UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Looking beyond parental reports: systematic biases in early word recognition assessment

Permalink

https://escholarship.org/uc/item/7511s972

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Chai, Jun Ho
McDonald, Margarethe
Ko, Eon-Suk

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Looking beyond parental reports: systematic biases in early word recognition assessment

Jun Ho Chai[1], Margarethe McDonald[2], Eon-Suk Ko[1]

[1]Chosun University, [2] University of Kansas

Abstract

This study examines convergence between parental reports and behavioral measures in assessing early word knowledge of twenty-eight 14-month-old Korean infants. We compared infants' word recognition patterns with parental reports using full and shorter versions of the Korean MacArthur-Bates Communicative Development Inventories (MCDI-K). Our analyses revealed three key patterns. First, while parents showed consistent judgment between the full CDI and the target-word checklist, the checklist demonstrated better convergence with eye-tracking measures, which accounted for baseline looking biases. Second, parents' reporting accuracy varied systematically with item difficulty: for early-acquired words, parents showed higher agreement with eye-tracking than for later-acquired words. Third, exploratory analyses suggested a possible asymmetry in word category recognition, with infants showing stronger recognition of nouns than verbs in the eye-tracking task, contrasting with more balanced verb-noun knowledge in parental reports. These findings show that assessment methods capture different aspects of early word knowledge.

Keywords: word comprehension; infants assessment; looking-while-listening; parental reports; measurement convergence

Introduction

Early language assessment during infancy presents unique methodological challenges due to infants' limited ability to produce explicit responses. While the MacArthur-Bates Communicative Development Inventories (MB-CDI) are valuable for tracking early vocabulary, their validity and reliability across languages remain uncertain. This study examines the convergent validity between parental reports and behavioral measures of word comprehension in infants learning Korean using eye-tracking technology. We investigate how different assessment methods capture early word knowledge, with particular attention to category-specific patterns and age of acquisition of words. This study contributes to our understanding of early language assessment methodologies while addressing the need for validated measurement tools in non-English speaking populations.

Early language assessments

Early language development studies have relied on observing children's spontaneous speech, a direct but time-consuming method (Bornstein & Haynes, 1998). Assessing infants presents additional challenges due to their limited speech production capacity and constrained motor skills, which further complicates obtaining verbal or motor responses (Fernald et al., 2008). In response to these limitations, researchers have turned to parental reports as indirect assessment methods. The MacArthur-Bates Communicative Development Inventories (MB-CDI; Fenson, 2007) offer structured measures of vocabulary development and have been shown to predict later language outcomes (Can et al., 2013; Marchman & Fernald, 2008). Although CDIs have advanced cross-linguistic research (Bleses et al., 2008; Braginsky et al., 2019; Frank et al., 2021), their subjective nature and potential caregiver bias (Law & Roy, 2008; Tomasello & Mervis, 1994) highlight the need for objective, quantifiable measures (Fenson et al., 2000).

The looking-while-listening eye-tracking paradigm shows that infants understand common words in their first year (Bergelson & Swingley, 2012; Fernald et al., 2008; Frank, Lewis, & MacDonald, 2016; Parise & Csibra, 2012; Von Holzen & Bergmann, 2019) and predicts later language outcomes (Fernald & Marchman, 2012; Hurtado, Marchman, & Fernald, 2008; Law & Edwards, 2015). However, interpreting covert eye-tracking data in these word comprehension tasks presents challenges, as it involves inferring word knowledge from gaze patterns rather than direct responses (Bergelson & Swingley, 2015). While alternative methods like tablet-based tasks offer more direct measures (Lo et al., 2021), they often exceed young infants' motor capabilities, making eye-tracking particularly valuable for early language assessment.

Alignment and convergence of assessments

While overall correlations between parental reports and behavioral measures are moderate to strong, item-level agreement is mixed. Eye-tracking studies have revealed varying degrees of alignment between parental reports and children's comprehension, with some studies finding underestimation (Houston-Price et al., 2007), overestimation (Weaver & Saffran, 2022), or good alignment (Styles & Plunkett, 2009; Syrnyk & Meints, 2017). Such inconsistencies may reflect immature lexical representations and infants' difficulty distinguishing similar items, as well as cross-linguistic differences (Arias-Trejo & Plunkett,

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

2010; Lo et al., 2021). Thirdly, although differences across languages may be a factor, not enough studies examining parental report and infant behavioral convergence have been examined in languages other than English to systematically make comparisons. Moderate to strong correlations between parent report and child looking or reaction time have been reported for infants and toddlers in French-English (López Pérez et al., 2025) and Spanish-English bilingual environments (Marchman et al., 2010). Given these complexities, there is a growing need for objective, quantifiable measures that complement parental assessments.

