

# Evaluating Language ENvironment Analysis (LENA™) System Performance for Russian: A Validation Study



## Adela Ulitsky Communication Sciences and Disorders

### Introduction

Early language input has long been shown to be an integral factor for proper language acquisition in typically developing children.

Hart and Risley (1995) ten-year longitudinal study revealed the importance of quality and quantity of early language exposure.

- Children's IQ and language abilities can be attributed to the amount parents speak to their children early in life.
- By the age of 3, the children of high income, professional families heard 30 million more words than those of welfare families.
- Differences in input were reflected in academic achievement; with the children of the low SES families demonstrating significantly lower academic success.

One tool that has been used to quantify, recode and analyze adult and child vocalizations is the Language ENvironmental Analysis (LENA) System.

- The LENA recordings are processed by the LENA analysis software (LENA Pro System, 2012) through sophisticated American-English based algorithms.
- The system categorizes human speech energy into accurate syllable distribution that, using the algorithm, equals to the word count.
- The resulting reports provide information about the number of words that were spoken to or near the child each hour, the amount of child vocalizations, the number of conversational turns and the amount of tv and electronic noises (Xu et al., 2009).

As of 2011, the U.S Census Bureau states that there are 906,000 Russian speakers in the United States and there will most likely be more children born into these households.

- LENA currently validated for English, Spanish (Xu et. al., 2009) and Chinese language (Gilkerson et. al., 2015)
- Sensitivity greater than 80% is considered acceptable.
- Russian, being a more multi-syllabic language could theoretically challenge the existing system's algorithms in the calculation of words spoken.
- Large number of suffixes (ru-ka → ru-ch-etch-ka)
- Heavy use of diminutives in "motherese"

### Purpose & Research Question

- Validate the use of the LENA within the environment of Russian-speaking households.
- Explore the accuracy of the system in order to answer the research question:
  - What is the sensitivity of the LENA for Russian?
  - To what extent does LENA Adult Word Count (AWC) match Human Word Count (HWC)?

### Subjects, Methods & Analysis

#### Participants:

- One participant age 10-12 months, and his Russian-speaking caretakers; grandfather, grandmother and nanny.
- The grandmother and grandfather were recorded together since they provided child care simultaneously while the nanny had her own set of recordings.

