Dynamic shift in isolating referents: From social to self-generated input

Hanako Yoshida (yoshida@uh.edu) and Joseph M. Burling (jmburling@uh.edu)
Department of Psychology, University of Houston
126 Heyne Bldg., Houston, TX 77204-5022 USA

Abstract—Infants as young as 6 months of age start comprehending some familiar words, yet there is little understanding of how young infants utilize information presented in their social environment in order to make sense of the world. As an initial step, we used recent technology that allowed us to narrow in on the point of view of the infant to explore how infants' visual input is dynamically synchronized with their own participation, as well as from social input in the context of parent-child word-learning play. The results reveal dynamic changes in the relation between infants' view and their level of participation.

I. EMBODIED VISUAL EXPERIENCES

A recent study has demonstrated that infants comprehend at least some words at 6 months of age—that is, before the traditional “first word” milestone in productive language [1]. What is the nature of this very early stage of word learning? During the first 2 years of life, infants learn and refine a whole set of new motor skills that dramatically change the ways in which the body moves and interacts with the environment, and their social interactions consequently also change. Six-month-old infants are relative novices at reaching for objects and do not sit steadily, and so for the most part, objects are brought to them or perceived from a distance; 12-month-olds, in contrast, walk and bring themselves to objects, and 18-month-olds are mobile, socially skilled, and capable of physically achieving their own desires. Do these changes in the ways infants physically interact in the world also change the way they socially interact, and determine the nature of their visual experiences that support word learning?

Recent studies that motivated the present study investigated how early learners create their own visual input, by studying the first-person view during toy play [2]–[4]. These studies typically used a small camera attached to young children’s foreheads and documented how this child-centered view provides insight into factors relevant for early word learning. In the study presented here, we sought to investigate moment-to-moment visual experiences and the developmental changes in this child-centered view via a head-mounted eye tracking device (Figure 1) by longitudinally following children from 6 to 18 months (Figure 2).

If infants shape the effective view of objects through their own physical growth, consequential actions, and mature social partners’ participation, the degree of referential ambiguity or perceptual accessibility may be partially addressed at the level of sensory motor coordination. In a recent study of 18-month-old infants in a naturalistic parentchild play context, both parent and infant physically participated in the play (holding objects) at the same level, but objects dominated the child’s view (much bigger, thus occluding other objects) when the object was held by the child [5], [6]. Infants of this age (18 months) are capable of producing smooth head turnings and can manage a stable posture, suggesting that their own body coordination may help optimize their view focus, and bodily movements having a relatively stronger role for the optimal view at this age. This quality of visual experiences was also related to learning object names [6] suggesting that how well an 18-month-old can zoom in on objects has important implications for their successful word learning. What happens if younger infants, who seem capable of learning words [1], do not yet have the physical capacity to coordinate their bodies and support their optimal view? We specifically targeted younger infants to further our understanding of the role of physical development in the organization of early visual input.

Fig. 1: An infant wearing a head-mounted eye-tracking device.

In the present study, we investigated two components that together influence an infant’s visual focus: (a) who (parent/social partner or the child him or herself) brought the object to the child’s view, and (b) how parents’ participation reflects their child’s physical constraints. Studying how visual and bodily experiences support language learning reveals what contributes to effective visual attention, such as joint attention, and how such effective attention may emerge through both the child’s own experiences and parental participation.

Fig. 2: Snapshot of third-person views of the same child across five sessions.
II. **Longitudinal Parent-infant Interactions**

At every visit, seven parent-infant (6 months to 18 months with 3 month intervals) were asked to sit on the designated seats where the room view (the third-person view) and child center view (first-person, head-mounted, eye-tracking) were recorded during their play session (6 min) in which parents were instructed to play and teach about toy objects (e.g., bunny, car, cup, etc.)

A. **Results**

The results suggest that a child’s view changes dramatically over the course of a year of development. Objects within an infant’s view during the first session take up a moderately sized proportion of the child-centered view, yet at 12 months (their 3rd visit), the average size of objects focused within their view is reduced dramatically, and reaches its lowest point during this age. At 6 months, the child’s view was moderately selective. The focus of the object is then reduced in size at 12 months, then the proportion of objects in terms of pixel area was greatest at 18 months of age (5th and final visit, see Figure 3a). These changes may be the development of optimal focus, and coders’ judgments about who was interacting with the object from the room view suggest that the infant’s frequency of reaching and holding of objects has a systematic influence on the infant’s own view (Figure 3b).

In particular, as early as 6 months, parents held the object of focus reliably more than infants ($\chi^2 = 6.25; p < .05$). This contrasts with later periods, in which children are interacting with objects much more frequently than adults. At 15 months, infants began to reliably hold the objects more often than adults ($\chi^2 = 4.9; p < .05$), and by 18 months, the magnitude of this difference in object interactions was most dramatic ($\chi^2 = 7.79; p < .01$). There were no significantly reliable differences found in object manipulation during the 9 month and 12 month time periods.

During the early stages of development, parental involvement seems to account for a majority of the visual experiences gained by infants, while at later stages, infants are responsible for their own experiences as they hold and manipulate objects with greater frequency. At the 12 month time point, there was no reliable difference in who was holding the object, which is the point where their center view quality drops to as small as 12.1%. This may be the time at which the parent begins to demonstrate less involvement in shaping the child’s visual experiences—and thus a drop in object focus—which allows for the infant to independently start shaping what comes into their view and determine their own optimal focus, an ability that they apparently are not immediately successful in controlling during early months.

III. **Discussion**

The present findings suggest that the contexts for early and later word learning are quite different, and that they shift from greater parental control to greater infant control, with parents ensuring early that the named object is visually dominant and the infant playing a more active role later in development. Clearly these findings are just a first step toward understanding the possible, developmentally changing pathways through which infants learn words. These findings point to the possibility of how parents may contribute to the isolating of objects (e.g., focusing on objects one at a time, zooming in on a single object) to help this type of learning. Later, this isolation may be most effectively done when the child actively engages in object manipulation (holding and bringing objects) as suggested by the Yu and Smith results [6]. Studying changes in visual experiences mediated through a child’s own physical growth and social experiences is a way to gain new perspectives on the nature of embodied language learning.

**Acknowledgment**

This research was supported by a National Institutes of Health grant (R01 HD058620), the Foundation for Child Development, and University of Houston’s Grants to Enhance and Advance Research (GEAR) program. We especially wish to thank the families who participated in this study. We thank the undergraduate research students in the Cognitive Development lab for their support in coding for the present study.

**References**


