# Embodied cognition in prelingually deaf children with cochlear implants: Preliminary findings

**Chapter** · January 2018



4 authors:

Irina Castellanos
The Ohio State University

David B Pisoni
Indiana University Bloomington

Chen Yu
Indiana University Bloomington

Derek M Houston
The Ohio State University

# Embodied Cognition in Prelingually Deaf Children with Cochlear Implants: Preliminary Findings

Irina Castellanos
Department of Otolaryngology – Head & Neck Surgery
The Ohio State University

David B. Pisoni
Department of Psychological and Brain Sciences
Indiana University

Chen Yu
Department of Psychological and Brain Sciences
Indiana University

Chi–hsin Chen
Department of Otolaryngology – Head & Neck Surgery
The Ohio State University

Derek M. Houston
Department of Otolaryngology – Head & Neck Surgery
The Ohio State University

Correspondence to:
Irina Castellanos, Ph.D.
Department of Otolaryngology – Head & Neck Surgery
The Ohio State University
915 Olentangy River Rd
Columbus, Ohio 43212
USA
e–mail: Irina.Castellanos@osumc.edu

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (T32 DC00012), the National Institute of Child Health and Human Development (R01 HD074601), and the Indiana University Collaborative Research Grant.

<1>Abstract

The theory of embodiment postulates that cognition emerges from multisensory interactions of an agent with its environment and as a result of multiple overlapping and time–locked sensory–motor activities. In this chapter, we discuss the complex multisensory system that may underlie young children's novel word learning, how embodied attention may provide new insights into language learning after prelingual hearing loss, and how embodied attention may underlie learning in the classroom. We present new behavioral data demonstrating the coordination of sensory–motor behaviors in groups of young children with prelingual hearing loss (deaf, early–implanted children with cochlear implants and hard–of–hearing children with hearing aids) and without hearing loss (two control groups of chronological–aged and hearing–aged matched peers). Our preliminary findings suggest that individual differences and variability in language outcomes may be traced to children's coordination of auditory, visual, and motor behaviors with a social partner.

**Keywords:** Cochlear Implants, Embodied Cognition, Multisensory Processing, Eye Tracking, Mother–Infant Dyads, Novel Word Learning

# Embodied Cognition in Prelingually Deaf Children with Cochlear Implants: Preliminary Findings

<1>Background

Cochlear implants (CIs) provide access to sound for many deaf children with severe–to–profound sensorineural hearing loss. CIs represent a significant engineering and medical milestone in the treatment of sensorineural hearing loss. Unfortunately, while many CI candidates display substantial benefits in recognizing speech and processing spoken language following implantation, a significant number of children have poor outcomes and often display less than optimal speech and language skills following implantation even after several years of experience with their CIs. The estimates of poor outcomes following cochlear implantation range between 25 to 30 percent, depending on what behavioral criteria are used to assess benefit and outcomes. Most CI users get some benefit from their implant in quiet listening conditions such as in the audiology clinic or research laboratory, although they commonly report significant difficulties in listening to speech in the presence of background noise, especially when multi–talkers are present, or listening to speech under conditions of high cognitive and attentional load.

The problem in understanding and explaining the underlying basis of poor speech–language outcomes, including language acquisition, following cochlear implantation is a very challenging research issue that has not received sufficient attention in the literature despite the pressing clinical significance. Why do some CI users do extremely well in the quiet and why do others struggle and very often fail to reach optimal levels of speech recognition performance even under ideal listening conditions in the clinic or laboratory? This question represents a significant gap in our current knowledge concerning outcomes following cochlear implantation and is an important barrier to progress in developing novel and personalized interventions to help low–functioning CI users improve their speech–language outcomes. Part of the difficulty with speech recognition stems from the nature of the signal CI users receive through their implant which is spectrally–degraded and significantly compromised relative to the original signal presented at the ear. The critical acoustic–phonetic cues in the signal that support speech recognition are coarsely–coded and the fine acoustic–phonetic and indexical details of the original speech are significantly reduced or often absent from the neural encoding presented to the auditory nerve and higher information processing centers. While some of the minimal acoustic–phonetic cues needed for speech recognition are preserved in CIs in coarsely–coded form, the critical speech cues are generally underspecified by the signal processing algorithms currently available for clinical use.

Several researchers are looking beyond the endpoint or product–based outcome measures traditionally used to assess performance following implantation and are focusing their efforts on identifying and explaining sources of individual variability and the underlying processes that may lead to them (e.g., Markman et al., 2011; Moeller, 2007). Similarly, we propose that a broader, developmental systems approach should be adopted using process–based measures that allow for the discovery of both macrostructural and microstructural developmental change. Developmental change is not an isolated process, but instead arises from interactions across coordinated sensory and cognitive systems. For example, the variability observed in conventional speech–language outcome measures not only reflects the early sensory registration and encoding of acoustic signals by the auditory nerve, but it also reflects the important and central contribution of the information processing system as a whole. The information processing system involves cognitive processing factors and contributions such as selective and sustained attention, sustained sequential processing, and inhibitory control, which are actively used by listeners to support encoding, processing, storage and retrieval of information (see Kronenberger & Pisoni, this volume). Speech scientists and acoustical engineers have known for more than 60 years that speech recognition and spoken language processing do not take place at the auditory periphery in the ear. Robust spoken word recognition and speech understanding reflects the final product of a long series of stages of information processing that routinely draw on multiple resources and the interactions of different sources of knowledge in long–term memory that are based on the listener's prior experiences and unique developmental histories.

In this chapter, we adopt a developmental systems approach to the study of novel word learning in prelingually deaf and hard–of –hearing children (D/HH). We summarize recent findings on a feasibility study using novel eye tracking process–based measures to examine how D/HH children with cochlear implants and hearing aids (HAs) coordinate sensory–motor behaviors during naturalistic joint play with a social partner and we contrast these findings with data collected from two control groups of normal hearing (NH) children. Lastly, we suggest methods for how knowledge about embodied attention may translate into novel interventions to support children's learning in the classroom.

