Semi-automated Linguistic Transcription of Daylong Audio Files

Haille N. Heid
Washington State University
haille.heid@wsu.edu
Showcase for Undergraduate Research and Creative Activities
Pullman, WA | 28 March, 2016

College of Medical Sciences
Washington State University Health Sciences Spokane

Main Research Objective

The goal of this work is to demonstrate that a combination of machine-automated processes and human judgements can be combined to amass vastly more linguistic data than was previously available, while also maintaining naturalness at a reasonable cost.

Background

Children acquire language by exposure. In order to better understand the effects of exposure, it is important to study the amount and types of language in natural family environments. Recently, there have been advances in technology that allow us to better study natural language development via daylong family audio recordings.

Materials

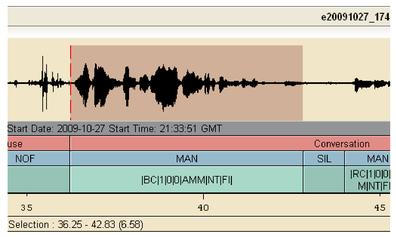
The LENA (LENA Research Foundation, Boulder, CO) is a small recording device that fits into a pocket on the front of specialized shirt. The LENA device records uninterrupted daylong audio from the child's auditory perspective. There are two principal advantages to the system.

1. The technology automates data collection, reducing the labor and financial expense.
2. The data is highly natural, collected in the home environments of families.

Automatic Speech Processing

LENA software processes the audio using automatic speech processing algorithms (ASP, similar to automatic speech recognition, ASR) to identify linguistically important segments in the child's auditory environment. Segment labels include:

1. Duration of total TV/electronic media
2. Number of words spoken by the child
3. Number of words spoken by adults (both fathers and mothers)
4. Number of conversational exchanges the child engages in
5. Other categories such as periods of silence or noise.

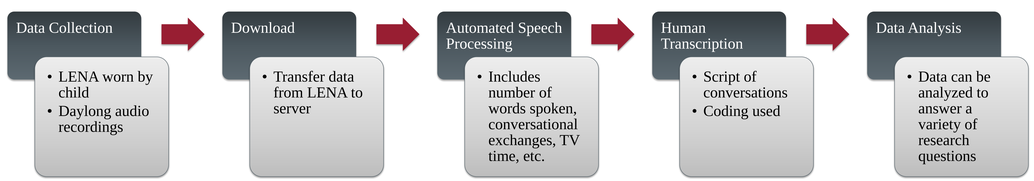


Human Transcription

This research project takes as input the audio files and machine-generated tagging from several daylong recordings that were automatically processed by the LENA system. That input is subsequently analyzed by trained human transcribers to augment the automatic processing, producing linguistic transcriptions of the audio files suitable for further analyses.

```
@Begin
@Languages:    eng
@Participants:  CHI Target Child, MOT Caregiver Adult
@Media:   123456, audio
*CHI:    owie mommy .
*MOT:  oops I'm sorry did that hurt ?
*MOT:  for some reason it's just not clicking .
*CHI:    yeah it hurt mommy .
*CHI:    it hurt .
*MOT:  almost done .
*MOT:  there now you're all buckled in .
*CHI:    buckled in ?
*MOT:  yep buckled in your car seat .
*MOT:  now mommy's gonna buckle up .
*CHI:    mommy were both buckled up .
*CHI:    both mommy and Kaitlyn .
*MOT:  yep mommy and Kaitlyn are buckled up .
*CHI:    and when we get to the store I'm gonna be not buckled .
@End
```

Semi-automated Linguistic Transcription Process



Conclusion

The hybrid approach, automatic speech processing coupled with human transcription, uses modern technology to collect larger and more useful datasets to study cognitive and language development in child populations that are both typically developing and at-risk.

Some recent work includes the following:

1. 'Fatherese,' or how fathers tend to use pitch differently from mothers.
2. Automatic classification of Autism spectrum disorders and specific language impairment from passive recordings.
3. Speech production differences between children with and without hearing loss.
4. Estimates of language and cognitive developmental milestones with passive recordings on a very large number of samples (e.g., an entire school system).
5. Amount of TV viewing and potential impact on developmental outcomes.
6. Examination of syntactic development in children with hearing loss and their typically-developing peers.
7. Screening for language or cognitive delay.