

## LENA computerized automatic analysis of speech development from birth to three

Li-Mei Chen

National Cheng Kung University, [leemay@mail.ncku.edu.tw](mailto:leemay@mail.ncku.edu.tw)

D. Kimbrough Oller

University of Memphis, USA, [koller@memphis.edu](mailto:koller@memphis.edu)

Chia-Cheng Lee

University of Wisconsin-Eau Claire, USA, [jiajiajanejane@gmail.com](mailto:jiajiajanejane@gmail.com)

Chin-Ting Jimbo Liu

National Chin-Yi University of Technology, [jimbo@ncut.edu.tw](mailto:jimbo@ncut.edu.tw)

### Abstract

This study investigated the relationship between the linguistic input children receive and the number of children's vocalizations by using computerized LENA (Language Environment Analysis) system software which is able to collect and analyze the data automatically, instantly, and objectively. Data from three children (two boys and one girl) at ages of 10, 20, and 30 months were analyzed. The results indicated that: 1) child vocalizations (CV) and child-caregiver conversational turns (CT) increased with time while adult word counts (AWC) children overheard in the background decreased; 2) CT is highly correlated with CV. LENA automatic device can substantially shorten the time of data analysis in the study of child language, and provide preliminary results for further analysis.

### 1. Introduction

This paper explores the relationship between children's vocalization and the linguistic input they received by using LENA (Language Environment Analysis <https://www.lena.org/>), a computer system of automatic, objective and inexpensive device that collects and analyzes data for up to 16-hour-long recordings. It has been shown that the audio input children exposed to might largely affect their vocalizations [1], [2], [3], which in turn serve as an essential indicator of their later language development [4], [5].

## 1.1 Volubility as early indicators of future language development

Infants' volubility can be defined as the number of speech-like utterances made by infants within a certain period of time. Vocalizations appear with a variety of forms, including but not limited to cooing, vocal play and canonical babbling (e.g., "ba", "ma"). The forms of vocalizations vary from time to time, depending on the maturation of the speech motor control ability.

Studies have shown that infants' vocalization was associated with their later speaking ability. Typically developing (TD) children produced 41% more vocalizations than those with cleft lip and palate [6]. Similarly, children with severe-to-profound hearing loss were reported to have their onset of canonical babbling later than those of TD children [7]. In short, it has been confirmed that infants' early vocalizations are an essential indicator of their later language development.

## 1.2 Audio input and later language development

Input from adults is identified as one of the most influential factors affecting children's verbal performances [8]. However, it is still inconclusive whether children can acquire language by overhearing. In other word, it is still debatable if the linguistic input in the background plays a significant role in children's language development. Many scholars claimed that language acquisition takes place even when the linguistic materials, which children are exposed to, are not addressed to them directly [9]. On the other hand, other scholars concluded that, in comparison to adult conversations overheard by children, speech addressed directly to children have stronger effect on children's vocabulary learning [10]. Studies investigating the relationship among adults' speech in the background, adults' speech directly addressed to children and children's vocalizations will shed light on our understandings of the relationship between different types of linguistic input and language development.

## 1.3 Coding efficiency and length of observation

Due to the laborious measurements required for measuring the linguistic input in the ambient environment, studies focusing on infant's vocal developments are usually based on limited data [11]. In view of this, LENA, an innovative recording device

with automatic analyzing capacities, is adopted for data collection in the present study. The device weighs about 70 grams and can be put into children's vest pocket. More importantly, the device is able to run for a consecutive of 16 hours to record the holistic picture of the input a child receives when s/he is awake. LENA is able to calculate the number of the words adults produce in the background (i.e. input by overhearing) and to calculate the number of the conversational turns between the child and the interlocutor (i.e. speech addressed directly to the child). Additionally, the vocalizations made by the child are counted automatically, too. It has been shown that the results from LENA are accurate and reliable [12].