## <1>Embodied Cognition and Joint Attention

One of the overarching questions in the field of cognitive development in children concerns how selective attention is organized during early development to facilitate the mapping of a word to a referent (novel word learning). During early development, word learning results from complex parent–child interactions involving joint attention. Parent–child joint attention occurs when the parent and child coordinate visual attention on the same object or event at the same time (Moore & Dunham, 1995). Embodied cognition provides an approach for describing and examining multiple pathways to the coordination of joint visual attention. Children may utilize their own sensory–motor skills, parental linguistic input and social cues to coordinate joint attention in the service of word learning. Consequently, novel word learning may be viewed as grounded and emerging from multisensory interactions between the child and his/her social–cultural environment and as a result of multiple overlapping and time–locked sensory–motor activities. Yu & Smith (2013) have described how visual and motor movements are spatially and temporally coupled during goal–directed actions, thereby redundantly specifying information about the agent's locus of attention. In fact, young NH children make use of information garnered from watching their parent's motor movements to engage in joint attention with their parent.

Engaging in joint visual attention is not accidental; instead, it relies on a foundation of multisensory functioning (coordinating visual, linguistic, and motor cues) with the goal of sharing a social experience/interest. The child may respond to episodes of joint attention by following a parent's gaze shifts to, touching/holding, or labeling of objects in the environment or the child may initiate episodes of joint attention with a caregiver or peer. Delays and/or disturbances in children responding to or initiating joint attention with a social partner may have cascading effects on neurocognitive development, particularly language learning.

Skilled parent–child coordination is a product of a complex system, with multiple degrees of freedom, relying on multiple solutions to the in–moment tasks of coordinating attention and behavior. For example, 18–month–old toddlers with normal hearing use their hand actions to select visual objects during mother–child play interactions, and parents' notice and use toddlers' actions on objects as behavioral cues to label objects for toddlers (Yu & Smith, 2012). This time–locked sequential pattern, from child manual handling to parent labeling, suggests an interpersonal coordination that jointly solves the referential uncertainty problem in early word learning — finding correct word referent mappings among many co–occurring words and referents.

Tomasello and colleagues' seminal work investigated macrostructural and microstructural changes in mother–child dyadic interactions during naturalistic play (Tomasello & Farrar, 1986; Tomasello, Mannle, & Kruger, 1986; Tomasello & Todd, 1983). The dyad was video recorded playing with novel toys supplied by the experimenters and no specific instructions were provided about how to engage with the toys (classic joint attention task). At the macrostructural level, they quantified the amount of mother–child joint attention episodes that occurred during play. At the microstructural level, they described the verbal and nonverbal behaviors of the mother–child dyad within episodes of joint attention. Several interesting findings were uncovered indicating that measures of joint attention may predict children's language knowledge.

For example, at the macrostructural level, Spencer, Meadow–Orlans, Koester, & Ludwig (2004) examined how the quality of mother–child interactions at age 12 months affect deaf and NH children's language learning at age 18 months. The data revealed associations between the amount of time the parent–child dyad engaged in joint attention and later language learning for both the deaf and hearing children. Similarly, Tomasello & Todd (1983) found the amount of time NH parent–child dyads jointly attended to novel toys was related to children's vocabulary size six months later.

Tomasello & Todd (1983) suggest that children's novel word learning may in part depend on parents' attunement or skills at determining and following the child's attentional focus. Parents may facilitate episodes of joint attention by providing a label for an object already in the child's attentional focus or may regulate episodes of joint attention by redirecting the child's attentional focus to labeled objects. Observational studies suggest that object labels provided by parents while following their NH child's attentional focus facilitated joint attention and larger vocabulary sizes (Tomasello & Farrar, 1986). Experimental studies that systematically manipulated how labels for objects are presented suggest that NH children learn novel words for objects more easily when the labels are presented in an attempt to follow the child’s attentional focus instead of when the labels are presented in an attempt to redirect the child’s attentional focus from one object to another (Dunham, Dunham, & Curwin, 1993; Tomasello & Farrar, 1986).

Several factors may influence the ease of coordinating joint attention between parent and child, one of which may be a shared sensory history. Mother–child dyads who share hearing status (deaf mother with deaf child; NH mother with NH child) spend more time in joint attention than mother–child dyads that do not share hearing status (deaf mother with NH child; NH mother with deaf child; Spencer, Swisher, & Waxman, 2004).

At the microstructural level, Tomasello et al. (1986) found that parents’ use of verbal and nonverbal actions to direct NH children’s visual attention and behavior was negatively correlated with shorter episodes of mother–child joint attention and less knowledge of object–word pairings, suggesting that highly directive parental styles are not conducive to early language learning in NH children. Research also indicates that NH mothers of deaf children engage in higher levels of directive parental styles, which result in lesser gains in language growth (Musselman & Churchill, 1992). Highly directive parental styles may also influence the kind of language children use. Parents who employ highly directive styles often have NH children with predominately expressive, rather than referential, language (Tomasello & Todd, 1983). Moreover, case studies suggest that NH children with predominately referential language hear more maternal talk about objects, have more labels for objects, and initiate episodes of joint attentional focus more often than children with predominately expressive language (Goldfield, 1986).

Cejas, Barker, Quittner, & Niparko (2014) have further differentiated between joint attention with and without accompanying parental symbols. The authors investigated parent–child joint attention in a group of prelingually deaf CI candidates (tested prior to implantation) and NH peers aged 0.75 – 5.09 years. All participating parents were hearing and children were divided into three age groups: under 18 months, 18–36 months, and over 36 months. The parent–child dyads were asked to play with age–appropriate toys and their behaviors were coded for engagement. Parental symbols were categorized as verbal (spoken language) or nonverbal (sign language). Child use of symbolic play was coded, as well as, their looking behaviors (looking toward parent, object, or both). If children were actively attending to symbols (e.g., verbal or nonverbal parental gestures or engaged in symbolic play) while maintaining joint attention with a parent they were coded as engaged in "symbol–infused" joint attention. This differs from parent–child joint attention without accompanying parental symbols.

The authors found no differences in joint attention versus "symbol–infused" joint attention between deaf and NH children aged under 18 months. In the two older age groups, however, CI candidates spent less time engaged in "symbol–infused" joint attention with their hearing parent when compared against hearing children with their hearing parents. Additionally, CI candidates spent more time engaged in joint attention (without any accompanying parental symbols), than their NH peers. These results are similar to those reported in an earlier study by Prezbindowski, Adamson, & Lederberg (1998), indicating that deaf children engaged in more joint attention, but less "symbol–infused" joint attention with their NH parents when compared to NH parent–NH child dyads.

