Augmenting Sign Language Exposure for Hearing Parent-Deaf Child Interaction

Ashely Tenesaca¹, Jung Yun Oh², Sizhe Li¹, Wanyin Hu¹, Zhen Bai¹
¹University of Rochester, ²Rice University

Keywords—Assistive Technology; American Sign Language; Early Language Development; Augmented Reality.

Above 90% of deaf children are born to hearing parents who are often non-fluent with sign language. The developmental consequences of early language deprivation experienced by deaf children can be witnessed in their later life outcomes. A significant portion of the deaf population are prone to cognitive delays, poor educational experiences, diminished career opportunities, and psychosocial difficulties (Hall 2017, 961). Given this, it becomes imperative to improve early childhood language experiences. To foster neuropsychological growth, language must be accessible to the child; in the case of deaf children, this would mean early exposure to a visual language.

There are two main challenges for hearing parents to offer linguistic-rich interaction with their DHH children. First, the learning curve of ASL makes it difficult to accommodate the infant’s critical language acquisition window. Second, the mismatch in communication modalities between the hearing parent and deaf child causes insufficient communication strategies for the child’s linguistic intake. Adjusting to utilizing a single visual modality to deliver language is a new situation for hearing parents who are accustomed to using spoken communication. In contrast, deaf parents, who tend to be fluent signers, intuitively adapt communicative behaviors that support the child’s visual attention (Spencer and Koester 2016, 183). With these in mind, we developed the following research questions: (1) How to develop assistive technologies that support hearing parents to carry out just-in-time and situation-aware ASL to their DHH child? (2) How to develop assistive technologies that support hearing parents to adapt to sufficient joint attention strategies for linguistic intake for their DHH child? (3) How to develop assistive technologies that support non-intrusive face-to-face interaction?

I. SYSTEM PROPOSAL

As an initial step to address the above research questions, we propose a proof-of-concept prototype using projection-based augmentation technology that aims to enhance joint free play episodes between hearing parents and their DHH child. The main goal of the prototype is to support real-time, context-aware, and non-intrusive ASL-based interaction, with adaptation to joint-attention strategies that match with the child's communication modality. It does so through the projection of on-demand ASL vocabulary videos next to the child’s object of visual attention. The ASL labeling will promote the parent’s sign language learning, while also serving as a visual feedback cue on child’s eye gaze. Our target infant population is between 6 to 18 months of age. We chose this as our focus age group since it’s during this time period that infants are learning social communicative behaviors that aid in language development (Carpenter et al. 1998, 176).

Figure 1 - Projection of words apple, fork and blue.

II. PROTOTYPING

The prototype consists of a portable projector, a gooseneck clamp holder, an HDMI cable, and a laptop computer hosting a user interface with the projection contents (see Figure 1). The web-based user interface allows the researcher to select the appropriate video to play that matches with the object that the DHH child is attending to. For automating this context understanding process, we are working on an object detection system, which is taken from the state-of-the-art model architecture named FasterRCNN (Ren et al. 2015). We’re using the publicly available pre-trained model provided by TensorFlow objection detection API. The model contains two stages, a region proposal network that generates bounding boxes and a box regressor for classifying each box. Instead of feeding the region proposals to the convolutional neural network (CNN) as previous attempts (Fast-RCNN), this approach feeds
input images directly into a CNN to produce corresponding feature maps, allowing faster inference speeds. In addition to the object recognition module, the gaze detection features are currently under development.

III. REFERENCES


