USC University of Southern California

USC Viterbi School of Engineering

Ming Hsieh Department of Electrical Engineering

Assessment of a child's engagement using sequence model based features

Rahul Gupta, Chi-Chun Lee, Sungbok Lee, Shrikanth Narayanan

Signal Analysis and Interpretation Lab (SAIL)

Motivation & Introduction

Hypotheses

- Engagement behavior of children in reflected in the speech of the participants in conversation
- Local patterns of features also carries information about the engagement behavior apart from global statistics

Goal

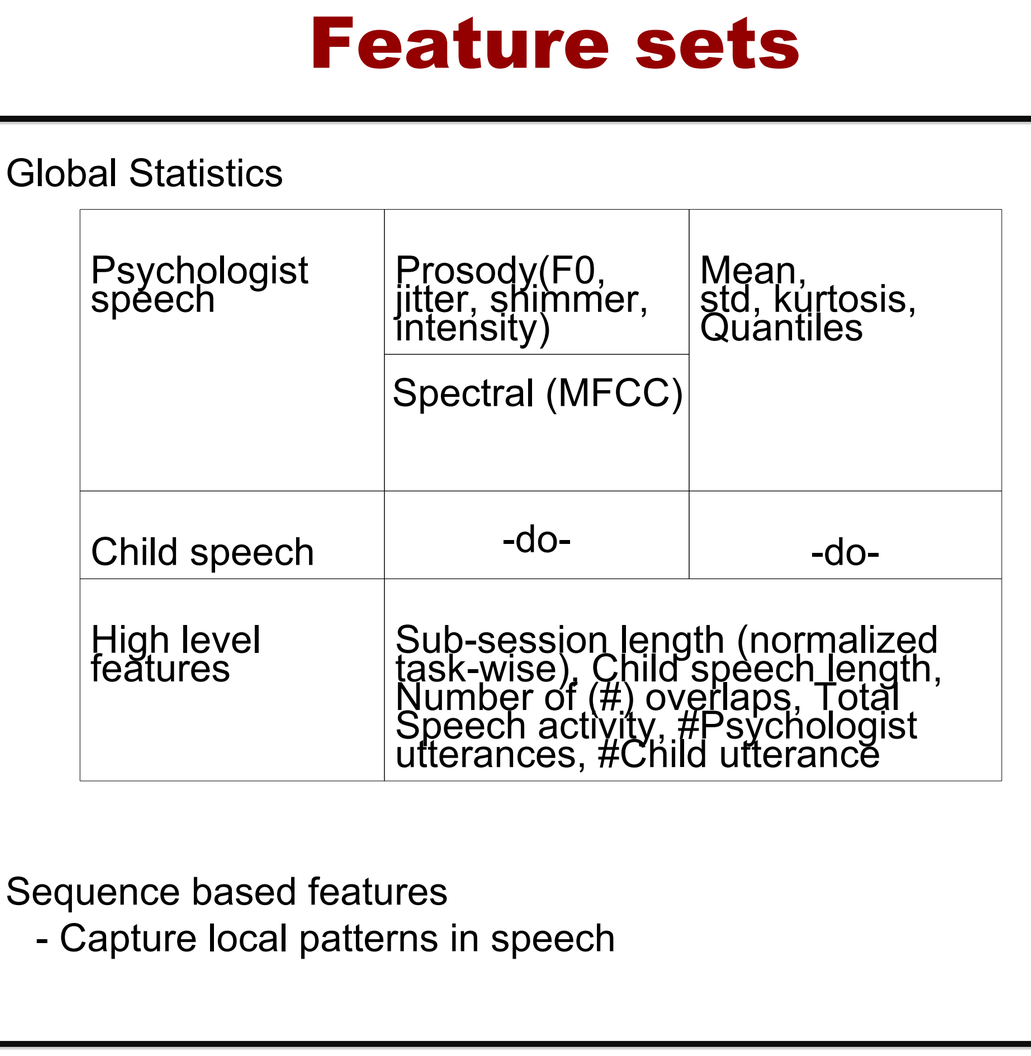
- Develop a purely speech based system to predict the engagement level of children in their interaction with a psychologist

-Exploring the phenomenon of joint attention and other developmental aspects in children using engagement

Experiment

Database:

- 65 sessions from R-abc database
- Predict the perceived child engagement from audio stream



Feature sets

- Global Statistics

Psychologist speech	Prosody(F0, jitter, shimmer, intensity)	Mean, std, kurtosis, Quantiles
	Spectral (MFCC)	
Child speech	-do-	-do-
High level features	Sub-session length (normalized task-wise), Child speech length, Number of (#) overlaps, Total Speech activity, #Psychologist utterances, #Child utterance	

- Sequence based features
  - Capture local patterns in speech

Sequence Features

- Train a sequence model on the training set
  - Normalized count of an “n-gram” over the training set

$$\mathrm{SM}(w_k/w_{k-1},..,w_{k-n+1}) = \frac{\sum_{t\in\text{Train Set}} \#(w_k^{S_t}/w_{k-1}^{S_t},..,w_{k-n+1}^{S_t})}{\sum_{t\in\text{Train Set}} N_t}$$

- Get the probability on the test set
  - # of occurrence of an “n-gram” X above normalized count

$$\mathrm{M}(w_k^S/w_{k-1}^S,..,w_{k-n+1}^S) = \mathrm{SM}(w_k/w_{k-1},..,w_{k-n+1}) \times \#(w_k^S/w_{k-1}^S,..,w_{k-n+1}^S)$$

Results

Feature Source	Selected model	Class-wise accuracy				Uw. accura-cy(2 class)	Uw accur-acy (3 class)
		'0' (%)	~'0' (%)	'1' (%)	'2'		
Global statistics on features	High level features	78.6	63.7			71.1	
		78.6		30.0	38.7		49.1
Features based on SM	None,child pitch, psyc.pitch, child intensity	77.2	69.2			73.2	
		77.2		33.3	32.3		47.6
Fused Model	None,child pitch, psyc.pitch, child intensity, high level features	77.2	69.2			73.2	
		79.0		33.3	41.9		51.4

- Apart from global statistics, local patterning of features also informative in capturing engagement levels

Future Work

- Incorporate visual features
- Better sequence models
  - Smoothing techniques