Taken together, these studies suggest that joint attention scaffolds children's attention and promotes the acquisition of words for objects. Episodes of joint attention are influenced by a number of endogenous and exogenous factors such as matched sensory history (hearing or deafness), parenting style (redirecting attentional focus versus following attentional focus), and linguistic input (directives). Dyads consisting of a NH mother and a deaf child are at a disadvantage when engaging in joint attention because they are unable to rely on linguistic input to direct or maintain episodes of joint attention. Several interesting questions follow from these findings, for example: Is it possible that NH parents are so attuned to their child's profound–to–severe hearing loss that they fully rely on visual attention before the child's cochlear implantation/hearing aid amplification? Does dyadic engagement in more advanced "symbol–infused" joint attention episodes increase following cochlear implantation? How are visual, auditory, and motor behaviors spatially and temporally coupled during episodes of parent–child joint attention when the child has a hearing loss? To study these questions and many others related to macrostructural and microstructural change, in the following section, we introduce a new multisensory eye tracking methodology that provides multiple streams of high–density frame–by–frame recordings, allowing us to identify the potential multiple pathways through which children (with and without hearing loss) align their attention with a social partner. Multiple pathways involving visual, auditory, and motor skills are hypothesized to regulate the coordination of joint attention skills that are critically important for the development of novel word learning.

<1>Multisensory Language Learning

Recently we began a collaborative research project involving the Departments of Otolaryngology–Head & Neck Surgery, Psychological and Brain Sciences, and Informatics and Computing to employ a multisensory experimental methodology to investigate parents and D/HH children's reciprocal roles in language acquisition and cognitive development. Specifically, we sought to investigate how D/HH children's multisensory (auditory, visual, and motor) functions lead to learning of novel words for toy objects during joint play activities with their parents. The ongoing goal of this research program is to achieve a deeper and more detailed understanding of the sensory–motor basis of early social coordination and its potentially critical role in later language learning and other development milestones.

The premise for the experiment is quite simple: the parent–child dyad is presented with novel toy objects and asked to engage in play. Up to this point, our experimental methodology is very similar to the traditional studies examining novel word learning described earlier in this chapter. We depart from previous studies examining deaf children's word learning with our use of sophisticated computer vision technology. The use of our state–of–the–art sensing and computing technology allows us to collect process–based measures (instead of only endpoint measures of word learning) of real–time microstructural change in how the mother–child dyad arrives at joint attention and, if word–referent learning occurs, we can pinpoint with frame–by–frame precision the visual, linguistic, or sensory–motor cues that facilitated learning.

In our experiment, the child and parent wear small head–mounted cameras and eye trackers throughout the joint play session so that precise time–locked measures of what they are looking at and what they are touching can be obtained. Before the development of these novel research methods, it was not possible to measure the fine temporal dynamics of parent–child interactions with this level of detail and precision. These new research methods allow us to collect fine–grained multimodal behavioral data from young children with time–synchronized eye gaze, action, and speech data.

The experimental setup requires that the testing room (walls, floors, and play furniture) be outfitted in white. The dyads are provided with white clothing (long sleeve shirts, pants, and socks), as well. The all–white experimental room is necessary so that the novel play toys (painted in primary colors: red, green, and blue) can be easily detected and extracted from the background by our computer vision algorithm. Employing computer vision technology, a three–step segmentation algorithm was designed to automatically identify the toy object and output its precise size and location in space (for detailed information, see Yu, Smith, & Pereira, 2008). With this information, we can make statements about how much space an object occupies in a child's visual field (e.g., object "Mobit" occupies 80% of the space in the child's visual field and is visually dominant over the other objects), and test specific hypotheses about how visual and motor behaviors are coordinated.

—– Insert Figure 1 About Here —–

Once dressed in white clothing, the dyad is led into the all–white experimental room which contains a white play table. On opposite sides of the play table, a white chair is available for the child and a white floor cushion is available for the parent. The child's custom chair measures 32 cm above the floor, which allows the child to be approximately eye–to–eye with their parent when the parent is seated on the floor. The child wears a cotton cap with the lightweight head–mounted camera and eye tracker securely attached (see Figure 1, Panel A), while the parents' equipment is mounted onto frameless glasses. The play session is also audio–recorded using microphones integrated into the parent and child's eye tracking equipment. The majority of the weight of the eye tracking equipment is placed on the back of the chair (for the child) or on the floor (for the parent). We have tested this eye tracking equipment on children before and after cochlear implantation, children with HAs, and with Bone Anchored Hearing Aids (BAHAs), and have not had any device issues or electrical interference. Additionally, a bird's eye view camera is positioned above the play session and two scene cameras are positioned on opposite corners of the room. These recording devices allow us to obtain multiple streams of high–density frame–by–frame gaze, action, and speech data throughout the free–flowing interactive play session.

At the beginning of the testing session, an experimenter calibrates the position of the head–mounted camera and the angle of the eye tracker relative to the child's head. A similar calibration procedure is performed for the parent head–mounted camera and eye tracker. A secondary experimenter controls the recording of all cameras and recommends adjustments to the positioning of cameras when necessary. After the calibration phase, the parent–child dyad is free to engage to play.

The dyad is presented with two sets of toys that are similar in size, with each set containing three novel objects paired with three novel labels. Each set is presented twice for a total of four 90–second trials. Each novel toy is a complex nonsense object, and many have functional parts (e.g., wheels). Parents are asked to play with their children as they normally would at home. However, if they are going to refer to the toy objects by name, we ask the parents to refer to toy objects by their novel labels. The parent is not tasked with remembering the object–word pairings; in fact, a cheat–sheet is attached to their side of the table for easy reference (see Figure 1, Panel B). Object labels follow a consonant–vowel–consonant–vowel structure, which can be easily said by the parent, and can be easily heard by the child.

Immediately following the parent–child play session, children's novel word learning is assessed using the intermodal preferential looking paradigm (IPLP). In the IPLP, two objects are presented side–by–side, but only one of the objects corresponds to an accompanying linguistic label (referred hereafter as the target object; Golinkoff, Hirsh–Pasek, Cauley & Gordon, 1987). The IPLP contains three training trials and twelve test trials. During each training trial, two familiar objects are presented, a target and a nontarget, side–by–side and the child is asked, *"Where is the (target object)? Can you find the (target object)?"* Familiar objects (bear, car, banana, cow, horse, fork) were selected from highly familiar words that appear on the MacArthur–Bates CDI Words and Gestures list (Fenson et al., 2007). If the child incorrectly identifies the incorrect (nontarget) object, the experimenter corrects the response by redirecting the child towards to target object. Test trials commence if the child demonstrates competence with the task by correctly identifying all three target objects during the training trials. During test trials, the experimenter presents the play session objects two at a time, and again asks the child, *"Where's the (target object)? Can you find the (target object)?"* After each test trial the experimenter praises the child for selecting a toy, but no corrections are provided. Each novel toy object serves twice as the target (six novel toy objects presented across twelve test trials). The order in which the toy objects are presented and the lateral position of the target object is counterbalanced.

