What Leads to Coordinated Attention in Parent-toddler Interactions? Children's Hearing Status Matters

Chi-hsin Chen[1], Irina Castellanos[1,2], Chen Yu[3], Derek M. Houston[1,2]

[1] Department of Otolaryngology – Head and Neck Surgery, The Ohio State University
915 Olentangy River Road, Columbus, Ohio 43212

[2]Nationwide Children's Hospital
700 Children's Dr, Columbus, Ohio 43205

[3]Department of Psychological and Brain Sciences, Indiana University
1101 E. 10th Street, Bloomington, Indiana 47405

Corresponding Author:
Chi-hsin Chen
Email: Chi-Hsin.Chen@osumc.edu

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Running Title: What leads to coordinated attention

Key words: parent-child interactions; coordinated attention; children with hearing loss; eye-tracking; gaze following

Research Highlights

- The current study used head-mounted eye-trackers to investigate how hearing parents and their children with and without hearing loss achieve coordinated attention in toy play.
- We found that children's hearing status did not affect how likely parents and children attended to the same object at the same time in play.
- When following parents' attention, children with hearing loss used both parents' gaze directions and hand actions as cues to establish coordinated attention with their parents.
- In contrast, children with normal hearing mainly relied on parents' hand actions to guide their own visual attention.

Conflict of Interest Statement

We declare that we have no conflict of interest.

Acknowledgements

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (T32 DC00012), the National Institutes of Health (R01 HD074601), and the Indiana University Collaborative Research Grant. We thank Heidi Neuberger, Steven Elmlinger, Charlene Ty, Mellissa Hall, and Seth Foster for help with data collection and members of the Computational Cognition and Learning Laboratory and the DeVault Otologic Research Laboratory for help with coding. We also thank Seth Foster and Tian (Linger) Xu for developing software for data management and processing.

Abstract

Coordinated attention between children and their parents plays an important role in their social, language, and cognitive development. The current study used head-mounted eye-trackers to investigate the effects of children's prelingual hearing loss on how they achieve coordinated attention with their hearing parents during free-flowing object play. We found that toddlers with hearing loss (age: 24 – 37 months) had similar overall gaze patterns (e.g., gaze length and proportion of face looking) as their normal hearing peers. In addition, children's hearing status did not affect how likely parents and children attended to the same object at the same time during play. However, when following parents' attention, children with hearing loss used both parents' gaze directions and hand actions as cues, while children with normal hearing mainly relied on parents' hand actions. The diversity of pathways leading to coordinated attention suggests the flexibility and robustness of developing systems in using multiple pathways to achieve the same functional end.

Introduction

The coordination of attention between young children and their parents during social interaction plays an important role in their social, language, and cognitive development (Brooks & Meltzoff, 2005; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Landry, Smith, Swank, & Miller-Loncar, 2000; Niedzwiecka, Ramotowska, & Tomalski, 2019). Social interactions are multimodal, but little is known about how atypical sensory experiences, such as hearing loss, affect the coordination of attention between children and their social partners. Investigating the effects of hearing loss on how attention is coordinated during parent-child social interaction can provide important insights into the role of sensory experiences in both typical and atypical development.

Most work on attention coordination has been conducted with normal-hearing children in the context of assessing joint attention skills (Bakeman & Adamson, 1984; Charman et al., 2000; Mundy et al., 2007; Mundy & Gomes, 1998; Scaife & Bruner, 1975; Tomasello & Farrar, 1986). In the rich literature on this topic, joint attention has been conceptualized and operationally defined in different ways (Siposova, & Carpenter, 2019; Tasker & Schmidt, 2008). Traditionally, many researchers view intentionality and/or (mutual) awareness of social partner's attentional state as critical components of joint attention (e.g., Bakeman & Adamson, 1984; Depowski, Abaya, Oghalai, & Bortfeld, 2015; Moll & Meltzoff, 2011; Tasker, Nowakowski, & Schmidt, 2010; Tomasello & Farrar, 1986). To successfully establish joint attention, one needs to look at an object of interest first, which is followed by a verification look to the partner's face to confirm that the other person is also looking at the same object. If the partner is attending somewhere else, one would intentionally guide the partner's attention to the object of interest to establish joint attention. Thus, according to this perspective, the social components of intentionality and/or awareness are crucial for joint attention. Indeed, the literature shows that young children's skills to achieve joint attention predict not only later social development, such as social understanding, but also their receptive and productive language development, self-regulation, and executive functions (Brooks & Meltzoff, 2005; Charman et al., 2000; Morales et al., 2000; Mundy et al., 2007; Van Hecke et al., 2012).

More recently, a growing body of research uses a "leaner" definition to study attention coordination -- sometimes termed "coordinated attention" -- and defines it as two social partners looking at the same object at the same time, with or without conscious awareness of the other person's attentional state (e.g., de Barbaro, Johnson, Forster, & Deak, 2016; Deák, Krasno, Jasso, & Triesch, 2018; Deak, Krasno, Treisch, Lewis, & Sepeta, 2014; Suarez-Rivera, Smith, & Yu, 2019; Yu & Smith, 2013; Yu, Suanda, & Smith, 2019). Research taking this perspective defines "leading" or "following" the other person's attention based solely on the temporal relationship of the social partners' gaze or behaviors (Chen, Castellanos, Yu, & Houston, 2019a; Deák et al., 2018; Deak et al., 2014; Piazza, Hasenfratz, Hasson, & Lew-Williams, 2018; Suarez-Rivera et al., 2019; Wass et al., 2018; Yu & Smith, 2013; Yu et al., 2019). The two partners may or may not intentionally lead or follow the other. In addition to definitional differences, studies taking this perspective usually analyze data at the micro-level, using high-resolution video recording or eye-tracking devices to study sensorimotor behaviors, with a temporal resolution within fractions of a second. In contrast, studies focusing on the social aspects of joint attention tend to analyze data at the macro-level with more coarse resolution, usually at the second to over a dozen of seconds level (e.g., Tasker et al., 2010). Using a leaner definition and focusing on high-resolution sensorimotor data is largely in line with contemporary approaches focusing on interpersonal coordination that may or may not be intentional or conscious (De Jaegher, Di Paolo, & Gallagher, 2010; Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012; Noy, Dekel, & Alon, 2011; Repp & Su, 2013; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). Past research on adult-to-adult interaction and conversation reveals that bodily adjustments -- head, hand, and posture -- occur at timescales of fractions of a second and that these rapid adjustments are essential to the negotiation of speaking turns and the maintenance of common ground (Shockley, Richardson, & Dale, 2009; Shockley, Santana, & Fowler, 2003). In the course of development, it is likely that parent-child interactions teach or entrain the sensorimotor dynamics that are characteristic in mature adult-to-adult interactions. If so, parent-child interactions may also depend on tight dynamic coupling. Subtle distortions or lags in partners' sensorimotor responses could significantly reduce the quality of these interactions and opportunities for learning. Recent studies using high-density behavioral data and micro-level analyses have shown that micro-level sensorimotor behaviors observed in parent-child joint play are predictive of children's later development (Yu et al., 2019). Following this line of research, the current study will take the micro-level sensorimotor analyses approach. However, to avoid confusion with studies focusing on the social components of joint attention, we use the term *coordinated attention*, which is operationally defined as two social partners looking at the same object at the same time, and investigate what leads to coordinated attention in free-flowing parent-child interactions in children with and without hearing loss.

Pathways to Coordinated Attention

There is more than one way for a young child and a parent to follow each other's attention and establish coordinated attention between the two. For example, they may use the social partner's gesture, such as pointing, to join the partner to attend to the same object (de Villiers Rader &

Zukow-Goldring, 2012). Another way to establish coordinated attention is via gaze following – looking at the partner's eyes to infer gaze direction (Brooks & Melzoff, 2005; Butterworth & Cochran, 1980; Doherty, Anderson, & Howieson, 2009; Gredeback, Fikke, & Melinder, 2010; Slaughter & McConnell, 2003). Gaze following has been posited to be the single most effective pathway leading to coordinated attention (Butterworth & Cochran, 1980; Slaughter & McConnell, 2003), which is evidenced by a large literature showing that even young infants are capable of following others' gaze direction in well-controlled experimental settings (Brooks & Melzoff, 2005; Butterworth & Cochran, 1980; Flom & Pick, 2005; Gredeback, Theuring, Hauf, & Kenward, 2008; Gredeback et al., 2010; Slaughter & McConnell, 2003). However, several recent studies conducted in more naturalistic contexts have found that infants and young children rarely look at their parent's face during several common daily activities, such as crawling, walking, and object play (Chang, de Barbaro, & Deak, 2016; de Barbaro et al., 2016; Deak et al., 2018; Deak et al., 2014; Franchak, Kretch, Soska, & Adolph, 2011; Yu & Smith, 2013, 2017a, 2017b). Instead, a more available and reliable cue infants use to establish coordinated attention is through following their partner's manual actions on objects (Chang et al., 2016; de Barbaro et al., 2016; Deak et al., 2018; Deak et al., 2014; Yu & Smith, 2013, 2017a, 2017b). Because gaze direction usually aligns with manual actions on objects during toy play, hand following has been shown to be a viable pathway for young children to achieve coordinated attention in parent-infant interactions (Deak et al. 2014; Deak et al., 2018; Yu & Smith, 2013, 2017a, 2017b).

