



Published in final edited form as:

*Otol Neurotol.* 2014 September ; 35(8): 1301–1305. doi:10.1097/MAO.0000000000000489.

## Feasibility Study to Quantify the Auditory and Social Environment of Older Adults using a Digital Language Processor

Lingsheng Li, M.H.S.<sup>1</sup>, Ami R. Vikani, M.P.H.<sup>2</sup>, Gregory C. Harris<sup>3,4</sup>, and Frank R. Lin, M.D., Ph.D.<sup>4,5</sup>

<sup>1</sup>University of Oklahoma College of Medicine, Oklahoma City, Oklahoma <sup>2</sup>The George Washington University School of Medicine and Health Sciences, Washington, District of Columbia <sup>3</sup>Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland <sup>4</sup>Center on Aging and Health, Johns Hopkins Medical Institutions, Baltimore, Maryland <sup>5</sup>Departments of Otolaryngology-Head & Neck Surgery, Geriatric Medicine, and Epidemiology, Johns Hopkins University, Baltimore, Maryland

### Abstract

**Objective**—To investigate the feasibility of using a digital language processor (DLP) to objectively quantify the auditory and social environment of older adults.

**Design**—Thirty-seven participants aged 64–91 residing in a retirement community were asked to wear a DLP to record their auditory and language environment during one waking day. Recordings were analyzed with specialized software to derive quantitative estimates such as the number of spoken words heard and percent of time spent around meaningful speech versus television/radio.

**Results**—Adequate DLP recordings that began before 10 A.M. and lasted for 10 hours were collected from 24 participants. The mean duration of recording was 13 hours and 13 minutes, and individuals spent a mean of 26.7% (range=4% – 58%) of their waking day near a television or other electronic sounds. The projected mean word count over a maximum of 16 hours was 33,141 with nearly a 14-fold range between the lowest and highest observed values (range=5120 – 77,882).

**Conclusions**—High-quality objective data on the auditory environment of older adults can be feasibly measured with the DLP. Findings from this study may guide future studies investigating auditory and language outcomes in older adults.

---

Address correspondence, reprint requests, and proofs to: Frank R. Lin, Johns Hopkins Center on Aging & Health, 2024 E. Monument St., Suite 2-700, Baltimore, MD 21205, Telephone: (443) 287-6509, Fax: (410) 502-6713, flin1@jhmi.edu.

**Conflict of Interest:** Dr. Lin reported serving as a consultant to Pfizer, Autifony, and Cochlear Corp; and receiving speakers fees from Amplifon. No other author reported any disclosures.

## Introduction

The LENA Language Environment Analysis system (LENA Foundation, Boulder, Colorado) was originally created as a tool to enable healthcare professionals and researchers to collect, process, and analyze language environment data in young children. This system is an automatic language collection and analysis tool that uses a digital language processor (DLP) to capture communication between the child and his or her caregivers in the child's daily environment. The audio file captured by the DLP can then be automatically analyzed using validated sound processing algorithms contained in the LENA software to produce detailed composite reports of the number of words spoken around the recorded individual, as well as the amount of time spent in the presence of televisions and other electronic media<sup>1</sup>. One study demonstrated the feasibility of using the DLP to help enhance the language environment of children who are deaf or hard of hearing in English and Spanish-speaking households<sup>2</sup>.

While studies have implemented the DLP to measure auditory perception and language acquisition in children<sup>1-3</sup>, no studies have used the processor to obtain information on the auditory and social environment of older adults. Metrics that could objectively quantify the auditory and social environment of older adults could be used to better understand the broader, downstream impact that hearing rehabilitative treatments (e.g., hearing aids, cochlear implants) have on the social functioning of older adults. Social engagement is a critical determinant of morbidity and mortality in older adult<sup>4</sup>, and currently is not an outcome that is routinely assessed after hearing rehabilitative treatment in adults.

The objective of this pilot study was to investigate the feasibility of using a digital language processor to quantify the auditory environment of community-dwelling older adults. Through our study, we aim to provide the preliminary data and experience to guide future research using the DLP to quantify the auditory and social environment of older adults. Such metrics will likely be increasingly important to utilize as we seek to demonstrate the benefits of services (e.g., hearing rehabilitative counselling after adult cochlear implantation; provision and training in assistive listening devices to work with a cochlear implant or hearing aid which remain poorly or not reimbursed by Medicare and third party payers). Importantly, researchers and healthcare professionals have long advocated for the importance of social engagement and connectedness in late-life. This feasibility study using the LENA system is an important first step in objectively quantifying social engagement in older adults and such outcome metrics could likely be utilized to demonstrate the benefits of many of the services provided in otology and audiology that are not captured in routine clinical metrics such as a speech understanding score.