Age of acquisition (AoA) norms provide crucial benchmarks for understanding these developmental patterns. These norms, indicating when words are typically acquired by children, have been documented through various methods including parental reports (Frank et al., 2021), adult ratings (Kuperman et al., 2012), and objective test-based measures (Dale & O'Rourke, 1981). While each method has its strengths, parental report data has been particularly valuable in establishing cross-linguistic patterns of early vocabulary development (Braginsky et al., 2019) and documenting the consistency of semantic category acquisition across languages (Frank et al., 2021). However, the relationship between AoA norms and real-time measures of word comprehension is complex. While parents show good implicit knowledge of typical acquisition patterns (DeAnda et al., 2016), their reporting accuracy varies with factors such as word frequency (Kartushina & Mayor, 2019) and semantic category (Houston-Price et al., 2007). These complex relationships highlight the need for careful examination of how different assessment methods converge in measuring early word knowledge.

Early word learning trajectory

Early word learning follows predictable category-specific trajectories, with nouns for concrete objects typically learned before verbs and abstract terms, which can affect both recognition and parental reporting. Research has consistently shown that animate objects are particularly salient for infant attention and processing (Rakison & Poulin-Dubois, 2001). This animacy effect has been documented in both parental reports and behavioral measures, though with varying degrees of consistency. In eye-tracking studies, Bergelson and Swingley (2012) found that 6-9 month-olds showed stronger recognition for body parts compared to other semantic categories, while DeLoache et al. (2010) demonstrated that infants' looking patterns were particularly sensitive to animate features in visual stimuli. However, parental report studies using the CDI have shown mixed results regarding animate-inanimate distinctions, with some finding earlier acquisition of animate concepts (Frank et al., 2021) and others showing more complex patterns influenced by factors such as input frequency (Kartushina & Mayor, 2019) and contextual salience.

Nouns are typically acquired before verbs, with this noun bias especially pronounced at 13–14 months, as verb learning involves mapping words to relational concepts and events rather than concrete objects, placing greater cognitive demands on developing lexical representations (Gentner, 1982; Waxman et al., 2013). While the classic "noun bias" has been documented across many languages (Frank et al., 2021; Gentner, 1982; Waxman et al., 2013), its universality has been challenged by cross-linguistic evidence. In verb-friendly languages like Korean and Mandarin, where verbs gain prominence through frequent pronoun-dropping and either sentence-final position (Korean) or standalone usage (Mandarin), parental reports show different patterns of noun and verb acquisition across development (Choi & Gopnik, 1995; Frank et al., 2021).

These developmental patterns raise important methodological considerations for early language assessment. While parental reports like the MB-CDI include both nouns and verbs, as well as words for animate and inanimate objects, the accuracy of reporting may vary systematically across these categories. Parents might be more attuned to their child's comprehension of certain word types based on the frequency and clarity of responses in daily interactions. For example, concrete nouns referring to visible objects might elicit more obvious recognition behaviors than verbs or relational terms (Frank et al., 2021). In Korean specifically, these reporting patterns may be influenced by cultural and linguistic factors, such as the prevalence of verbs in child-directed speech (Choi, 2000) and the grammatical properties that make certain categories more or less prominent in everyday interactions.

Language-specific patterns in word learning

The Korean adaptation of the MB-CDI (MCDI-K; Pae & Kwak, 2011) has played a crucial role in documenting early language development patterns in Korean-learning children. Cross-linguistic comparisons using CDI data have revealed that Korean-learning children show distinct vocabulary composition patterns compared to English learners. For instance, Frank et al. (2021) documented that Korean-learning children's vocabulary composition varies across development, with distinct patterns in early comprehension (verb-bias) versus later production (noun-bias). Korean's verb-final word order and frequent argument omission may influence word learning patterns, potentially facilitating earlier verb acquisition compared to English, where verbs often appear in less salient positions (Choi & Gopnik, 1995; Kim et al., 2000). However, how these patterns manifest in behavioral measures like eye-tracking, and whether they align with parental reports, remains an open empirical question that requires further investigation.

