INFANCY

THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES

Infancy, **24**(4), 589–612, 2019

ISSN: 1532-7078 online
DOI: 10.1111/infa.12291

WILEY Blackwell

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

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

Irina Castellanos
Department of Otolaryngology–Head & Neck Surgery
The Ohio State University and
Nationwide Children's Hospital

Chen Yu
Department of Psychological and Brain Sciences
Indiana University

Derek M. Houston
Department of Otolaryngology–Head & Neck Surgery
The Ohio State University and
Nationwide Children's Hospital

Parent–child interactions are multimodal, often involving coordinated exchanges of visual and auditory information between the two partners. The current work focuses on the effect of children's hearing loss on parent–child interactions when parents and their toddlers jointly played with a set of toy objects. We compared the linguistic input received by toddlers with hearing loss (HL) and their chronological age-matched (CA) and hearing age-matched (HA) normal-hearing peers. Moreover, we used head-mounted eye trackers to examine how different parental linguistic input affected children's visual attention on objects when parents either led or followed children's attention during joint object play. Overall, parents of children with HL provided comparable *amount* of linguistic input as parents of the two normal-hearing groups. However, the *types* of linguistic input produced by parents of children with HL were similar to the CA group in some ways and similar to the HA group in other ways. Interestingly, the

Correspondence should be sent to Chi-hsin Chen, Department of Otolaryngology–Head & Neck Surgery, The Ohio State University, 915 Olentangy River Rd, Columbus, OH 43212. E-mail: Chi-Hsin.Chen@osumc.edu

effects of different types of linguistic input on extending the attention of children with HL qualitatively resembled the patterns seen in the CA group, even though the effects were less pronounced in the HL group. We discuss the implications of these results for our understanding of the reciprocal, dynamic, and multi-factored nature of parent–child interactions.

Infants and toddlers learn about the world mostly through social interactions, particularly through interactions with their parents. What parents say or do affects what children attend to, process, and learn (e.g., Tamis-LeMonda, Kuchirko, & Song, 2014; Tomasello & Farrar, 1986; Tomasello & Todd, 1983; Yu & Smith, 2012). Parent–child interactions are multimodal, often involving coordinated exchanges of auditory and visual information between the two partners (Bornstein, Tamis-LeMonda, Hahn, & Haynes, 2008; Frank, Tenenbaum, & Fernald, 2013; Gogate, Bahrick, & Watson, 2000; Gogate, Bolzani, & Betancourt, 2006; Harris, Jones, & Grant, 1983; Karasik, Tamis-LeMonda, & Adolph, 2014; Messer, 1978; Suanda, Smith, & Yu, 2016; West & Iverson, 2017; Yu & Smith, 2017a, 2017b). These coordinated, multimodal interactions are important for early cognitive, social, and language development (Charman et al., 2000; Tamis-LeMonda, Song, Leavell, Kahana-Kalman, & Yoshikawa, 2012; Yu & Smith, 2012). But what happens during development if signals from one modality are degraded or lost? In this study, we focus on the role of sensory systems in coordinated interactions by studying young children with hearing loss and compare them to normal-hearing children. Specifically, we examine the effects of children's hearing experience on the linguistic input parents provide. Furthermore, we investigate how parents' linguistic input, in turn, affects children's attention to objects—which is often viewed as crucial to the learning of object concepts and object names (e.g., Baldwin & Markman, 1989; Rohde & Frank, 2014; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Suanda et al., 2016; Tomasello, 1988; Tomasello & Farrar, 1986; Yu & Smith, 2012).

Parental linguistic input to children with and without hearing loss

Previous studies have shown that hearing parents of children with hearing loss (HL) interact with their children differently compared to hearing parents with their normal-hearing children. For example, the interactions between hearing mothers and their children with HL tend to be less synchronized, in that mothers' utterances are more likely to overlap with children's vocalizations (Fagan, Bergeson, & Morris, 2014). Another major interactional difference lies in the linguistic input parents provide. Hearing mothers of children with HL tend to use simpler language and more directives or verbal controls than hearing mothers of normal-hearing children (Cross, Johnson-Morris, & Nienhuys, 1980; Fagan et al., 2014; Henggeler, Watson, & Cooper, 1984). These patterns lead to the following questions: (1) Why are there differences in parental linguistic input for children with and without HL; and (2) what are potential consequences of these input differences?

One possible reason for the differences in linguistic input is related to children's language ability, or at least parents' perception of children's ability. Parents tend to talk to children in accord with their language experience or ability (DesJardin et al., 2014; Fagan et al., 2014; Pellegrini, Brody, & Sigel, 1985; Snow, 1972). For example, parents of children with lower language skills or less language experience are more likely to take a leading role in conversation and use more directives and simpler language (e.g., Bornstein et al., 2008; Cruz, Quittner, Marker, & DesJardin, 2013; DesJardin et al., 2014; Fagan et al., 2014; Musselman & Churchill, 1992; Snow, 1972). These studies suggest that hearing parents' use of more directives and simpler language with their children with HL is likely a way of adjustment to their children's language ability or experience.

Differences in linguistic input can have potential consequences on children's language processing skills and their development of both receptive and expressive vocabularies. The more the utterances, word tokens, and different word types parents provide, the larger the children's receptive and expressive vocabularies tend to be (e.g., Hurtado, Marchman, & Fernald, 2008; Rowe, 2008, 2012; Weisleder & Fernald, 2013). In addition to the amount of input, studies with both children with and without HL have shown that the ways in which parents talk to their children also affect children's language development (Cruz et al., 2013; DesJardin & Eisenberg, 2007; Girolametto, Weitzman, Wiigs, & Pearce, 1999; Tomasello, 1988; Tomasello & Todd, 1983; Yoder, Kaiser, Alpert, & Fischer, 1993). For example, parents' use of advanced facilitative language techniques (e.g., open-ended questions or expansion on children's vocalization) is positively associated with children's expressive language development (Cruz et al., 2013; DesJardin & Eisenberg, 2007; Girolametto et al., 1999). In contrast, parents' naming and re-directing their child to an object different from the child's focus of attention is negatively associated with children's learning of object names (Tomasello, 1988; Tomasello & Todd, 1983). These studies have established a relationship between parents' language use and children's language outcomes. However, far less is known about the mechanism(s) by which different parental language styles affect children's word learning.

Effects of linguistic input on visual attention in children with and without hearing loss

One possible pathway by which parental language use may affect children's word learning is through real-time attentional mechanisms. Early word learning, particularly learning object names, requires young children to attend to the target object when parents name it. The role of attention in language acquisition is supported by studies showing that children's sustained attention—the stabilization of visual attention to an object for long durations—predicts later language learning and cognitive development (Kannass & Oakes, 2008; Lawson & Ruff, 2004; Ruff, Lawson, Parrinello, & Weissberg, 1990). Furthermore, a growing number of studies indicate that toddlers are more likely to remember a name-object mapping if they visually attend longer to an object when it is named than when visual attention to the named object is brief (MacRoy-Higgins & Montemarano, 2016; Pereira, Smith, & Yu, 2014; Yu & Smith, 2012). Thus, one mechanistic explanation for the effects of parents' language use during joint play on children's word learning is that their language affects children's real-time attention to objects. A recent study by Suarez-Rivera and colleagues (Suarez-Rivera, Smith, & Yu, 2019) showed that parents' speech in joint play sustains the attention of children with normal hearing. The extension effect of parental language input on the visual attention of children with normal hearing, however, may be different from that of children with HL, as previous studies found that children with HL and their normal-hearing peers showed different attentional patterns in visual selection tasks (Dye & Hauser, 2014; Quittner, Smith, Osberger, Mitchell, & Katz, 1994; Smith, Quittner, Osberger, & Miyamoto, 1998). Moreover, because different utterance types serve different functional purposes and elicit different types of responses from listeners (O'Brien & Nagle, 1987; Spekman & Roth, 1985; Tyack & Ingram, 1977), different utterance types may elicit different effects on visual attention and such effects may also vary depending on children's age and hearing status. Thus, the present study will be the first to examine the effects of different parental utterance types on children's real-time attention in free-flowing toy play and compare the effects between children with HL and children with normal hearing.