## Materials and Methods

### Study participants

We recruited residents from Stadium Place, a retirement community for low and moderate income older adults 62 years in Baltimore City, to attend one of three informational sessions. Recruitment and retention of study participants were coordinated with the Director of Senior Services and Service Coordination Administrator at Stadium Place. Of the 52

residents who attended an information session, 44 (85%) residents agreed to participate in the study by completing questionnaires and wearing the DLP for one day. Seven participants dropped out of the study due to medical issues before using the DLP. Written informed consents were obtained from participants by trained research staff, and participants were compensated \$20. All study procedures were reviewed and approved by the Johns Hopkins Institutional Review Board.

### Digital Language Processor

The DLP (LENA Research Foundation, Boulder, CO) is a body-worn digital processor that records 16 hours of audio data (Figure 1). Audio data are analyzed using a computer-based software package that utilizes proprietary auditory processing algorithms to quantify and summarize various features of the audio data<sup>5</sup>. The DLP is approximately the size of a pager with a simple interface including a power button, record button, display screen, small microphone, and a USB port for transferring and processing audio data. To ensure optimal acoustic quality, the processor was originally designed by the LENA research foundation to be worn in custom-made clothing for children. For Stadium Place participants, we purchased black nylon cases (Timbuk2 I.D Money Clip –B008VSZXL4, Timbuk2, San Francisco) and adapted them to accommodate the DLP and to be worn around the neck. The case had a clear window through which participants could operate the power and record buttons, as well as the display screen used to identify when the DLP is idle or recording (Figure 2). Participants were asked to wear the device around their neck and outside of any clothing to ensure optimal sound quality.

At each informational session, each participant received detailed written and verbal instructions on how to use the DLP. A trained staff member showed each participant how to turn the device on and off, and how to begin and end recording. Participants were instructed to (1) turn on the DLP when they woke up in the morning, (2) wear the device during waking hours, and (3) turn off the DLP before bedtime. Participants were also asked to complete a brief questionnaire on processor usage (i.e. what time they turned on/turned off the DLP) for a quality assurance check. All participants were issued the device to use at the information session or within two weeks of enrollment in the study depending on the available number of DLP's and the number of participants enrolled at each session. To mitigate risks associated with loss of privacy, we asked participants to let others know that they may be recorded. Visible signage stating "This Recording Device is On" was affixed to each DLP before use. All participants were informed that the audio recording captured by the DLP would be permanently deleted. We systematically expunged all audio files after quantitative data had been extracted and processed.

The following quantitative outcome variables were extracted automatically by the LENA software<sup>3,5-7</sup>:

- **Projected adult word count (AWC):** Number of adult words spoken to and near the participant, as well as the number of words vocalized by the participant in natural, spontaneous speech environments. (We used projected values that were extrapolated and averaged to a maximum of 16 hours to match up to the LENA system's 12-hour normative data set and to standardize across individuals and

allow comparability between participants with variable recording durations). The AWC is an estimate based on vowel durations and syllable lengths, which allows the LENA system to filter out coughing, sneezing, laughing, and hiccupping<sup>8</sup>. Non-speech sounds such as coughing and vegetative sounds are filtered out and statistical models are used to estimate the number of words spoken in each adult segment<sup>5</sup>.

- **Meaningful speech:** Speech sounds within a six-foot radius that are algorithmically-defined as being “close and clear vocalizations”.
- **TV time:** Sound from a television, radio, or other electronic devices.
- **Distant speech:** Speech sounds coming from six or more feet away from the DLP, including overlapping speech, speakers who are not near the participant or segments that do not match auditory processing algorithms.

The LENA system’s accuracy and reliability for coding of auditory recordings have been previously tested and validated against coding of audio files performed by trained individuals<sup>1,3,5</sup>.

### Statistical Analysis

Based on the distribution of recording duration for participants during the waking day (Figure 3), we classified an adequate recording as recordings that began in the morning (before 10 A.M.) and lasted ≥ 10 hours. Independent group t-tests and Fisher’s Exact Test were used to compare the means of different demographic characteristics between participants with and without adequate DLP recordings. Significance testing for all analyses was 2-sided with a type I error of 0.05. The statistical software used was STATA 12 (STATA Corp, College Station, TX).

