**Title**: **A Systematic Comparison of the Stability and Validity of Indices of Caregiver Talk from Caregiver Child Free Play Samples and Language ENvironment Analysis Software in Infant Siblings of Autistic and Non-autistic Children**

**Authors**: Insung Kim1, Jennifer E. Markfeld2, 3, Grace Pulliam2,5, Bahar Keceli-Kaysili1, Jacob I. Feldman1,3, Caroline Braun4, Ruoxi Guo4, S. Madison Clark1, Kristen Bottema-Beutel4, & Tiffany G. Woynaroski1,3,5-7

**Introduction**: Caregiver-child free play (CCFP) samples are commonly used to measure caregiver talk. However, it is unclear at present whether these commonly used, conventional methods for measuring caregiver talk yield scores that are stable, particularly (a) when children are in the earliest stages of development and (b) when child-caregiver transactions may not yet be consolidated. Stability is an important consideration, as it mathematically places an upper bound on the validity of the construct, and thus impacts our ability to observe relations of interest (e.g., predicting later language) in populations of interest, including infant siblings of autistic children (Sibs-autism) and infant siblings of non-autistic children (Sibs-NA). Recently developed Language ENvironment Analysis (LENA) technology now offers an alternative method for measuring caregiver talk. Scores on constructs derived via this novel automated method may be valid for predicting language in Sibs-autism and Sibs-NA (Swanson et al., 2019; Markfeld et al., 2023), though relatively little is known about the stability of caregiver talk constructs derived from LENA. This study aims to compare the relative stability and validity of scores indexing caregiver talk as derived via human-coded CCFP samples versus LENA for predicting later language across Sibs-autism and Sibs-NA.

**Method**: We recruited 30 Sibs-autism and 30 Sibs-NA between 12–18 months of age, matched at the group level on chronological age and sex assigned at birth, and their caregivers for this study. Each caregiver–child dyad completed (a) two 15-minute CCFP samples collected in the laboratory and coded by the research team using a previously developed coding scheme and (b) two day-long LENA recordings in children’s natural settings at study entry (i.e., Time 1). Expressive and receptive language of Sibs-autism and Sibs-NA were measured 9 months later, at 21–27 months of age (i.e., at Time 2), via the Mullen Scales of Early Learning; the MacArthur-Bates Communicative Development Inventories, Words and Sentences checklist; and the Vineland Adaptive Behavior Scales, second edition. Our indices of caregiver talk were (a) caregiver follow-in utterances (FIU) derived from CCFPs (i.e., comments or directives that relate to a child’s current focus) and (b) adult word count (AWC), the estimated number of adult words spoken in the presence of the infant during a daylong recording derived from the LENA prorated (i.e., divided by) the number of hours in the LENA recording. The stability of these scores indexing caregiver talk was evaluated in a series of Generalizability and Decision (G & D) studies. We set our a priori threshold for acceptable stability at *g* > .8. Prior to the Gatlinburg Conference, we will conduct a series of correlation analyses for each candidate index of caregiver talk as measured at Time 1 with aggregates of child expressive and receptive language as measured at Time 2, and Steiger’s *Z*-tests will be utilized to compare the relative predictive validity of scores derived via CCFPs and LENA recordings.

**Results**: When analyzing the stability of caregiver FIUs from the caregiver–child free plays across sibling groups, as well as within Sibs-autism and Sibs-NA, acceptable stability was reached (i.e., a *g* coefficient of at least .8) with two CCFP sessions (*g* for all participants = .86, *g* for Sibs-autism = .84, and *g* for Sibs-NA = .88). Across all participants, as well as within Sibs-autism and Sibs-NA, caregiver FIUs did not reach the threshold for acceptable stability with one CCFP session (*g =* .76, *g =* .73, and *g =* .78, respectively). When analyzing the stability of prorated AWC as derived via LENA from daylong recordings, only Sibs-autism reached our threshold for acceptable stability with two daylong recordings (*g =* .83). Within Sibs-NA and across sibling groups, this variable did not reach the threshold for acceptable stability with two daylong recordings (*g =* .72 and *g =* .78, respectively). The D-study for this variable indicated that four daylong audio recordings would be required to surpass acceptable stability in Sibs-NA (projected *g* = .84), and three daylong audio recordings would be required to achieve acceptable stability across all participants (projected *g* = .84).

**Discussion:** This study provides insights into the psychometrics of novel versus more conventional measures of caregiver talk. Our preliminary results suggest that indices of caregiver talk derived from CCFPs are more stable than indices from LENA across infants at high and general-population level likelihood for autism. Findings for predictive validity will be shared at the Gatlinburg Conference. Implications of this work for research and clinical practice will be discussed.

**References:**

Markfeld, J. E., Feldman, J. I., Daly, C., Santapuram, P., Bowman, S. M., Dunham-Carr, K., Suzman, E., Keçeli-Kaysılı, B., & Woynaroski, T. G. (2023). The stability and validity of automated indices of vocal development in infants with autistic and non-autistic siblings. *Journal of Speech, Language, and Hearing Research*, *66*(12), 4934–4948. <https://doi.org/10.1044/2023_JSLHR-23-00069>

Swanson, M. R., Donovan, K., Paterson, S., Wolff, J. J., Parish-Morris, J., Meera, S. S., Watson, L. R., Estes, A. M., Marrus, N., Elison, J. T., Shen, M. D., McNeilly, H. B., MacIntyre, L., Zwaigenbaum, L., St John, T., Botteron, K., Dager, S., Piven, J., & IBIS Network (2019). Early language exposure supports later language skills in infants with and without autism. *Autism Research*, *12*(12), 1784–1795. <https://doi.org/10.1002/aur.2163>

Woynaroski, T., Oller, D. K., Keçeli-Kaysılı, B., Xu, D., Richards, J. A., Gilkerson, J., Gray, S., & Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder. *Autism Research, 10*(3), 508–519. <https://doi.org/10.1002/aur.1667>

Vanderbilt University Medical Center

2 Vanderbilt University

3 Frist Center for Autism and Innovation in the Workplace, Vanderbilt University

4 Boston College

5 Vanderbilt Brain Institute

6 Vanderbilt Kennedy Center

7 University of Hawaii at Manoa