Following data collection, the parent and child's gaze (e.g., where and what the parent/child is looking at, how much space does an object occupy in the parent/child's visual field) and motor (e.g., what toy is the parent/child holding in their left and right hand) data are coded frame–by–frame, coupled together with the speech transcription, and analyzed for patterns of dyadic synchronization. Audio recordings from the parent and child during the play session are first transcribed using Systematic Analysis of Language Transcripts (SALT) transcription conventions. These transcriptions are then used to identify the number of topics discussed per utterance. A topic refers to a label that is correctly provided for an object. Adjacent utterances containing pronouns referring to the labeled object are considered a topic expansion. The following is an example of two utterances on one topic: Mother: *"Is that a Zeebee? What will you do with it?"* And an example of two utterances with two topics: Mother: *"Is it a Dodi? Is that the Zeebee?"* Finally, transcriptions from the play session are used to classify utterances as declarative, open–ended questions, directives, or other.

<*2*>Feasibility Study

A feasibility study was conducted to examine if D/HH children with CIs/HAs display differences in visual skills (gaze duration, gaze shifting), motor skills (duration and frequency of object holding), and linguistic input (numbers of utterances and topics, utterance forms) during joint play interactions with their parents. Additionally, we sought to determine the viability of using an embodied eye tracking system with D/HH children who received different forms of audiological hearing intervention (CIs and HAs). The long–term goal of this project is to identify the underlying processes that regulate and contribute to individual differences in joint attention skills and novel word learning before and after audiological hearing intervention.

Five D/HH children (three with CIs and two with HAs) were recruited from a large university hospital–based speech and hearing clinic and local advertisements. CI and HA users were required to have used their hearing device for 10 months or more, use a currently available state–of–the–art hearing device, live in a home with spoken English as the primary language, and have no additional developmental, neurological, or cognitive conditions other than hearing loss. Demographics and hearing history variables obtained for the D/HH sample are provided in Table 1. Etiology of hearing loss included unknown ($N = 4$, 80%) and congenital ($N = 1$, 20%). Participants in the D/HH sample averaged 32.80 (*4.21)* months old at time of testing with 16.60 *(6.54)* months of device use. At time of testing, all children reported using oral communication strategies.

Participants in the NH control sample were 10 children who reported no significant developmental, neurological, or cognitive delays. NH peers were recruited from advertisements placed in the community. Characteristics of the NH sample are also summarized in Table 1. Five NH participants were recruited as chronological–age matched controls and averaged 31.40 *(5.86)* months old at time of testing. An additional five NH participants were recruited as hearing–age matched controls and averaged 17.20 *(5.77)* months old at testing. Two NH control samples were necessary to disentangle effects of age from effects of access to sound. All D/HH and NH children participated in this feasibility study with their NH mothers.

––– Insert Table 1 About Here –––

Study procedures were followed as described above, except because of the large amount of individual variability and our small sample size, children’s novel word learning results are omitted. Figure 2 depicts a comparison between one deaf child with CIs (top panel) and their NH chronological–age matched peer (bottom panel). This figure shows the duration of each novel object (represented in the three primary colors: red, green, blue) and face (represented in pink) in the child's and mother's visual field over the course of 30 seconds, respectively. Figure 2 (top panel) provides evidence of decoupling between what the child and mother viewed. However, there are also several episodes in time where there is clear visual evidence of the mother–child dyad jointly attending to the same object at the same time (bottom panel).

–– Insert Figure 2 About Here ––

Children's ability to deploy selective attention and maintain sustained attention is a hallmark of self–regulation (Ruff & Rothbart, 2001), is influenced by joint play with a social partner (Yu & Smith, 2016), and is highly predictive of later executive functioning skills (Barkley, 1997). As such, we examined the effect of hearing status and age on gaze duration and gaze shifting. Gaze data were analyzed with respect to four regions of interest (ROIs): each of the three novel toys and the face. Children's gaze duration to the ROIs was equivalent across hearing status and age; however, the frequency of gaze shifting across ROIs was significantly different ($F(2, 12) = 6.381$, $p < .05$; see Figure 3, Panel A and B respectively). D/HH children shifted gaze across ROIs (18.47 switches per minute) more often than their NH chronological–age (11.12 shifts per minute) and hearing–age (8.94 shifts per minute) matched peers. Is increased gaze shifting indicative of CI/HA users' increased perceptual flexibility or an inability to use linguistic information to constrain and sustain visual attention? The latter interpretation is in line with previous research indicating that preschool CI users are at high risk for delays in executive functioning, namely, attention and sustained sequential processing skills (Kronenberger, Beer, Castellanos, Pisoni, & Miyamoto, 2014).

Analyses were also conducted to determine if hearing status and age influenced children's and parent's toy manipulation during joint play. There were no differences across hearing status or age in children's duration or frequency of object holding. NH parents' duration and frequency of object holding also did not differ as a function of their child's hearing status or age.

–– Insert Figure 3 About Here ––

With respect to linguistic input, there were no differences in mothers' total number of utterances and the total number of topics discussed. However, differences across children's hearing status began to emerge when mothers' utterances were classified by form. NH mothers of D/HH children were more likely to use words to describe object features (e.g., color, shape) than NH mothers of hearing–age matched peers. NH mothers of D/HH children used more directives than NH mothers of chronological–age matched controls. NH mothers of D/HH children and hearing–aged matched controls used similar amounts of directives, suggesting that mother's speech to D/HH children may be specially tailored by their child's hearing experience instead of their child's chronological age. Previous studies have similarly shown that NH mother tailor their infant–directed speech according to CI users' hearing age (Bergeson, Miller, & McCune, 2006). Alternatively, mother's use of more directives may suggest that mothers in this study needed to exert more control over D/HH children's physical behavior. NH mothers of NH chronological–age and hearing–age matched children used more open–ended questions than NH mothers of D/HH children. This finding is particularly interesting given the work by DesJardin and colleagues suggesting that parent's use of open–ended questions is positively associated with growth in expressive language skills in pediatric CI users (Cruz, Quittner, Marker, & DesJardin, 2013; DesJardin & Eisenberg, 2007).