### Results

The demographic characteristics of study participants using the DLP are summarized in Table 1. Of the 37 participants, 24 (65%) participants had an adequate recording consisting of a recording that began before 10 A.M. and lasted for ≥ 10 hours. All participants were African American and the mean age for the study population was 75.9 years (SD=6.4, range=64.5–91.3). Participants were primarily women (89.2%) and living alone (91.9%). A majority of participants did not complete college education and had an annual household Income <\$25,000. We did not observe any differences in demographic characteristics comparing participants with and without adequate DLP recordings.

The duration of DLP recordings for the 37 participants are shown in Figure 3. For individuals without an adequate recording, 5 participants attempted to use the device (based on self-reported usage questionnaire) but did not have recordings stored on the DLP; 2 participants had < 10 hours of recording; and 6 participants did not begin recording until late afternoon which was several hours after their self-reported waking time.

The auditory processing algorithms contained in the software allow for quantitative temporal segmentation of an individual’s audio file into projected Adult Word Count and Audio

Environment Data. The duration of each day spent around meaningful speech and TV time varied widely between individuals (Figure 4). Estimates for time spent around meaningful speech varied from 23 minutes to over 6 hours, while TV times also varied from 40 minutes to nearly 7 hours. Composite summary metrics of the auditory environment of the 24 participants with adequate recordings are provided in Table 2. Adult word count (AWC) indicates the number of words spoken by the participant and words spoken to or in the vicinity of the participant that is considered to be “close and clear” or within a 6 foot radius of the participant by the processing software. The projected mean word counts for the 24 participants was 33,141 with nearly a 14-fold range between the lowest (AWC=5120) and highest observed values (AWC = 77,882) (range=5120–77,882). The mean duration of recording for these 24 participants was 13 hours and 13 minutes. On average, participants spent approximately 26.7% of their recording session near a television or other electronic devices.

For the 24 participants with adequate LENA recordings, we explored the preliminary face and construct validity of the derived metrics by examining the intercorrelations of the respective measures. We observed a negative correlation (coefficient=  $-0.55$ ,  $p < 0.01$ ) between projected Adult Word Count (AWC) and percentage of time spent in front of TV. Additionally, the duration of meaningful speech was negatively correlated with total TV time (coefficient=  $-0.44$ ,  $p < 0.05$ ). These findings have to be interpreted with caution, however, due to the small sample size of participants with adequate recordings.

## Discussion

Our study explored the feasibility of using a digital language processor as a novel method to characterize the auditory and social environment of older adults residing in a mixed-income retirement community. We found that the DLP could be successfully implemented to gather high-quality data on the auditory and language environment of older adults. In this cohort of individuals residing in a low socioeconomic and disadvantaged inner city neighborhood, we were able to successfully gather data from 24 of 37 participants with minimal participant training. Our results indicate substantial inter-individual variability in summary measures of time spent around meaningful speech, time spent in front of the TV and other electronic media, and projected adult word count.

Previous studies have also attempted to use electronic sensors to assess individuals' current behaviors and experiences in real time, giving rise to the current concept of “ecological momentary assessment”<sup>9</sup>. One study found that using electronically activated recorders (EAR) to periodically record ambient sounds from the participants' environment yielded highly naturalistic and informative data on health-relevant behaviors<sup>10,10</sup>. Another feasibility study using ecological momentary assessment demonstrated the importance of detailed and real-life measures in examining rehabilitation outcomes in hearing aid users; however, the results were still based on self-reported surveys<sup>11</sup>. The fundamental advantage of using the DLP is the ability to capture large quantities of real-time data that are capable of providing highly descriptive, objective, and noninvasive reports on auditory engagement in participants' daily lives.