Typically developing children are capable of using both gaze following and hand following to establish coordinated attention with their parents, but little is known about whether children growing up with different sensory experiences also use more than one pathway to achieve coordinated attention with their parents. The current study investigates this larger issue by focusing on a specific group of children – children with prelingual hearing loss who either receive acoustical information via hearing aids or cochlear implants. Compared with hearing children, children with prelingual hearing loss spend their first several months to over a year of life with limited-to-no access to acoustical information. The goal of the present study was to examine for this group of children, whether, and if so how often, and in what ways they achieve coordinated attention with their hearing parents. Given that coordinated attention is fundamental for early development of typically developing children, better understanding of how children with hearing loss and their parents achieve coordinated attention could lead to important insights into their development of social interaction and language.

Some previous studies with children with hearing loss suggest that they may have different coordinated attention patterns with their parents compared to children with normal hearing. For example, Bortfeld and Oghalai (2018) compared the abilities to initiate and respond to attentional bids (e.g., pointing, speech, gaze switching, or waving) between 4 toddlers with hearing loss (age range: 18.2 - 36.7 months) and their age-matched hearing peers. They found that toddlers with hearing loss were less successful in initiating or responding to their hearing parents' attentional bids than their hearing peers. Moreover, hearing parents of toddlers with hearing loss have been found to be more likely to touch the child when trying to elicit or maintain their attention (Depowski et al., 2015). Prezbindowski, Adamson, and Lederberg (1998) found that, compared to age-matched hearing peers, 22-month-old children with hearing loss spent more time alternating attention between their mother's face and the objects they played with, suggesting that those children may be more likely to look at their mother's face than hearing children did. The face looking behavior potentially gives children with hearing loss more chances to obtain and use parents' gaze direction as a cue to detect and follow parents' focus of attention. In a recent study, Lieberman, Hatrak, and Mayberry (2014) compared the naturalistic interactions between 4 deaf children between the ages of 1;9 and 3;7 and their deaf mothers and the interactions between hearing children and their hearing mothers. They found that deaf children spent more time looking at their mother's face than hearing children did and they tended to shift their gaze between their mother's face and the object of interest. More frequent looks to their mother's face reported in these studies suggest that gaze cues may be more available to children with hearing loss and this looking pattern may affect how likely and the ways in which they establish coordinated attention with their parents, which is the focus of the present study.

Current Study

The current study used head-mounted eye-trackers worn by both parents and toddlers to investigate how hearing parents and their toddlers with and without hearing loss achieve coordinated attention during naturalistic object play. We were particularly interested in the two pathways that have been previously shown to be used by typically developing children in joint object play: gaze following and hand following. We collected real-time gaze and manual data from both social partners during object play and examined the temporal unfolding of parents' and toddlers' gaze and other behavioral patterns. With the fine-grained real-time gaze and manual data, we were able to examine how parents and children used different pathways to achieve coordinated attention. Two groups of children with normal hearing were also recruited, one matched to the hearing loss group in chronological age and the other matched in hearing experience. Including these two groups of children with normal hearing allowed us to investigate whether children's chronological age, hearing experience, or hearing status *per se*, affect how they achieve coordinated attention with their parents. We asked two specific questions by comparing children with and without hearing loss. The first question is whether children's chronological age, hearing experience or hearing status affect how well parents and children both attend to the same object during interactions. Children with hearing loss have been shown to be more easily distracted compared to children with normal hearing (Dye & Hauser, 2014). One possibility is that children's hearing status may influence their visual attention and consequently they may not be able to establish coordinated attention with their parents as effectively as children with normal hearing do. Alternatively, they may be able to create their own solution to successfully establish coordinated attention with their parents. To answer this question, we will first report data on parents' and toddlers' overall looking behaviors in object play. We will then examine the coordinated attention episodes created by both partners and report how often parents follow children to establish coordinated attention and how often children follow parents. The second question is whether children with hearing loss use both gaze and hand following pathways to follow the parent's attention in similar ways as children with normal hearing do. The finding of more face looks from children with hearing loss suggests that there are more opportunities for them to use gaze following compared with hearing children who rarely look at the parent's face. To test this hypothesis, we will examine whether parents and children in different groups rely on different pathways to achieve coordinated attention.

Method

Participants

Participants were 21 parent-child dyads. One of the parents was a father (in the HA group, definition see below); the rest were mothers. Children in 7 dyads had hearing loss (subsequently termed HL group) and were between the ages of 24 and 37 months (for detailed information see Table 1). Children in another 14 dyads had normal hearing; 7 of them were matched to the HL group in chronological age (subsequently termed CA group) and the other 7 children were matched to the HL group in hearing age (subsequently termed HA group). The hearing age of the HL group was calculated based on how long they had their hearing device (cochlear implants or hearing aids), while the hearing age of the children with normal hearing was calculated based on their chronological age. Recruitment and experimental procedures were approved in advance by the University Institutional Review Board and all parents gave informed consent prior to participation. The entire sample of participants was broadly representative of the State of Indiana (86% European American, 10% African American, 4% Asian, Hispanic, and other), and consisted of predominantly working and middle-class families.

Table 1. Participant Information

<table>
<tr><th colspan="8">HL</th><th>CA</th><th>HA</th></tr>
<tr><th>Participant #</th><th>Chronological Age</th><th>Hearing age</th><th>Sex</th><th colspan="2">Degree of Hearing Loss</th><th colspan="2">Hearing Device</th><th>Age</th><th>Age</th></tr>
<tr><th></th><th></th><th></th><th></th><th>Left</th><th>Right</th><th>Left</th><th>Right</th><th></th><th></th></tr>
<tr><td>1</td><td>24</td><td>15</td><td>F</td><td colspan="2">profound</td><td>Cochlear implant</td><td>Cochlear implant</td><td>24</td><td>15</td></tr>
<tr><td>2</td><td>27</td><td>22</td><td>M</td><td colspan="2">severe</td><td>Hearing aid</td><td>Hearing aid</td><td>25</td><td>23</td></tr>
<tr><td>3</td><td>28</td><td>12</td><td>F</td><td colspan="2">profound</td><td>Cochlear implant</td><td>Cochlear implant</td><td>26</td><td>12</td></tr>
<tr><td>4</td><td>30</td><td>10</td><td>F</td><td colspan="2">severe to profound</td><td>Cochlear implant</td><td>Cochlear implant</td><td>28</td><td>12</td></tr>
<tr><td>5</td><td>34</td><td>14</td><td>F</td><td colspan="2">severe</td><td></td><td>Cochlear implant</td><td>35</td><td>14</td></tr>
<tr><td>6</td><td>36</td><td>25</td><td>F</td><td>profound</td><td>mild-moderate</td><td></td><td>Hearing aid</td><td>36</td><td>24</td></tr>
<tr><td>7</td><td>37</td><td>12</td><td>M</td><td>profound</td><td>severe</td><td>Cochlear implant</td><td>Hearing aid</td><td>36</td><td>13</td></tr>
<tr><td>Mean age</td><td>30.9</td><td>15.7</td><td></td><td></td><td></td><td></td><td></td><td>30.0</td><td>16.1</td></tr>
</table>

Note: All ages are reported in months. HL: children with hearing loss, CA: chronological-age-matched children with normal hearing, HA: hearing-age-matched children with normal hearing.

Design

Parents and their toddlers sat across from each other at a small table (61 cm * 91 cm * 64 cm) and played with novel objects (see Fig. 1). During the play session, both participants wore a head-mounted eye-tracker (Positive Science, http://www.positivescience.com/, also see Franchak et al., 2011). Each eye-tracker was composed of an eye camera that recorded eye movements and a scene camera that recorded the first-person view. Two additional cameras were used to record from third-person views. The experiment was divided into 4 "trials;" each lasted 1.5 minutes. Two sets of three objects (average size: 288 cm^3) -- one red, one blue, and one green -- were used. Participants played with each set twice in an alternating order. The whole experiment lasted 6 minutes.