Study objectives

This study examines the alignment between parental reports and behavioral measures of early word comprehension in Korean-learning infants. Using items

from the Korean MB-CDI (MCDI-K; Pae & Kwak, 2011), we investigate three key aspects of measurement convergence. First, we assess the correlation between different formats of parental report and infant measures. Specifically, we compared the full CDI and a shorter checklist that targets the words used in the eye-tracking task and examined how each aligns with different ways of calculating the infants' looking patterns in a word recognition paradigm. Second, we evaluate how this alignment varies according to item difficulty by examining age of acquisition norms, to determine whether parents are more accurate in reporting early-acquired words compared to later-acquired ones. Third, as an exploratory analysis, we examine whether the alignment between parental reports and behavioral measures differs across semantic categories, with a particular focus on noun versus verb and animate versus inanimate distinctions. Through this multi-faceted analysis, we aim to identify both the strengths and limitations of parental report and gaze-based measures, and contribute to the growing body of research on early language assessment in non-English-speaking populations.

Methods

Participants

This study involved 28 typically developing Korean children (12 female) around 13 months old (M = 13.2, SD = 0.41) in an infant-directed assessment of word comprehension using an eye-tracking test. The sample size is adequate according to a power analysis for the planned correlation test between infants' word comprehension measured by eye-tracking and parental report. With an aimed power level of 0.8, a significance level of 0.05 (two-tailed), and an expected correlation coefficient of 0.55 based on previous studies (Fernald et al., 2006; Houston-Price et al., 2007; Styles & Plunkett, 2009), the calculated sample size required for this correlation test was N = 24 participants.

Administration of parental report

To assess parental reports of word comprehension, we sent parents two survey links before their lab visit. The two forms administered were: (1) a full version of the MCDI-K language assessment tool (full-CDI), and (2) a 40-item checklist (target-word checklist) containing only the target words used in the eye-tracking task. We measured parent-reported word comprehension using both the full-CDI and the target-word checklist, focusing on parents' reports of target words in each version.

Procedure

We used the Looking-While-Listening paradigm (Fernald et al., 2008) with a Tobii Pro X3-120 eye-tracker. Each 9-second trial presented a yoked pair of static images and the target word (in isolation and two carrier sentences; 3000 ms silence, 4000 ms audio, 1000 ms silence). For example, with the target word koyangi 'cat': "Koyangi! Koyangi boyeo? Koyangi ceki iss-ne!" 'cat! cat see-Q cat there exist-EXCL' ("Cat! Can you see the cat? There is the cat!"). To mitigate potential boredom from repetitive presentations, the first carrier sentence was alternated with "Koyangi chaca polk'a?" 'cat find try-Q' ("Shall we find the cat?"). Notably, target words always appeared in the initial position of the carrier sentences, reflecting the typical Object-Verb word order in Korean and the frequent elision of subjects and objects. When the target word was a verb, it was embedded in a naturalistic sentence frame that omitted the explicit object, consistent with Korean discourse patterns that allow argument omission. For example, with the target word an-ass-ne 'hug': An-ass-ne! Nuga an-ass-ji? Ce-ki an-ko iss-ne! 'hug-EXCL! who hug-Q? there hug-PROG exist-EXCL!' ("Hugged! Who hugged? Someone is hugging over there!"). This approach maintained natural and consistent sentence structure across trials. Each trial concluded with a 1000 ms silence interval.

Visual Stimuli

The visual stimuli consisted of static color images presented side-by-side on a monitor. We selected 40 words (30 nouns, 10 verbs) understood by at least 40% of 13–14-month-olds, based on WordBank norms (Frank et al., 2017). Images depicted familiar objects and actions, with pairs drawn from different semantic categories to minimize interference. Although many actions are dynamic, static images were used to ensure consistency across noun and verb trials and to allow precise timing of gaze measurements.

Gaze features extraction

Task performance was measured by focusing on the proportion of looking time to the target during the 300-2000 ms window after audio onset (see Figure 1).

Trials with <50% valid gaze data were excluded and gaze to target was analyzed during the 300–2000 ms window. Trials were excluded from analysis if gaze direction could not be reliably coded (due to missing data, blinks, or looking away), or if less than 50% of the gaze data for a trial were valid.