Our results demonstrated substantial variability in potential indirect indicators of auditory modality, such as Adult Word Count (AWC), Meaningful Speech, and TV time. In a recent study, 24 hearing aid users providing a total of 991 completed ecological momentary assessments reported that “face-to-face” conversation was the most common hearing difficulty that they encountered, followed by problems experienced during telephone conversation, TV, and environmental sounds<sup>11</sup>. Therefore, calculated metrics based on the processing output could also be potentially used to characterize the auditory perception of older adults and the effects of hearing interventions. One substantial advantage of the DLP is that the software’s algorithms automatically calculate quantitative variables such as AWC, Meaningful Speech, and TV time and provide detailed composite reports of the participants’ auditory environment (similar to figure 4). In implementing the DLP, we found that asking participants to complete diary entries was helpful in assessing participant compliance and as a quality assurance check in comparison with the quantitative output. Based on the diaries, we were able to determine whether participants attempted to use the device even though they might have forgotten or had difficulty depressing the record button.

A limitation of using the DLP to capture auditory environment in the current pilot study is the reliance on a single-day recording session. Due to limited availability of devices, we were unable to provide more than one device per participant at each informational session. Repeated measurements of auditory data may provide a more optimal and reliable assessment of the participants’ daily experiences<sup>10</sup>. Further studies using the DLP should consider capturing multiple days of data to ensure reliable estimates of performance. Our metric of Adult Word Count is also limited to capturing the total number of adult words spoken in the auditory environment, rather than being able to distinguish between the words spoken by the study participant versus other individuals. In contrast, because the DLP speech processing algorithms were originally designed for children and caretakers, the system software can distinguish between words spoken by a child wearing the device (designated as the “key child”) versus other children or adults in the environment. Future versions of the DLP may also allow for identification of the “key adult” thereby allowing researchers to quantify the number of words spoken by the key adult and the number of conversational turns between adult speakers, which may be better indicators of actual social engagement.

There are also limitations to the current design of the DLP which may limit its use in older adults. We found that the two most common difficulties for participants operating the DLP and contributing to individuals providing inadequate recordings were (1) being unable to depress the ON button through the thick nylon case, and (2) remembering to press the record button to begin data capture. Future versions of the DLP could possibly address these issues through incorporating larger, “easy to press” buttons and/or automated start times for recording.

Our study sample consisted of mostly low-income African American women residents living in a subsidized retirement community and thus our experiences in using the DLP may not be generalizable to other older adult study populations. However, we note that our ability to successfully implement use of the DLP with minimal training and instruction in this relatively disadvantaged study population is encouraging. Our preliminary data support the

use of DLP data in providing meaningful indications of an individual's degree of social engagement (e.g., observed variance in adult word count, meaningful speech, TV time, etc.). We would expect that studies involving cohorts of higher-functioning older adults and clinical populations providing greater instruction in device use would lead to higher rates of successful data collection.

This feasibility study using a digital language processor is an important first step in developing methods to objectively quantify the auditory and social environment of older adults. These types of outcome measure may provide critical, real world insights into the benefits of hearing rehabilitative strategies and interventions that conventional clinical metrics (e.g., a speech understanding score) would not capture. We believe that our data and experience in using the DLP as a naturalistic observation method will help advance future studies investigating the role of real-world auditory and language profiles in the functional well-being of older adults.

## Acknowledgments

We thank the staff and residents of Stadium Place for their assistance and participation in this pilot study. We also thank the LENA Foundation for kindly lending us LENA DLPs for use in this study.

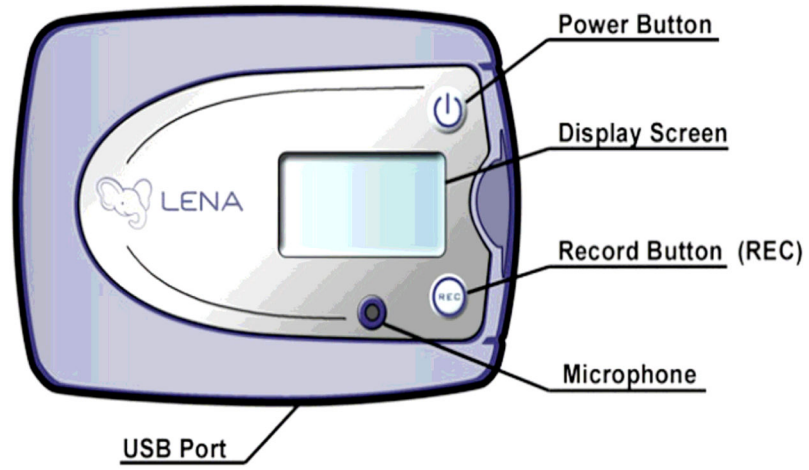
**Funding:** This study was supported by grant 1K23DC011279 from the National Institute on Deafness and Other Communication Disorders, with further funding from the Triological Society and the American College of Surgeons through a clinician scientist award, and funding from the Eleanor Schwartz Charitable Foundation.