## 1.4 Present study

The goal of the present study is to examine the contributions from two sorts of input (i.e. overhearing and speech addressed directly to the child) to children's vocalizations. The specific questions are:

- a. Do adult word counts in the background (AWC), conversational turns (CT) and number of child vocalizations (CV) produced by children increase as they grow older?
- b. Are both AWC and CT effective contributors to the number of CV at the ages of 10, 20 and 30 months?

## 2. Method

### 2.1 Participants

Data from three children (two boys and one girl) growing up in a Chinese-speaking family were recorded and analyzed by LENA. Each child participated in the 16-hour-long recording session twice in a month. In the current study, data at tenth, twentieth, and thirtieth months were reported. According to the reports from the parents, all the children were full-term and were without hearing and speech-related disorders.

### 2.2 Equipment, procedures and data analysis

LENA was the equipment used to collect and analyze the data. The recording device, digital language processor (DLP), was put in a vest specially designed for children (Figure 1). The caretaker turned the device on to start a session. The

caretakers were instructed and understood how to operate, interrupt and recharge the device. After the recording sessions were completed, DLP was sent back to the experimenters and the device was connected with a computer of LENA system software which automatically uploaded and processed the audio file (Figure 2) and developed the reports (Figure 3).

The adult word counts measure the total number of adult words the child hears per hour, per day and per month. The conversational turns display the total number of conversational turns the child engages in with an adult per hour, per day and per month. The child vocalizations array the total number of vocalizations the child produces per hour, per day and per month.

The core design of LENA system is the algorithmic models which is capable to segment and identify sounds of different amplitude and intensity. The audio data were categorized into eight categories with different sources: the key child (wearing LENA DLP), other child, adult male and adult female, overlapping sounds, noise, electronic sounds (e.g., TV), and silence. Each category is further identified as clear and unclear (or quiet and distant) subcategories. In order to identify a variety of audio signals efficiently and accurately, LENA audio processing models are trained by audio transcriptions of each category first. For example, the speech processing algorithms need to recognize the differences between adult speech and child speech, and between the key child and other children [13].



Figure 1. The LENA digital language processor (DLP) placed in the pocket of LENA vest.

