

Infants' gesture-speech combinations elicit more verbal and multimodal responses from caregivers during early play

Anika van der Klis¹, Frans Adriaans¹, René Kager¹

¹Institute for Language Sciences, Utrecht University

a.vanderklis@uu.nl, f.w.adriaans@uu.nl, r.w.j.kager@uu.nl

Before they learn how to speak, infants start to communicate by producing vocalisations and gestures. These early behaviours, including gesture-speech combinations, have been found to relate to children's lexical development (e.g., Brooks & Meltzoff, 2008; Rowe et al., 2022; Wu & Gros-Louis, 2014). Caregivers individually differ in their verbal responsiveness to these behaviours which has been found to correlate with children's lexical development (e.g., Donnellan et al., 2019; McGillion et al., 2013; Wu & Gros-Louis, 2014). Caregivers' verbal responses could mediate the relationship between infants' behaviours and their vocabulary outcomes (Olson & Masur, 2015). Previous studies have mostly focused on verbal responses provided by caregivers. However, there is ample evidence that nonverbal behaviours – such as handing over a toy, pointing, or smiling – play a facilitative role in lexical development (Pearson et al., 2011; Ruddy & Bornstein, 1982). Children appear to rely on visual information when speech is novel (e.g., a label for an unfamiliar object) or unclear (e.g., a label that could refer to multiple objects) (e.g., Baldwin et al., 1996; Grassmann & Tomasello, 2010; Chen et al., 2021). Until recently, we did not know to what extent caregivers produce multimodal responses and whether specific infant behaviours tend to elicit multimodal responses. In a recent study, we analysed infants' vocalisations, gestures, and gesture-speech combinations and caregivers' verbal, nonverbal, and multimodal responses (van der Klis et al., 2023). In this abstract, we focus primarily on infants' gesture-speech combinations and their role in eliciting caregiver responses.

The data are derived from YOUth, an ongoing longitudinal cohort study in the Netherlands. We included 117 infants (66 females) around 9-11 months of age ($M = 10.5$ months, $SD = 0.9$) and their caregivers (92 mothers). All dyads participated in a caregiver-child interaction task which included six minutes of free play with a standard set of toys. The sessions were filmed with four cameras surrounding the playing rug and annotated in ELAN. We developed a coding scheme to first annotate infants' vocalisations and gestures (including pointing, showing, giving, reaching, and requesting) and then caregivers' verbal, gestural (e.g., pointing and passing), facial responses (e.g., smile and surprise), and other bodily behaviours (e.g., turning and leaning closer), within two seconds after the offset of the infant behaviour. For verbal responses, we annotated whether it was semantically contingent (i.e., on the infant's focus of attention), onomatopoeia or sound effect, infant imitation, or non-contingent. When a verbal response was combined with at least one nonverbal behaviour, this was a multimodal response. After several steps of training and improving the coding scheme, we achieved high inter-rater reliability ($\kappa = .81$).

In total, we annotated 2036 infant behaviours (1892 vocalisations; 207 gestures) of which only 63 were bimodal. At least one bimodal behaviour was produced by a quarter of the infants in our study. Infants' bimodal pointing gestures were surprisingly often coordinated with more advanced Consonant-Vowel (CV) vocalisations ("dada") (58.3%), while all other bimodal gestures were typically coordinated with non-CV vocalisations. In the full data set only 2.8% of all vocalisations were CV. Of all infant behaviours, 87.1% were responded to by caregivers. Caregivers mostly responded verbally ($n = 1566$), but 39.7% ($n = 704$) were multimodal. Caregivers' gestural responses were most often multimodal (80.9%), especially pointing gestures (98.4%). We found that multimodal gestures are often coordinated with semantically contingent responses. Then, we examined whether caregiver responses differed as a function of infant behaviour. The elicitation rates of all infant behaviours are depicted in Figure 1. We fitted mixed-effects logistic regression models predicting a caregiver response with infant behaviour (vocalisation, gesture, or gesture-speech combination) as the predictor and a random intercept for subjects. Infants' gesture-speech combinations are 6.42 times (95% CI [2.39, 17.23]) more likely to receive a verbal response compared to infant vocalisations ($p < .001$). We also found that infants' gesture-speech combinations are 3.48 (95% CI [2.00, 6.06]) times more likely to elicit a multimodal response compared to infant vocalisations ($p < .001$). The results overall show that infants' gesture-speech combinations elicited higher verbal and multimodal response rates from caregivers compared to unimodal behaviours.

Infants show large variability in the frequency and types of behaviours they produce during early interactions, which in turn affects the frequency and types of responses they receive from their caregivers. Although infants around 9-11 months of age do not yet produce many gesture-speech combinations, the results of our study suggest that they could be an important developmental milestone which shapes infants' early learning environments. Infants who produce gesture-speech combinations are likely more advanced in their language development, supported by the fact that many bimodal gestures are coordinated with more advanced CV vocalisations. Infants who produce gesture-speech combinations elicit more verbal and multimodal responses from their caregivers. The latter usually consists of verbal responses coordinated with manual or deictic gestures. These types of responses could be useful for infants to disambiguate novel or unclear speech – subsequently facilitating word learning. Future studies should investigate whether caregivers' verbal and multimodal responses mediate the relationship between infants' gesture-speech combinations and their vocabulary outcomes. For intervention studies, it would then not only be useful to increase caregivers' overall responsiveness, but also to assess whether specific behaviours reinforce infants' gesture-speech combinations during early interactions. When examining caregiver-child interactions, analysing caregivers' verbal responses alone undermines the multimodal richness and the bidirectionality of early communication.