## References

1. Christakis DA, Gilkerson J, Richards JA, et al. Audible television and decreased adult words, infant vocalizations, and conversational turns: a population-based study. *Arch Pediatr Adolesc Med.* 2009; 163(6):554–8. [PubMed: 19487612]
2. Aragon M, Yoshinaga-Itano C. Using Language Environment Analysis to improve outcomes for children who are deaf or hard of hearing. *Seminars in Speech Language.* 2012; 33(04):340–53. [PubMed: 23081793]
3. Yapanel, U.; Gray, S.; Xu, D. Reliability of the LENA Language Environment Analysis System in Young Children's Natural Home Environment. Boulder, CO: Infoture; 2007. Available at: <http://www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2> [Accessed June 10, 2013.]
4. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. *Soc Sci Med.* 2000; 51(6):843–57. [PubMed: 10972429]
5. Ford, M.; Baer, C.; Xu, D.; Yapanel, U.; Gray, S. [Accessed June 10, 2013.] The LENA language environment analysis system: Audio specifications of the DLP-0121. 2008. Available at: [http://www.lenafoundation.org/TechReport.aspx/Audio\\_Specifications/LTR-03-2](http://www.lenafoundation.org/TechReport.aspx/Audio_Specifications/LTR-03-2)
6. Gilkerson, J.; Coulter, K.; Richards, J. [Accessed June 10, 2013.] Transcriptional analyses of the LENA Foundation Natural Language Corpus. 2008. Available at <http://www.lenafoundation.org/TechReport.aspx/Transcription/LTR-06-2>
7. Xu, D.; Yapanel, U.; Gray, S. [Accessed June 10, 2013.] Reliability of the LENA Language Environment Analysis system in young children's natural home environment. Available at: <http://www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2>
8. Yoshinaga-Itano, C.; Gilkerson, J. [Accessed June 10, 2013.] Paradigm shifting: Automatic assessment of natural environments. 2010. Available at: <http://www.lenafoundation.org/pdf/agbell-2010.pdf>
9. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* 2008; 4:1–32. [PubMed: 18509902]

10. Mehl MR, Robbins ML, Deters FG. Naturalistic observation of health-relevant social processes: The electronically activated recorder methodology in psychosomatics. *Psychosom Med.* 2012; 74(4):410–17. [PubMed: 22582338]
11. Galvez G, Turbin MB, Thielman EJ, Istvan JA, Andrews JA, Henry JA. Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users. *Ear Hear.* 2012; 33(4):497–507. [PubMed: 22531573]