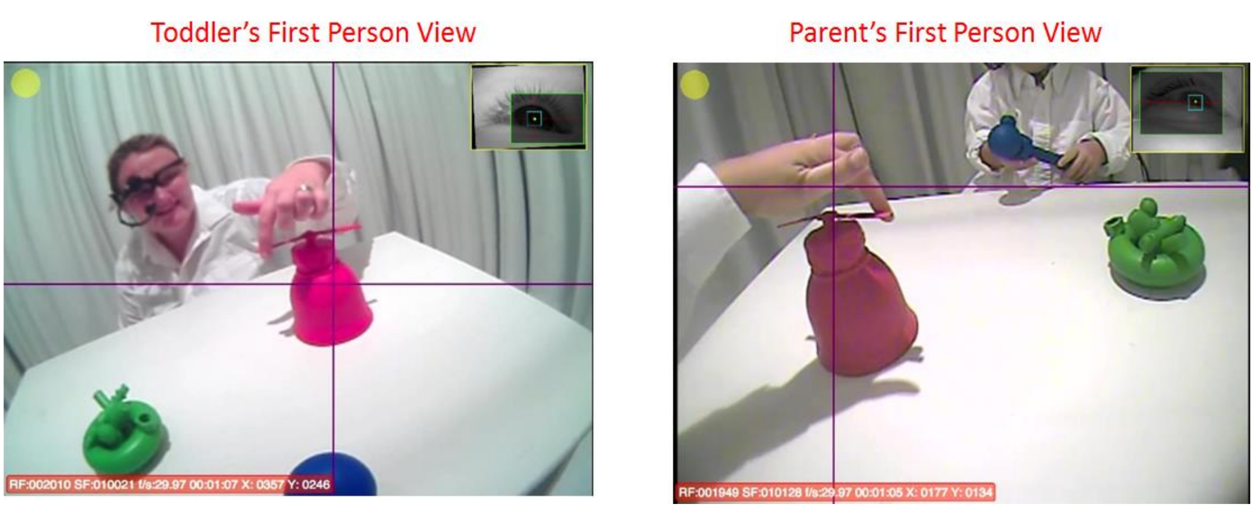

Fig. 1. *The two first-person views from the toddler's and parent's perspectives. The parent and child sat across from each other at a small table and played with 3 novel objects in each trial. Both participants wore a head-mounted eye-tracker, which was composed of an eye camera that recorded eye movements and a scene camera that recorded the first-person view. The cross-hair in each image indicates where the participant looked in the first-person view.*

Data Coding

Coding Eye Gaze

The eye-trackers recorded at a sampling rate of 30 Hz. Therefore, we obtained approximately 10,800 frames from each camera during the six-minute interaction. Gaze data was coded frame by frame. Four regions of interest (ROIs) were identified, the three objects in each trial and the partner's face. Trained coders went through each frame and coded whether the participant's gaze direction fell on any of the ROIs, and if so, which ROI. In total, children across all three groups generated 3,989 looks to the ROIs and parents generated 7,830 looks to the ROIs. Looks to these four ROIs served as our dependent variables in the following analyses (detailed information about ROI coding can be seen in Appendix B in Yu & Smith, 2017b).

Reliability was computed between the coding of two independent coders on 10 randomly selected participants. We calculated Cohen's kappa based on frame-by-frame coding. Because each frame lasted approximately 33 ms in real time (sampling rate of eye-trackers: 30 Hz), the coding inherently took the onset and offset of each gaze and gaze duration into account. Inter-rater reliability was good with an average Cohen's kappa of 0.77 (range: 0.71-0.86, Landis & Koch, 1977).

Coding Hand Contact with Objects

Parents' and children's hand contact with objects were coded separately. In addition, each hand of the same participant was also coded separately. Participants' manual contact with an object was coded frame-by-frame from the two participants' scene cameras and third-person-view cameras. Trained coders went through each frame and determined whether a participant's left or right hand was in contact with an object, and if so, which object. If the left and right hands were in contact with different objects at the same time, both objects were counted as a target of hand contact at that moment. It should be noted that hands were not a region of interest in the gaze coding. In the gaze coding, only the objects (and partner's face) were coded, regardless of whether or not they were touched by the hands.

A second coder independently coded data from 8 randomly selected participants. The average Cohen's kappa based on frame-by-frame coding was 0.94 (range: 0.90-0.97)[1], indicating that there was near-perfect inter-rater agreement on the coding of hand contact (Landis & Koch, 1977).

Data Analyses

Parent-Led vs. Child-Led Coordinated Attention Episodes

[1] The reason why the kappa score for ROI coding was lower than hand coding was that from the eye-trackers (and third-person view cameras), we can only tell the direction of an eye gaze, but not the exact distance of a gaze point. When objects are cluttered together, with one closer to the viewer and the other partially behind the former in the 3-D space, it can be challenging to identify the exact object the participant was looking at. However, for hand contact, coders can use both participants' scene videos and third-person-view videos to judge whether a hand was in contact with any of the objects. Because these different videos were recorded from different angles and allowed checking from different views, it was easier to identify the target of a hand contact than the target of a gaze.

We operationally defined coordinated attention (CAtt) based on the temporal overlap between the parent's and child's looks on the same object (see Fig. 2, CAtt #1-4). A coordinated attention episode was defined as the parent and child looking at the same object at the same time, which could include looks briefer than 300 ms elsewhere. This 300 ms threshold allowed one brief look away from the target before switching back. Because the present study examined both how well and in what ways children followed parents' attention and also how parents followed children's attention, we further divided episodes of coordinated attention as either parent-led or child-led based on who is leading and who is following. A parent-led coordinated attention episode was defined as the parent starts looking at an object before the child joins the parent to look at the same object (See CAtt #2 and CAtt #4 in Fig. 2). A child-led coordinated attention episode was defined as the child starts looking at an object before the parent joins the child to look at the same object (see CAtt #1 and CAtt #3 in Fig. 2). Thus, a parent-led coordinated attention episode was identified as the onset of the parent's look to an object preceding the onset of the child's look to the same object, while a child-led coordinated attention episode was identified as the onset of a child's look to an object preceding the onset of the parent's look. The moment from which the leader's look starts until the follower joins in to look at the same object is termed the *leading moment* (see the double-headed arrows in Fig. 2). In this way, both parent-led and child-led coordinated attention episodes were defined objectively based on the *temporal relationship* between the parent's and child's gazes (Deák et al., 2018; Deak et al., 2014; Suarez-Rivera et al., 2019; Yu & Smith, 2013; Yu et al., 2019). Parents or children may or may not intentionally or consciously lead or follow the other person's attention.

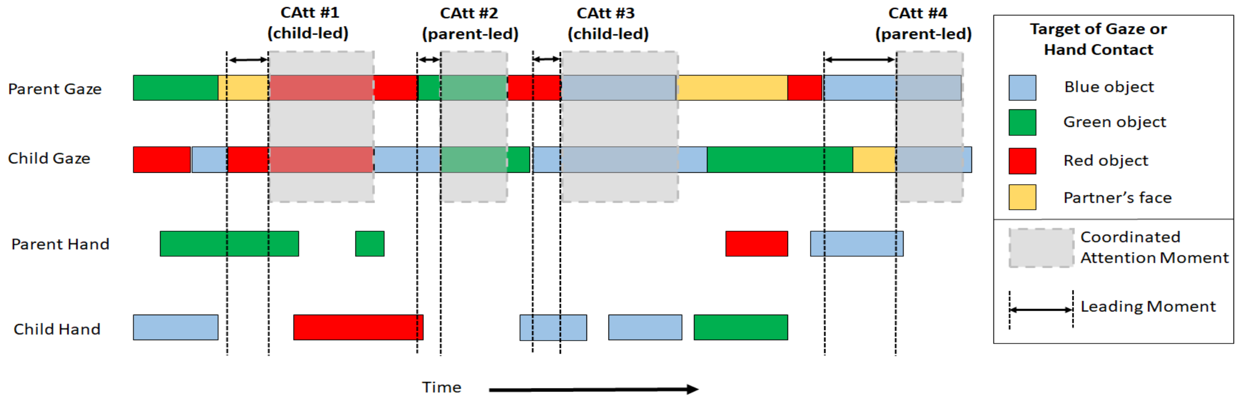

Fig. 2. *Representative time series of parent's gaze, child's gaze, parent's hand contact and child's hand contact. Coordinated attention was objectively defined as the temporal overlap between parent's look at an object and the child's look at the same object. A leading moment is the time in between the coordinated attention leader's gaze onset and the follower's gaze onset.*

Gaze Following and Hand Following

To answer the question of what leads to coordinated attention, we zoomed into the leading moments – the moments right before coordinated attention -- and examined during that period of time whether the follower looked at the leader's face and whether the leader was touching the target object (i.e., the object the two participants both attend within the coordinated attention episode). Following previous studies (Yu & Smith, 2013; 2017a), we defined a coordinated attention episode as being established through gaze following as long as the follower looks at the leader's face during the leading moment (e.g., the leading moments before CAtt #1 and CAtt #4 in Fig. 2). Because of the geometric setup and participants' sitting positions, when they look at their social partner's face, the eyes of the social partner are usually in the field of view and can be used to infer the partner's gaze direction. In contrast, if the follower does not look at the leader's face, then gaze following is not possible. For the rest of the coordinated attention instances without face looking (and therefore not counted as gaze following), if the leader touches *the target object* of attention during the leading moment, we define this coordinated attention event as hand following (e.g., the leading moment before CAtt #3; Yu & Smith, 2013; 2017a). It should be noted that, by our definition, the gaze following pathway takes precedence over the hand following pathway. As long as there is a face look by the follower during the leading moment, we code the episode as gaze following, even if the leader's hand cues are available (e.g., leading moment before CAtt #4). In addition to the gaze and hand following pathways, episodes of coordinated attention may occasionally be established when neither gaze nor hand cues are available (e.g., leading moment before CAtt #2).