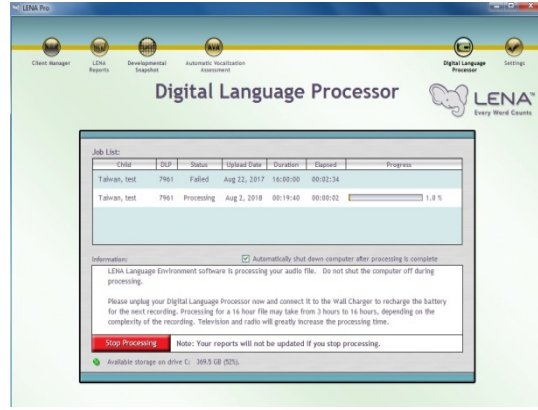
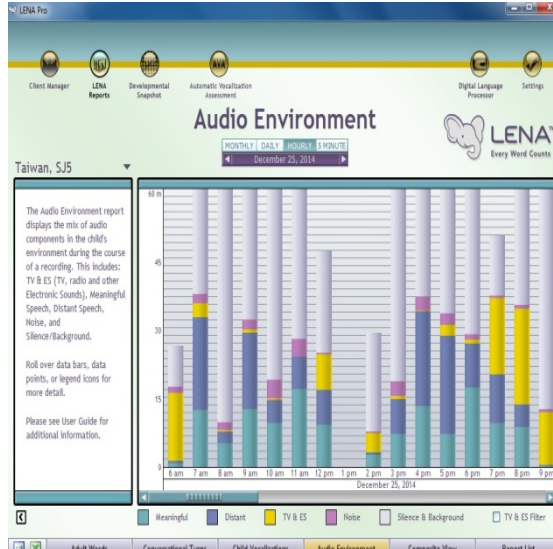
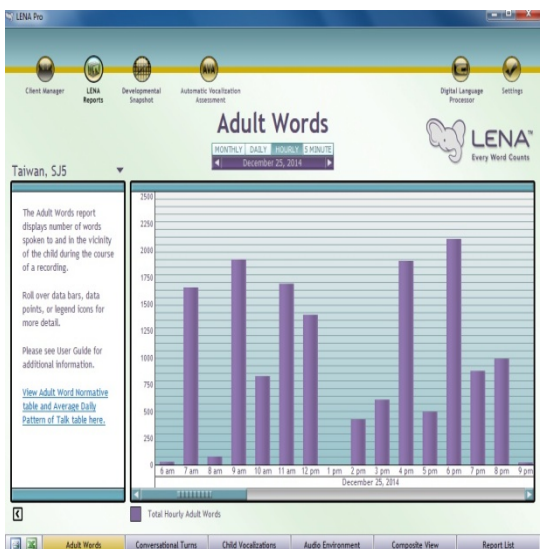
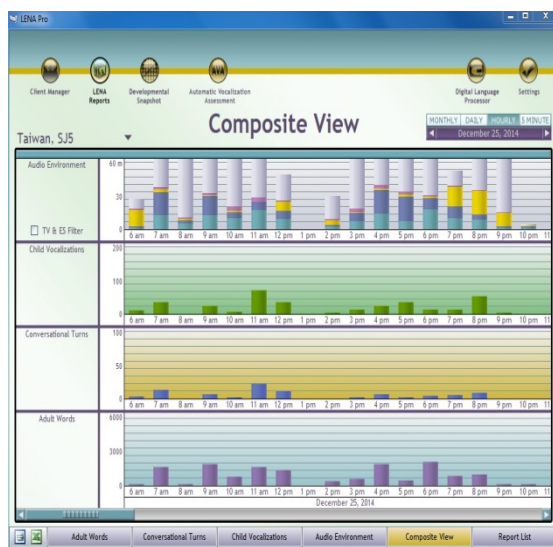
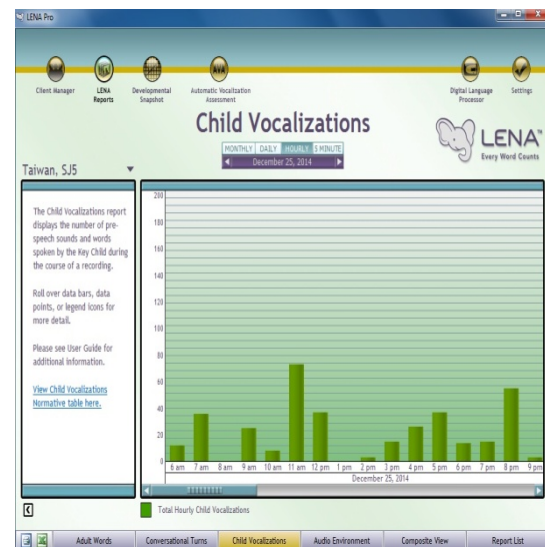


Figure 2. Data transfer from DLP to LENA system software



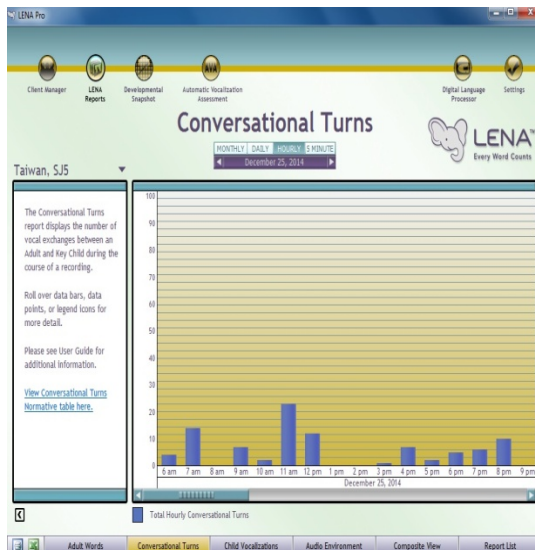


Figure 3. Reports from LENA system software

In order to answer the research questions, two statistical measures were performed, First, three one-way repeated measures ANOVAs were performed to explore if there were any changes of the three variables, (adult word counts, conversational turns, and child vocalizations) along the time. Next, multiple regressions were performed at the ages of 10, 20 and 30 months to examine how much AWC and CT contribute to CV at each age.