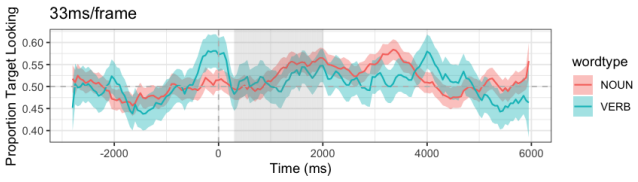


Figure 1: Average proportion of looks to the target during baseline (pre-audio: before 0 ms) and test trials (after 0 ms), aggregated by time point. The dotted line at 0 ms represents the onset of the target audio stimulus. The shaded region (between 300 and 2000 ms) indicates the commonly used analysis window for assessing target word recognition.

The current study employed three methods to measure the proportion of looking time: method 1) calculated the

proportion of looking time to the target relative to the distractor during the 300-2000 ms window after audio onset. Scores above .5 were coded as 1 (indicating word comprehension), and below .5 as 0; method 2: used a yoked-pair approach (Bergelson & Swingley, 2012), comparing looking times when an image served as the target versus the distractor.

Results

Overall convergence

The image-adjusted method showed the strongest correlation with the target-word checklist (r = .66; Table 2). Other eye-tracking measures also correlated moderately with the checklist (r = .54–.66), but not with the full-CDI. Binarizing gaze data increased correlations.

target-word checklist (r = -.57, CI = [-.75, -.32], t(38) = -4.322, p < .001).

To explore how agreement between parental reports and eye-tracking measures varied with item difficulty, we compared performance across early- and late-acquired words. For easy items (AoA < 13 months), parents and eye-tracking measures showed strong agreement when children demonstrated word knowledge (90% agreement, 27/30 cases), but poor agreement when children showed no recognition (21.1% agreement, 4/19 cases). This asymmetry suggests a tendency for parents to over-report knowledge of early-acquired words, with parents claiming knowledge in 78.9% of cases (15/19) where eye-tracking showed no recognition.

For more difficult items (AoA > 13 months), agreement between measures was more balanced. Agreement remained

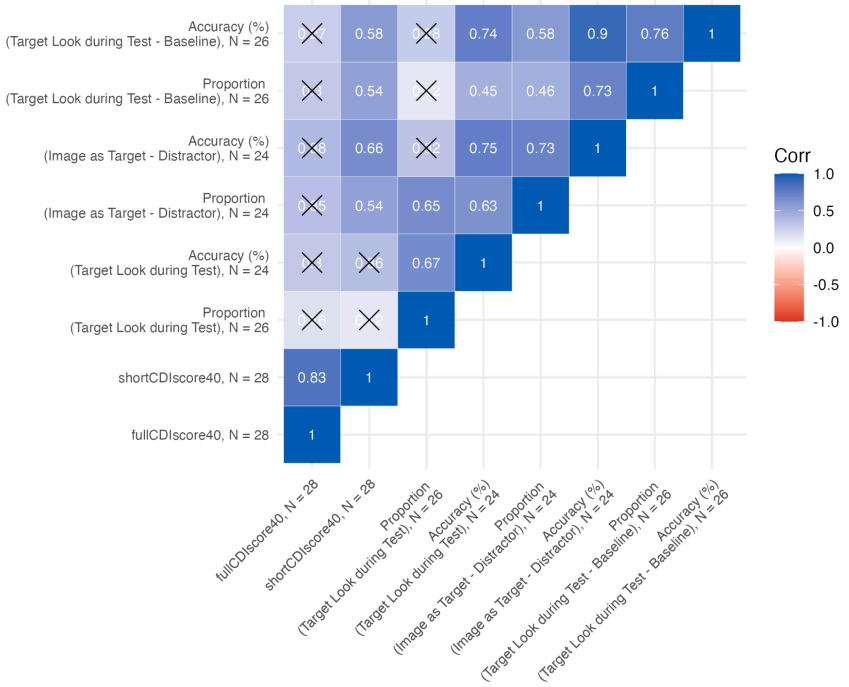


Figure 2: Correlation coefficient between word comprehension based on parent-reported and task-based measures.

Convergence based on item difficulty

The age of acquisition (AoA) for these words, defined as the point at which 50% of infants understood them, ranged from 11 to 16 months. Analysis of parental reports revealed significant correlations between the method of parent report and AoA. Parents' reports of known target words showed a strong positive correlation between the full-CDI and target-word checklist formats (r = .70, CI = [.50, .83], t(38) = 6.109, p < .001). Both CDI versions showed significant negative correlations with AoA, indicating parents were more likely to report knowledge of earlier-acquired words. This relationship was stronger for the full-CDI (r = -.69, CI = [-.83, -.49], t(38) = -5.918, p < .001) than for the

high when children showed word recognition (80% agreement, 8/10 cases) and improved substantially when children showed no recognition (52.4% agreement, 11/21 cases). While parents were more conservative in reporting knowledge of later-acquired words, some over-reporting persisted, with parents reported comprehension in 47.6% of cases where eye-tracking showed no recognition.