Results

In the following, we will first focus on parents' and children's overall ROI looks and examine whether there was any group difference in how often they looked at their social partner's face. After presenting the overall results of the parents' and children's ROI looking behaviors, we will then focus on coordinated attention episodes and investigate the pathways parents used to follow children's gazes and the pathways children used. Because each participant contributed more than one gaze and each dyad created more than one episode of coordinated attention, we used Generalized Estimating Equations (GEE) to account for the non-independence of events within each data set (Liang & Zeger, 1986).

Parents' and Children's ROI looks

Parents across the 3 groups generated 7830 ROI looks in total (HL: 2580, CA: 2577, HA: 2673). There was no group difference in the mean number of ROI looks per minute (HL: 61.43, CA: 61.36, HA: 63.64; Wald χ^2 = 0.54, p = .97). Of these ROI looks, roughly 35% to 40% were looks toward the child's face (HL: 40.1%, CA: 34.2%, HA: 41.1%), and the rest were looks to the objects in play. There was no between-group difference in the mean ROI length (HL: 0.83 s, CA: 0.87 s, HA: 0.83 s; Wald χ^2 = 0.19, p = .91) or proportions of face looks (Fig. 3A, Wald χ^2 = 3.45, p = .18). These results suggest that parents tend to look at their children's faces fairly frequently in free play. Moreover, parents in all three groups had very similar overall gaze patterns.

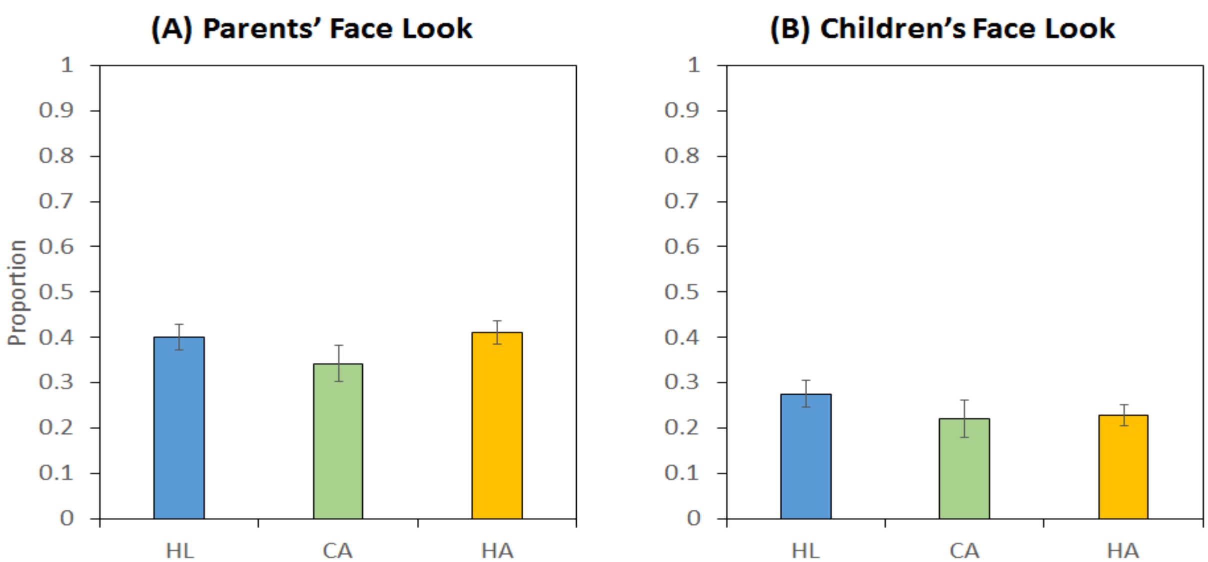

Fig. 3. *Proportion of face looks among all ROI looks. (A) Proportion of parents' face looks in different groups. (B) Proportion of children's face looks in different groups.*

Children generated 3989 ROI looks in total (HL: 1561, CA: 1343, HA: 1085). Children in the HL group generated more ROIs per minute than the younger HA group, but neither of these two groups was different from the CA group (HL: 37.17, CA: 31.98, HA: 25.83, group difference: Wald χ^2 = 7.08, p = .029; HL > HA at p = .01; HL = CA at p =.31; CA = HA at p = .16). Of children's ROI looks, fewer than 30% were looks toward the parent's face (HL: 27.6%, CA: 22.1%, HA: 22.8%) and the rest were looks toward the objects. There was no between-group difference in the mean ROI length (HL: 1.43 s, CA: 1.66 s, HA: 1.75 s; Wald χ^2 = 1.96, p = .38) or proportions of face looks (Fig. 3B, Wald χ^2 = 2.69, p = .26). All three groups of children had a lower proportion of face looks than their parents did (at Wald χ^2 > 8, p < .01). These results show that, HL children's overall gaze patterns were not different from their age-matched CA children, even though HL children generated more ROI looks than the younger HA children. Moreover, the children in all three groups spent less time on attending to the parent's face compared with the proportion of time that their parents spent on attending to the child's face.

Coordinated Attention Episodes

As can be seen in Fig. 4, the three groups generated a total of 1045 coordinated attention episodes (HL: 346, CA: 372, HA: 327). There was no group difference in the mean number of coordinated attention episodes across groups (HL: 49.43, CA: 53.14, HA: 46.71; Wald χ^2 = 1.49, p = .47). The mean duration of coordinated attention episodes was approximately 2.5 seconds (HL: 2.28 s, CA: 2.64 s, HA: 2.62 s). There was no group difference in the mean duration of coordinated attention episodes (Wald χ^2 = 2.82, p = .25).

With regard to parent-led and child-led coordinated attention episodes, all three groups had over half of all the episodes of coordinated attention led by children (HL: child-led = 61.2%, parent-led = 38.8%; CA: child-led = 55.0%, parent-led = 45.0%; HA: child-led = 61.9%, parent-led: 38.1%). There was no significant group difference (Wald $\chi^2 = 3.81$, $p = .15$). These results suggest that the 3 groups had similar overall coordinated attention patterns. Children with hearing loss were able to establish coordinated attention as well as their hearing peers. Thus, children's hearing status did not affect how likely parents and children attended to the same object at the same time in play.

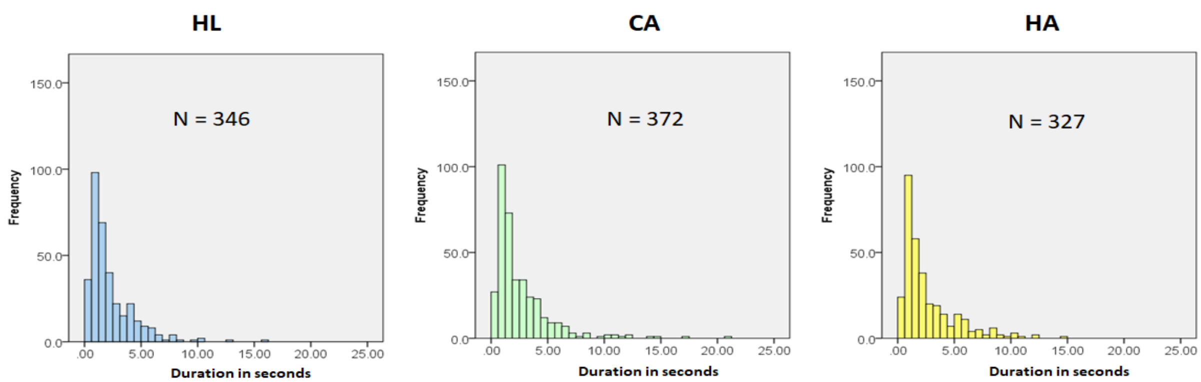

Fig. 4. *Histogram of coordinated attention episodes across groups.*

Different Pathways to Coordinated Attention

The next few sets of analyses aim at clearly identifying the pathways through which the follower used to join the leader to establish a coordinated attention episode. To do so, we excluded ambiguous cases in which the follower's hand was already in contact with the target object during the leading moments[2]. Those cases were ambiguous in our analyses because the followers might look at the object of coordinated attention by following their *own hand action*, but not through

[2] To examine whether removing those episodes affected the general conclusions, we ran the same analyses on all coordinated attention episodes that included the follower's hand actions. The results were qualitatively the same and did not affect the overall pattern of findings.

following gaze or hand cues provided by the leader of coordinated attention. After removing those episodes (423 in total), there were 622 episodes (397 child-led and 225 parent-led) included in the following analyses. We examined whether there were between- and within-group differences in the proportions of using gaze following and hand following pathways. It is noteworthy that coordinated attention episodes can sometimes be established when neither gaze nor hand cues are available (see CAtt#2 in Fig. 2). This type of episodes accounts for less than 15% of the data. Since this type of episodes is not the main focus of our study and the proportions do not differ across groups, we will only report group differences in gaze and hand following pathways in the following analyses. In the next two sections, we will first report parents' use of gaze and hand following pathways to join children to create child-led coordinated attention and then examine the pathways children used to join parents to create parent-led coordinated attention.