### 3. Results and Discussion

Figure 4 shows the average number of child vocalizations (CV), adult word counts (AWC), and conversational turns (CT) from the recordings made at 10, 20, and 30 months. The average number of CV gradually increased from 10, 20, and 30 months. The average number of CT also showed a gradual increase from 10, 20, and 30 months although the increase was small. However, the average number of AWC showed a decrease from 10, 20, and 30 months. Results of the three one-way repeated measures ANOVAs did not show significant differences among the three ages (AWC:  $F(2, 10)=1.89, p = 0.20$ ; CT:  $F(2,10)=0.43, p = 0.67$ ; CV:  $F(2,10)=1.03, p = 0.39$ ), suggesting that the results of the mean of AWC, CT, CV were similar across the three ages.

The results of multiple regressions showed that CT and CV were highly correlated at all the three ages (10 months  $p<.001, \beta = 3.746$ ; 20 months  $p<.001, \beta = 3.600$ , 30 months  $p<.01, \beta = 4.829$ ). However, AWC and CV were not correlated at

any of the ages (10 months  $p = 0.08$ ,  $\beta = -0.15$ ; 20 months  $p = 0.06$ ,  $\beta = -0.08$ ; 30 months  $p = 0.1$ ,  $\beta = -0.18$ ). Since a CT can be calculated when a child speaks and an adult responds, or when an adult speaks and the child responds, it is hard to identify if the increase in CT reported in Figure 4 was a result of the adult's responses to the child's vocalizations or vice versa.

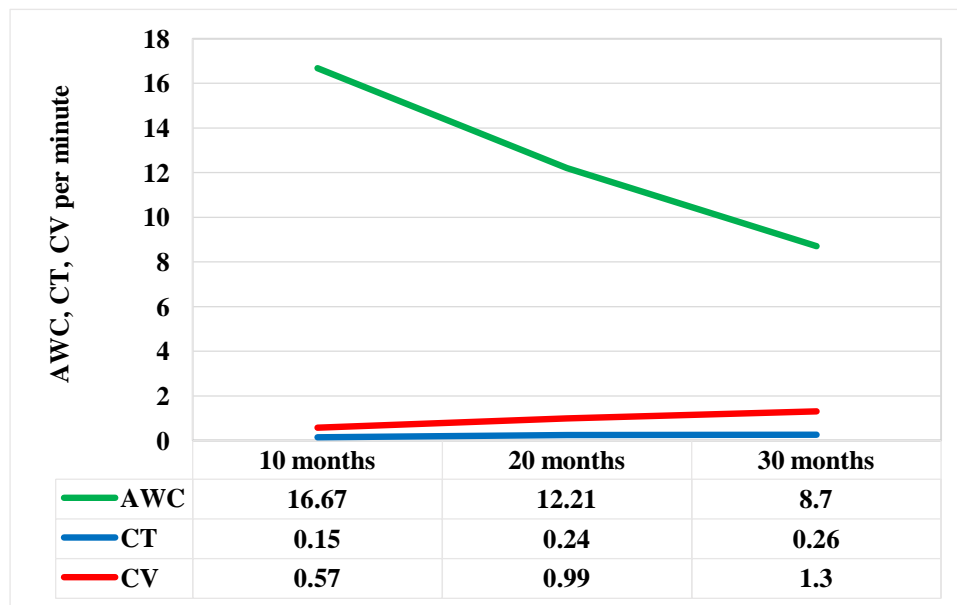


Figure 4. Average adult words (AWC), conversational turns (CT), child vocalizations (CV) per minute at 10, 20, and 30 months

It was also found that AWC decreased with child's ages. One possible explanation is that parents or other people may have been very voluble when addressing their 10-month-old child even when the child's utterances were not meaningful words or when the child did not vocalize. As the child grew up and started to produce words and sentences, parents may talk or respond only when their child produced meaningful utterances.