#### Data Collection and Selection:

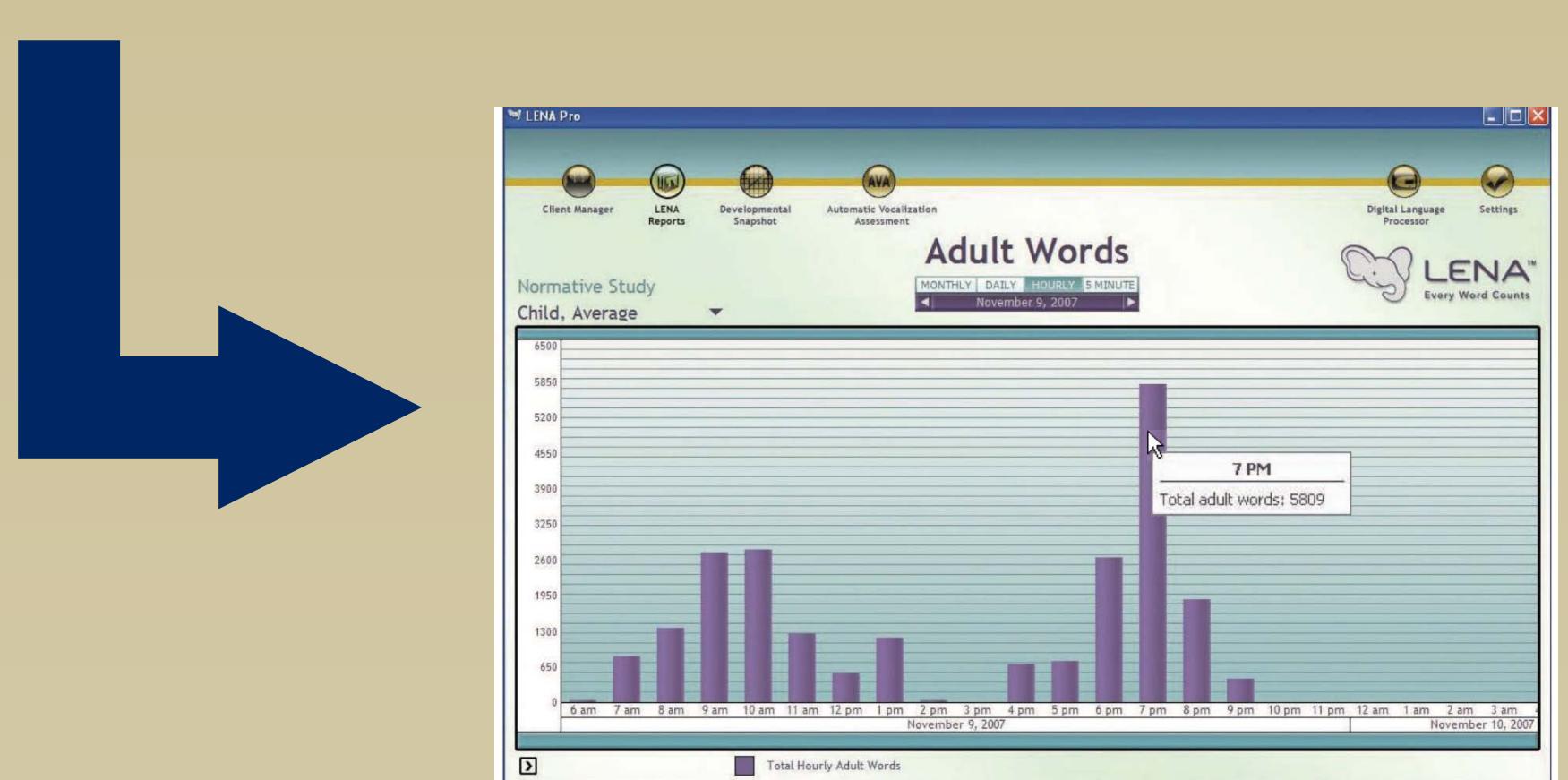
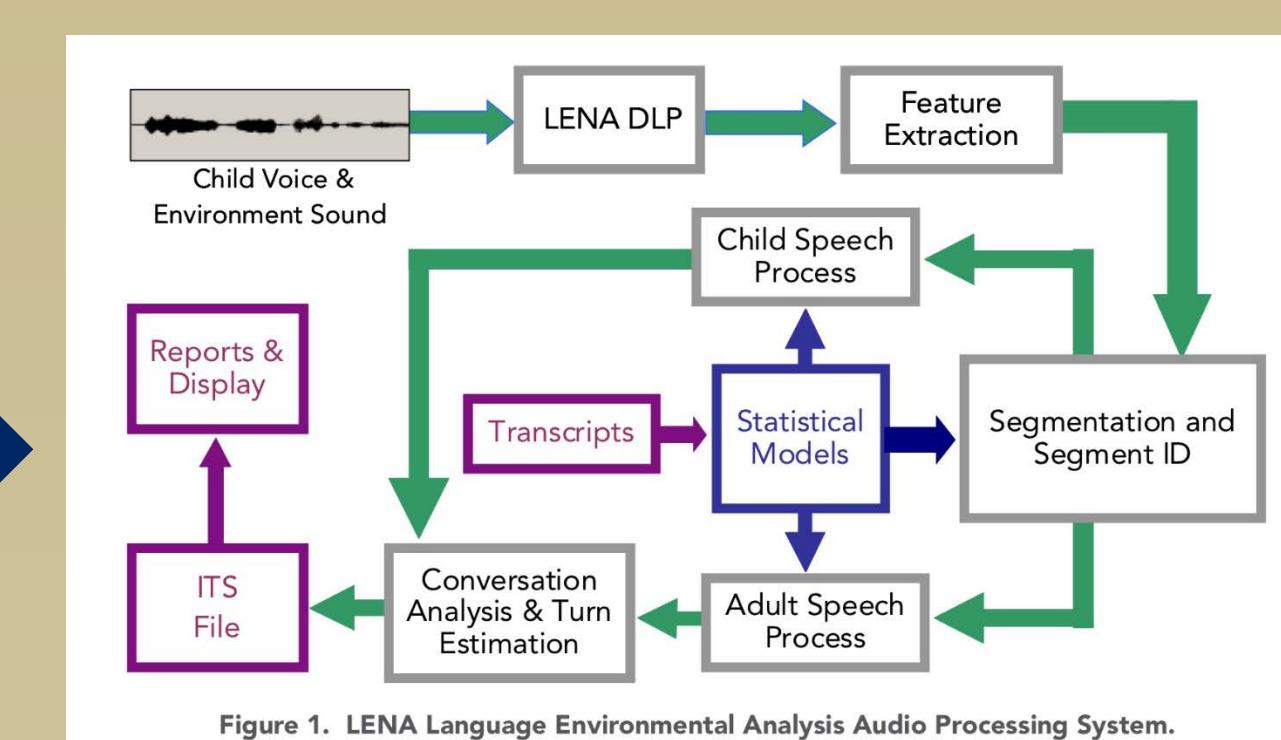
- Audio recording was conducted through the digital language processor (DLP). The DLP records the communication of the child and that of all other participants in his or her environment, including other children and adults, as well as non-speech environmental noises, such as television and electronic sounds.
- Recordings: 12 hours from 8 a.m. to 8 p.m. for 4 weeks
  - The day was divided into three 4-hour zones: morning, afternoon, and evening.
  - From each of the 3 zones, a 5-minute sample with the highest adult word count was selected.
- The final data sample consisted of 120 minutes of recordings (5 minutes from each of the two sets of caregivers for 4 weeks).

#### Data Transcription and Analyses:

- Completed recordings were processed by LENA software to produce the AWC estimates used in the current study.
- A native speaker counted the number of Russian words heard in each 5-minute sample.
- Reliability was assessed by having a second human coder count the number of adult words in the sample.
- Each zone's HWC was then divided by the LENA estimated AWC to produce a sensitivity percentage for each zone.
- Sensitivities were then averaged to equal the overall sensitivities of the total of the grandparents recordings and the total of the nanny recordings.



Figure 3. The LENA digital language processor (DLP-0121), actual size.



### Results



### Conclusion and Implications

1. Human-based word-counting and the LENA system's algorithms fluctuated in sensitivity throughout the day, weeks, and participants.
2. The use of the LENA system in Russian is **not yet applicable** for tracking the language environment of Russian-speaking households.
3. The LENA system does not presently match the perceptual skills of human listeners, especially not if listeners have the possibility to listen to segments several times.
4. LENA's counts should be **considered more as estimations rather than accurate counts**. This by no means lessens the importance of the system, being that approximations are more beneficial than having no insight into a child's natural, everyday environment.
5. The LENA system is a valuable resource to have in order to educate and involve families with their child's development.
6. Having a tool like the LENA can provide clinician's a more thorough and dynamic early intervention model as well as the opportunity to provide intervention in a sustainable, realistic way.