The predictability of naturalistic evaluation of all-day recordings for speech and language development

Seunghee Ha[1], Tae-Jin Yoon[2], Jungmin So[3]

[1] Div. of Speech Pathology and Audiology, Hallym University, Korea, shha@hallym.ac.kr
[2]Dept. Of English Language and Literature, Sungshin Women's University, Korea, tyoon@sungshin.ac.kr
[3]Dept. of Computer Science and Engineering, Sogang University, Korea, jso1@sogang.ac.kr

ABSTRACT

This research assessed the predictability of naturalistic evaluation of all-day recordings for speech and language development in children who acquire Korean language. One-day home recordings from 20 Korean children were collected using a Language Environment Analysis (LENA) recorder at 6-8, 12-14, 18-20 months. Both LENA's automated measures and measures from human coding were obtained from recordings at 6-8 and 12-14 months. A number of different words (NDW), consonant inventory, and utterance structure inventory were identified from recordings of 18-20 months. The results showed that conversation turn count (CTC) and canonical babbling ratio (CBR) at 6-8 months could significantly predict NDW at 18-20 months. Also, CTC and CBR at 12-14 months could significantly predict consonant inventory size at 18-20 months. This study highlights that rather than the amount of speech to children or overheard by children, conversational interactions between infants and caregivers positively affect children's later speech and language outcomes.

Keywords: LENA's automated measures, human coding, vocalization, speech & language development.

1. INTRODUCTION

Since Hart & Risley (1995) [1] provided strong evidence that the amount of caregiver speech to children plays a critical role for child cognitive and language development later in life, many studies have focused on the properties of early language environment and their relationship with various aspect of child development. These research efforts have been enhanced in the aspect of research methodology with the advent of automatic speech processing technologies. In particular, the LENA system enables us to collect daylong recordings of a child's language environment and obtain automated analyses of speech data. The LENA system makes it possible to capture children's naturalistic full language environment and overcome methodological limitations associated with language environment research.

The LENA software segments and analyzes the daylong acoustic data, generating automatic estimates related to child vocalization and language environment. The measures include child vocalization count (CVC), adult word count (AWC), and conversation turn count (CTC), which are the most relevant to Hart & Risley (1995)' [1] s measures to reflect the child language environment. Also, the LENA system provides measures that represent child auditory environment such as durations for TV and media, background noise, overlapping speech, and silence. Many studies quantified interactions between children and their caregivers using these LENA automated measures. They made an effort to shed light on the effect of language environment on various aspects of child development and explain individual differences and variability in child development [2-3]. In addition, the research efforts have expanded to the at-risk population, including preterm infants, autism spectrum disorder, and hearing loss [4-6].

Although there is no doubt that LENA automated measures reflect a naturalistic language environment, many researchers from various language backgrounds have tested and discussed the accuracy and reliability of LENA automated measures [7-8]. The research suggests that LENA automated measures have predictability for child language development. According to normative studies of the LENA Research Foundation [9], the segmentation agreement percentages between LENA automated measures and human transcribers were 82% and 76% for adults and children, respectively. This issue related to the validity and reliability of the LENA system is very important so that researchers and clinicians with a variety of language backgrounds can apply the LENA system to their own language settings. As the algorithm used to provide the LENA automated measures was trained on American English, research to verify the reliability of the LENA system in languages other than English should be preceded to adopt the LENA system in their respective languages. For Korean language, the two studies [10-11] evaluated the accuracy of the LENA system applied to Korean. Specifically, McDonald et al. (2021) [10] showed that the correlation between

LENA and human coding for AWC was *r*= .78 or .79 in Korean language, which was identical or similar to that of other studies for other languages [9] as well as Korean [11]. The correlation between LENA and human coding for CVC and CTC was relatively lower (*r*= .34 - .47). However, it fell within ranges found in non-European languages, and Korean language may not lead to more significant error rates than English and other languages. These findings indicate that caution should be needed for research and clinical use of LENA automated measures. For better research and clinical condition, using measures from human coding as a supplementary approach would be ideal.

Furthermore, the LENA automated measures tend to provide only broadly quantified information about the language environment and infant volubility. Specifically, it is difficult to determine to what extent of AWC would involve child-directed speech and to what extent of CVC included speech-like vocalizations or canonical babbling which were considered as more advanced forms of vocalization and precursors to mature phonology. Therefore, many recent studies investigated infant vocalizations utilizing human coding on the LENA recordings [12-13]. Given the importance of the quality of infant-caregiver interactions over the quantity of language input, further investigations are needed to be based on accurate coding of infant-caregiver interactions and as well as valid language environment sampling.

The purpose of this study was to investigate the predictability of measures from LENA recordings reflecting vocal development status and language environment in the prelinguistic stage for later language and speech development. Specifically, this study aimed to explore whether both measures from human coding and the LENA automated measures at 6-8 months and 12-14 months of age predict vocabulary and phonological development at 18-20 months.