Figure 5 shows AWC per minute from recordings of individual child at three ages. SJ5's AWC was overall higher than YC6's and CC8's AWC. SJ5's and YC6's AWC decreased from 10 to 30 months. However, CC8's AWC showed an increase from 10 to 20 months and a slight decrease from 20 to 30 months.

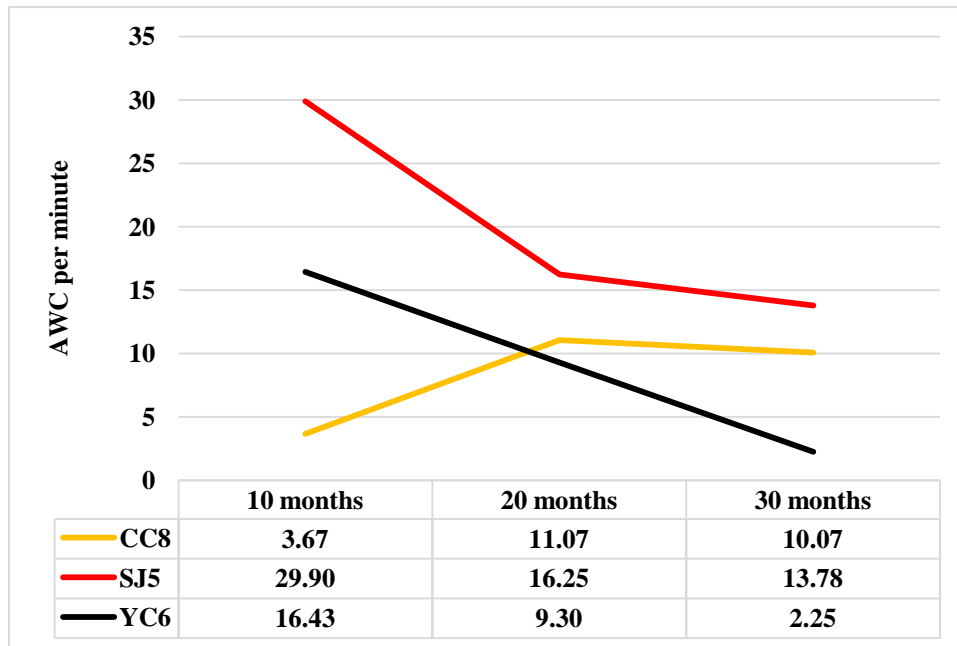


Figure 5. Adult words (AWC) per minute by individual child and age

Figure 6 and 7 shows CT and CV per minute from recordings of each individual child at three ages. The results of CT and CV showed a similar pattern. For example, CC8's and SJ5's CT and CV increased from 10 to 30 months. YC6's CT and CV decreased from 10 months to 30 months. Although SJ5's AWC was the highest of the three children, his CT and CV were the lowest among the three children.

