Towards Inclusive Automatic Assessments of Children’s Oral Language Development

Alexander Johnson

Advised by Prof. Abeer Alwan

UCLA, Speech Processing and Auditory Perception Lab
Children’s Oral Language Assessments

Tests of children’s speaking, language, and comprehension skills designed to:

- Identify children who may be falling behind their appropriate grade level
- Monitor for speech difficulties that need intervention through speech-language pathology

Focus of the study is on automatic assessments for African American English-speaking children
Children’s Oral Language Assessments

Example Assessment Prompt:

- Please describe the picture as completely as possible
Children’s Oral Language Assessments

Example Assessment Prompt:

- Please describe the picture as completely as possible

Sample responses:

1. That boy, he ain’t got no umbrella
2. Dhose kids takin’ dhey time walkin’ cross da street
3. She had aksed her mom take her home
Children’s Oral Language Assessments

- Assessments need to account for dialectal differences to indicate if an intervention is needed
- Can machines also be trained to do this?

- Unstressed syllable deletion
- Inconsistent use of verb tense
- Overregularization in conjugation of irregular verb
- Metathesis of final /s/+stop (eg. “ask” as “aks”)

Is this normal?
Children’s Oral Language Assessments

Challenges for a machine

- Lack of training in -domain training data
- Variability in children’s speech
- Variability across and within dialects (eg. code switching)
Our Framework

A possible pipeline for oral language assessment scoring for African American English and Southern American English Speaking kids from Georgia, USA
1. **Extract known linguistic features of dialect**
   - Phonetic Modeling (ASR features)
   - Language Modeling (Perplexity and surprisal)
   - Speaker Identification (X-vectors)
   - Paralinguistics (ComParE16 feature set)
   - Prosodic Modeling (Energy, Duration, Pitch)

1. **Train a classifier to predict dialect and strength of dialect from the input features**

---

Improved Dialect and Age-specific ASR

We fine-tune child-specific ASR models and apply data augmentation for low-resource dialects.

Schematic of LPCAugment - a data augmentation method created to shift vowel formant frequencies in a manner characteristic of southern dialectal shifts.

NLP for Oral Assessments Scoring

Given the best transcripts, we then extract NLP features and train a classifier to give a score for the assessment.

Groundtruth Test item: Jake and Todd asked their mom for pizza, hamburgers, and ice cream.

ASR Transcript: Jake and Todd asked they mom for pizza and ice cream.

Semantic Similarity Score: 82%
Verb tense matching: 100%
Recall of keywords: 67%
Recall of Characters: 100%

Classifier: Penalize or accept language differences?

Diagnostic score: 84%
Language skills: Average
Thank you for listening

Questions welcome