Gaze and Hand Following Pathways Parents Used to Establish Child-Led Coordinated Attention

Of all the child-led coordinated attention episodes, parents in all three groups used the gaze following pathway over 60% of the time (i.e., numerator: total number of children's leading moments with parent's look to child's face; denominator: total number of child-led episodes). As can be seen in Fig. 5A, there was no group difference in how often they used the gaze following pathway (HL: 0.66, CA: 0.63, HA: 0.70; Wald χ^2 = 0.60, p = .74). Fig. 5B shows that parents followed their children's hand actions much less frequently as they used the gaze following pathway and there was no group difference in their hand following (HL: 0.26, CA: 0.23, HA: 0.20; Wald χ^2 = 0.81, p = .67). We next asked whether parents preferred the gaze cue given the presence of the hand cue at the same time. To answer this question, we calculated the proportion of parents' gaze following given children's hand cues available. When children's hands were touching the target object during the leading moments, parents still checked children's face over 60% of the time. Again, there was no significant group difference (HL: 0.67, CA: 0.63, HA: 0.71; Wald χ^2 = 0.596, p = .742). Within-group analyses suggest that all three groups of parents preferred the gaze following pathway over the hand following pathway (HL: Wald χ^2 = 11.37, p = .001; CA: Wald χ^2 = 4.11, p = .043; HA: Wald χ^2 = 22.79, p < .001).

To determine whether parents' face looking during the leading moments were different from their overall face looking patterns, we used the proportion of face looks outside of all coordinated attention moments and leading moments as the baseline (HL: 0.39; CA: 0.35; HA: 0.42) and compared the proportion of face looks within the leading moments (i.e., proportions of child-led episodes established through gaze following) against the baseline[3]. All three groups of parents had a significant higher proportion of face looks right before they followed children's attention than their own baseline (HL: 0.66, Wald χ^2 = 22.58, p < .001; CA: 0.63, Wald χ^2 = 17.94, p < .001; HA: 0.70, Wald χ^2 = 38.76, p < .001).

Together, these results suggest that the main cue parents used to follow children's attention was children's gaze direction, regardless of whether or not the children's hand cues were available. Parents' higher proportions of face looks right before they followed suggest that their face looking served a functional purpose to lead parents to join children's attention.

[3] As coordinated attention episodes can include brief looks elsewhere which were shorter than 300 ms, parents sometime had a quick look to the children's face during the coordinated attention episodes. To provide a complete picture of the proportions of face looks, we also calculated the proportions of coordinated attention episodes which contained of parents' brief looks to children's face *during* the coordinated attention moments. The mean proportion of face looks was 0.35 for the HL group, 0.23 for the CA group, and 0.28 for the HA group.

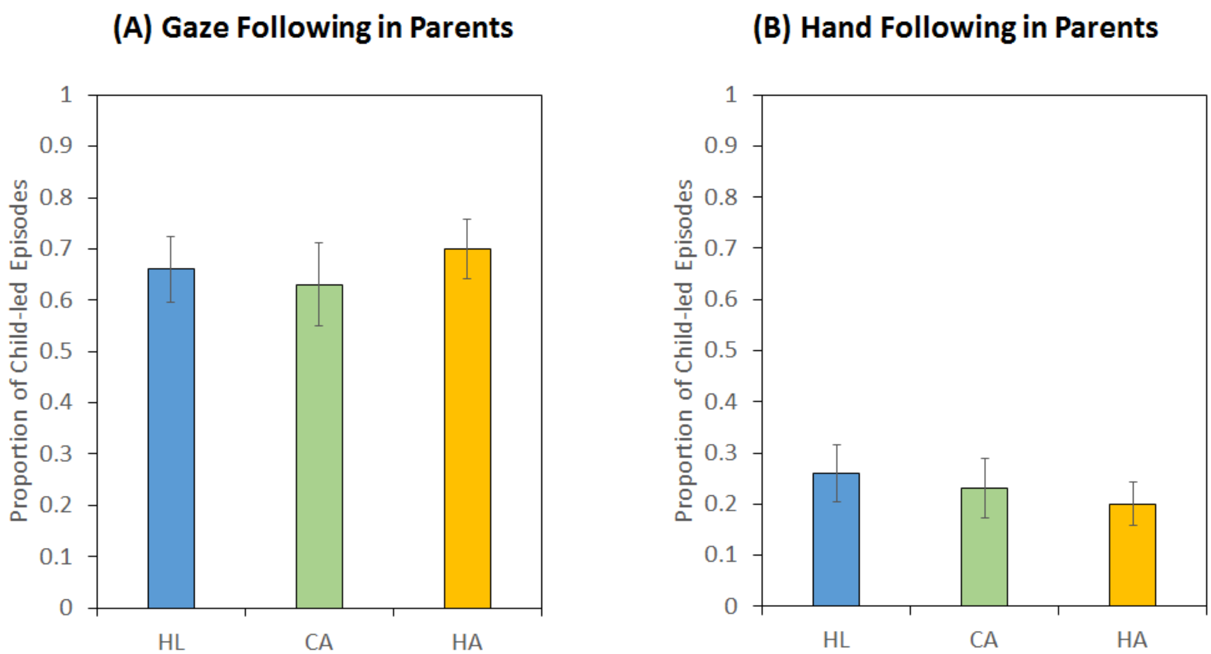

Fig. 5. *Pathways parents used. (A) Proportion (and SE) of parents' gaze following in child-led coordinate attention episodes. (B) Proportion and (SE) of parents' hand following in child-led coordinate attention episodes.*

Gaze and Hand Following Pathways Children Used to Establish Parent-Led Coordinated Attention

As can be seen in Fig. 6A, children in the HL group followed parents' gaze directions almost half of the time while the CA and HA children rarely did so (numerator: total number of parents' leading moments with child's look to parent's face; denominator: total number of parent-led episodes; HL: 0.47, CA: 0.19, HA: 0.21; Group difference: Wald $\chi^2 = 12.76$, $p = .002$). Pairwise comparisons suggest that the HL group were more likely to use gaze following than their CA and HA peers while the two hearing groups had comparable performance (HL > CA, $p = .001$; HL > HA, $p = .004$; CA = HA, $p = .77$). In contrast, the CA and HA groups were more likely to follow parents' hand actions than the HL group (HL: 0.37, CA: 0.62, HA: 0.53; Group difference: Wald

$\chi^2 = 12.35$, $p = .002$; HL < CA, $p < .001$; HL < HA, $p = .035$; CA = HA, $p = .28$; shown in Fig. 6B). We next calculated the proportions of children's face look when parents' hand cues were available. When parents' hands were touching the target object of the following coordinated attention episode, children in the HL group were significantly more likely to look at parent's face right before joining the parent compared to the CA and HA groups (HL: 0.49, CA: 0.15, HA: 0.27; Group difference: Wald $\chi^2 = 19.01$, $p < .001$; HL > CA, $p < .001$; HL > HA, $p = .024$). Within-group analyses indicate that the HL group used the gaze and hand following pathways equally frequently (Wald $\chi^2 = 0.97$, $p = .33$). In contrast, the CA and HA children both preferred the hand following pathway over the gaze following pathway (CA: Wald $\chi^2 = 15.64$, $p < .002$; HA: Wald $\chi^2 = 7.12$, $p = .008$).

We again calculated each groups' proportion of face looks outside of all coordinated attention moments and leading moments and then used it as the baseline (HL: 0.24; CA: 0.21; HA: 0.20) to compare with each group's proportion of face looks right before following[4]. Compared to their baseline, the HL group had a significantly higher proportion of face looks right before they followed parents' attention (HL: 0.47, Wald $\chi^2 = 13.38$, $p < .001$). In contrast, the proportions of the CA and HA groups' face looking before following parents' attention were not different from their own baseline (CA: 0.19, Wald $\chi^2 = 0.11$, $p = .75$; HA: 0.21, Wald $\chi^2 = 0.08$, $p = .78$). The average duration of parents' leading moments was similar across groups (HL: mean = 0.78 s, *SD* = 0.19; CA: mean = 0.87 s, *SD* = 0.25; HA: mean = 0.77 s, *SD* = 0.17, Wald $\chi^2 = 0.88$, $p = .65$).

[4] To provide a complete picture of children's proportions of face looks, we also calculated proportions of coordinated attention episodes which contained children's brief looks to parents' face *during* the coordinated attention moments. The mean proportion of face looks was 0.22 for the HL group, 0.16 for the CA group, and 0.11 for the HA group.

Therefore, the HL group's higher proportion of face looking before following was not due to longer leading moments and thus having more time to look at parents' face.