The current study

We recruited three groups of children, toddlers with hearing loss (HL group), normal-hearing toddlers matched to the HL group in chronological age (CA group), and normal-hearing toddlers matched to the HL group in hearing age (HA group), to participate in free-flowing toy play. We chose toy play because it is a common activity that parents and their toddlers engage in and the linguistic input in parent–child toy play conducted in laboratories mirrors the input observed in peak interactions during naturalistic daily routines (Tamis-LeMonda, Kuchirko, Luo, Escobar, & Bornstein, 2017). With these three groups of children participating in toy play, the present study focused on two goals: one on examining the similarities and differences of linguistic input produced by parents in different groups and the other on examining the effects of parental linguistic input on children's attention within and across groups. More specifically, the first goal was to investigate whether parents in the HL group provided different linguistic input about objects they played with than parents of normal-hearing children in general and whether the differences arose as a function of children's age or hearing experience. To achieve this goal, we investigated quantitative and qualitative properties of object-related utterances to children with HL compared to the CA and HA groups. There are three hypotheses. If parents talk to children in accord with their age, we should expect that the linguistic input received by children with HL would be similar to the input received by CA children. Alternatively, if parents talk to children in ways that are consistent with their hearing experience (e.g., Pellegrini et al., 1985; Snow, 1972), we should expect that the linguistic input the HL group received would be similar to the HA group. It could also be that the patterns seen in the HL group are similar to the CA group in certain ways (e.g., amount of linguistic input, Fagan et al., 2014; VanDam, Ambrose, & Moeller, 2012) but similar to the HA group in other ways (e.g., frequent use of directives, e.g., Fagan et al., 2014). Including both CA and HA matches allowed us to have a more complete picture of the input received by normal-hearing children and to examine whether differences in linguistic input arise as a function of children's hearing status (i.e., hearing vs. deaf), chronological age, or hearing experience.

The second goal was to investigate the effects of different types of parental utterances on children's real-time attention allocation during joint object play. We were interested in whether and how auditory experiences influenced the relationship between parental language input and children's visual attention in joint object play. Some prior studies suggest that children's ability to sustain their attention is associated with the development of their cognitive systems, such as self-regulatory ability (e.g., Colombo, 2001; Ruff & Capozzoli, 2003). In general, the older the children are, the more mature their cognitive systems are. According to this perspective, we should expect that the HL group and CA group show similar attentional patterns. It has also been found that infants' sustained attention to an object is influenced by parents' visual attention to the same object (Yu & Smith, 2016). When parents talk about an object, they most likely also attend to the object. It is thus expected that parents' linguistic input about an object also affects children's attention to it, not only because of their shared attentional focus, but also because of the additional information provided by the utterances. If so, different types of parental linguistic input may differentially affect children's visual attention to objects. Due to their limited hearing experience, children with HL may be less responsive to parent's utterances compared with the CA group. From this perspective, we predict that the effects of parental utterances on children's visual attention seen in the HL group may be more similar to the effects seen in their HA peers. In addition to utterance types, the timing of parental utterances may also matter. Prior research suggests that parental utterances following children's attention facilitate children's sustained attention on the object (e.g., Miller, Ables, King, & West, 2009; Pereira et al., 2014; Yu & Smith, 2012). Therefore, we expect to find more pronounced sustaining effects of parental utterances on children's attention when parents talk about an object already attended to by their children (i.e., child-led, details see the Method section), compared with when they talk about an object not yet attended to by their children (i.e., parent-led, see the Method section). We expect that this pattern would be the same for all three groups of children. To answers these questions on the effects of parental utterance types and leading–following interactional patterns, we used a dual-eye-tracking paradigm developed by Yu and Smith (Yu & Smith, 2013, 2017a, 2017b) to examine children's moment-to-moment attention in relation to their parents' utterances. The results will enhance our understanding of the attentional mechanism(s) by which parental linguistic input affects children's language outcomes.

METHOD

Participants

Participants were 15 parent–child dyads. Five dyads consisted of children with hearing loss (subsequently termed HL group) and their parents. These children were between 27 and 37 months of age (three female; mean age: 32.8 months). Demographic information of the children can be found in Table 1. The other 10 dyads consisted of normal-hearing children and their parents. Children in five of the dyads were matched to the children with HL on their chronological age, plus or minus 2 months (subsequently termed CA group; three female; mean age: 32.0 months). Children in the other five dyads were matched to the children with HL on their hearing age (subsequently termed HA group; three female; mean age: 17.2 months). The hearing age of the children with HL was calculated based on the amount of time after fitting with hearing aids or after cochlear implantation, while the hearing age of the children with normal hearing was calculated based on their chronological age. The present study was conducted according to Declaration of Helsinki guidelines, with written informed consent obtained from a parent or guardian for each child before data collection. All procedures in this study were approved by the Human Subjects and Institutional Review Boards at Indiana University.

TABLE 1
Participant Information

HL Participant No.	Chronological Age	Hearing Age	Sex	*Hearing Device* Left	Right	Etiology	*CA* Age	*HA* Age
1	27	22	M	Bilateral hearing aids		unknown	25	23
2	30	10	F	Bilateral cochlear implants		unknown	28	12
3	34	14	F		Cochlear implant	unknown	35	14
4	36	25	F		Hearing aid	unknown	36	24
5	37	12	M	Cochlear implant	Hearing aid	LVA	36	13
	32.8	16.6					32.0	17.2

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; LVA, large vestibular aqueduct disorder. Children's hearing age was calculated based on the amount of time after fitting with hearing aid(s) or after cochlear implantation.

Stimuli

The stimuli were six novel toys constructed in the laboratory (examples can be seen in Figure 1b). The toys were complex objects made from wood and plastic, and each was colored red, blue, or green. The objects were of comparable overall size (average size: 288 cm^3). Each object was assigned a novel name that followed the phonological rules of English (e.g., *wawa*, *dodi, mapoo*). The objects were divided into two sets, and each set contained one object of each color.

Experimental setup

During the experiment, parents and children sat across from each other at a small table (61 cm x 91 cm x 64 cm). Both participants wore a head-mounted eye tracker (Positive Science, http://www.positivescience.com/, also see Franchak, Kretch, Soska, & Adolph, 2011). Each eye tracker consisted of one infrared camera that pointed to the right eye of the participant and one scene camera that was placed on the forehead of the participant and captured their first-person view (see Figure 1a). The eye trackers recorded participants' gaze directions and their first-person view, with a sampling rate of 30 Hz. Two additional cameras were placed in the room to record the parent–child interaction from third-person views. The parent's speech during the interaction was recorded through a microphone incorporated into the parent's eye tracker.

Procedure

Prior to the play session, two experimenters helped the parent and child put on the eye trackers. To calibrate the eye trackers, an experimenter directed the parent's and child's attention to a small toy on the table. The toy was moved to several different pre-determined locations on the table (corners, center of table, and a few locations in between), which served as calibration points. The procedure was

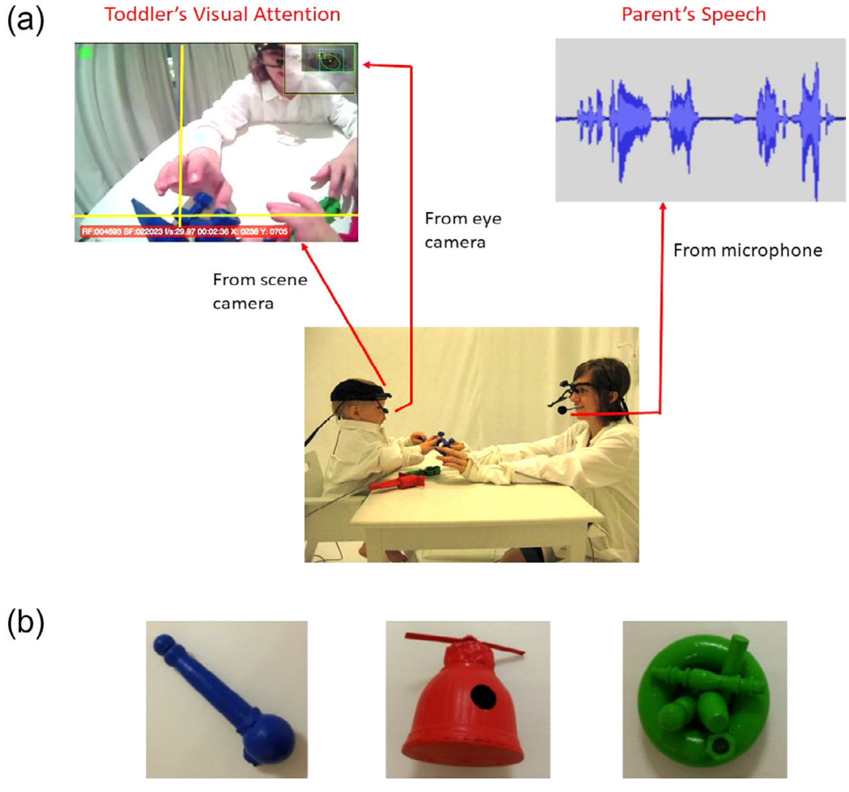

Figure 1 (a) Parent and child sat at a table playing with a set of novel toys. Both participants wore a head-mounted eye-tracking device, which was composed of an eye camera that recorded eye movements and a scene camera that recorded what's in front of the participant from the first-person view. In the current project, we only analyzed the gaze data from the child. Parent's speech was recorded through a microphone incorporated in the parent's eye tracker. (b) Example toy set. Each set of toys contained three novel objects (one blue, one red, and one green) constructed in the laboratory.

repeated until we collected 15 calibration points that covered the entire workspace (i.e., the table). The calibration procedure was administered once before and once after the play session. If, however, the eye trackers was bumped during experiment, it was adjusted thereafter and an additional calibration procedure was administered (more detailed information about how eye trackers were placed on participants and how calibration was conducted can be found in Appendix A in Yu & Smith, 2017b).

