

Introduction

- The Language ENvironment Analysis (LENA) is a wearable audio device that includes software for automated vocal analysis (Ganek & Erics-Brophy, 2018).
- LENA classifies chunks of audio as speech vocalizations (adults and children), non-speech/vegetative noises, or background noises.
- LENA provides automated metrics such as adult word count, child vocalizations, and conversational turn count.
- LENA is in widespread use by clinicians to get an accurate representation of the natural linguistic environment of children (Xu et al, 2009).
- Parental responsiveness and turn-taking are integral to children's early language acquisition.
 - There is little existing independent data on the device's accuracy of conversational turn count.
- Misclassifications of speakers and discarding overlapping speech can lead to errors in LENA's computation of the conversational turn counts (Ganek & Erics-Brophy, 2018).
- **Purpose of the study:**
 - The purpose is to conduct an independent study on the reliability of LENA.
 - i. How accurate are its classifications of conversational turns?
 - Work out a method of computing LENA's accuracy in regards to its reliability in calculating conversational turn counts.

Methods

- LENA recordings were obtained from Ohio State University where 15 families of children with various hearing statuses recorded a typical 10-16 hour day.
- LENA identifies when speech occurs and segments these blocks further into the following sound source codes:
 - *Class 1 segments*
 - Female adult, Male adult, key child, other child
 - *Class 2 segments*
 - Overlapping speech, TV/electronic media, noise, uncertain/fuzzy, and silence
- We obtained a Conversational Turn Count (CTC) from LENA's output.
 - Human coders quantified conversational turn count through identifying instances of (1) adult and child vocalizations as well as (2) when an adult and a key child (i.e., the child wearing the LENA device) were speaking to each other.
- We are interested in analyzing how often LENA's conversational turn count agreed with human coders, which heavily depends on if LENA's classification of speakers was accurate.

LENA's Classifications

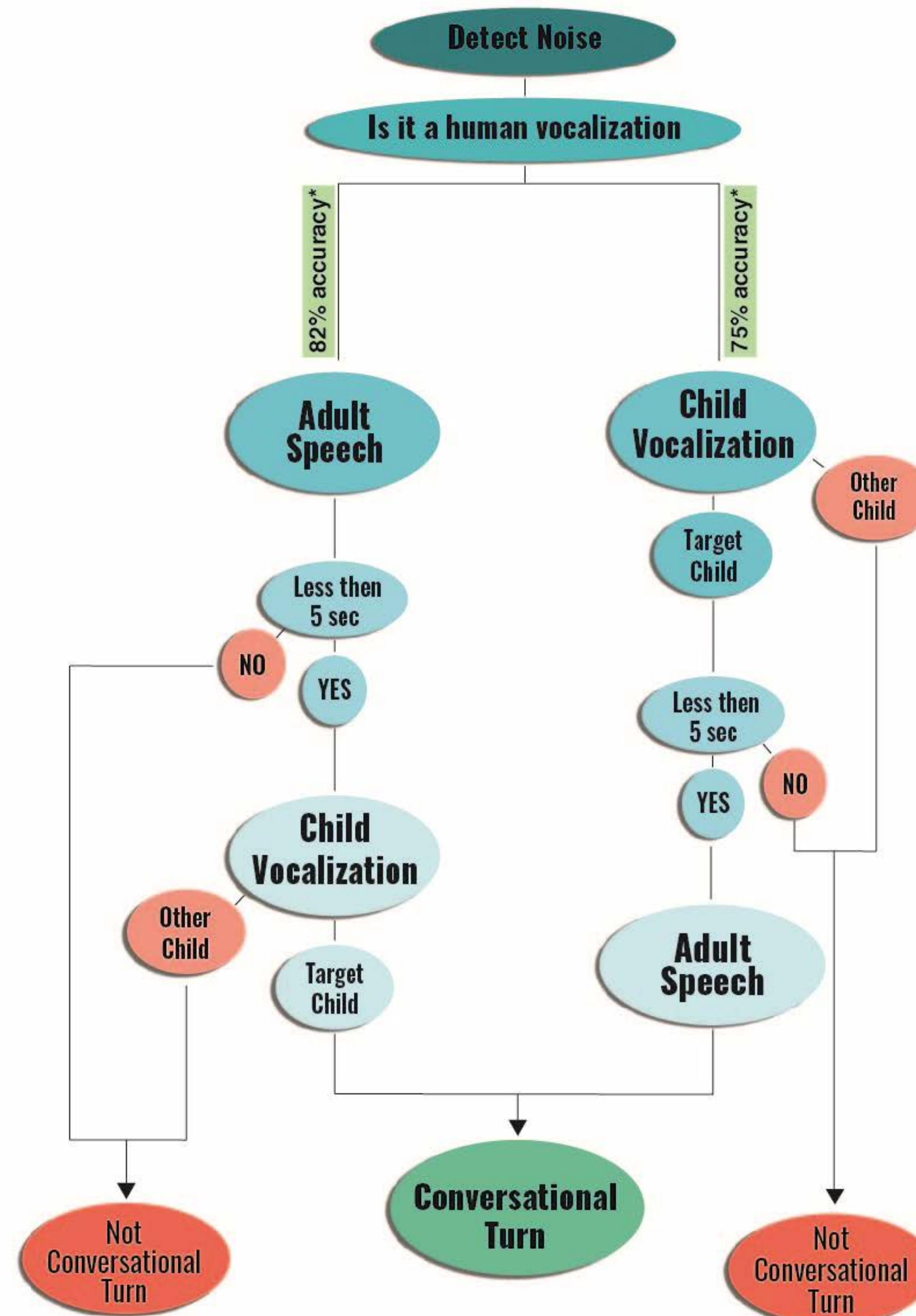


Diagram 1: above demonstrates the process LENA uses to classify all possible conversational turns with in a chunk of audio
 *Classification agreement between humans and LENA

Where does LENA falter?