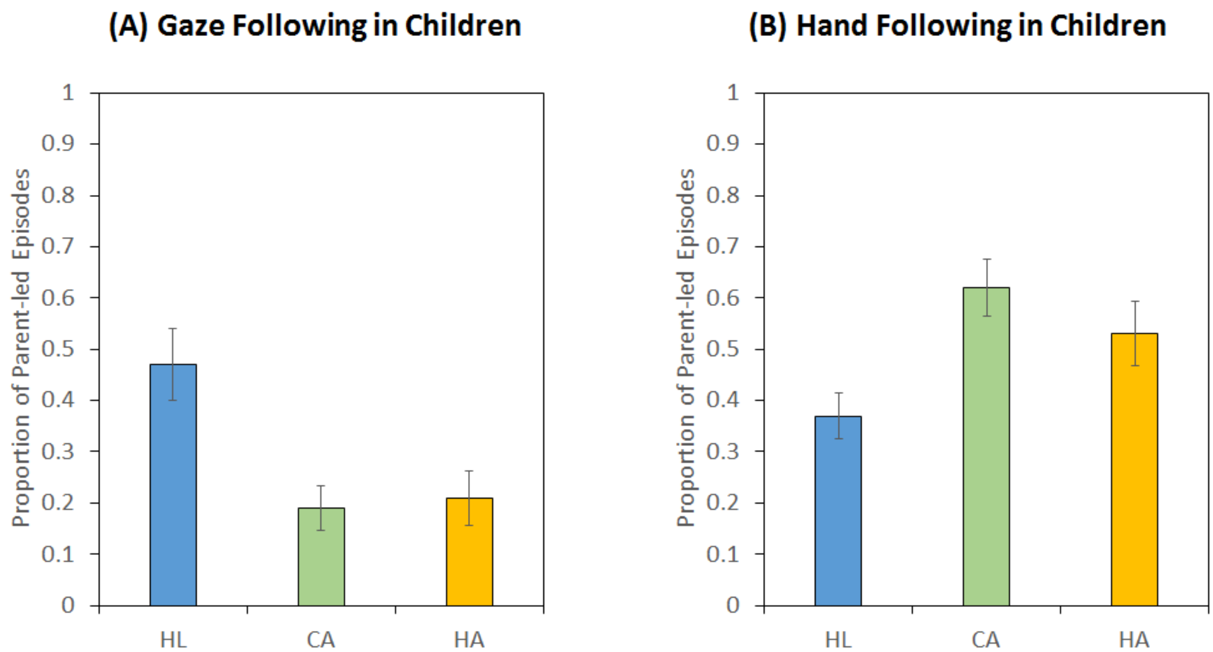

Fig. 6. *Pathways children used. (A) Proportion (and SE) of children's gaze following in parent-led coordinate attention episodes. (B) Proportion and (SE) of children's hand following in parent-led coordinate attention episodes.*

To illustrate the temporal dynamics of children's face looking patterns right before coordinated attention, we used a data analysis method that has been invented and widely used in psycholinguistic research (Allopenna, Magnuson, & Tanenhaus, 1998, Yu & Smith, 2013). Using this approach, we aligned the onsets of all parent-led coordinated attention episodes and calculated at each moment (i.e., every 33 ms) how likely children looked at the parent's face at that particular moment in time right before coordinated attention (i.e., numerator: pre-parent-led coordinated attention moments contained child's look to parent's face; denominator: total number of parent-led episodes). Fig. 7 shows the temporal profiles of children's face looking patterns from 3 seconds prior to the onsets of coordinated attention episodes. As mentioned previously, the mean durations of parents' leading moments lasted less than 1 second for all 3 groups (with 95% of all leading moments < 2.12 seconds). This 3-second window covered the change of face looking from the baseline to the onsets of coordinated attention episodes. As can be seen in Fig. 7, for the HL children, the proportion of parent-led episodes that contained children's face look increased at around 2 seconds prior to the onset of coordinated attention. Then their proportion of face looks dropped right before the onset of coordinated attention, as they switched from looking at parent's face to the target object of the coordinated attention episode. In contrast, the CA and HA children's face looking patterns remained relatively stable within the 3-second window (though they also dropped right before they joined parents to look at the same object). These patterns again confirm that the HL children increased their looks to their parent's face prior to coordinated attention while their CA and HA peers looked at parents' face less frequently before following parents' attention.

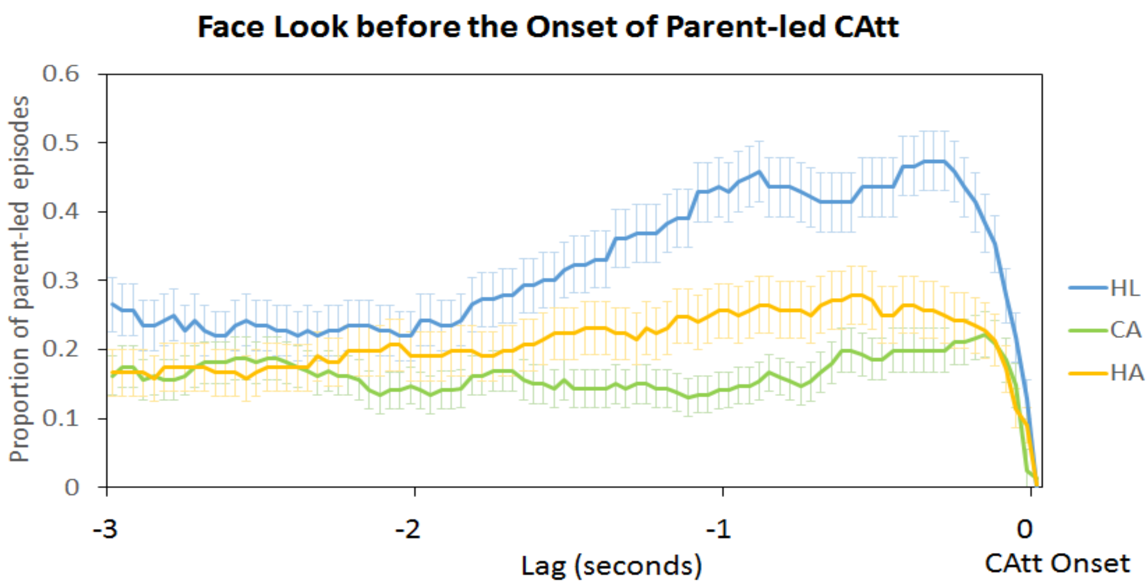

Fig. 7. *Temporal dynamics of children's face looking behaviors prior to parent-led coordinated attention episodes. Each data point represent data from one frame prior to the onset (i.e., 33 ms, because the sampling rates of the cameras was 30 Hz). Error bars represent standard errors.*

In sum, these results suggest that children in the HL group were more likely to follow parents' gaze directions in achieving coordinated attention than children in the CA and HA groups. Their increase of face looking right before following was similar to the patterns seen in parents. In contrast, children in the CA and HA groups mainly used the hand following pathway as reported previously (Yu & Smith, 2017a).

Discussion

In the current study, we investigated the effect of toddlers' hearing loss on how they coordinated attention with their parents. We found that parents in all three groups had very similar looking patterns. Toddlers' hearing status did not affect their overall gaze lengths, proportions of face looks (versus object looks), and how likely parents and children both attended to the same object at the same time during play. Children with hearing loss used gaze and hand following pathways equally frequently. In contrast, children with normal hearing used hand following pathway more frequently than the gaze following pathway. These results suggest that children's hearing status affects the pathways they use to achieve coordinated attention with their parents.

Multiple Pathways to Achieve Coordinated Attention

According to the dynamic systems theory, there are usually multiple pathways to the same functional end (Thelen, 1992). Differences in the initial conditions and the component(s) of a developmental system can cause different behaviors and re-organization of the system (Karmiloff-Smith, 1998; Smith & Thelen, 2003; Thelen, 1992). Children in the HL group spent their first few months to over a year of life with degraded or no auditory signal. Even after cochlear implantation or fitting with hearing aids, the signals they receive are still often suboptimal and can affect not only their sound processing but also their language learning (Bergeson, Houston, & Miyamoto, 2010; Houston, Stewart, Moberly, Hollich, & Miyamoto, 2012). The (lack of) auditory experiences they have from early life likely creates different patterns and trajectories when they interact with their hearing parents, who mostly rely on auditory/verbal mode of communication (Bortfeld & Oghalai, 2019; Depowski et al., 2015). These children did not only have a different start point from their hearing peers, they also have a different history in interacting with their parents which may have cascading effects on their social interaction with their parents.