2. METHOD

2.1. Participants

Nineteen children in Korean monolingual environments completed all the recording sessions among children recruited for this longitudinal study, and one child will complete the last recording session in January 2023. They consisted of 12 boys and 8 girls. All parents in the study were categorized as middle-class or above and college-educated in two-parent families. All children were recruited when they were between 6 and 8 months of age. According to their parents' reports, none of the children had significant physical, developmental, or hearing problems. This study was approved by the Hallym University Institutional Review Board. Signed informed consent forms were obtained from the children's parents.

2.2. Data Collection

Complete all-day home recordings using LENA recorders were obtained from each child at 3 different time blocks with 6-month intervals. The initial data collection occurred between 6 and 8 months of age, and subsequent data were collected at 12-14 months and 18-20 months of age. Research assistants visited the infants' home, interviewed the primary caregiver regarding the child's general developmental status, and provided detailed instructions on the LENA recorder use. The parents were asked to have their child wear a specially designed vest in a chest pocket with the LENA recorder. The parents were instructed to turn on the LENA recorder as soon as the child awoke in the morning and to maintain the recorder on all day until switching off automatically. Each recording provided up to 12-16 hours of data outside infants' sleep time in their natural environment.

2.3. Data Analysis

2.3.1. LENA automated analysis

Complete LENA recordings were automatically processed with LENA Pro software. Automated measures were obtained for one day-long recording collected at 6-8 months and 12-14 months of age. The measures included (1) AWC (the number of adult words spoken in the child's auditory environment), (2) CTC (the number of conversational turn-taking interactions between the child and adults), (3) CVC (the number of child vocalizations), and (4) the total duration of all segments classified as electronic media exposure. The LENA software supplies data for each measure in terms of both values from the actual recording duration and The 12-hour projected values. The measures were selected from 12-hour projected values as each recording duration was different across sessions and children.

2.3.2. Human coding

In-depth recording analyses on the LENA recordings were conducted based on human coding with the Action Analysis Coding Training (AACT) software. For the human coding, 20 5-minute segments with the child's highest vocalization rate determined by the LENA automated analysis were selected from one-day home recordings of each child.

In this study, an utterance is a vocalization occurring within one breathing cycle and consists of single or several syllables. A syllable in infant vocalization indicates the minimal rhythmic unit of an

utterance, and coders can count the number of syllables in an utterance by listening to "beats" in the rhythm of vocalization during coding [13]. The coders in this study were intensively trained to focus on syllables within an utterance, count the number of syllables, and determine syllable types reliably. They could observe changes in amplitudes of the waveform in the acoustic display of the AACT, by which they could supplement their judgments on the number of syllables. The coders classified syllables as one of four types: true canonical syllables, non-canonical syllables, syllables with glottal consonants only, or syllables with glide consonants only. The coders were trained to mark each syllable as one of the four types in real-time by keystrokes that were automatically recorded in the log of the AACT software. They also were trained to determine utterances of persons other than the child whose voices could be heard on the recordings. They determined whether utterances were directed toward the infant or adults or toward someone other than the infant wearing the record while listening in real-time.

Measures from human coding on the LENA recordings of 6-8 months and 12-14 months included the number of total syllables and the canonical babbling ratio (CBR). CBRs were calculated as the number of true canonical syllables plus syllables with glides by the number of all syllables. Also, the number of utterances directed toward the infant (infant-directed speech [IDS]) was included as a potentially predictive measure.

2.3.3. *Vocabulary and phonological measures*

Vocabulary and phonological measures were analyzed from the LENA recordings obtained at 18-20 months. The coders evaluated vocabulary size from 20 5-minute segments with the child's highest vocalization rate selected from one-day home recordings of each child in the same manner as the LENA recordings of 6-8 and 12-14 months. The coders identified words based on criteria proposed by [14], determinative context, phonetic match, and relationship with other vocalizations and counted the number of different words for each child. The coders then performed broad phonetic transcription on the identified words to determine the consonant inventory and the utterance shapes. The consonant had to occur at least three times in the sample to be considered an "in inventory" consonant for the size of the consonant inventory. In addition, to examine whether children produce various syllable sequences within utterances, the number of different utterance shapes was determined by counting the number of utterance shapes occurring at least three times in the sample.

The two main coders who received training sessions on human coding of vocalization participated in the analysis of recordings from the LENA recordings obtained at 18-20 months of age. They are also involved in the examination of agreements. Approximately 10% of the entire data set was randomly selected and reanalyzed for inter- and intra-coder/transcriber agreements on the vocabulary and phonological measures. Cohen's Kappa coefficients were calculated for inter-coder and intra-coder agreements on the word identification from recordings of 18-20 months, which revealed $\kappa = .752$ and $\kappa = .997$, respectively. The reliability of phonetic transcription was determined on consonants and word shapes. Inter-transcriber and intra-transcriber agreements on consonants were 87.19% and 81.68%, respectively. Inter-transcriber and intra-transcriber agreements on word shapes were 80.07% and 81.80%, respectively.