Parents were told that the goal of the study was to investigate how parents and children interacted with objects during play. They were asked to play with their children as they normally would. However, if parents were going to refer to the toy objects by name, we asked them to use their novel names. The whole interaction session was divided into 4 "trials." Participants played with one set of objects (one red, one blue, and one green) in each trial (Figure 1b). Each trial lasted 1.5 min. Participants played with each set twice. This resulted in four trials, with a total of 6 min of play data for each dyad.

Data coding

Linguistic input

Speech transcription and target object coding. Following SALT (Systematic Analysis of Language Transcripts, Miller & Chapman, 1985) conventions, parents' speech was first segmented into utterances[1]. One coder transcribed and segmented the utterances for all participants and a second coder independently transcribed and segmented the utterances for five randomly selected participants (33.3% of the sample, one from the HL group, one from the CA group, and three from the HA group). For the two coders, the difference in the total numbers of utterances segmented for these five participants was 13 utterances (Coder 1: 686 utterances; Coder 2: 673 utterances; range of difference in utterance numbers across participant: 1–6). All utterances resulting in disagreement in segmentation had an unintelligible element either at the beginning or at the end. None of these utterances was an object-related utterance (for details, see the *Quantitative and Qualitative Speech Measures* section) and, therefore, were not included in the following analyses related to gaze. The number of the segmented utterances that both coders agreed upon was 670. These utterances will be used for the reliability check for the timing of utterance onsets and offsets.

The timing information of the utterances was obtained by using the open-source program Audacity (https://www.audacityteam.org/). The onset and offset of an utterance were identified based on the waveform of the utterance, without inclusion of adjacent pauses. To assess the reliability of the timing for the segmented utterances based on the waveforms, we calculated the timing differences between the above-mentioned 670 utterances segmented by the two coders. The mean onset timing difference was 11.39 milliseconds ($SD = 33.45$), while the mean offset timing difference was 23.94 milliseconds ($SD = 43.62$). Because the utterances were analyzed along with the gaze data in our study and each frame in the gaze data stream lasted 33 milliseconds (for details, see the following *Gaze Data* section), the average timing differences in the utterance onsets and offsets were shorter than the duration of one frame in the gaze data stream. Because of that, we found the timing differences in the utterance onsets and offsets negligible and should not affect the following utterance-associated gaze analyses.

After speech segmentation, each utterance was transcribed and categorized as object-related vs. non-object-related based on whether they were about the objects in play (for more details, see *Quantitative and Qualitative Speech Measures*). Following that, a coder used Excel to annotate the target object (i.e., the red, blue, or green object) of each object-related utterance. The target object was determined based on 1) the semantics of the utterance (e.g., "Give me the blue one") and 2) the information recorded from the parent's scene and the third-person view cameras (e.g., the parent points to the red object and says "there's a hole in there."). A second coder independently coded the target objects for object-related utterances for five randomly selected participants (33.3% of the sample, two from the HL group, two from the CA group, and one from the HA

[1]According to the SALT convention, utterances were segmented into communication units (C-units) based on both grammatical and phonological information. The grammatical definition of a C-unit is "an independent clause with its modifiers," which is generally composed of a subject and a predicate. The phonological information that is used includes the intonation contour of an utterance (e.g., falling for statements and rising for yes-no questions). Detailed SALT utterance segmentation rules can be found in http://saltsoftware.com/media/wysiwyg/tranaids/CunitSummary.pdf. Additional silence information was included in our segmentation. A silence lasting 400 ms or longer was used as a gap in between two utterances (Pereira et al., 2014; Yu & Smith, 2012).

TABLE 2
Description and Examples of Utterance Types

Types	*Description*	*Example*
Information	Providing information, such as label, features, or actions, about the toy object	"That's a wawa." "Wawa is green." "It goes round, round, round."
Question	Asking a question about the toy object	"Why do you like this one?" "Oh, is that a hammer?"
Directive	Telling or directing child to do something	"Spin it." "Show me the wawa."
Other	Utterances that do not fit in the above-mentioned types	"I like this one." "Oh, well, that one."

group). To assess the inter-rater reliability, we calculated Cohen's kappa coefficients for the coded participants. The average Cohen's kappa was 0.98 (ranged: 0.97–0.99), which was near-perfect based on conventional guidelines (Landis & Koch, 1977).

Following speech transcription and target object coding, we then used quantitative and qualitative analyses to examine the amount and types of input children received during the play session.

Quantitative and qualitative speech measures. The quantitative measures used in the current study focused on the total number of utterances (Hurtado et al., 2008; Klee, 1992) and total number of object-related utterances parents produced in play. In addition to quantitative measures, we further focused on object-related utterances and examined whether parents of different groups talked about objects in different ways. Based on the syntax and semantics of the utterances, we divided object-related utterances into four subtypes: utterances providing information about objects (including labeling, as in DesJardin & Eisenberg, 2007, and utterances commenting on or describing an object feature/property or an object action/movement, as in Bornstein et al., 2008), questions (Tomasello & Farrar, 1986), directives (Cruz et al., 2013; DesJardin & Eisenberg, 2007), and other (for a description and examples of each type, see Table 2). Similar to the general definition in previous studies (e.g., Cruz et al., 2013; DesJardin & Eisenberg, 2007), the directives included in the current study were utterances with which parents instructed children to do something. However, one difference between the directives used in the current study and previous studies is that we only focused on directives that were related to the objects in play. Regulatory directives (e.g., "Sit down!", "Don't touch the camera!") were not included in the current analyses. One coder coded the utterance types for all participants. Two other coders independently coded the data for one-third of the participants (two from the HL group, two from the CA group, and one from the HA group). The inter-rater reliability was near-perfect across coders with an average Cohen's kappa score of 0.94 (range: 0.92–0.99, Landis & Koch, 1977).

Gaze data. As the sampling rate of the eye trackers was 30 Hz, we obtained approximately 10,800 frames of gaze data for each participant during the 6-minute toy play session. Even though we had gaze data for both parents and children, we only focused on children's data for the purposes of the current project. The regions of interests (ROIs) were the three objects the dyads played with in each trial. Gaze data were coded frame-by-frame.[2] Trained coders used images captured by the eye tracker and coded whether children's gaze direction overlapped with any of the ROIs, and if so, which ROI[3] (more detailed information about ROI coding procedure can be found in Appendix B in Yu & Smith, 2017b). In total, children generated 2,036 looks to the three ROIs during the play session. These 2,036 ROI looks served as the gaze data in the following analyses.

To assess inter-coder reliability, a second coder independently coded data for seven participants (46.7% of the sample, three from the HL group, two from the CA group, and two from the HA group). For these seven participants, four were randomly selected to have the whole interaction coded and three had the first half of the interaction coded. The reliability checks reported in the following did not differ across groups or whether the whole interaction or the first half was coded. We conducted two sets of reliability checks. The first set of analyses included frame-by-frame comparisons from the entire coded data, which contained all looks that landed both within and outside of ROIs (i.e., both gazes that will be included in the following analyses and looks that will not be included in the analyses). The second set of analyses focused on frames which coders had determined to be ROI looks (i.e., gazes that will be included in the following analyses). Since we calculated Cohen's kappa based on frame-by-frame coding and that each frame lasted approximately 33 ms in real time, the scoring inherently took the duration of each ROI look into account. If the coders disagreed on the onset and offset frames of a gaze, the differences in coding each single frame were taken as disagreement. We first calculated Cohen's kappa based on the entire coding. This analysis took gaze onset–offset into account as well as coders' determination on whether or not each look landed on one of the ROIs and if so, which one. Inter-coder reliability of the entire coding was good with an average Cohen's kappa of 0.75 (range: 0.70–0.78, Landis & Koch, 1977). After that, we focused on the frames in which both coders had determined to be an ROI look and then assessed coders' agreement on which target object it was. For the second set of analyses, the average kappa score was 0.93 (range: 0.90–0.96), indicating that there was near-perfect inter-rater agreement on the target object of the ROI looks (i.e., gazes included in the following analyses).

[2]In the current study, gaze was defined as one (continuous) look to an ROI. We did not use pre-determined duration criteria, such as only counting looks being stable for at least a certain number of frames or milliseconds to be counted as gazes. As long as there was a shift in looks to a different ROI or to a region outside of the previous ROI, it was counted as a different look. In theory, a gaze could be as short as lasting only one frame. However, in practice, only 1.3% of gazes in our analyses lasted fewer than 3 frames. The majority of gazes included in our analyses lasted longer than 100 milliseconds.

