Expeditions in Computing

AUDIO PARSING SEMI-STRUCTURED CHILD-EXAMINER INTERACTIONS USING KEYWORD SPOTTING

Hrishikesh Rao and Mark A. Clements
Georgia Institute of Technology

Introduction

Rapid ABC consists of a semi-structured play interaction with an adult examiner and a child. The stages of the Rapid ABC consist of greeting, ball play, book reading activity, placing the book on the head and pretending it is a hat, and a tickle game. It is of interest to examine the child's visual and verbal response to the speech stimuli from the examiner. Analysis of examiner's speech can be done using state-of-the-art keyword spotting tools.

Goal

- Generate timestamps of phrases of interest and export them to CSV.
- Export the results in an ELAN, a video coding tool, compatible format for analysis

Method

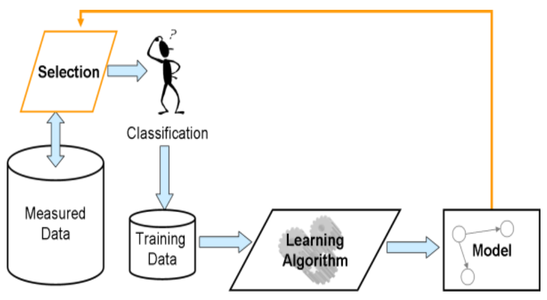
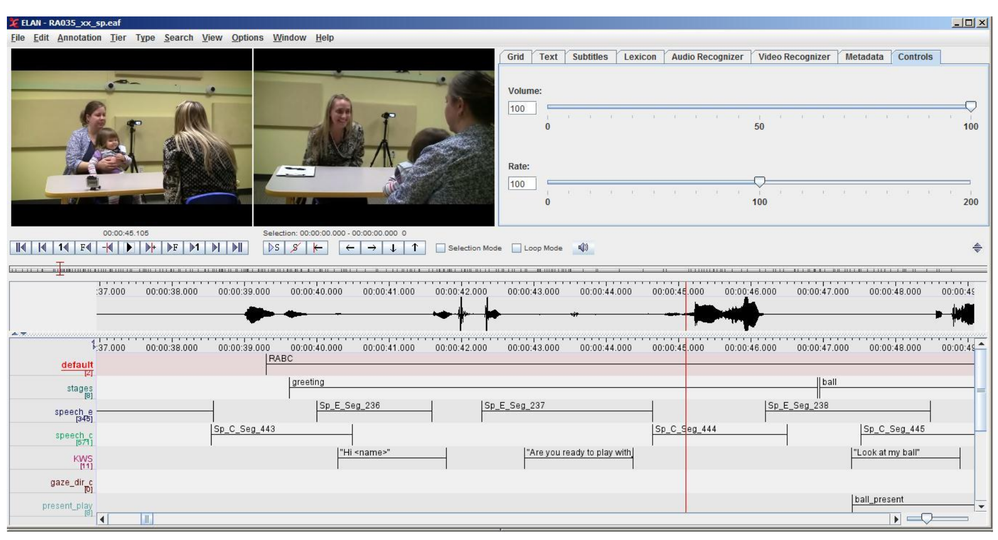


Fig. 1 Active learning model

- Keyword spotting tools are commercially available from Nexidia.
- Use Hidden Markov Model (HMM) for developing acoustic models
- Phonetically search for phrases of interest
- Refine results using pronunciation optimization
- Pronunciation optimization is an active learning approach with user intervention that retrains models



Data

25 semi-structured Rapid ABC sessions consisting of 2 examiners

Results

Keyword	Stage	Number of correct timestamps
Hi <name>	Greeting	15/25
Are you ready to play with some new toys?		22/25
Look at my ball	Ball play	20/25
Let's play ball		17/25
Ready…set…go!!		23/25
Look at my book	Book activity	19/25
Where's the yellow duck?		19/25
Can you turn the page?		18/25
Let's see what's next		21/25
It's on my head, it's a hat!	Surprise	20/25
I'm gonna get you	Tickle	23/25

Conclusions

- Keyword spotting of phrases of interest in interactions between child and examiner is an effective tool for segmentation of sessions.