Index Terms: caregiver-child interactions, responsiveness, multimodal language, free play

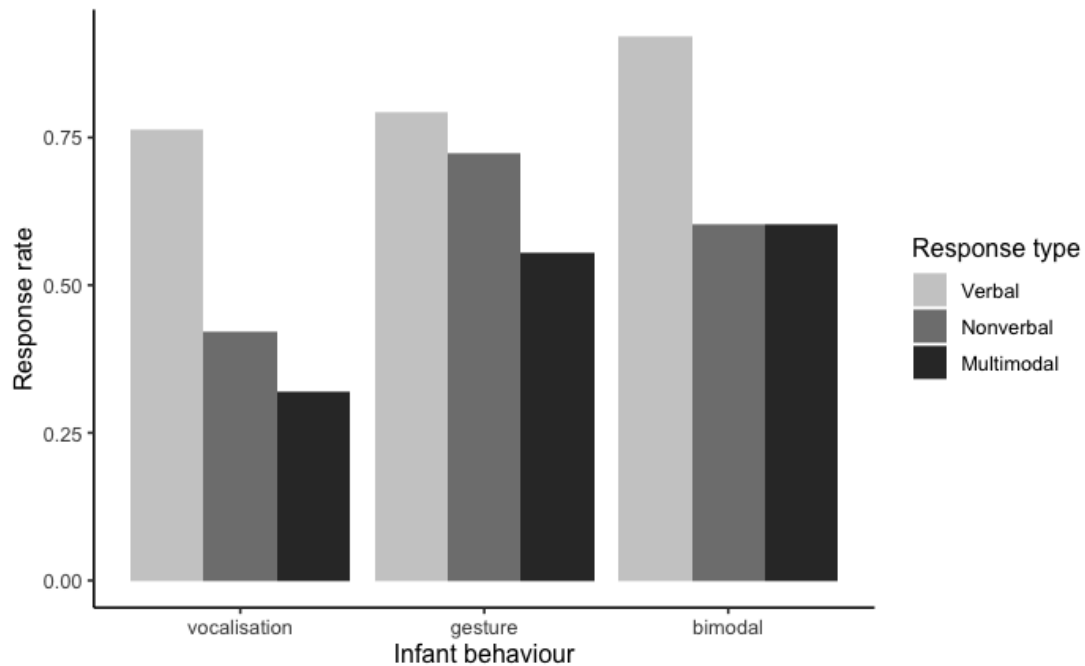


Figure 1: Caregivers' verbal, nonverbal, and multimodal response rates to infants' vocalisations, gestures, and bimodal (gestural-speech) combinations.

References

- Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development*, 67(6), 3135–3153. <https://doi.org/10.2307/1131771>
- Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, 35(1), 207–220. <https://doi.org/10.1017/S030500090700829x>
- Chen, C., Houston, D. M., & Yu, C. (2021). Parent–Child Joint Behaviors in Novel Object Play Create High-Quality Data for Word Learning. *Child Development*, 92(5), 1889–1905. <https://doi.org/10.1111/cdev.13620>
- Donnellan, E., Bannard, C., McGillion, M. L., Slocombe, K. E., & Matthews, D. (2019). Infants' intentionally communicative vocalizations elicit responses from caregivers and are the best predictors of the transition to language: A longitudinal investigation of infants' vocalizations, gestures and word production. *Developmental Science*, 23(1), 1–21. <https://doi.org/10.1111/desc.12843>
- Grassmann, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, 13(1), 252–263. <https://doi.org/10.1111/j.1467-7687.2009.00871.x>
- McGillion, M. L., Herbert, J. S., Pine, J. M., Keren-Portnoy, T., Vihman, M. M., & Matthews, D. E. (2013). Supporting early vocabulary development: What sort of responsiveness matters? *IEEE Transactions on Autonomous Mental Development*, 5(3), 240–248. <https://doi.org/10.1109/TAMD.2013.2275949>
- Olson, J., & Masur, E. F. (2015). Mothers' labeling responses to infants' gestures predict vocabulary outcomes. *Journal of Child Language*, 42(6), 1289–1311. <https://doi.org/10.1017/S0305000914000828>
- Pearson, R. M., Heron, J., Melotti, R., Joinson, C., Stein, A., Ramchandani, P. G., & Evans, J. (2011). The association between observed non-verbal maternal responses at 12 months and later infant development at 18 months and IQ at 4 years: A longitudinal study. *Infant Behavior & Development*, 34(4), 525–533. <https://doi.org/10.1016/j.infbeh.2011.07.003>
- Rowe, M. L., Wei, R., & Salo, V. C. (2022). Early gesture predicts later language development. In *Gesture in language: Development across the lifespan* (pp. 93–111). American Psychological Association. <https://doi.org/10.1037/0000269-004>
- Ruddy, M. G., & Bornstein, M. H. (1982). Cognitive correlates of infant attention and maternal stimulation over the first year of life. *Child Development*, 53(1), 183–188. <https://doi.org/10.2307/1129651>
- van der Klis, A., Adriaans, F., & Kager, R. (2023). Infants' behaviours elicit different verbal, nonverbal, and multimodal responses from caregivers during early play. *Infant Behavior and Development*, 71, 101828. <https://doi.org/10.1016/j.infbeh.2023.101828>
- Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, 34(1), 72–90. <https://doi.org/10.1177/0142723714521925>