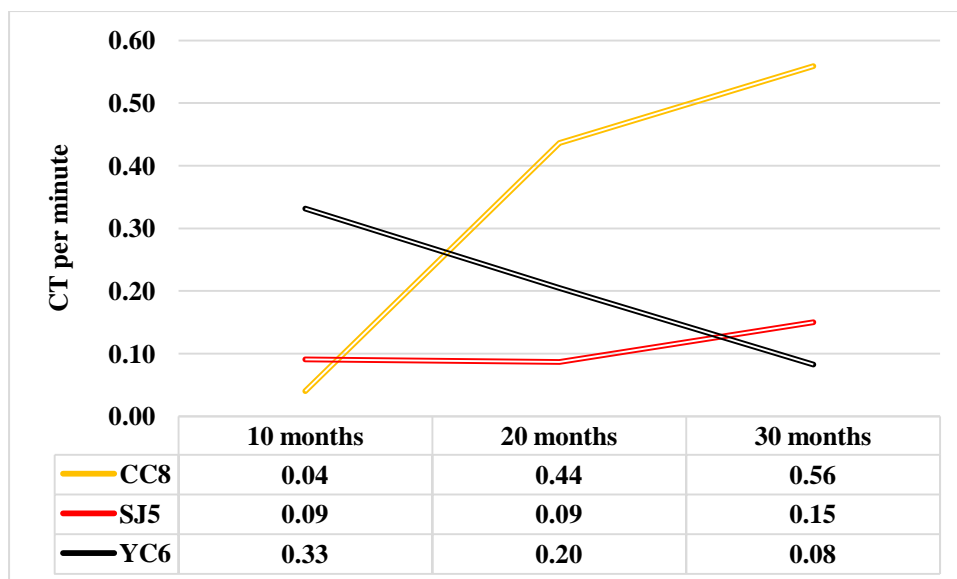


Figure 6. Conversational turns (CT) per minute by individual child and age



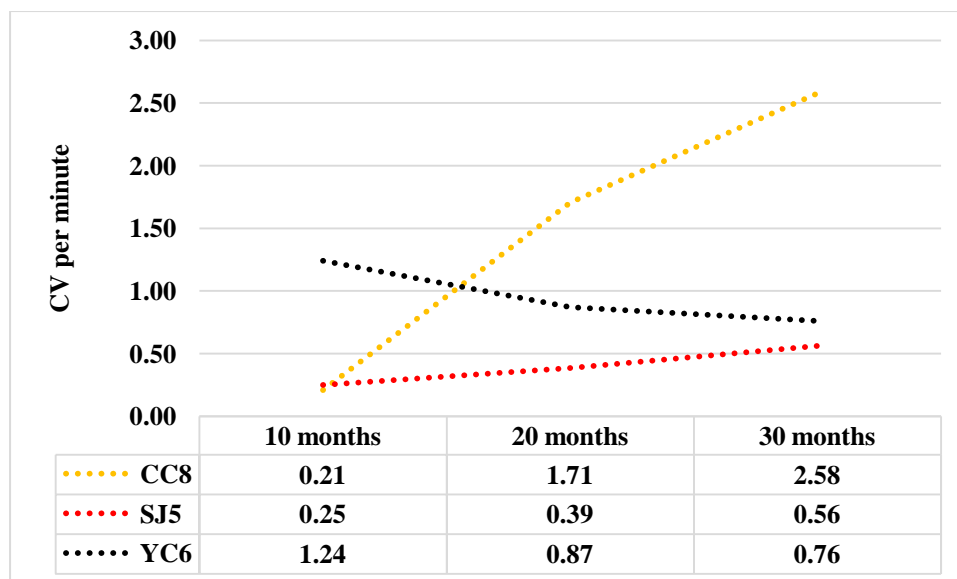


Figure 7. Child vocalizations (CV) per minute by individual child and age

#### 4. Summary

- (1) LENA system software is an objective, efficient, and reliable device to display profile of child's speech and the linguistic inputs.
- (2) Child vocalizations and conversational turns increased with time, but the numbers among the three ages were not statistically different.
- (3) Adult word counts decreased as the child grows older, the counts among the three ages were not statistically different either.
- (4) There was a high correlation between child vocalizations and conversational turns.
- (5) No correlation was found between child vocalizations and adult word counts.
- (6) The preliminary findings suggested that speech directed to children (i.e. CT) exerted much stronger influence to children's vocalizations than speech overheard by children (i.e. AWC) did. That is, the findings supported the view proposed in [10], but not in [9]. However, cautions should be taken in this interpretation as only limited participants were included. Therefore, more data from more ages and from more children are crucial to verify the current findings.

