**Title**: Test-Retest Reliability of Automatic Language ENvironment Analysis (LENA) System Analyses with Young Children with Down Syndrome

**Authors**: Miriam Kornelis1, Marianne Elmquist2, Claudia Schabes2, Audra Sterling2, & Lizbeth Finestack1

**Introduction**: Researchers and clinicians are challenged to identify appropriate outcome measures to evaluate language growth of young children with Down Syndrome (DS). Traditional, standardized norm-referenced assessments compare a child’s performance to typically developing peers, thus the results are often invalid. Additionally, these tests often do not consider outcomes in the context of a child’s typical environment and rely heavily on caregiver reporting, which is subject to bias. Thus, it is important to identify feasible and reliable outcome measures derived from caregiver-child interactions to evaluate child language ability and growth.

An alternative home-based measure is language sample analysis with Language ENvironment Analysis System (LENA), which records and analyzes everything that the child hears and says within a 16-hour period. Automatic analyses report three measures of communicativeness: adult word count (AWC), child vocalization count (CVC), and conversational turn count (CTC). LENA is a promising and effective method for measuring early language (Wang et al., 2017). Prior investigations have reported the reliability of the automatic analyses with human counts of comparable measures (Canault et al., 2016; Pae et al., 2016). However, few investigations have reported the test-retest reliability of these measures. Gilkerson et al. (2017) investigated the test-retest reliability with 329 families with typically developing infants and toddlers between recordings 0-16 weeks apart, reporting Pearson correlation coefficients (*r*) of .60-.69. Nadwodny (2024) investigated the test-retest reliability of automatic LENA measures with 128 toddlers with social communication delays. Test-retest reliability were reported for recordings from a two-week period, with intraclass correlation (ICC) coefficients of .66-.75. Despite moderate test-retest reliability for young typically developing children and children with social communication delays, the test-retest reliability of these measures with young children with Down syndrome is unknown.

**Methods:** Fifteen caregivers of children with DS between the ages of 2- and 5-years-old and their child participated in this study to date. At two timepoints, four-to-six weeks apart, caregivers recorded two, 16-hour LENA recordings over the course of one week. Participants completed the four recordings independently and in the home environment. ICC coefficients were computed for daily AWC, CVC, and CTC values for both recordings completed within one week, and recordings completed within four- to six-weeks. Recruitment and data collection for additional participants are in progress and may be incorporated into the analysis prior to the symposium.

**Results**: Participants (*n* = 15) successfully completed 100% (*n* = 60) of assigned LENA recordings. LENA recordings were 14-hours and 5-minutes in length, on average.

*Descriptive Statistics*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Time 1** | | **Time 2** | |
| **Day 1**  *n =* 15 | **Day 2**  *n =* 15 | **Day 1**  *n =* 15 | **Day 2**  *n =* 15 |
| Daily AWC | 15,066 (5,575) | 14,540 (3,796) | 10,736 (5,242) | 12,371 (6,367) |
| Daily CVC | 2,550 (1,168) | 2,423 (690) | 1,944 (1,047) | 2,005 (989) |
| Daily CTC | 708 (301) | 670 (200) | 489 (330) | 533 (256) |

*Note.* Mean (SD)

Comparisons between recordings completed within one week yielded ICC coefficients of .56 (AWC), .75 (CVC), and .64 (CTC). Comparisons between recordings completed within four to six weeks yielded ICC coefficients of .49 (AWC), .74 (CVC), and .61 (CTC).

**Discussion:** With moderate to strong correlations across recording times, satisfactory reliability is demonstrated for automated LENA analyses. Language sampling with LENA takes place in the child’s home environment without examiners present. This increases the ecological validity of the measures and has the potential to increase the reach of study teams by reducing family burden. Results from our study have the potential to change how researchers and clinicians can collect caregiver-child interactions in a manner that is reliable and convenient for families.

**References:** Canault, M., Le Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENATM) in European French. Behavior Research Methods, 48(3), 1109–1124. https://doi.org/10.3758/s13428-015-0634-8

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the Early Language Environment Using All-Day Recordings and Automated Analysis. American Journal of Speech-Language Pathology, 26(2), 248–265. https://doi.org/10.1044/2016\_AJSLP-15-0169

Nadwodny, N. (2024). The Language Environmental Analysis (LENA) System in Toddlers with Early Indicators of Autism: Test-Retest Reliability and Construct Validity [Graduate Masters Theses]. University of Massachusetts Boston.

Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J. A., Ma, L., & Topping, K. (2016). Effects of feedback on parent–child language with infants and toddlers in Korea. First Language, 36(6), 549–569. https://doi.org/10.1177/0142723716649273

Wang, Y., Hartman, M., Aziz, N. A. A., Arora, S., Shi, L., & Tunison, E. (2017). A Systematic Review of the Use of LENA Technology. American Annals of the Deaf, 162(3), 295–311.

University of Minnesota – Twin Cities

2 University of Wisconsin, Madison