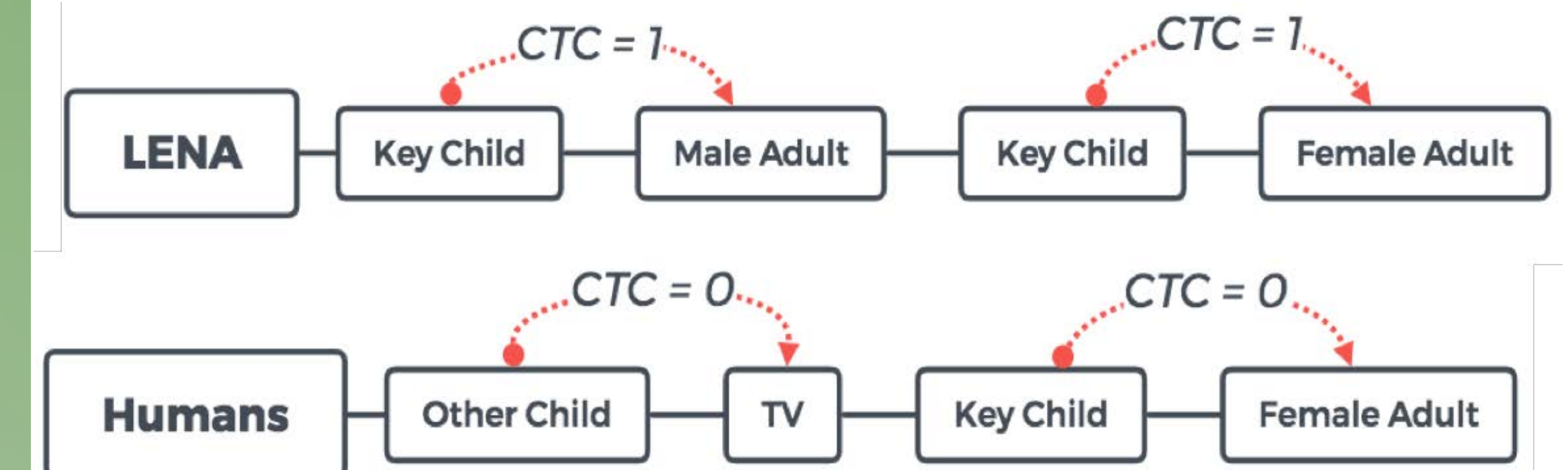


Diagram 2: An example of how LENA might classify speakers and quantify conversational turn counts (CTC) for a given audio, and how humans might code the same

As depicted in the diagram:

- LENA can misclassify other child as key child, and vice versa
- LENA can misclassify TV as adults, and vice versa
- LENA may classify speakers correctly, but the context of speech would render it a coincidence, not a conversation

Next Steps

- We want to gain an understanding of how LENA works as a machine.
 - Acoustic properties that drive its decision making
- We want to analyze LENA's reliability.
 - Given a human's decisions, how often does LENA agree with them?
- A chunk of audio may fulfill LENA's rules for counting a conversational turn, but LENA still does not register a conversational turn. Why is that?
- Ideally, these analyses would be used to help improve the technology.

References

- Ganek, H., & Erics-Brophy, A. (2018). Language ENvironment analysis (LENA) system investigation of day long recordings in children: A literature review. *Journal of communication disorders*, 72, 77-85.
- Gilkerson, J., & Richards, J. A. (2010). *Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development*. Technical Report LTR-01-2. 2008. Available at: https://www.lenafoundation.org/wp-content/uploads/2014/10/LTR-01-2_PowerOfTalk.pdf. Accessed September 6.
- Gilkerson, J., & Richards, J. A. (2008). The LENA natural language study. Boulder, CO: LENA Foundation. Retrieved March, 3, 2009.
- Gilkerson, J., Coulter, K. K., & Richards, J. A. (2008). Transcriptional analyses of the LENA natural language corpus. Boulder, CO: LENA Foundation. Retrieved January, 12, 2013.
- Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., Yapanel, U., & Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, 107(30), 13354-13359.
- "Quick Reference Sheet: Codes Used in the ITS File and in ADEX Version 1.1."
- Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in young children's natural home environment. Boulder, CO: LENA Foundation, 1-16.

How does LENA define a conversational turn?

First, LENA assigns speaker codes to a chunk of audio containing human speech. Then, it determines if there is an **adult talker** and **key child** speaking within 5 seconds of each other. If there are no other class 1 (live speaker) codes intervening, then LENA makes the call that a conversational turn has occurred. In table 1, the green represents two conversational turn counts as decided by LENA, and the red represents a class 1 speaker intervening in what would otherwise have been classified as a conversational turn.

Female Adult	Other Child	Female Adult	Key Child
Male Adult	Key Child	Male Adult	Key Child

Table 1: Example of two count conversational turn and non-conversational turn as categorized by LENA

Acknowledgements

Research reported in this poster was supported by the National Institutes on Deafness and other Communication Disorders (NIDCD) Grant R01-DC008581 to D. Houston and L. Dilley.