With only 15 mother–child dyads, our preliminary findings demonstrate the feasibility of collecting real–time eye tracking data with young D/HH children who use CIs and HAs. Although we are only beginning to scratch the surface, using this multisensory eye tracking system we are able to mine the frame–by–frame data for patterns of dyadic synchronization. We are also able to examine differences across individual mother–child dyads, as well as, within episodes of joint attention.

## <1>Embodied Cognition in the Classroom

> "In the normal environment there is always more information than the organism is capable of registering. There is a limit to the attentive powers of even the best educated human perceiver" (Gibson, 1969, p. 75).

Neurobiological research on the visual cortex indicates that the brain is unable to process all the visual information ($10^8$–$10^9$ bits per second) entering the retina (Deco, Pollatos, & Sihl, 2002). Since all properties of our multimodal environment cannot be encoded and processed simultaneously, attention is allocated to some properties while others are ignored. This information processing bottleneck has a great influence on deaf children with CIs, who have to rely on highly degraded auditory information and often report being mentally exhausted by the richness of the multimodal environment. As discussed by Marschark & Leigh (2016), educators need to acknowledge that numerous factors, including hearing and neurocognitive abilities, influence deaf children's learning and academic performance in the classroom. In this final section, we focus on strategies for how educators may reduce uncertainty and ambiguity in the learning environment in order to support the development of deaf children's focused attention and inhibition–concentration skills.

In terms of neurocognitive abilities, our eye tracking data, presented in the previous section, suggests that when interacting with parents during play, D/HH toddlers who use CIs/HAs display more gaze shifting per minute, a potential early behavioral manifestation of distractibility. Indeed, by preschool age, parent–reported data suggests that a larger portion of CI users are at risk for clinical delays in controlled and automatic attention when compared to their NH peers (Kronenberger et al., 2014). Parents also report that preschool children with CIs have greater difficulty with working memory and inhibiting prepotent behaviors (Beer, Kronenberger, Castellanos, Colson, Henning & Pisoni, 2014). Performance–based data corroborates these findings in preschool and school age CI users. On a nonverbal visual task measuring inhibition and concentration skills that required participants to identify and eliminate target items from a larger display of items, preschool CI users performed significantly poorer than their NH peers (Beer et al., 2014). Similarly, Castellanos et al. (2015) reports that when compared to a sample of NH peers matched on nonverbal intelligence, school age CI users performed significantly poorer on tasks measuring inhibition–concentration: the Test of Variables of Attention (TOVA; Leark, Dupuy, Greenberg, Corman, & Kindschi, 1996) and the Trail–Making Test (Delis, Kaplan, & Kramer, 2001). On the TOVA, participants were asked to press a button when presented with a target stimulus (a square at the top of a screen) but not when presented with a distractor stimulus (a square at the bottom of a screen). On the Trail–Making Test, participants were asked to connect a series of numbers and letters on a page by drawing a line alternating between numbers and letters (e.g., a line connecting the number 1 to the letter A). Results from the TOVA demonstrated that school age CI users were significantly slower to respond, failed to respond to a target more often, and responded inaccurately to a distractor more often than their NH peers. CI users also performed more poorly on the Trail–Making Test than their NH peers. Together, performance on these two inhibition–concentration tasks predicted

CI users' conceptual knowledge and use of linguistic labels to specifyi size, color, shape, and quantity, basic concepts that are necessary for the understanding of science and mathematics (Castellanos et al., 2015).

Classroom learning environments are often cluttered with acoustic noise (multiple talkers speaking at the same time in background noise that challenge listening), and visual noise (large amounts of visually salient stimuli that challenge focused visual attention). Classroom babble, internal classroom noise generated by peers, has been found to have negative effects on NH children's learning and academic performance. For example, Shield & Dockrell, (2008) reported that higher amounts of classroom babble was associated with NH children's (aged 7 years on average) poorer performance on reading, writing, and mathematics tests. Group learning activities and semicircular seating arrangements may present challenges for reducing classroom babble. Another aspect to consider is that determining the localization of sound sources during group activities may be difficult for unilateral CI users. Dedicated use of personal frequency modulation (FM) systems to amplify the foreground sound (teacher's voice) while attenuating background noise would be one strategy for reducing competing noise caused by classroom babble.

Fisher, Godwin, & Seltman (2014) describe how the visual environment of a typical classroom can effect NH children's attention and learning. Two classroom environments were constructed to investigate the effect of visual distractions on children's learning of science concepts: one classroom contained visually salient stimuli such as wall decorations (e.g., posters, student drawings) typically found in kindergarten classrooms, while the second classroom was void of wall decorations. The data suggest that NH children instructed in classrooms with visually salient wall decorations spent more time off task on environmental distractions and displayed less learning than NH children instructed in classrooms void of visual decorations. The visual noise available in the classroom (both in central and peripheral visual field) may be far more stimulating for CI users than their NH peers. Dye, Hauser, & Bavelier (2008) caution educators from constructing classroom–seating arrangements such that deaf children are in the front of the classroom with their peers behind them, because it may lead to deaf children becoming more distractible. Instead of altering classroom–seating arrangements, one strategy for reducing visual noise may be to simply limit visual decorations so as not to attract attention away from the primary learning tasks.

Acoustic and visual noises make deploying selective attention, maintaining sustained attention, inhibiting prepotent responses, and responding to joint attention more difficult for CI users. Group learning activities, which are typical in primary school settings, can make it difficult for CI users to follow the flow of ideas and conversation. Frequent adult monitoring of CI users' attention during tasks, especially peer–to–peer and group tasks is paramount to promoting learning and academic success. We suggest the need for the frequent assessment of CI users' joint attentional skills throughout early development and school entry. CI users' joint attentional skills may be assessed during short play interactions with their parents, peers, and teachers. Also, during one–to–one instruction, parents, speech language pathologists, and educators should introduce novel words and concepts by following CI users' attentional focus instead of attempting to redirect attentional focus.