Word recognition performance based on semantic categories

We calculated normalized verb (relative to total words) and animate item (relative to total nouns) proportions, with

Table 1: Descriptive statistics of word recognition and short target-word checklist by animacy and word type.

	Child word recognition			Parental target-word checklist		
Measure	Mean (SD)	Range	N	Mean (SD)	Range	N
Word type						
Verb-noun proportion	0.271 (0.205)	0 - 0.667	20	0.413 (0.254)	0.111 - 1	20
Verb count	1.545 (1.335)	0 - 4	22	2.227 (1.412)	0 - 5	22
Animacy						
Animate-inanimate proportion	0.418 (0.248)	0 - 1	20	0.421 (0.162)	0 - 0.6	20
Animate count	1.909 (1.342)	0 - 5	22	3.227 (2.525)	0 - 8	22
Total count	7.136 (5.566)	0 - 19	22	5.318 (3.51)	0 - 14	22

Note. The measures are not mutually exclusive, that is, nouns are the sum of animate and a portion of inanimate pairs that are not verbs.

0.5 indicating equal category representation. Raw counts were also recorded.

One-sample tests comparing verb-to-noun proportion (Table 1) to 0.5 showed that only the verb-to-noun proportion in word recognition was significantly below 0.5 (β = -0.229, p < .001), indicating relatively stronger recognition of nouns compared to verbs. However, parents reported relatively more balanced verb-to-noun knowledge (β = -0.087, p = .083). Infants show a noun recognition bias in looking patterns that parents do not reflect in checklist reports.

In contrast, the animate-inanimate proportions – both in recognition (β = -0.082, p = .102), and CDI reports (β = -0.079, p = .115) – did not significantly differ from 0.5, indicating similar representation across these categories, as well as alignment between parental reports and children's recognition patterns for animate versus inanimate items.

Discussions

Parental reports align most strongly with eye-tracking when using the image-adjusted method. This suggests that specific methodological adjustments can improve the reliability of behavioral assessments. Parents often over-reported early-acquired words, but were more accurate for later-acquired words, possibly influenced by their expectations of typical development trajectories. Noun-verb discrepancies indicate that parental reports may miss finer distinctions, underscoring the value of complementary behavioral measures.

Our analysis revealed that assessment format and item difficulty systematically affect alignment between parental reports and behavioral measures. Parents showed consistent judgment across longer and shorter checklist formats but were notably more conservative in the target-word checklist, which showed better convergence with eye-tracking measures. Parents' strong alignment with normative acquisition patterns indicates good implicit understanding of typical vocabulary trajectories, though this knowledge might paradoxically contribute to over-reporting of early-acquired words, when comparing the CDIs. Even with the more highly correlated target-word checklist format, parents showed asymmetric reporting patterns for early-acquired words: strong agreement when eye-tracking confirmed word knowledge, but substantial over-reporting when eye-tracking showed no recognition. This asymmetry was less pronounced for later-acquired words, where parents showed more balanced reporting patterns. While Houston-Price et al. (2007) found parental underestimation and we found overestimation, especially for early-acquired words, both studies point to assessment format and item difficulty as key drivers of reporting accuracy.

Developmental patterns in early comprehension

Our findings revealed distinct patterns across word types and assessment methods that warrant careful theoretical consideration. While Korean is traditionally characterized as a verb-friendly language (Choi & Gopnik, 1995) with mothers typically emphasizing action-oriented language in child-directed speech (Choi, 2000), our eye-tracking data showed stronger recognition for nouns compared to verbs, contrasting with more balanced verb-noun knowledge in parental reports. This discrepancy with previous findings of early verb advantage in Korean (Frank et al., 2021) likely reflects the interaction between linguistic input and cognitive development at 13-14 months. At this age, the cognitive demands of processing dynamic actions in real-time recognition tasks may temporarily override the advantages of verb-friendly input patterns (Gentner, 1982; Golinkoff et al., 2008), particularly given our use of static images to represent action concepts. This pattern suggests that while language-specific input characteristics are important, their influence on early word recognition may be modulated by factors such as word imageability and processing demands (Kartushina & Mayor, 2019).