## 5. Acknowledgement

This research was supported by Chiang Ching-Kuo Foundation for International Scholarly Exchange in Taiwan to Li-Mei Chen. A special thank you is extended to the families of the children in this longitudinal study for their support of this project.

## 6. References

- [1] D. Suskind, K. Leffel, M. Hernandez, S. Sapolich, E. Suskind, E. Kirkham and P. Meehan, "An Exploratory Study of 'Quantitative Linguistic Feedback'", *Communication Disorders Quarterly*, vol. 34, no. 4, pp. 199-209, 2013.
- [2] M. Rowe, "A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development", *Child Development*, vol. 83, no. 5, pp. 1762-1774, 2012.
- [3] M. Caskey, B. Stephens, R. Tucker and B. Vohr, "Importance of Parent Talk on the Development of Preterm Infant Vocalizations", *Pediatrics*, vol. 128, no. 5, pp. 910-916, 2011.
- [4] B. Franklin, A. Warlaumont, D. Messinger, E. Bene, S. Nathani Iyer, C. Lee, B. Lambert and D. Oller, "Effects of Parental Interaction on Infant Vocalization Rate, Variability and Vocal Type", *Language Learning and Development*, vol. 10, no. 3, pp. 279-296, 2014.
- [5] E. Patten, K. Belardi, G. Baranek, L. Watson, J. Labban and D. Oller, "Vocal Patterns in Infants With Autism Spectrum Disorder: Canonical babbling status and vocalization frequency", *Journal of Autism and Developmental Disorders*, vol. 44, no. 10, pp. 2413-2428, 2014.
- [6] N. Scherer, A. Williams and K. Proctor-Williams, "Early and Later Vocalization Skills in Children With and Without Cleft Palate", *International Journal of Pediatric Otorhinolaryngology*, vol. 72, no. 6, pp. 827-840, 2008.
- [7] S. Iyer and D. Oller, "Fundamental Frequency Development in Typically Developing Infants and Infants With Severe To Profound Hearing Loss", *Clinical Linguistics & Phonetics*, vol. 22, no. 12, pp. 917-936, 2008.
- [8] M. Caskey, B. Stephens, R. Tucker and B. Vohr, "Importance of Parent Talk on the Development of Preterm Infant Vocalizations", *PEDIATRICS*, vol. 128, no. 5, pp. 910-916, 2011.

- [9] N. Akhtar, J. Jipson and M. Callanan, "Learning Words Through Overhearing", *Child Development*, vol. 72, no. 2, pp. 416-430, 2001.
- [10] A. Weisleder and A. Fernald, "Talking to Children Matters: Early language experience strengthens processing and builds vocabulary", *Psychological Science*, vol. 24, no. 11, pp. 2143-2152, 2013.
- [11] D. Oller, P. Niyogi, S. Gray, J. Richards, J. Gilkerson, D. Xu, U. Yapanel and S. Warren, "Automated Vocal Analysis of Naturalistic Recordings from Children With Autism, Language Delay, and Typical Development", *Proceedings of the National Academy of Sciences*, vol. 107, no. 30, pp. 13354-13359, 2010.
- [12] J. Gilkerson, Y. Zhang, D. Xu, J. Richards, X. Xu, F. Jiang, J. Harnsberger and K. Topping, "Evaluating Language Environment Analysis System Performance for Chinese: A pilot study in Shanghai", *Journal of Speech Language and Hearing Research*, vol. 58, no. 2, p. 445, 2015.
- [13] M. Ford, C. T. Baer, D. Xu, U. Yapanel and S. Gray, "The LENA Language Environment Analysis System: Audio specifications of the DLP-0121, LTR-03-2", LENA Foundation, 2008.