[3]Due to the nature of viewing in naturalistic environments—which includes depth information, partial object occlusion, or object movement—we did not adopt a set criteria of spatial tolerance based on a certain number of pixels, which is a practice used in many screen-based eye-tracking studies (e.g., Blair, Watson, Walshe, & Maj, 2009; Vassallo, Cooper, & Douglas, 2009). Instead, we included contextual information in our coding. When objects were stationary and a crosshair indicating the center of gaze direction landed on one of the objects in an individual frame, then that object was coded as the target of the gaze. In cases when the center of the crosshair did not land inside an object in a frame, we included contextual information in the coding. For example, if a child was pointing to or reaching for an object but the center of the crosshair landed a few pixels outside of the object, we still coded it as landing on the object. Or when the eye movement across frames indicated tracking a certain object, even if the center of the crosshair appeared a few pixels outside of the object in the frame, we coded the tracked object as the target object for the frame. The reasoning is that our visual system's focus (foveation) cannot be reduced to a few pixels in the real world. Even if the center of the crosshair does not land precisely on the object in a certain frame, the object is still in focus.

Utterance-associated gaze types

To understand the potential different effects of parents' speech on children's attention, we examined child's attention in two distinct situations—we divided children's gazes to objects based on their temporal relation to parents' utterances: parent-led gazes and child-led gazes. A parent-led gaze was a child's gaze to an object that started within 3s after the onset of a parent's utterance about the *same* object (Figure 2a). The reason for choosing a 3s window was that the average length of utterances in our data was 1.16 sec (SD = .21) and we wanted to have a window larger than the mean utterance duration to allow for processing, but not too long to include gazes not driven by the utterances. A child-led gaze was a child's gaze to an object that 1) started before the onset of a parent's utterance about the *same* object and 2) overlapped with the utterance (Figure 2b). In addition to the above two categories, any gazes that did not belong to the parent-led or child-led categories were defined as Other as a comparison. The Other category included gazes outside of the 3s window or gazes overlapping with an utterance about a *different* object (Figure 2c). In the latter case, the target object of the parent's utterance and the target object of the child's fixation were mismatched. It is noteworthy that the parent-led and child-led gazes in our analyses were defined solely based on the timing relationships between parent's utterances and children's gazes. Parents or children may or may not be aware of or intentionally lead or follow the other. This coding allows us to examine whether parents' utterances were effective in extending children's attention to objects when the utterances preceded children's attentional focus and when the utterances followed children's attention.

RESULTS

We begin by reporting the amount and types of linguistic input children received in the interaction. We then ask whether parents' linguistic input affected children's visual attention. We are particularly interested in whether parents' utterances about objects extended children's gazes toward objects when parents' utterances led children's gazes

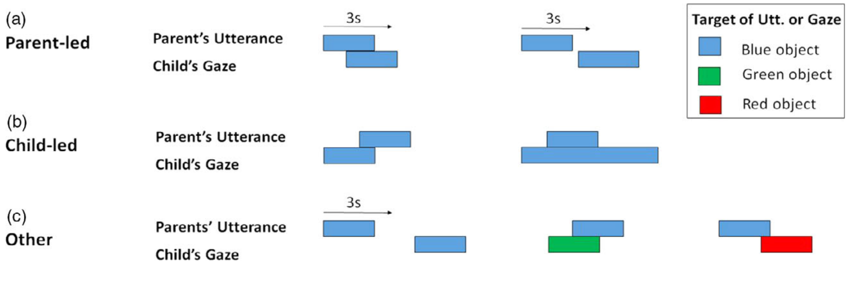

Figure 2 Gaze types. (a) A parent-led gaze was defined as a child's gaze to an object that started within 3s after the onset of a parent's utterance about the same object. (b) A child-led gaze was defined as a child's gaze to an object that 1) started before the onset of a parent's utterance about the same object and 2) overlapped with the utterance. (c) An Other gaze was defined as a gaze that did not belong to the first two types. This category included gazes outside of the 3s window and gazes overlapping with an utterance about a different object (i.e., mismatch between the target of the parent's utterance and the target of the child's gaze).

and when parents' utterances followed children's gazes. All following analyses were conducted using SPSS version 24. We used an alpha level of .05 for all statistical tests. For multiple comparisons, we will report both the raw p values and compare them against the adjusted alpha levels based on Bonferroni corrections.

Amount and types of parental linguistic input

The measures on the amount of input included the total number of utterances parents produced during the 6 min of play interaction as well as the total number of object-related utterances. Children received approximately 150 parental utterances during the play session (HL: $M = 155.20$, $SD = 31.30$; CA: $M = 160.20$, $SD = 40.44$; HA: $M = 144.00$, $SD = 25.29$). Of those parental utterances, over half were related to the novel objects at play (HL: $M = 85.80$, $SD = 22.30$; CA: $M = 90.80$, $SD = 34.14$; HA: $M = 90.00$, $SD = 24.49$). Between-group ANOVAs showed that there was no significant group difference in terms of the total number of utterances ($F(2, 12) = .32$, $p = .73$) or the number of object-related utterances children received in the play session ($F(2, 12) = .05$, $p = .95$). In sum, parents in these three groups provided children with similar amount of linguistic input.

The next set of analyses focused on object-related utterances and examined whether parents in different groups used different utterance types during the play session. Because different parents produced different amount of utterances, instead of using raw numbers, we calculated the proportions of: 1) utterances providing information about the toy objects, 2) questions about the toy objects, 3) directives, and 4) other utterances. Since we were interested in the characteristics of the linguistic input children received, the following analyses excluded the "other" type, for which the characteristics were difficult to define by nature and which only accounted for a small proportion of utterances in each group (HL: 0.05; CA: 0.07; HA: 0.10). In what follows, we compare the types of utterances children received across groups (Figure 3).

Between-group ANOVAs showed that there were significant group differences in the proportions of utterances providing information about objects ($F(2, 12) = 9.77$, $p = .003$) and directives ($F(2, 12) = 7.05$, $p = .009$). Post hoc comparisons using Bonferroni corrections (alpha level = .017) revealed that parents in the HL and CA groups produced significantly more utterances providing object information than parents in the HA group (HL > HA at $p = .007$; CA > HA at $p = .001$). Parents of HL and HA

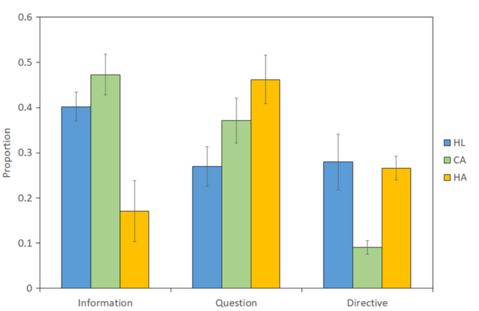

Figure 3 Proportion of different utterance types for each group. The error bars represent standard errors.

groups used more directives than parents of CA group (HL > CA at $p = .006$; HA > CA at $p = .009$). There was also a trending group difference in parents' production of questions ($F(2, 12) = 3.81$, $p = .052$). Parents in the HA group tended to ask more questions than parents in the HL group ($p = .017$).

These results suggest that parents of older children (HL and CA groups) were more likely to provide information about objects than parents of younger children (HA). In addition, parents of children with hearing loss and younger normal-hearing children were more likely to take a directive role in the joint play than parents of older normal-hearing children did. We next examine whether similar or different linguistic patterns observed among the three groups had different effects on children's attention.

Parent-led gazes

The first set of gaze analyses tested whether parent-led gazes were longer than Other gazes. Because each participant contributed multiple gazes in the analyses, we used generalized estimating equations (GEE) to account for the non-independence of data (Liang & Zeger, 1986). In the following analyses, each of the three utterance-related (i.e., Information, Question, and Directive) parent-led gaze was dummy-coded as 1 and each Other gaze was dummy-coded as 0. The gaze type (e.g., Information vs. Other) was included in the GEE model as the predictor, and the gaze length was entered as the dependent variable. We tested whether gaze type was a significant predictor of gaze length[4]. For each group, gazes associated with each utterance type (i.e., Information, Question, and Directive) were compared against the Other type separately. Therefore, we used an adjusted alpha level of .017 (.05/3) for the following within-group analyses. As shown in Figure 4a, the HL group's parent-led gazes associated with three utterance types were all significantly longer than Other gazes (Other: $M = 1.23$; Information: $M = 2.24$, Wald $\chi^2 = 10.92$, $p = .001$; Question: $M = 1.55$; Wald $\chi^2 = 5.75$, $p = .016$; Directive: $M = 2.14$, Wald $\chi^2 = 44.21$, $p < .001$). The results suggest that all three types of parents' utterances caused children to sustain their attention on the object.