Interestingly, while our baseline looking patterns showed systematic preferences for animate items, we found no significant animacy effects in word recognition measures. This dissociation between visual preferences and word recognition suggests that early attentional biases toward animate objects do not necessarily facilitate word learning at this age. The lack of animacy effects in recognition, contrary to previous findings (Rakison & Poulin-Dubois, 2001), might reflect the development of more balanced processing abilities by 13-14 months, though animacy

advantages might be more pronounced at earlier ages (Bergelson & Swingley, 2012) or in different task contexts.

Parental reports and eye-tracking capture complementary aspects of early word knowledge, but parental reports may miss real-time processing differences. Additionally, the lack of significant animacy effects in word recognition, despite baseline visual preferences, indicates that attentional biases do not directly translate to word learning. These findings emphasize the value of integrating multiple assessment methods to obtain a more comprehensive understanding of early language development.

Methodological considerations

Different adjustment methods for looking-to-target metrics in eye-tracking studies reflect distinct assumptions about infant attention and visual biases. The traditional baseline adjustment method compares looking times relative to a pre-stimulus period, assuming trial-specific visual biases. In contrast, the image-adjustment method assumes persistent visual biases across trials by using yoked pairs, where the same images serve as both target and distractor across different trials. These methodological choices significantly affect how we interpret infant word comprehension data, particularly in accounting for systematic looking preferences.

Our analysis revealed that the image-adjusted method provided better convergence with parental reports, likely due to its effective control of inherent looking biases. We observed systematic preferences for animate items and certain household objects in baseline patterns, which the image-adjusted method controlled for through the yoked-pair analysis. This improved control of visual biases may explain the method's stronger alignment with parental reports, as it better isolated word recognition from general visual preferences. However, the image-adjusted approach has limitations in providing only pair-level rather than item-level resolution, as items were consistently paired in fixed combinations. While more dynamic item pairing could offer finer resolution, such an approach would necessarily increase task duration and potentially affect participant engagement. Future studies should consider incorporating a broader range of item pairings while maintaining reasonable task duration, perhaps through careful manipulation of semantic relationships between items.

Another limitation concerns the use of static images to represent verbs. While this approach ensured consistency across trials and precise timing of gaze measurement, it may have disadvantaged recognition of action words, which are often better conveyed through dynamic visual stimuli. Representing actions such as eating or clapping in a single static frame may have reduced the salience or interpretability of the verb referent, especially for younger infants. Future studies could explore the use of video stimuli or dynamic animations to better capture the meaning of action concepts and assess whether this improves alignment with parental reports.

In addition to the visual format, the composition of our stimulus set also warrants consideration. We selected 30 nouns and 10 verbs, proportionally reflecting their distribution in the Korean CDI, where nouns greatly outnumber verbs. While this approach enhances ecological validity, it introduces an asymmetry that may limit direct comparisons between word classes. Future research could adopt a more balanced design with equal numbers of nouns and verbs to more directly assess category-specific differences in early word comprehension, especially in relation to task demands and stimulus characteristics.

More broadly, future research could enhance measurement convergence by providing detailed guidance to parents about behavioral indicators of word knowledge, particularly for early-acquired words where over-reporting is common. Online assessment platforms like CDI-CAT (Kachergis et al., 2022), e-Babylab (Lo et al., 2021), and Web-CDI (deMayo et al., 2021) offer promising avenues to integrate environmental measures and nuanced reporting guidelines efficiently. Additionally, large-scale repositories such as Peekbank (Zettersten et al., 2022) will advance understanding of early word recognition through cross-linguistic comparisons and meta-analyses of looking patterns across diverse populations.

Conclusions

This study highlights how assessment format and item difficulty influence the alignment between parental reports and eye-tracking measures of early word knowledge in Korean-learning infants. The parent reported target-word checklist showed stronger convergence with behavioral data than the full CDI, while discrepancies across word types and visual preferences reveal that different methods capture distinct facets of language development. These findings underscore the need for integrated assessments that address cognitive and language-specific factors.

Acknowledgements

We express our gratitude to the participating families and research assistants Hyunji Kim, Jihyo Kim, Hyeonah Jeong, and Suhan Kim, who administered the experiments. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5A2A01073515).