3. RESULTS

The correlation and linear regression analyses were conducted with measures obtained from 19 children who had completed the longitudinal study. Table 1 presents the results from correlation analyses between speech and language measures at 18-20 months and measures from LENA automated analysis and human coding at 6-8 months and 12-14 months. NDW, consonant inventory, and utterance structure inventory at 18-20 months showed the highest positive correlations with CTCs at 6-8 months and 12-14 months. The amount of IDS also had higher positive correlations with measures from LENA automated analysis and human coding at 6-8 months and 12-14 months.

Linear regressions were conducted to examine the influence of predictor variables on speech-language development at 18-20 months. The regression analyses showed that at 6-8 months, CTC and CBR accounted for 65.8% and 10.9 % of the variance in NDW at 18-20 months, respectively. Also, CTC accounted for 32.8 % and 28.3% of variances in consonant inventory and utterance structure inventory, respectively. At 12-14 months, CTC, AWC, IDS, and CBR significantly accounted for 59.7%, 16.4%, 5.6%, and 5.4% of the variance in NDW at 18-20 months, respectively. CTC and CBR accounted for 38.6% and 16.8% of the variance in consonant inventory, and CTC accounted for 41.8% of the variance in utterance structure inventory at 18-20 months.

Table 1. Pearson coefficients between measurements at 6-8, 12-14months, and speech-language development at 18-

20months

LENA measurement	Speech-language development at 18-20 months		
	NDW	Consonant inventory	Utterance structure inventory
6-8months			
AWC	.525*	.403	.279
CTC	.811**	.573*	.532*
CVC	.795**	.549*	.512*
IDS	.550*	.483*	.452
CBR	.456*	.354	.447
12-14months			
AWC	.356	.338	.396
CTC	.773**	.622**	.646**
CVC	.666**	.487*	.441
IDS	.611**	.347	.590**
CBR	.411	.526*	.421

AWC=adult word counts; CTC=conversational turn count; CVC=child vocalization count; NDW=number of different words;CDS = the amount of child directed speech identified by human coding; CBR = canonical babbling ratio

*p< 05, **p< .01.

4. DISCUSSIONS

The findings provide strong evidence for the predictive validity of the LENA system to analyze language environments with human coding automatically. The results highlight that conversational interactions between infants and caregivers positively affect children's later speech and language outcomes rather than the amount of speech to children or overheard by children. This longitudinal study has significant implications for identifying children at risk for poor speech-language development and providing appropriate early interventions to improve the early home language environment.

5. REFERENCES

[1] Hart, B. and Risley, T.R., 1995. *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

[2] Gilkerson, J., Richards, J.A., Warren, S.F., Oller, D.K., Russo, R. and Vohr, B., 2018. Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*(4).

[3] Wang, Y., Williams, R., Dilley, L. and Houston, D.M., 2020. A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review*, *57*, p.100921.

[4] Ambrose, S.E., VanDam, M. and Moeller, M.P., 2014. Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear and hearing*, *35*(2), p.139.

[5] Caskey, M., Stephens, B., Tucker, R. and Vohr, B., 2014. Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics*, *133*(3), pp.e578-e584.

[6] Sulek, R., Smith, J., Bent, C.A., Hudry, K., Trembath, D., Vivanti, G. and Dissanayake, C., 2022. The utility of LENA as an indicator of developmental outcomes for young children with autism. *International Journal of Language & Communication Disorders*, *57*(1), pp.103-111.

[7] Cristia, A., Bulgarelli, F. and Bergelson, E., 2020. Accuracy of the language environment analysis system segmentation and metrics: A systematic review. *Journal of Speech, Language, and Hearing Research*, *63*(4), pp.1093-1105.

[8] Gilkerson, J., Zhang, Y., Xu, D., Richards, J.A., Xu, X., Jiang, F., Harnsberger, J. and Topping, K., 2015. Evaluating language environment analysis system performance for Chinese: A pilot study in Shanghai. *Journal of Speech, Language, and Hearing Research*, *58*(2), pp.445-452.

[9] Gilkerson, J., Richards, J.A., Warren, S.F., Oller, D.K., Russo, R. and Vohr, B., 2018. Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*(4).

[10] McDonald, M., Kwon, T., Kim, H., Lee, Y. and Ko, E.S., 2021. Evaluating the Language ENvironment Analysis System for Korean. *Journal of Speech, Language, and Hearing Research*, *64*(3), pp.792-808

[11] Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J.A., Ma, L. and Topping, K., 2016. Effects of feedback on parent–child language with infants and toddlers in Korea. *First Language*, *36*(6), pp.549-569.

[12] Lee, C.C., Jhang, Y., Relyea, G., Chen, L.M. and Oller, D.K., 2018. Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings. *Infant Behavior and Development*, *50*, pp.140-153.

[13] Ha, S. and Oller, K.D., 2022. Longitudinal Study of Vocal Development and Language Environments in Infants With Cleft Palate. *The Cleft Palate-Craniofacial Journal*, *59*(10), pp.1286-1298.

[14] Vihman, M.M. and McCune, L., 1994. When is a word a word?. *Journal of child language*, *21*(3), pp.517-542.