As suggested by Dye, Hauser, & Bavelier (2008), increased research efforts need to be placed on delineating the optimal learning environments necessary to support the instruction of deaf children. The majority of the research conducted on CI users, our research included, is visually and acoustically sterile (e.g., occurring in the sound booth, a quiet room with an experimenter, or in an all white room with no visual distractors). Therefore, there is a pressing clinical and educational need for research to be conducted on CI users' learning in mock classrooms with peers or in naturalistic classroom environments under more adverse and challenging conditions. We hope that the easy functionality of our eye tracking equipment will allow us to do just that: to obtain multiple measures of real–time learning in speech–language therapy sessions and in the classroom. Early knowledge about CI users' basic learning strategies and skills during naturalistic interactions will help inform decisions about assessment, intervention, and educational placement.

<1>References

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94.

Bergeson, T. R., Miller, R. J., & McCune, K. (2006). Mothers' speech to hearing–impaired infants and children with cochlear implants. *Infancy*, *10*(3), 221–240.

Castellanos, I., Kronenberger, W. G., Beer, J., Colson, B. G., Henning, S. C., Ditmars, A., Pisoni, D. B. (2015). Concept formation skills in long–term cochlear implant users. *Journal of Deaf Studies and Deaf Education*, *20*(1), 27–40.

Cejas, I., Barker, D. H., Quittner, A. L., & Niparko, J. K. (2014). Development of joint engagement in young deaf and hearing children: Effects of chronological age and language skills. *Journal of Speech Language and Hearing Research*, *57*(5), 1831–20.

Cruz, I., Quittner, A. L., Marker, C., & DesJardin, J. L. (2013). Identification of effective strategies to promote language in deaf children with cochlear implants. *Child development*, *84*(2), 543–559.

Deco, G., Pollatos, O., & Sihl, J. (2002). The time course of selective visual attention: theory and experiments. *Vision Research*, *42*, 2925–2945.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis– Kaplan executive function system*. San Antonio, TX: The Psychological Corporation.

DesJardin, J. L., & Eisenberg, L. S. (2007). Maternal contributions: Supporting language development in young children with cochlear implants. *Ear and hearing*, *28*(4), 456–469.

Dunham, P. J., Dunham, F., & Curwin, A. (1993). Joint–attentional states and lexical acquisition at 18 months. *Developmental Psychology*, *29*(5), 827–831.

Dye, M. W. G., Hauser, P., & Bavelier, D. (2008). Visual attention in deaf children and adults: Implications for learning environments. In M. Marschark & P. Hauser (Eds.), *Deaf Cognition* (pp. 250–263). New York: Oxford University Press.

Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual environment, attention allocation, and learning in young children. *Psychological Science*, *25*(7), 1362–1370.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur–Bates Communicative Development Inventories* (Second Edition) (pp. 1–12). Baltimore, MD: Brookes.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton–Century–Crofts.

Goldfield, B. A. (1986). Referential and expressive language: a study of two mother–child dyads. *First Language*, *6*(17), 119–131.

Golinkoff, R. M., Hirsh–Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14,* 23–45.

Kronenberger, W. G., Beer, J., Castellanos, I., Pisoni, D. B., & Miyamoto, R. T. (2014). Neurocognitive risk in children with cochlear implants. *JAMA Otolaryngology–Head & Neck Surgery*, *140*(7), 608–8.

Leark, R. A., Dupuy, T. R., Greenberg, L. M., Corman, C. L., & Kindschi, C. L. (1996). *Test of variables of attention professional manual version 7.0*. Los Alimitos, CA: Universal Attention Disorders.

Markman, T. M., Quittner, A. L., Eisenberg, L. S., Tobey, E. A., Thal, D., et al. (2011). Language development after cochlear implantation: an epigenetic model. *Journal of Neurodevelopmental Disorders*, *3*(4), 388–404.

Marschark, M., & Leigh, G. (2016). Recognizing diversity in deaf education: Now what do we do with it?! In M. Marschark, V. Lampropoulou, & E. Skordilis (Eds.), *Diversity in deaf educatio*n (pp. 507–535). New York, NY: Oxford University Press.

Moeller, M. P. (2007). Current state of knowledge: psychosocial development in children with hearing impairment. *Ear and Hearing, 28*(6), 729–739.

Moore, C., & Dunham, P. J. (Eds.). (1995). *Joint attention: Its origins and role in development*. New York, NY: Lawrence Erlbaum Associates.

Musselman, C., & Churchill, A. (1992). The effects of maternal conversational control on the language and social development of deaf children. *Communication Disorders Quarterly*, *14*(2), 99–117.

Prezbindowski, A. K., Adamson, L. B., & Lederberg, A. R. (1998). Joint attention in deaf and hearing 22 month–old children and their hearing mothers. *Journal of Applied Developmental Psychology*, *19*(3), 377–387.

Ruff, H. A., & Rothbart, M. K. (2001). *Attention in early development: Themes and variations.* New York, NY: Oxford University Press.

Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *The Journal of the Acoustical Society of America*, *123*(1), 133–144.

Spencer, P. E., Meadow–Orlans, K. P., Koester, L. S., & Ludwig, J. (2004). Relationship across developmental domains and over time. In K. P. Meadow–Orlans, P. E. Spencer, & L. S. Koester (Eds.), *The world of deaf infants: A longitudinal study* (pp.205–217). New York, NY: Oxford University Press.

Spencer, P. E., Swisher, M. V. & Waxman, R. P. (2004). Visual attention: Maturation and specialization. In K. P. Meadow–Orlans, P. E. Spencer, & L. S. Koester (Eds.), *The world of deaf infants: A longitudinal study* (pp.205–217). New York, NY: Oxford University Press.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463.

Tomasello, M., & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, *4*(12), 197–211.

Tomasello, M., Mannle, S., & Kruger, A. C. (1986). Linguistic environment of 1– to 2–year–old twins. *Developmental Psychology*, *22*(2), 169–176.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition, 125*(2), 244–262.

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye–hand coordination. *PLoS ONE*, *8*(11), e79659.

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one–year–old human infants. *Current Biology*, 1–12.

Yu, C., Smith, L. B., & Pereira, A. F. (2008, August). *Embodied solution: The world from a toddler's point of view*. Paper presented at the 7th IEEE International Conference on Development and Learning.