Similar patterns can be seen for the CA group (Figure 4b). Overall, different types of parent-led gazes were longer than Other gazes (Other: $M = 1.25$; Information: $M = 2.54$; Wald $\chi^2 = 11.12$, $p = .001$; Question: $M = 2.39$; Wald $\chi^2 = 18.07$, $p < .001$; Directive: $M = 2.63$; Wald $\chi^2 = 19.13$, $p < .001$). For the HA group (Figures 4c), parent-led gazes associated with Questions and Directives were overall longer than Other gazes (Other: $M = 1.59$; Question: $M = 3.03$; Wald $\chi^2 = 8.60$, $p = .003$; Directive: $M = 2.47$; Wald $\chi^2 = 5.65$, $p = .017$). In contrast, the mean length of information-related parent-led gazes were not different from Other gazes ($M = 1.84$; Wald $\chi^2 = .80$, $p = .37$). These results suggest that, similar to the HL group's patterns, all utterance types were effective in extending CA children's attention to objects. However, only questions and directives extended younger normal-hearing children's visual attention to objects whereas utterances providing object information did not.

[4]Each GEE model included an intercept term and one predictor "utterance type" (e.g., Information vs. Other). Because the dependent variable was gaze length, we used a linear link function. Model-based covariance estimators were used in the GEE models to account for the small sample size (Lu et al., 2007). The significance of the predictor and the goodness of the fit of the model were determined by the p value and whether the inclusion of the predictor reduced the quasi-likelihood under the independence model criterion (QIC).

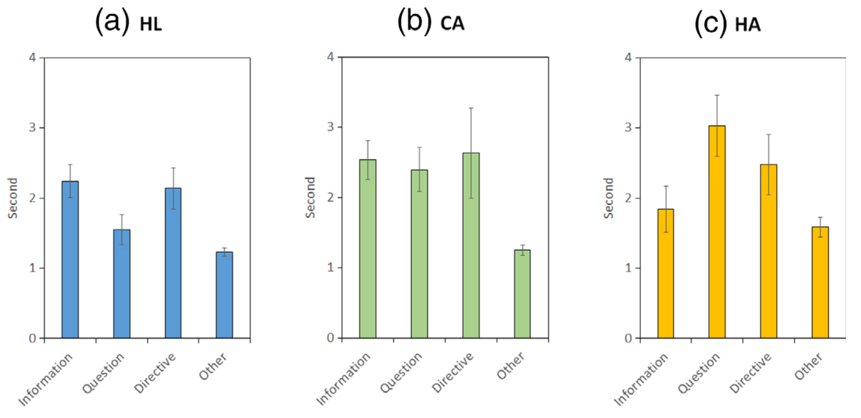

Figure 4 Parent-led and Other Gazes across Groups. (a) Mean duration of HL group's parent-led gazes associated with different utterance types and Other gazes. (b) Mean duration of CA group's parent-led gazes associated with different utterance types and Other gazes. (c) Mean duration of HA group's parent-led gazes associated with different utterance types and Other gazes. The error bars represent standard errors.

The above-mentioned analyses compared each group's parent-led gazes against Other gazes. To further examine the similarities and/or differences of how parents' speech affected children's gazes across groups[5], we directly compared the lengths of gazes associated with different utterance types between HL and CA groups, and between HL and HA groups (at adjusted alpha level .025). Compared to the CA group, the HL group's gazes associated with utterances providing information and directives were similar in length (Information: Wald χ^2 = .29, p = .59; Directive: Wald χ^2 = 1.96, p = .16). However, HL group's gazes associated with questions were significantly shorter than the ones in the CA group (Wald χ^2 = 5.97, p = .015). Similar patterns were found in the comparisons between HL and HA groups. These two groups had comparable lengths of gazes associated with utterances providing information and directives (Information: Wald χ^2 = 1.07, p = .30; Directive: Wald χ^2 = .78, p = .38). Yet, the HL group's question-related gazes were shorter than the HA group's (Wald χ^2 = 9.90, p = .002).

In summary, for the HL and CA groups, all types of parent-led gazes were significantly longer than Other gazes. These results suggest that in parent-led situations, utterances were effective in extending the older children's gazes to objects. For the HA group, questions and directives were effective in extending their attention. However, utterances providing information about objects, which is critical for word learning, did not have a significant effect on the younger toddlers' visual attention to objects. When compared across groups, the HL group's gazes associated with questions were not as long as the ones in the two normal-hearing groups, even though they were longer than the HL group's Other gazes. These results suggest that even though questions extended HL group's gazes to objects, the effect may not be as pronounced as in the two normal-hearing groups.

[5]Each GEE model included an intercept term and one predictor "group" (e.g., HL vs. CA). We used a linear link function in the models, because the dependent variable was gaze length. To account for the small sample size, model-based covariance estimators were used in the GEE models (Lu et al., 2007).

Child-led gazes

We next present three sets of analyses, which jointly allow us to answer the question "When children already look at an object, does parent's utterance about the same object also extend their attention?" The first set of analyses tests whether child-led gazes were overall longer than Other gazes. In the second set of analyses, we ask whether child-led gazes were long even before parents started to talk about the object. In the third set of analyses, we examine whether children's gazes extended till after the end of parents' utterances.

As can be seen in Figure 5a, all HL group's child-led gazes were significantly longer than Other gazes (adjusted alpha level = .017; Other: $M = 1.23$; Information: $M = 3.96$, Wald $\chi^2 = 14.65$, $p < .001$; Question: $M = 3.99$, Wald $\chi^2 = 19.97$, $p < .001$; Directive: $M = 3.19$, Wald $\chi^2 = 39.49$, $p < .001$. In the CA group (Figure 5b), child-led gazes were also significantly longer than Other gazes (Other: $M = 1.25$; Information: $M = 5.48$, Wald $\chi^2 = 39.77$, $p < .001$; Question: $M = 5.54$, Wald $\chi^2 = 284.45$, $p < .001$; Directive: $M = 4.44$, Wald $\chi^2 = 67.32$, $p < .001$). Figure 5c shows similar patterns in the HA group (Other: $M = 1.59$; Information: $M = 5.00$, Wald $\chi^2 = 19.91$, $p < .001$; Question: $M = 5.46$, Wald $\chi^2 = 77.92$, $p < .001$; Directive: $M = 4.41$, Wald $\chi^2 = 11.73$, $p = .001$). Taken together, these results suggest that, for all three groups, child-led gazes associated with all three types of object-related utterances were longer than Other gazes.

Interestingly, between-group analyses (HL vs. CA and HL vs. HA; adjusted alpha level = .025) indicated that the HL group's child-led gazes associated with Questions and Directives, overall, were shorter than CA group's (Information: Wald $\chi^2 = 2.90$, $p = .09$; Question: Wald $\chi^2 = 6.26$, $p = .012$; Directive: Wald $\chi^2 = 8.04$, $p = .005$). Their gaze lengths associated with different utterance types were overall comparable to the HA group's (Information: Wald $\chi^2 = 1.75$, $p = .19$; Question: Wald $\chi^2 = 3.21$, $p = .07$; Directive: Wald $\chi^2 = 2.21$, $p = .14$).

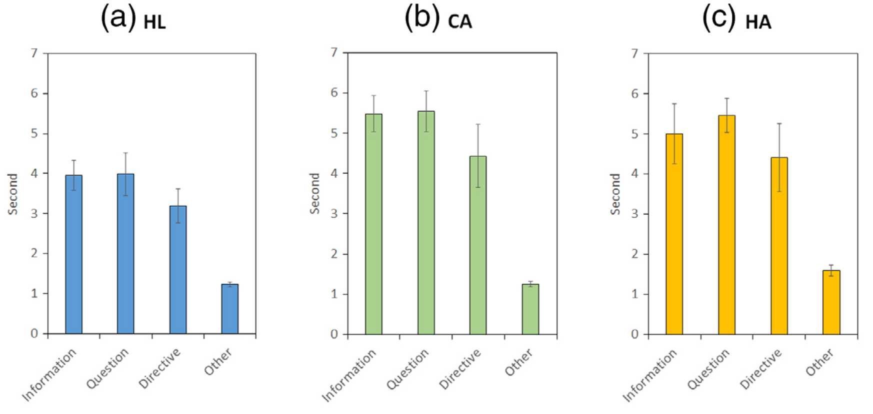

Figure 5 Child-led and Other Gazes across Groups. The mean durations of Other gazes illustrated in Figure 5 are identical to the ones in Figure 4. They are repeated here for ease of comparisons. (a) Mean duration of HL group's child-led gazes associated with different utterance types and Other gazes. (b) Mean duration of CA group's child-led gazes associated with different utterance types and Other gazes. (c) Mean duration of HA group's child-led gazes associated with different utterance types and Other gazes. The error bars represent standard errors.

