer software). Madison: Waisman Research Center Computing Facility, University of Wisconsin—Madison.
- [2] Peppé, S., McCann, J., 2003. Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics and Phonetics*, 17, 345–354.
- NIH 1R01DC007129 (van Santen, PI); Autism Speaks: Mentor-based Fellowship (Prud'hommeaux); Autism Speaks: Computerized Interactive Game for Remediation of Prosody in Children with Autism (Black, PI); Autism Speaks: ITA: Automated Measurement of Dialogue Structure in Autism (Roark, PI).