Table 1

*Participant Demographics and Hearing History*

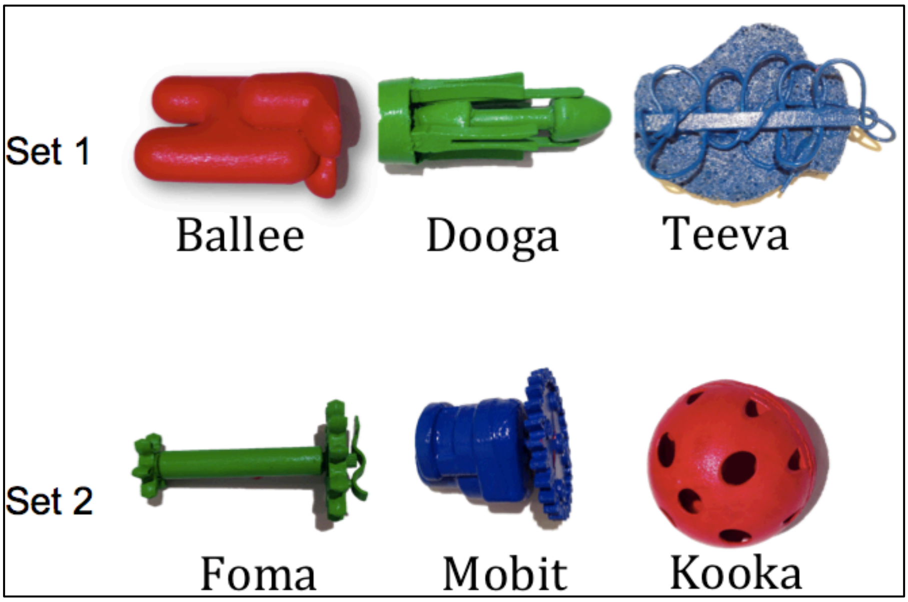
| | Hearing Status | | |
|---|---|---|---|
| | Deaf/ Hard–of –Hearing | Normal Hearing Chronological–Age Match | Normal Hearing Hearing–Age Match |
| *N* | 5 | 5 | 5 |
| | M *(SD)* | | |
| Age at Test (mos.) | 32.80 *(4.21)* | 31.40 *(5.86)* | 17.20 *(5.77)* |
| Duration of CI/HA use (mos.) | 16.60 *(6.54)* | | |
| | Count (% of Sample) | | |
| Hearing Device | | | |
| Bilateral, Sequential CIs | 1 (20%) | | |
| Unilateral CI | 1 (20%) | | |
| CI and HA | 1 (20%) | | |
| Bilateral HAs | 1 (20%) | | |
| Unilateral HA | 1 (20%) | | |
| Etiology of Hearing Loss | | | |
| Unknown | 4 (80%) | | |
| Congenital | 1 (20%) | | |
| Gender | | | |
| Female | 3 (60%) | 3 (60%) | 3 (60%) |
| Male | 2 (40%) | 2 (40%) | 2 (40%) |

*Note.* CI = Cochlear Implant; HA = Hearing Aid

Figure 1

*Panel A: Head–mounted camera and eye tracking system*

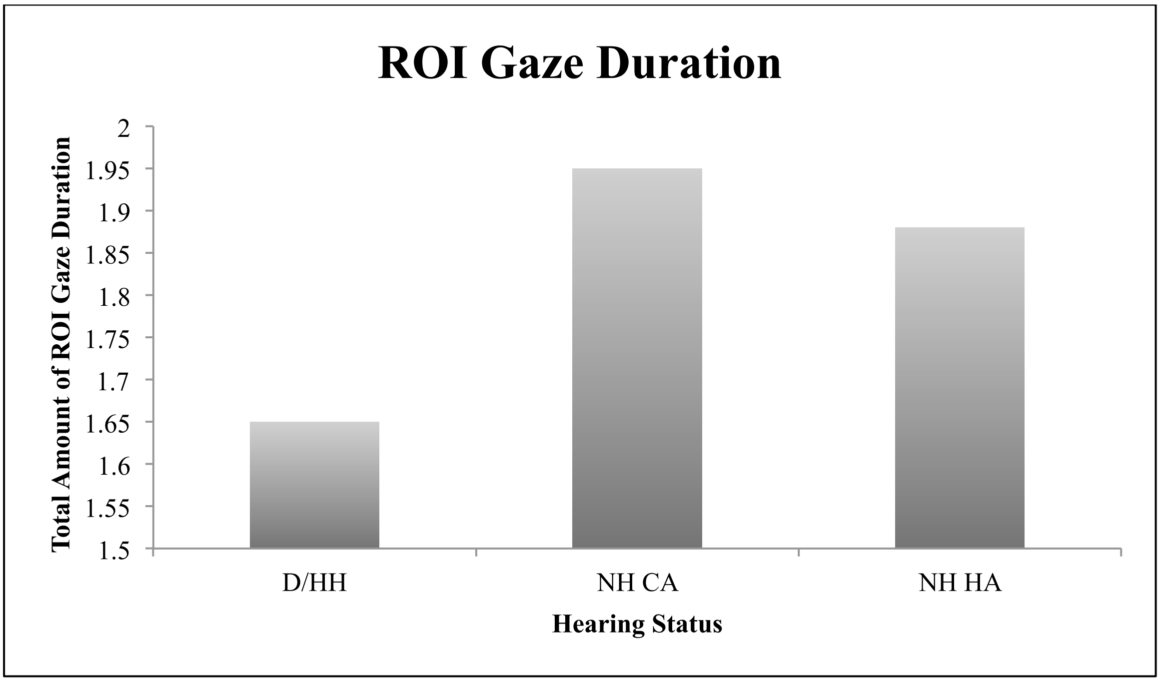
*Panel B: Example of two sets of novel toys*