In studying parent-child social interaction, Yu and Smith (2017a) suggest that coordinated attention in free-flowing interaction can be achieved through multiple pathways in a coupled dynamic system between parents and children. Consistent with their findings, our study also showed multiple pathways that parents and toddlers employed to achieve coordinated attention, which was true both *within* and *between* groups. Within each group, both parents and toddlers used gaze and hand following pathways, even though there were differences in how often parents and children used different pathways. This ability to use different pathways suggests parents and children's flexibility of using different solutions to achieve the same goal. Between groups, children's hearing status affected how often they relied on parents' gaze information. Children with normal hearing rarely followed parents' gaze direction during joint play, which was consistent with the findings from previous studies (Chang et al., 2016; Deak et al., 2018; Deak et al., 2014; de Barbaro et al., 2016; Franchak et al., 2011; Yu & Smith, 2013, 2017a, 2017b). Compared to their hearing peers, children in the HL group were significantly more likely to actively look up and check their parents' face before they followed. They did so even when parents' hand cues were available. Previous studies showed that children with hearing loss look at parents' face to seek additional visual information to better understand parents' speech (Bergeson, Pisoni, & Davis, 2005; Summerfield, 1992). One positive side effect of looking at parents' face is that it would give them information about parents' head orientation. Since gaze direction tends to align with head orientation (Bambach, Smith, Crandall, & Yu, 2016; Yoshida & Smith, 2008), this face looking pattern also potentially gives them chances to follow parents' head, and therefore, gaze directions. By doing so, looks to their parents' face create multimodal effects on facilitating parent-child social interaction.

Overall, the HL children used gaze and hand following pathways equally frequently. This suggests that they used gaze following in combination with hand following – a preferred pathway by the hearing children -- to coordinate attention with their parents. This also indicates their ability to use different cues (i.e., gaze and hand) to rapidly adjust their behaviors and follow in to create coordinated attention moments with their parents. These results are in line with the proposal that differences in initial conditions or components of the developing system can lead to differences in behaviors (Smith & Thelen, 2003; Thelen, 1992). Future research is needed to investigate whether these behavioral differences have any cascading effects on children's learning and long-term social, language, and cognitive development. For example, does hand following or gaze following lead to better learning of words or object concepts? Or are these different pathways equally good as long as coordinated attention is established? It is also important to investigate whether the flexibility demonstrated here can be leveraged to help children with hearing loss overcome the challenges they face in early social interaction and early language learning.

Our study suggests that even though children with different hearing statuses tend to use different pathways, children with hearing loss were still able to coordinate visual attention with their parents. More importantly, they did so as frequently as their peers with normal hearing. The between-group similarity in how likely parents and children achieved coordinated attention and group differences in gaze and hand following indicate that there is not one optimal solution to the same end (Yu & Smith, 2017a). Instead, depending on children's sensory system and the dynamics in parent-child interactions, different dyads were able to adjust to the specific contexts of the learning environment and create their own pathways to achieve the same goal. These findings also suggest that, despite the constraints of hearing loss, the robust flexibility in the developmental systems allowed children with hearing loss and their parents to achieve fairly fluid coordination of visual attention in joint object play.

Another interesting finding is the similarities in how children in the three groups distributed their visual attention in free-flowing toy play. Even though the mean age of the HA children was nearly half of the CA and HL children (16.1 vs. 30.0 & 30.9 months) and the HL group and the two hearing groups differed in hearing status, there was no group difference in their overall ROI lengths, proportions of looks to objects in play versus parent's face, or measures related to coordinated attention. Along with previous studies using similar measures, the results suggest that young children's ability to control their visual attention during object play and to coordinate their attention with their parents develop at an early age (Yu & Smith, 2013, 2017a, 2017b). Prior research using screen-based visual selection tasks suggests that children with hearing loss have more distributed attention than their hearing peers (Dye & Hauser, 2014; Quittner, Smith, Osberger, Mitchell, Katz, 1994; Smith, Quittner, Osberger, & Miyamoto, 1998). Our study suggest that HL children's overall ROI lengths and proportions of object looks versus face looks in dynamic parent-child object play were very similar to their hearing peers (see also Chen, Castellanos, Yu, & Houston, 2019b). One future direction is to examine whether children's hearing status also affects their visual selection in naturalistic parent-child interaction contexts.

It is also important to note that, in all three groups, over half of the coordinated attention episodes were led by children. Parents play an active role in joint play by detecting and following their children's attentional states. Parents in different groups showed very similar patterns when they followed in. Some prior studies have shown that parents of children with hearing loss produced more directives in speech compared to parents of age-matched hearing children (e.g., Chen et al., 2019a; Fagan, Bergeson, & Morris, 2014). Our study suggests that, like parents of hearing children, parents of children with hearing loss were also sensitive to their children's attentional states and were able to successfully use children's gaze directions (as well as hand actions) to follow in and create coordinated attention moments with their children (Tamis-LeMonda, Kuchirko, & Song, 2014). Parent-child interaction is reciprocal as the two social partners work together to create smooth interactions.

Take Timing into Account

Analyzing real-time data from different behavioral streams (gaze & hand contact) allowed us to study the temporal unfolding of behavioral patterns. Unlike previous observational studies showing higher overall face looking in children with hearing loss (Lieberman et al., 2014; Prezbindowski et al., 1998), using eye-tracking measures, our study showed that children across groups had very similar overall gaze patterns. Fewer than 30% of their ROI looks were face looks. This suggests that children in the HL group were *not* more likely to look at parents' face in general. For children in the CA and HA groups, their proportion of face looks did not change before they followed parents in the coordinated attention episodes. However, high-resolution gaze data collected in the present study allowed us to zoom into and examine micro-level behaviors with high temporal resolution. By doing that, we found that the HL group had a higher proportion of face looks *right before following*. This finding was very similar to the patterns seen in parents. Their face looking pattern likely serves a functional purpose to lead to coordinated attention. These results highlight the dynamics of parent-child interactions and underscore the importance of using high-density data to examine how naturalistic behaviors unfold over time as a useful way to understand the mechanisms supporting free-flowing interaction (de Barbaro et al., 2016; Deak et al., 2018; Smith & Thelen, 2003; Thelen, 1992; Yu & Smith, 2013, 2017a, 2017b).

Limitations and Future Directions

It is noteworthy that the present study focused on one specific measure of manual action – which object was in either the child's or parent's hand(s) -- which didn't contain any information about hand movements. This is because in free-flowing toy play, the object held by one's hand(s) at a moment usually aligns well with their focus of attention and therefore provides a reliable and complementary (in addition to gaze) measure of one's attentional focus through manual action (Bambach et al., 2016; Yoshida & Smith, 2008). Previous studies have shown that moving an object while talking about it at the same time can increase the saliency of the object and facilitate young children's learning of object names (Gogate, Bolzani, & Betancourt, 2006; Gogate, Maganti, Bahrick, 2015; Lund & Schuele, 2015). It is very likely that moving an object attracts children's attention more than simply holding an object still and makes it easier for young children to follow in and look at the object of parents' focus of attention. Moreover, parents' speech about objects also extends children's attention to objects (Chen et al., 2019a; Suarez-Rivera et al., 2019). One future direction would be to study how hand movement and speech contribute to the establishment and maintenance of coordinated attention in parent-child interactions.

Another future direction is to take individual differences into account and examine how real-time coordination of attention within-dyads affects and/or is affected by children's long-term language, cognitive, and social development. For example, for children who are faster in coordinating visual attention with their parents, do they have better social skills and language outcomes? Do children who show more flexibility in using different pathways have better cognitive control? Another question is whether children's real-time speech processing abilities influence or are influenced by their face looking or gaze following behaviors. These findings will not only further our understanding of children's general language, cognitive, and social development, but also shed light on the development of intervention programs for clinical populations.

Some previous studies focusing on intentionality or mutual awareness aspects of joint attention have shown that, compared to their hearing peers, children with hearing loss were less successful in initiating and responding to their parents' attentional bids, and therefore, less likely to achieve joint attentional states with their parents (Bortfeld & Oghalai, 2019; Depowski et al., 2015). Using a leaner definition and focusing on the micro-level gaze data, our study showed that children with hearing loss had similar coordinated attention patterns as their hearing peers. However, children's hearing status affected their looking patterns prior to achieving coordinated attention with their parents. The differences in these findings are likely due to different definitions and different experimental paradigms used to examine different aspects of parent-child social interactions. Therefore, they are not necessarily contradictory in terms of empirical evidence. However, these differences in findings likely result in different conclusions and have different implications. For example, one possible implication from studies focusing on the intentionality or awareness components may be to develop intervention programs that train children with hearing loss to better detect parents' communication intention and to respond to their attentional bids. On the other hand, based on the results of our study, the next step for future research is to investigate whether using different pathways to achieve coordinated attention affects later development. To address the discrepancies of findings and interpretations, one direction for future work is to unify micro- and macro-level understandings of parent-child interaction. For example, one way to do so is to code parent-child interactions from the same dataset using both macro-level analyses that incorporate social components, such as intentionality and/or mutual awareness, and micro-level analyses that focus on sensorimotor coding, and then investigate the similarities and differences of results generated by these two types of analyses on the same dataset (see also Gabouer, Oghalai & Bortfeld, 2018). The findings will provide insights for the development of theories on coordinated attention and children's social development.

Conclusion

Children's hearing status does not affect how likely they coordinate attention with their parents. However, when following parents' attention, children with hearing loss relied on both parents' gaze directions and hand actions to guide their attention, while children with normal hearing mainly relied on parents' hand actions. Our study highlights the dynamics in daily social interactions and the flexibility and robustness of the developing systems.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, *38*(4), 419-439.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*(4), 1278-1289.