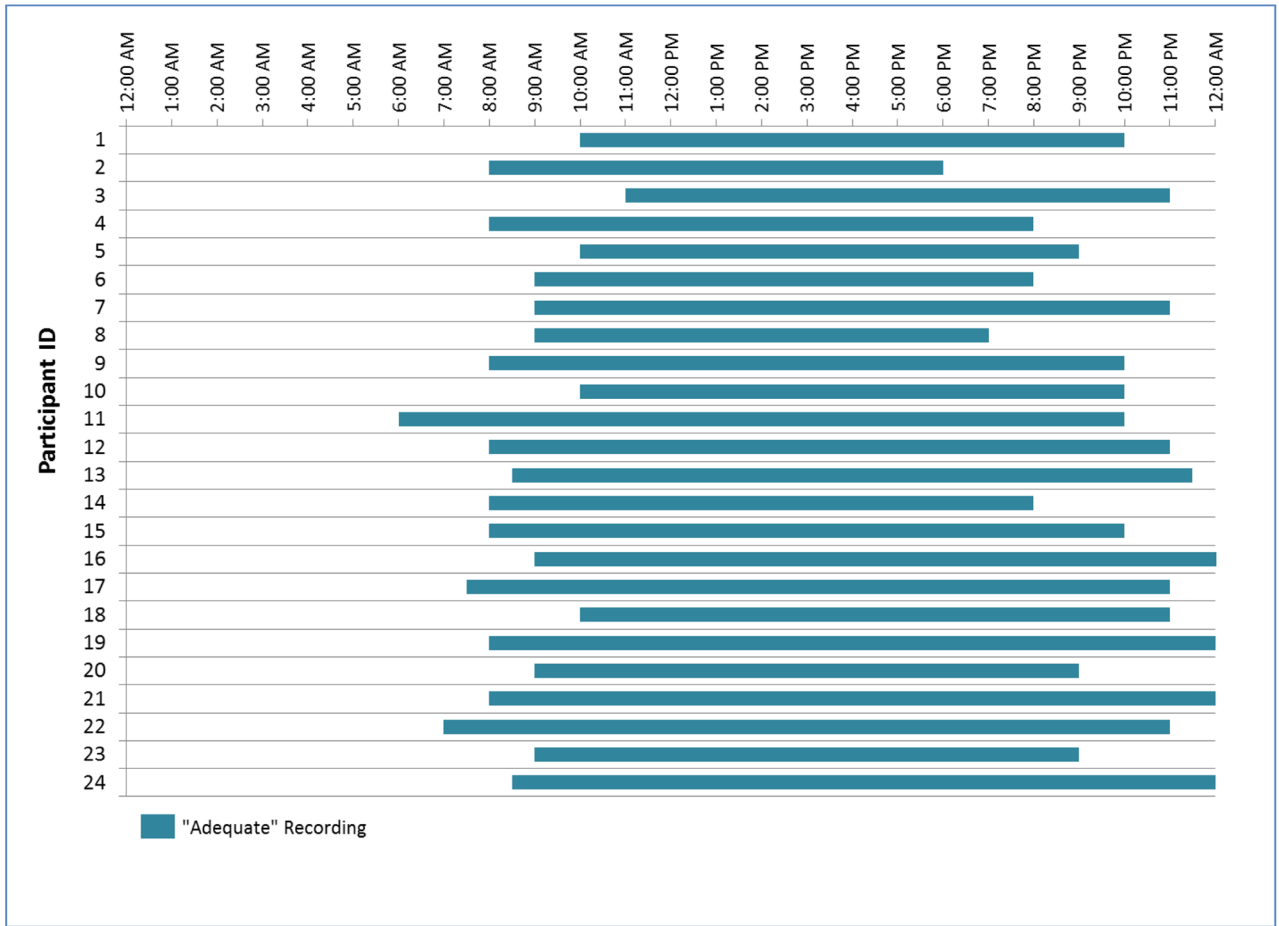




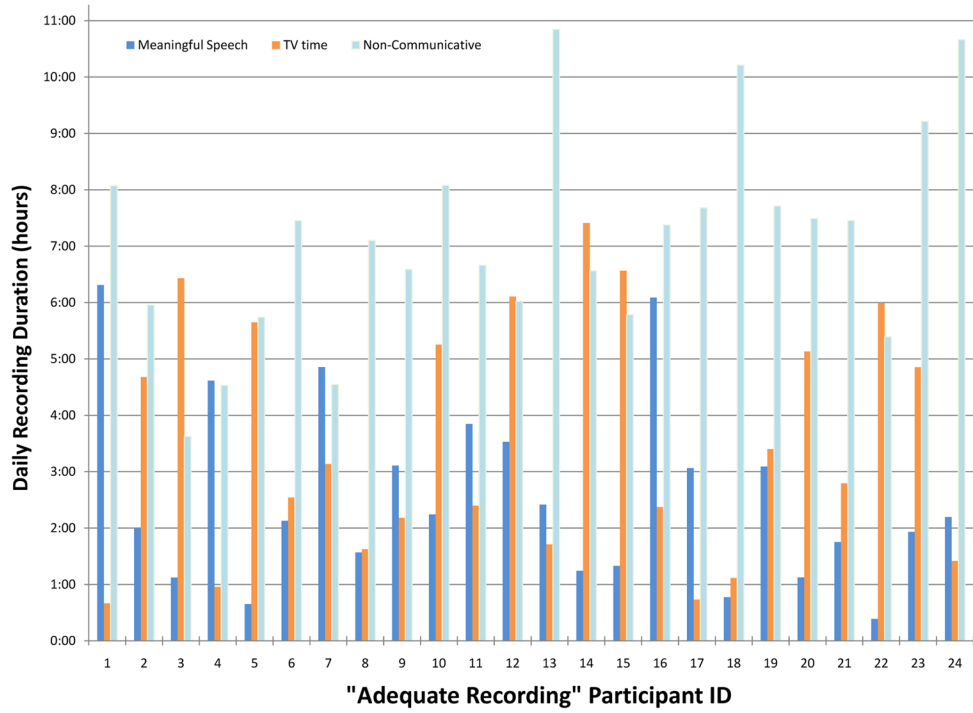
**Figure 1.**  
LENA Digital Language Processor (DLP) Interface