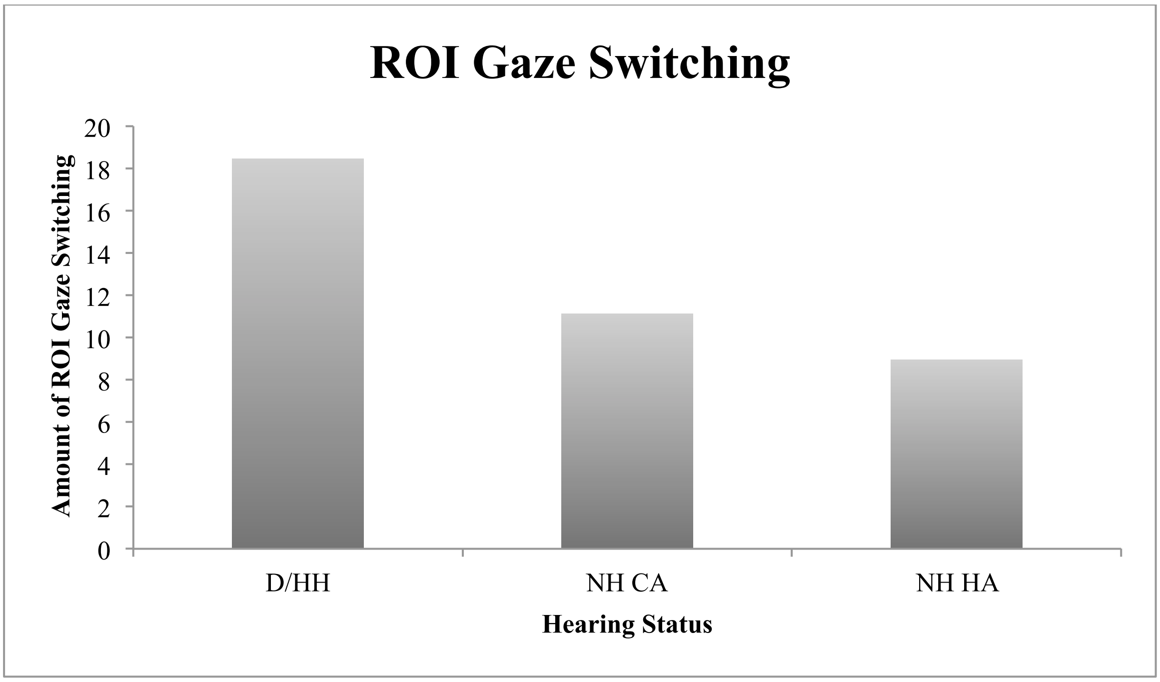

*Note.* (A): A head–mounted camera and an eye tracker are placed on the child (pictured) and the mother in order to collect visual information from their egocentric views. (B): The parent–child dyad is provided with two sets of novel toys. Each set contains three toys painted in primary colors: red, green, and blue.

Figure 3

*Panel A: Gaze Duration for Regions of Interest (ROIs: 3 novel toy objects, face)*



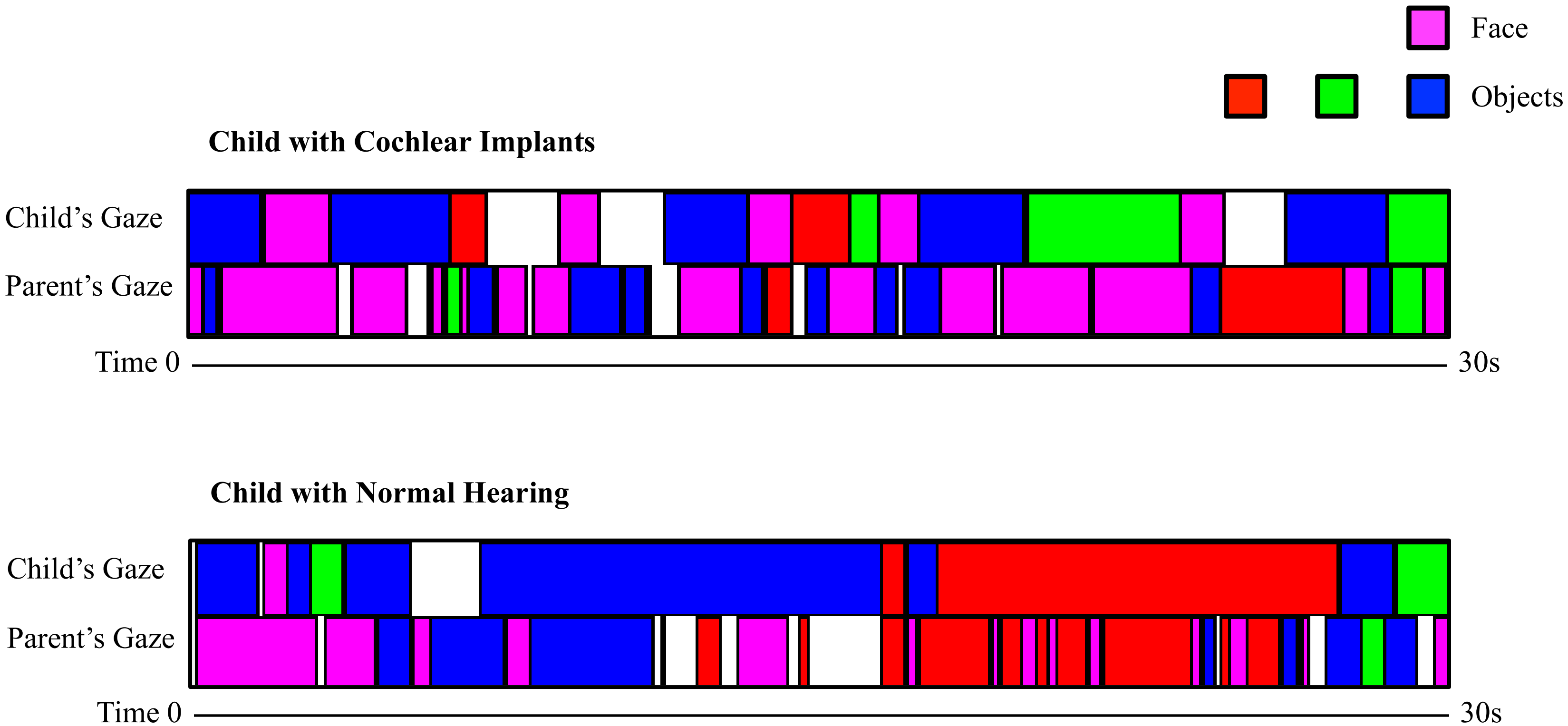

*Panel B: Gaze Switching across Region of Interests (ROIs: 3 novel toy objects, face)*




*Note.* D/HH = Deaf or Hard–of–Hearing; NH CA = Normal Hearing Chronological–Age Match; NH HA = Normal Hearing Hearing–Age Match

Figure 2

*An example of the synchronization of gaze data across time*




*Note.* The child's and parent's gaze data stream are presented in each panel on the vertical axis. Looking to the face is presented in pink, while looking to the three novel objects is presented red, green, and blue. Thirty seconds of elapsed time appears on the horizontal axis.