There are two possible explanations to why child-led gazes were longer than Other gazes. First, the child-led gazes were long even before parents started to talk about the objects. Alternatively, parents' utterances extended children's looks to the objects. If the gazes were long prior to the onset of utterances, parents might have more chance to join in and talk about the objects. If parents' object-related utterances extended children's attention, we should expect to see gazes last until after the offset of the utterances. These two explanations are not mutually exclusive. Some gazes may be long for one reason while others are long for both reasons.

We next checked whether child-led gazes were long prior to parents' utterances by examining each group's pre-utterance durations associated with different utterance types and comparing them against Other gazes. We calculated the pre-utterance durations by taking the time difference between the onset of a child-led gaze and the onset of the parent's utterance (Figure 6a). If the pre-utterance durations are longer than Other gazes, this suggests that the long durations of child-led gazes are partially caused by long pre-utterance durations. If, however, the pre-utterance durations are not different from Other gazes, this suggests that the longer child-led gazes (compared to Other gazes) are not caused by longer duration prior to parents' utterances.

As can be seen in Figure 6b, the mean pre-utterance durations for all utterance types in the HL group were comparable to the duration of Other gazes (adjusted alpha level = .017; Information: M = 1.50, Wald χ^2 = .50, p = .48, Question: M = 1.63, Wald χ^2 = .37, p = .53; Directive: M = 1.53, Wald χ^2 = .97, p = .33). In the CA group (Figure 6c), the pre-utterance durations associated with Questions and Directives were significantly longer than Other gazes (Question: M = 2.44, Wald χ^2 = 39.68, p < .001; Directive: M = 1.54, Wald χ^2 = 18.62, p < .001). The pre-utterance durations associated with utterances providing information also tended to be longer than Other gazes (Information: M = 1.80, Wald χ^2 = 3.76, p = .053). In the HA group (Figure 6d), the pre-utterance durations associated with questions tended to be longer than Other gazes

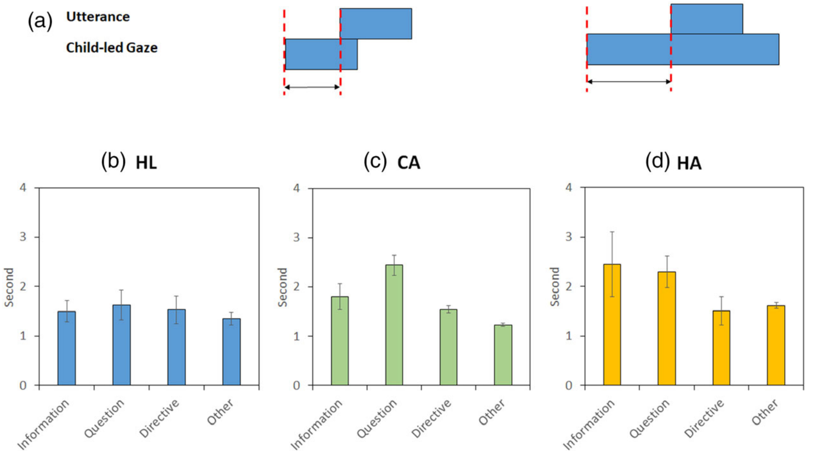

Figure 6 Length of pre-utterance duration. (a) The length was calculated by taking the time difference between the onset of a child-led gaze and the onset of an utterance on the same object. We then compared the pre-utterance duration associated with different utterance types against the length of Other gazes. (b) Pre-utterance durations in the HL group. (c) Pre-utterance durations in the CA group. (d) Pre-utterance durations in the HA group. The error bars represent standard errors.

(Question: M = 2.30, Wald χ^2 = 4.05, p = .044), while the pre-utterance durations associated with utterances providing information or directives were not different from Other gazes (Information: M = 2.45, Wald χ^2 = 1.44, p = .23; Directive: M = 1.50, Wald χ^2 = .10, p = .75). These results suggest that, at least for children in the HL and HA groups, their long child-led gazes were not caused by long pre-utterance durations. In the CA group, their long child-led gazes were caused partially by the long pre-utterance durations, which may have provided parents more of an opportunity to talk about the objects.

In the last set of analyses, we tested whether child-led gazes continued until after the end of the utterances. We calculated the time difference between the offset of an utterance and the offset of a child-led gaze (subsequently termed post-utterance duration, see Figure 7a). If the gaze ends before the offset of the utterance, the value for the post-utterance duration would be negative. In contrast, if the gaze ends after the offset of the utterance, the value for the post-utterance duration would be positive. If the values are positive and significantly different from zero[6], then it suggests that the gazes tend to extend until after the utterances end.

As shown in Figure 7b, the HL group's values of post-utterance durations associated with questions were positive and significantly different from zero (adjusted alpha level = .017; Question: M = 1.21, Wald χ^2 = 8.71, p = .003). A similar trend was found for utterances providing object information and directives (Information: M = 1.31, Wald χ^2 = 4.27, p = .039; Directive: M = 0.51 Wald χ^2 = 4.17, p = .041). For the CA group (Figure 7c), the post-utterance durations associated with different utterance

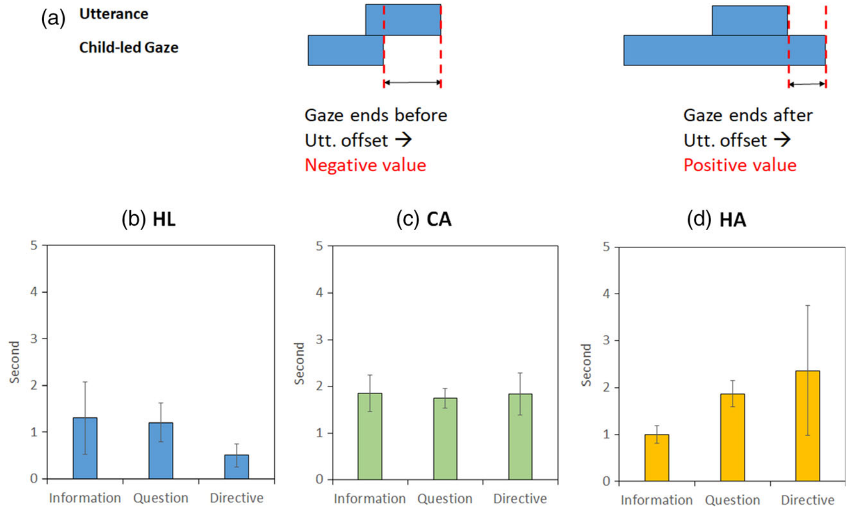

Figure 7 Length of post-utterance duration. (a) The length was calculated by taking the time difference between the offset of a child-led gaze and the offset of an utterance on the same object. If the gaze ended before the offset of the utterance, the value would be negative. If the gaze ended after the offset of the utterance, the value would be positive. (b) Post-utterance durations for HL group. (c) Post-utterance durations for CA group. (d) Post-utterance durations for HA group. The error bars represent standard errors.

[6]Each GEE model included only an intercept term and was compared against 0. A linear link function and a model-based covariance estimator were used.

types were all positive and significantly different from zero (Information: $M = 1.85$, Wald $\chi^2 = 21.91$, $p < .001$; Question: $M = 1.74$, Wald $\chi^2 = 71.90$, $p < .001$; Directive: $M = 1.84$, Wald $\chi^2 = 16.70$, $p < .001$). Figure 7d shows similar patterns for the HA group (Information: $M = 1.00$, Wald $\chi^2 = 30.03$, $p < .001$; Question: $M = 1.87$, Wald $\chi^2 = 45.07$, $p < .001$; Directive: $M = 2.36$, Wald $\chi^2 = 3.86$, $p = .049$). These results suggest that child-led gazes did not only start before the onset of utterances, they also tended to last until after the end of the utterances.

We next examined the extension effects across groups by comparing the post-utterance durations (HL vs. CA and HL vs. HA; adjusted alpha level .025). Compared to their CA peers, the HL group's post-utterance durations associated with directives did not last as long (Wald $\chi^2 = 7.24$, $p = .007$), even though their post-utterance durations associated with utterances providing information and questions were overall similar (Information: Wald $\chi^2 = .80$, $p = .37$; Question: Wald $\chi^2 = 1.80$, $p = .18$). The overall gaze extension effects for the HL and HA groups were comparable to each other (Information: Wald $\chi^2 = .05$, $p = .83$; Question: Wald $\chi^2 = 1.76$, $p = .18$; Directive: Wald $\chi^2 = 2.55$, $p = .11$). These results suggest that directives were more effective in extending child-led gazes in the CA group than in the HL children.