References

Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word–referent identification. *Journal of Experimental Child Psychology, 105*(1–2), 63–80.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253–3258.

Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary

development in Danish and other languages: A CDI-based comparison. *Journal of Child Language, 35*(3), 619–650.

Bornstein, M. H., & Haynes, O. M. (1998). Vocabulary competence in early childhood: Measurement, latent construct, and predictive validity. *Child Development, 69*(3), 654–671. https://doi.org/10.1111/j.1467-8624.1998.tb06235.x

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind, 3*, 52–67.

Can, D. D., Ginsburg-Block, M., Golinkoff, R. M., & Hirsh-Pasek, K. (2013). A long-term predictive validity study: Can the CDI Short Form be used to predict language and early literacy skills four years later? *Journal of Child Language, 40*(4), 821–835.

Choi, S. (2000). Caregiver input in English and Korean: Use of nouns and verbs in book-reading and toy-play contexts. *Journal of Child Language, 27*(1), 69–96.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language, 22*(3), 497–529.

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary: A national vocabulary inventory.* World Book-Childcraft International.

DeAnda, S., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Lexical processing and organization in bilingual first language acquisition: Guiding future research. *Psychological Bulletin, 142*(6), 655–667.

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., … & O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science, 21*(11), 1570–1574.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Brookes.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*(1), 203–222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). John Benjamins.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language, 44*(3), 677–694.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project.* MIT Press.

Frank, M. C., Lewis, M., & MacDonald, K. (2016). A performance model for early word learning. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2261–2266). Cognitive Science Society.

Gentner, D. (1982). *Why nouns are learned before verbs: Linguistic relativity versus natural partitioning* (BBN Report No. 4854). Bolt Beranek and Newman Inc.

Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language, 34*(4), 701–724.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish‐learning children. *Developmental Science, 11*(6), F31–F39.

Kartushina, N., & Mayor, J. (2019). Word knowledge in six‐to nine‐month‐old Norwegian infants? Not without additional frequency cues. *Royal Society Open Science, 6*(9), 180711.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English- and Korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language, 27*(2), 225–254.

Kuperman, V., Stadthagen-González, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990.

Law, F., & Edwards, J. R. (2015). Effects of vocabulary size on online lexical processing by preschoolers. *Language Learning and Development, 11*(4), 331–355.

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the communicative development inventories. *Child and Adolescent Mental Health, 13*(4), 198–206.

Lo, C. H., Rosslund, A., Chai, J. H., Mayor, J., & Kartushina, N. (2021). Tablet assessment of word comprehension reveals coarse word representations in 18–20-month-old toddlers. *Infancy, 26*(4), 596–616.

López Pérez, M., Moore, C., Sander‐Montant, A., & Byers‐Heinlein, K. (2025). Infants' Knowledge of Individual Words: Investigating Links Between Parent Report and Looking Time. *Infancy*, *30*(1), e12641.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science, 11*(3), F9–F16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Marchman, V. A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish–English bilinguals. *Journal of child language*, *37*(4), 817-840.

Pae, S., & Kwak, K. (2011). *Korean MacArthur-Bates Communicative Development Inventories User's Guide and Technical Manual.* Mindpress Publishing Co.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by

9-month-old infants. *Psychological Science, 23*(7), 728–733.

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate–inanimate distinction. *Psychological Bulletin, 127*(2), 209–228.

Styles, S., & Plunkett, K. (2009). What is 'word understanding' for the parent of a one-year-old? Matching the difficulty of a lexical comprehension task to parental CDI report. *Journal of Child Language, 36*(4), 895–908.

Syrnyk, C., & Meints, K. (2017). Bye-bye mummy–word comprehension in 9-month-old infants. *British Journal of Developmental Psychology, 35*(2), 202–217.

Tomasello, M., & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in Child Development, 59*(5), 174–179.

Von Holzen, K., & Bergmann, C. (2021). The development of infants' responses to mispronunciations: A meta-analysis. *Developmental Psychology, 57*(1), 1–18.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. J. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives, 7*(3), 155–159.

Weaver, H., & Saffran, J. (2022). Eye-gaze as a window into word level variability in toddlers. Presentation at the Boston University Conference on Language Development 47, Boston, MA.

Zettersten, M., Yurovsky, D., Xu, T. L., Uner, S., Tsui, A. S. M., Schneider, R. M., … & Frank, M. C. (2022). Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition. *Behavior Research Methods*, 1–16.