Bambach, S., Smith, L. B., Crandall, D. J., & Yu, C. (2016). Objects in the center: How the infant's body constrains infant scenes. In *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics* (pp. 132-137). IEEE.

Bergeson, T. R., Houston, D. M., & Miyamoto, R. T. (2010). Effects of congenital hearing loss and cochlear implantation on audiovisual speech perception in infants and children. *Restorative Neurology and Neuroscience*, *28*(2), 157-165.

Bergeson, T. R., Pisoni, D. B., & Davis, R. A. (2005). Development of audiovisual comprehension skills in prelingually deaf children with cochlear implants. *Ear and Hearing*, *26*(2), 149-164.

Bortfeld, H., & Oghalai, J. (2018). Joint Attention in Hearing Parent-Deaf Child and Hearing Parent-Hearing Child Dyads. *IEEE Transactions on Cognitive and Developmental Systems*. DOI: 10.1109/TCDS.2018.2877658.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535-543.

Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, *3*(3), 253-272.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*, i-174.

Chang, L., de Barbaro, K., & Deák, G. (2016). Contingencies between infants' gaze, vocal, and manual actions and mothers' object-naming: Longitudinal changes from 4 to 9 months. *Developmental Neuropsychology*, *41*(5-8), 342-361.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive development*, *15*(4), 481-498.

Chen, C., Castellanos, I., Yu, C., & Houston, D. M. (2019a). Parental linguistic input and its relation to toddlers' visual attention in joint object play: A comparison between children with normal hearing and children with hearing loss. *Infancy*, *24*(4), 589-612.

Chen, C. H., Castellanos, I., Yu, C., & Houston, D. M. (2019b). Effects of children's hearing loss on the synchrony between parents' object naming and children's attention. *Infant Behavior and Development*, *57*, 101322.

de Barbaro, K., Johnson, C. M., Forster, D., & Deák, G. O. (2016). Sensorimotor decoupling contributes to triadic attention: A longitudinal investigation of mother–infant–object interactions. *Child Development*, *87*(2), 494-512.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441-447.

de Villiers Rader, N., & Zukow-Goldring, P. (2012). Caregivers' gestures direct infant attention during early word learning: the importance of dynamic synchrony. *Language Sciences*, *34*(5), 559-568.

Deák, G. O., Krasno, A. M., Jasso, H., & Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses during object play. *Infancy*, *23*(1), 4-28.

Deak, G. O., Krasno, A. M., Triesch, J., Lewis, J., & Sepeta, L. (2014). Watch the hands: infants can learn to follow gaze by seeing adults manipulate objects. *Developmental Science*, *17*(2), 270-281.

Depowski, N., Abaya, H., Oghalai, J., & Bortfeld, H. (2015). Modality use in joint attention between hearing parents and deaf children. *Frontiers in Psychology*, *6*, 1556.

Doherty, M. J., Anderson, J. R., & Howieson, L. (2009). The rapid development of explicit gaze judgment ability at 3 years. *Journal of Experimental Child Psychology*, *104*(3), 296-312.

Dye, M. W., & Hauser, P. C. (2014). Sustained attention, selective attention and cognitive control in deaf and hearing children. *Hearing Research*, *309*, 94-102.

Fagan, M. K., Bergeson, T. R.,& Morris, K. J. (2014). Synchrony, complexity and directiveness in mothers' interactions with infants pre-and post-cochlear implantation. *Infant Behavior and Development*, *37*(3), 249–257.

Flom, R., & Pick, A. D. (2005). Experimenter affective expression and gaze following in 7-month-olds. *Infancy*, *7*(2), 207-218.

Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-mounted eye tracking: A new method to describe infant looking. *Child Development*, *82*(6), 1738-1750.

Gabouer, A., Oghalai, J., & Bortfeld, H. (2018). Hearing Parents' Use of Auditory, Visual, and Tactile Cues as a Function of Child Hearing Status. *International Journal of Comparative Psychology*, 31. Retrieved from https://escholarship.org/uc/item/98q9n3dc

Gogate, L. J., Bolzani, L. H., & Betancourt, E. A. (2006). Attention to maternal multimodal naming by 6-to 8-month-old infants and learning of word–object relations. *Infancy*, *9*(3), 259-288.

Gogate, L., Maganti, M., & Bahrick, L. E. (2015). Cross-cultural evidence for multimodal motherese: Asian Indian mothers' adaptive use of synchronous words and gestures. *Journal of experimental child psychology*, *129*, 110-126.

Gredeback, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science*, *13*(6), 839-848.

Gredeback, G., Theuring, C., Hauf, P., & Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy*, *13*(5), 533-543.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114-121.

Houston, D. M., Stewart, J., Moberly, A., Hollich, G., & Miyamoto, R. T. (2012). Word learning in deaf children with cochlear implants: Effects of early auditory experience. *Developmental Science*, *15*(3), 448-461.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Science*, *2*(10), 389-398.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

Landry, S. H., Smith, K. E., Swank, P. R., & Miller-Loncar, C. L. (2000). Early maternal and child influences on children's later independent cognitive and social functioning. *Child Development*, *71*(2), 358-375.

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13-22.

Lieberman, A. M., Hatrak, M., & Mayberry, R. I. (2014). Learning to look for language: Development of joint attention in young deaf children. *Language Learning and Development*, *10*(1), 19-35.

Lund, E., & Schuele, C. M. (2015). Synchrony of maternal auditory and visual cues about unknown words to children with and without cochlear implants. *Ear and hearing*, *36*(2), 229-238.

Moll, H., & Meltzoff, A. N. (2011). Perspective-taking and its foundation in joint attention. In J. Roessler, H. Lerman, & N. Eilan, (Eds.), *Perception, Causation, and Objectivity* (pp. 286–304). Oxford, England: Oxford University Press.

Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6-through 24-month age period and early language acquisition. *Journal of applied developmental psychology*, *21*(3), 283-298.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*(3), 938-954.

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development*, *21*(3), 469-482.

Niedźwiecka, A., Ramotowska, S., & Tomalski, P. (2018). Mutual gaze during early mother–infant interactions promotes attention control development. *Child Development*, *89*(6), 2230-2244.

Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, *108*(52), 20947-20952.

Piazza, E. A., Hasenfratz, L., Hasson, U., & Lew-Williams, C. (2018). Infant and adult brains are coupled to the dynamics of natural communication. *bioRxiv*, 359810.

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.

Quittner, A. L., Smith, L. B., Osberger, M. J., Mitchell, T. V., & Katz, D. B. (1994). The impact of audition on the development of visual attention. *Psychological Science*, *5*(6), 347-353.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*(3), 403-452.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867-891.Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265-266.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265-266.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305-319.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326-332.

Siposova, B., & Carpenter, M. (2019). A new look at joint attention and common knowledge. *Cognition*, *189*, 260-274.

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology*, *164*(1), 54-71.

Smith, L. B., Quittner, A. L., Osberger, M. J., & Miyamoto, R. (1998). Audition and visual attention: the developmental trajectory in deaf and hearing populations. *Developmental Psychology*, *34*(5), 840-850.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343-348.

Suarez-Rivera, C., Smith, L. B., & Yu, C. (2019). Multimodal parent behaviors within joint attention support sustained attention in infants. *Developmental Psychology*, *55*(1), 96-109.

Summerfield, Q. (1992). Lipreading and audio-visual speech perception. *Philosophical Transactions of the Royal Society B*, *335*(1273), 71-78.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*(2), 121-126.

Tasker, S. L., Nowakowski, M. E., & Schmidt, L. A. (2010). Joint attention and social competence in deaf children with cochlear implants. *Journal of Developmental and Physical Disabilities*, *22*(5), 509-532.

Tasker, S. L., & Schmidt, L. A. (2008). The “dual usage problem” in the explanations of “joint attention” and children’s socioemotional development: A reconceptualization. *Developmental Review*, *28*(3), 263-288.

Thelen, E. (1992). Development as a dynamic system. *Current Directions in Psychological Science*, *1*(6), 189-193.

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

Van Hecke, A. V., Mundy, P., Block, J. J., Delgado, C. E., Parlade, M. V., Pomares, Y. B., & Hobson, J. A. (2012). Infant responding to joint attention, executive processes, and self-regulation in preschool children. *Infant Behavior and Development*, *35*(2), 303-311.

Wass, S. V., Noreika, V., Georgieva, S., Clackson, K., Brightman, L., Nutbrown, R., Covarrubias, L. S., & Leong, V. (2018). Parental neural responsivity to infants’ visual attention: how mature brains influence immature brains during social interaction. *PLoS Biology*, *16*(12), e2006328.

Yoshida, H., & Smith, L. B. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy*, *13*(3), 229-248.

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. (2017a). Multiple sensory-motor pathways lead to coordinated visual attention. *Cognitive Science*, *41*, 5-31.

Yu, C., & Smith, L. B. (2017b). Hand-eye coordination predicts joint attention. *Child Development*, *88*(6), 2060-2078.

Yu, C., Suanda, S. H., & Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental Science*, e12735.