**Figure 2.**  
Digital Language Processor (DLP) positioned in the neck-worn protective sleeve



**Figure 3.** DLP recording duration in a 16-hour waking day for Stadium Place study participants with adequate recording, N=24



**Figure 4.** Composite DLP report showing TV time versus Meaningful Speech for Stadium Place study participants with adequate recordings, N=24

**Table 1**Demographic characteristics of participants with and without adequate DLP recordings<sup>a</sup>, Stadium Place Study

| Characteristics                     | Total<br>N=37           | With Adequate Recording<br>N=24 | Without Adequate Recording<br>N=13 | p-value |
|-------------------------------------|-------------------------|---------------------------------|------------------------------------|---------|
| <b>Mean age, SD [range]</b>         | 75.9, 6.4 [64.5 – 91.3] | 75.5, 6.1 [64.5 – 91.3]         | 77.1, 6.0 [69.4 – 87.7]            | 0.43    |
| <b>Female, n (%)</b>                | 33 (89.2)               | 21 (87.5)                       | 12 (92.3)                          |         |
| <b>Race, n (%)</b>                  |                         |                                 |                                    |         |
| African American                    | 37 (100.0)              | 24 (100.0)                      | 13 (100.0)                         |         |
| <b># living in household, n (%)</b> |                         |                                 |                                    |         |
| 1                                   | 34 (91.9)               | 23 (95.8)                       | 11 (84.6)                          | 0.28    |
| 2                                   | 3 (8.1)                 | 1 (4.2)                         | 2 (15.4)                           |         |
| <b>Education</b>                    |                         |                                 |                                    |         |
| <12 <sup>th</sup> grade             | 14 (37.8)               | 8 (33.3)                        | 6 (46.2)                           | 0.82    |
| High School graduate                | 2 (5.4)                 | 2 (8.3)                         | ---                                |         |
| Some College                        | 16 (43.2)               | 11 (45.8)                       | 5 (38.5)                           |         |
| College                             | 5 (13.5)                | 3 (12.5)                        | 2 (15.4)                           |         |
| <b>Income</b>                       |                         |                                 |                                    |         |
| < \$25,000                          | 27 (73.0)               | 18 (75.0)                       | 9 (69.2)                           | 0.54    |
| \$25,000 & <\$49,999                | 3 (8.1)                 | 1 (4.2)                         | 2 (15.4)                           |         |
| Refused                             | 2 (5.4)                 | 2 (8.3)                         | ---                                |         |
| Don't know                          | 5 (13.5)                | 3 (12.5)                        | 2 (15.4)                           |         |
| <b>Employment status</b>            |                         |                                 |                                    |         |
| Working now                         | 1 (2.7)                 | 1 (4.2)                         | ---                                | 0.26    |
| Retired                             | 29 (78.4)               | 20 (83.3)                       | 9 (69.2)                           |         |
| Disabled                            | 5 (13.5)                | 3 (12.5)                        | 2 (15.4)                           |         |
| Keeping house                       | 2 (5.4)                 | ---                             | 2 (15.4)                           |         |

<sup>a</sup>DLP recordings starting before 10am and lasting for 10 hours

**Table 2**

Composite DLP report for participants with adequate DLP recordings

| <b>Composite report</b>          | <b>Adequate Recording<sup>a</sup><br/>N=24</b> |
|----------------------------------|--|
| Mean Adult Word Count, Projected | 33141 [5120–77882]                             |
| Mean Duration of Recording       | 13 hrs 13 min [10 hrs 6 min – 16 hrs 0 min]    |
| Meaningful Speech                | 2 hrs 33 min [23 min – 6 hrs 18 min]           |
| Distant Speech                   | 2 hrs 58 min [21 min – 6 hrs 47 min]           |
| Silence & Background             | 3 hrs 33 min [1 hr 10 min – 8 hrs 35 min]      |
| TV & Electronic Sounds           | 3 hrs 32 min [40 min – 7 hrs 24 min]           |
| Mean % TV time                   | 26.7 % [4%–58%]                                |

<sup>a</sup>Values indicate the mean and [range] of quantitative DLP output for participants with adequate recording