

The Role of Input Quantity vs. Language Background in Word Learning among Bilingual Korean Infants

Joo Kyeong Kim¹ & Eon-Suk Ko^{1,2}

¹Center for Data Science in Humanities (CDS), Chosun University

²Department of English Language and Literature, Chosun University

anitajookyeong@chosun.ac.kr; consukko@chosun.ac.kr

Many studies examining bilingual language input have focused on measuring language exposure by the proportion of one language relative to the other (e.g., Hoff et al., 2012). However, proportional measures may not accurately reflect the amount of linguistic input available for learning (De Houwer, 2014). For example, a child receiving 50% Korean exposure in a highly talkative household may hear more Korean words overall than a child receiving 80% Korean exposure in a less talkative household. The present study uses Language ENvironment Analysis (LENA) technology to move beyond proportion-based estimates of language exposure and directly quantify the amount of linguistic input available to young learners. Specifically, we ask: (i) Is absolute input quantity a significant predictor of early vocabulary development? (ii) When actual input quantity is taken into account, does bilingual status remain a meaningful predictor of lexical outcomes?

We collected daylong recordings of approximately 16 hours from children aged 8 to 20 months, comprising 97 monolinguals and 35 bilinguals (Mandarin: 19, Vietnamese: 16). This number reached the target sample size of 34, as suggested by a simulation-based power analysis. Total adult word counts (AWC) were extracted from these recordings using LENA ADEX (LENA Research Foundation, 2011). To estimate Korean-language input, total AWC values were weighted by caregivers' reported proportion of Korean use at home. Additionally, caregivers completed the K-M CDI (i.e., a Korean-adapted version of the MacArthur-Bates Communicative Development Inventory; Fenson et al., 1993; Pae & Kwak, 2011). Bilingual caregivers completed the Mandarin- or Vietnamese-translated version of the K-M CDI to obtain a conceptual vocabulary measure, calculated as the union of words acquired in Korean and Mandarin/Vietnamese. Raw CDI scores were converted into age-adjusted percentiles.

To evaluate whether input quantity predicts vocabulary development above and beyond bilingual status, we constructed two linear regression models: one using the Korean CDI percentile (formula: $\text{korean_cdi_percentile} \sim \text{scale(AWC_Korean)} + \text{biling_status} + \text{age_diff} + \text{infant_gender} + \text{age} + \text{cdi_test_type}$) and another using the conceptual vocabulary CDI percentile as the dependent variable (formula: $\text{conceptual_cdi_percentile} \sim \text{scale(AWC_Total)} + \text{biling_status} + \text{age_diff} + \text{infant_gender} + \text{age} + \text{cdi_test_type}$). The conceptual vocabulary analysis further allowed us to determine whether any observed group differences reflected language-specific vocabulary measures or broader lexical knowledge across languages.

Korean AWC significantly predicted CDI percentile ($\beta = 5.239, p = .028$), whereas bilingual status did not ($\beta = 0.278, p = .962$) (Figure 1a). A similar pattern emerged for the conceptual vocabulary measure, with a significant effect of Total AWC ($\beta = 5.055, p = .019$) but no reliable effect of bilingual status ($\beta = 7.302, p = .177$) (Figure 1b).

These findings suggest that absolute input quantity may provide a more informative account of early vocabulary development than bilingual status alone. Unlike previous research that reported a delay in bilingual children (Bialystok & Feng, 2009; Hoff et al., 2012), our data showed no evidence that the relationship between input quantity and vocabulary differed by bilingual status. These results underscore the importance of directly quantifying children's everyday linguistic experience. Although additional data collection is ongoing, the present findings raise the possibility that some differences previously attributed to bilingual status may instead reflect variation in the amount of language input available to the child.

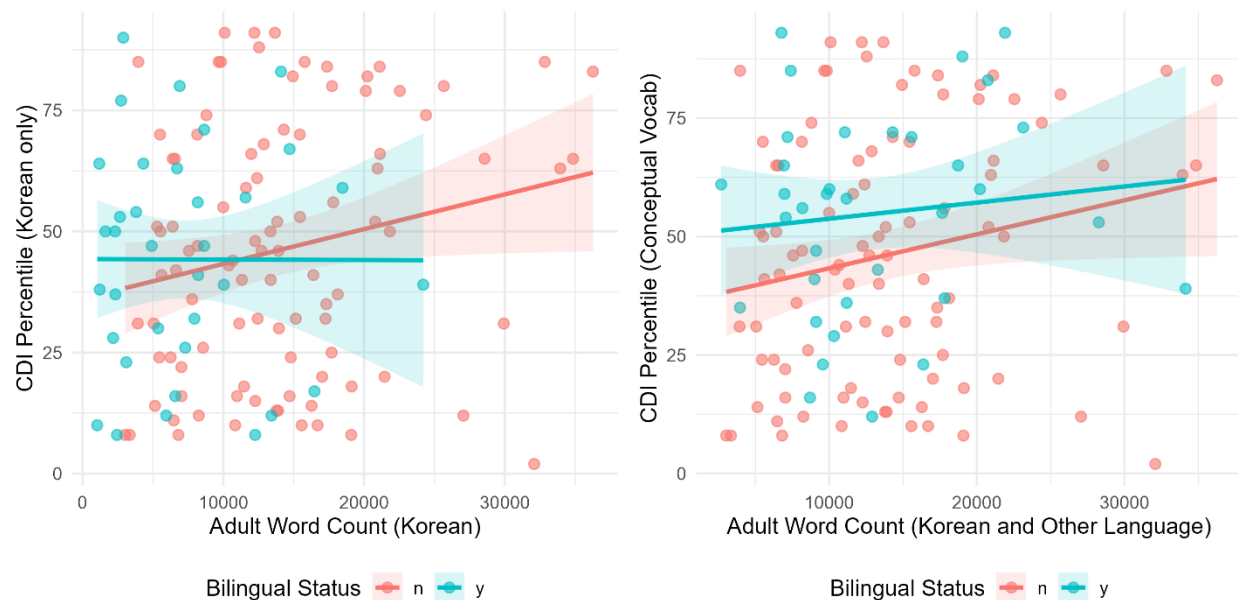


Figure 1. (a) Relationship between Korean AWC and Korean CDI percentile, and (b) relationship between total AWC and conceptual vocabulary CDI percentile.

References

- Bialystok, E. & Feng, X. (2009). Language proficiency and executive control in proactive interference: evidence from monolingual and bilingual children and adults. *Brain and language*, 109(2-3), 93–100. <https://doi.org/10.1016/j.bandl.2008.09.001>
- De Houwer, A. (2014). The absolute frequency of maternal input to bilingual and monolingual children. In T. Grüter & J. Paradis (Eds.), *Input and Experience in Bilingual Development* (pp. 37-58). John Benjamins Publishing Company.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. Singular Publishing Group, Inc.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(1), 1–27.
- LENA Research Foundation. (2011). *Advanced Data Extractor (ADEX) User Guide*.
- Pae, S. & Kwak, K. C. (2011). *Korean MacArthur-Bates Communicative Development Inventories (K M-B CDI)*. Seoul: Mindpress.