To summarize, these gaze analyses showed that overall, parent-led and child-led gazes were longer than Other gazes, and that all types of parents' utterances seemed to help extend HL and CA children's gazes. However, various between-group analyses suggest that the effects of utterances on gaze extension were less pronounced in the HL group than in the CA group.

DISCUSSION

The current study investigated parent–child interactions in children with hearing loss and children with normal hearing. Overall, parents of toddlers with hearing loss produced comparable *amount* of linguistic input as parents of normal-hearing toddlers. With regard to the *types* of linguistic input, parents of children with hearing loss (HL) produced a similar amount of utterances providing object information as parents of children matched for chronological age (CA). They also provided a similar amount of directives as parents of children matched by hearing experience (HA). Similar to the patterns seen in their CA peers, HL children's gazes were extended by all object-related utterances. However, the extension effects of parental utterances on children's attention were less pronounced in the HL group.

Linguistic input across groups

Overall, parents in different groups provided their children with comparable amounts of input during the joint play session, which is consistent with findings from previous studies (Fagan et al., 2014; VanDam et al., 2012). The types of utterances parents produced seemed to be associated with *both children's chronological age and hearing experience*. Similar to the CA group, over 40% of the utterances produced by parents of children with HL were to provide information about objects, while parents in the HA group were significantly less likely to do so. These results suggest that parents of older toddlers tended to provide labels for objects and describe or comment on the objects' features or actions when they engaged in joint object play with their children. Learning information about objects has been shown to be important for word and concept learning (e.g., Landau, Smith, & Jones, 1998; Madole & Oakes, 1999). These differences in HA parents' linguistic input compared to the two older groups may be related to parents' belief or estimate about what their children could understand (Miller, 1988).

Many prior studies suggest that parents of children with hearing loss tend to be more controlling and dominant in interactions than parents of children with normal hearing (Fagan et al., 2014; Henggeler et al., 1984; Meadow, Greenberg, Erting, & Carmichael, 1981; Spencer, Bodner-Johnson, & Gutfreund, 1992; Wedell-Monnig & Lumley, 1980). In the current study, parents in the HL group tended to produce more directives than CA parents. However, parents of HA children also produced more directives than their CA counterparts. These results suggest that parents' use of directives may be related to children's hearing experience, a finding consistent with the proposal that parents of children with less language experience or poorer language abilities are more likely to take a leading role in interactions (e.g., Bornstein et al., 2008; DesJardin et al., 2014; Musselman & Churchill, 1992; Snow, 1972). All together, these linguistic patterns suggest that parents of HL children provided their children with the same amount of opportunities to learn about object information as CA parents did. On the other hand, like the parents of younger HA children, they also gave more verbal guidance in interactions.

Another noteworthy aspect of this study is that our analyses of utterance types only included utterances related to the objects in play. Therefore, "directives" in our study are utterances that inform or direct the child to do something with the objects, which may or may not already be in the child's focus of attention (as we included both parent-led and child-led gazes in our analyses). This type of directives is different from the measures of directiveness used in many previous studies (e.g., Henggeler et al., 1984; Spencer et al., 1992; Tomasello, 1988; Yoder et al., 1993), which focused on parents' "re-directing" children's attention to a different object or regulatory directives (e.g., directives to prohibit or ask children to stay on task, such as "sit down!" or "don't touch the camera!"). Several previous studies suggest that the use of these re-directing or regulatory directives has a negative effect on children's vocabulary development (e.g., Tomasello, 1988; Yoder et al., 1993). Interestingly, in a study that correlated maternal language in a toy play session with children's vocabulary size, Tomasello and Farrar (1986) found that the only type of directives that correlated positively with children's vocabulary was the one that *successfully* guided children's attention to the object their mother referred to. This type of directive is similar to the one we studied in the parent-led scenarios, in which parents produced a directive about an object and children followed the parent's direction by looking at the same object. Our findings suggest that using object-related directives may allow parents to guide their children through the give-and-take in interactions. Or put it in another way, object-related directives may give children, particularly younger children or children with less language experience, needed guidance, which may in turn facilitate the maintenance of interaction fluency (Vaccari & Marschark, 1997).

Contribution of parental linguistic input on children's attention

One interesting finding of the current study is that the extension effects of different utterance types on the HL group's visual attention were qualitatively similar to the ones seen in the CA group. Despite the fact that HL and CA children had different hearing experiences, all three types of object-related utterances extended these two groups' attention to objects in both parent-led and child-led scenarios. There are at least two possible reasons for this result, one related more to the general developmental mechanisms, while the other related to the nature of parent–child interactions. First, the gaze extension effects of different utterance types do not depend solely on children's hearing or language experience. It is possible that children's cognitive maturity or more general cognitive processing skills, which are often related to children's age (e.g., Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Gerhardstein & Rovee-Collier, 2002; Rose, Feldman, & Jankowski, 2002, 2004; Rose, Feldman, Jankowski, & Van Rossem, 2011), also play a role in how they visually attend to objects. The more mature age and cognitive skills of the HL and CA children (relative to the HA group) may contribute to their ability in sustaining visual attention to objects. Second, many recent studies have demonstrated that parent–child interactions are multimodal, in that parents' speech is accompanied by parents' and children's own manual actions on objects (e.g., Suanda et al., 2016; Yu & Smith, 2012, 2017a, 2017b) or co-occurs with parents' pointing, looking, and gestures (Frank et al., 2013; Gogate et al., 2000, 2006; Yu & Smith, 2016). In addition to linguistic information, these multimodal cues could work together in increasing the salience of the referred objects and lead to children's sustained attention toward these objects. These two possibilities are not mutually exclusive. It is likely that children's real-time attentional states are influenced by multiple factors. They also point to the possibility that there may be multiple pathways, solutions, or action points for intervention that can be used to support learning. For example, encouraging parents to use multimodal cues, such as gestures, pointing, or manual movement of objects, in order to highlight the target of their utterances may increase the salience of objects. This may then help children with hearing loss to better figure out the topic of conversation and sustain their attention to objects, which could have a positive effect on their language and concept learning.

Another interesting finding is that, in spite of the qualitative similarity, the gaze extension effects were less pronounced in the HL group compared to the CA group. One possible explanation is related to the proposal that children with hearing loss may need to use vision to monitor the world and, therefore, have more distributed visual attention (Mitchell & Maslin, 2007; Smith et al., 1998). Previous studies using visual selection tasks suggest that children with hearing loss show less sustained attention and are more likely to be distracted by non-task-related information compared to normal-hearing children (Dye & Hauser, 2014; Quittner et al., 1994; Smith et al., 1998). Our findings suggest that the attentional differences can be found in naturalistic, real-time interactions as well. These differences warrant further investigation to better understand whether and how they affect long-term language, cognitive, and social development.

Leading vs. following toddlers' attention

Past research with one-year-olds suggested that parents' utterances following children's attention are associated with children's better word learning and larger vocabulary size compared to parents' utterances leading children's attention (Pereira et al., 2014; Tomasello & Farrar, 1986; Tomasello & Todd, 1983; Yu & Smith, 2012). In our current paradigm, for the HL and CA groups, whose ages were older than the children in previous studies, all utterance types were effective in extending their gazes to objects, no matter whether parents' utterance followed or preceded the onset of children's gaze to an object. For the HA group, whose ages were similar to the children in previous studies, gazes associated with utterances providing information were longer than Other gazes only in the child-led but not in parent-led scenarios. Consistent with findings from previous studies (Miller et al., 2009; Yu & Smith, 2016), our research provides new evidence that, for these very young toddlers (i.e., HA group), following their attentional state may be a more efficient way in sustaining their attention than directing their focus of attention.

Reciprocal and dynamic nature of parent–child interactions

One implicit underlying assumption of many previous studies focusing on parental language styles is that the same type of utterances has roughly the same overall positive or negative effects across children (e.g., Cruz et al., 2013; DesJardin & Eisenberg, 2007; DesJardin et al., 2014; Girolametto et al., 1999; Tomasello, 1988; Tomasello & Todd, 1983; Yoder et al., 1993). However, our study suggests that, for children of different groups, the same utterance type can differentially affect their visual attention depending on contexts. The effects of parental linguistic input on children's attention are modulated by child characteristics. This work underlines the importance of studying the reciprocal and dynamic nature of parent–child interactions.

By showing the temporal unfolding of children's gaze behaviors in relation to parents' language behaviors, our study offers a process account of how different types of parental utterances affect children's attention. The sustained attention moments driven by different types of utterances may be an important component to explain why and how different parental language use affects children's long-term learning, as attention to objects is critical for learning object concepts and object names (e.g., Baldwin & Markman, 1989; Kannass & Oakes, 2008; Lawson & Ruff, 2004; Ruff et al., 1990; Smith et al., 2002; Tomasello, 1988; Tomasello & Farrar, 1986).

CONCLUSIONS

The current study demonstrates that the overall amount of input received by children with hearing loss was not different from their chronological age-matched peers and hearing age-matched peers. Like their chronological age-matched peers, children with hearing loss received a significant amount of linguistic input providing information about objects. However, similar to their younger hearing age-matched peers, they also received more directives than older normal-hearing children. Importantly, all types of parents' object-related utterances extended the hearing loss group's attention to objects and the effects were qualitatively similar to the ones seen in their chronological age-matched normal-hearing peers. Our study is the first one to examine the dynamic effects of different utterance types on children's visual attention in joint object play. Future work will focus on how the real-time extension on children's attention observed in the present study creates long-term outcomes in language and cognitive development.

ACKNOWLEDGMENTS

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (T32 DC00012), the National Institutes of Health

(R01HD074601), and the Indiana University Collaborative Research Grant. We thank Heidi Neuberger, Charlotte Wozniak, Melissa Elston, Steven Elmlinger, and Seth Foster for helping with data collection. We also thank Seth Foster and Tian (Linger) Xu for developing data management and processing software.

CONFLICT OF INTEREST

The authors declare no conflicts of interest with regard to the funding source for this study.

REFERENCES

Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: A first step. *Child Development*, *60*(2), 381–398.

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely selective attention: Eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1196–1206.

Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C. S., & Haynes, O. M. (2008). Maternal responsiveness to young children at three ages: Longitudinal analysis of a multidimensional, modular, and specific parenting construct. *Developmental Psychology*, *44*(3), 867–874.

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.

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*(1), 337–367.

Cross, T. G., Johnson-Morris, J. E., & Nienhuys, T. G. (1980). Linguistic feedback and maternal speech: Comparisons of mothers addressing hearing and hearing-impaired children. *First Language*, *1*(3), 163–189.

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.

DesJardin, J. L., Doll, E. R., Stika, C. J., Eisenberg, L. S., Johnson, K. J., Ganguly, D. H., … Henning, S. C. (2014). Parental support for language development during joint book reading for young children with hearing loss. *Communication Disorders Quarterly*, *35*(3), 167–181.

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.

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.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

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.

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development*, *9*(1), 1–24.

Gerhardstein, P., & Rovee-Collier, C. (2002). The development of visual search in infants and very young children. *Journal of Experimental Child Psychology*, *81*(2), 194–215.

Girolametto, L., Weitzman, E., Wiigs, M., & Pearce, P. S. (1999). The relationship between maternal language measures and language development in toddlers with expressive vocabulary delays. *American Journal of Speech-Language Pathology*, *8*(4), 364–374.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878–894.

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.

Harris, M., Jones, D., & Grant, J. (1983). The nonverbal context of mothers' speech to infants. *First Language*, *4*(10), 21–30.

Henggeler, S. W., Watson, S. M., & Cooper, P. F. (1984). Verbal and nonverbal maternal controls in hearing mother-deaf child interaction. *Journal of Applied Developmental Psychology*, *5*(4), 319–329.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31–F39.

Kannass, K. N., & Oakes, L. M. (2008). The development of attention and its relations to language in infancy and toddlerhood. *Journal of Cognition and Development*, *9*(2), 222–246.

Karasik, L. B., Tamis-LeMonda, C. S., & Adolph, K. E. (2014). Crawling and walking infants elicit different verbal responses from mothers. *Developmental Science*, *17*(3), 388–395.

Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders*, *12*(2), 28–41.

Landau, B., Smith, L., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memory and Language*, *38*(1), 1–27.

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

Lawson, K. R., & Ruff, H. A. (2004). Early focused attention predicts outcome for children born prematurely. *Journal of Developmental & Behavioral Pediatrics*, *25*(6), 399–406.

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

Lu, B., Preisser, J. S., Qaqish, B. F., Suchindran, C., Bangdiwala, S. I., & Wolfson, M. (2007). A comparison of two bias-corrected covariance estimators for generalized estimating equations. *Biometrics*, *63*(3), 935–941.

MacRoy-Higgins, M., & Montemarano, E. A. (2016). Attention and word learning in toddlers who are late talkers. *Journal of Child Language*, *43*(5), 1020–1037.

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review*, *19*(2), 263–296.

Meadow, K. P., Greenberg, M. T., Erting, C., & Carmichael, H. (1981). Interactions of deaf mothers and deaf preschool children: Comparisons with three other groups of deaf and hearing dyads. *American Annals of the Deaf*, *156*(4), 454–468.

Messer, D. J. (1978). The integration of mothers' referential speech with joint play. *Child Development*, *49*(3), 781–787.

Miller, S. A. (1988). Parents' beliefs about children's cognitive development. *Child Development*, *59*(2), 259–285.

Miller, J. L., Ables, E. M., King, A. P., & West, M. J. (2009). Different patterns of contingent stimulation differentially affect attention span in prelinguistic infants. *Infant Behavior and Development*, *32*(3), 254–261.

Miller, J., & Chapman, R. (1985). *Systematic analysis of language transcripts*. Madison, WI: Language Analysis Laboratory.

Mitchell, T. V., & Maslin, M. T. (2007). How vision matters for individuals with hearing loss. *International Journal of Audiology*, *46*(9), 500–511.

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.

O'Brien, M., & Nagle, K. J. (1987). Parents' speech to toddlers: The effect of play context. *Journal of Child Language*, *14*(2), 269–279.

Pellegrini, A. D., Brody, G. H., & Sigel, I. E. (1985). Parents' book-reading habits with their children. *Journal of Educational Psychology*, *77*(3), 332–340.

Pereira, A. F., Smith, L. B., & Yu, C. (2014). A bottom-up view of toddler word learning. *Psychonomic Bulletin & Review*, *21*(1), 178–185.

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.

Rohde, H., & Frank, M. C. (2014). Markers of topical discourse in child-directed speech. *Cognitive Science*, *38*(8), 1634–1661.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2002). Processing speed in the 1st year of life: A longitudinal study of preterm and full-term infants. *Developmental Psychology*, *38*(6), 895–902.

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2004). Infant visual recognition memory. *Developmental Review*, *24*(1), 74–100.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2011). The structure of memory in infants and toddlers: An SEM study with full-terms and preterms. *Developmental Science*, *14*(1), 83–91.

Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1), 185–205.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774.

Ruff, H. A., & Capozzoli, M. C. (2003). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, *39*(5), 877–890.

Ruff, H. A., Lawson, K. R., Parrinello, R., & Weissberg, R. (1990). Long-term stability of individual differences in sustained attention in the early years. *Child Development*, *61*(1), 60–75.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19.

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.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, *43*(2), 549–565.

Spekman, N. J., & Roth, F. P. (1985). Preschool children's comprehension and production of directive forms. *Journal of Psycholinguistic Research*, *14*(3), 331–349.

Spencer, P. E., Bodner-Johnson, B. A., & Gutfreund, M. K. (1992). Interacting with infants with a hearing loss what can we learn from mothers who are deaf? *Journal of Early Intervention*, *16*(1), 64–78.

Suanda, S. H., Smith, L. B., & Yu, C. (2016). The multisensory nature of verbal discourse in parent–toddler interactions. *Developmental Neuropsychology*, *41*(5–8), 324–341.

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.

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. H. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science*, *20*(6), e12456.

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.

Tamis-LeMonda, C. S., Song, L., Leavell, A. S., Kahana-Kalman, R., & Yoshikawa, H. (2012). Ethnic differences in mother–infant language and gestural communications are associated with specific skills in infants. *Developmental Science*, *15*(3), 384–397.

Tomasello, M. (1988). The role of joint attentional processes in early language development. *Language Sciences*, *10*(1), 69–88.

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.

Tyack, D., & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language*, *4*(2), 211–224.

Vaccari, C., & Marschark, M. (1997). Communication between parents and deaf children: Implications for social-emotional development. *Journal of Child Psychology and Psychiatry*, *38*(7), 793–801.

VanDam, M., Ambrose, S. E., & Moeller, M. P. (2012). Quantity of parental language in the home environments of hard-of-hearing 2-year-olds. *Journal of Deaf Studies and Deaf Education*, *17*(4), 402–420.

Vassallo, S., Cooper, S. L., & Douglas, J. M. (2009). Visual scanning in the recognition of facial affect: Is there an observer sex difference? *Journal of Vision*, *9*(3), 1–10.

Wedell-Monnig, J., & Lumley, J. M. (1980). Child deafness and mother-child interaction. *Child Development*, *51*(3), 766–774.

Weisleder, A., & Fernald, A. (2013). Talking to children matters early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

West, K. L., & Iverson, J. M. (2017). Language learning is hands-on: Exploring links between infants' object manipulation and verbal input. *Cognitive Development*, *43*, 190–200.

Yoder, P. J., Kaiser, A. P., Alpert, C., & Fischer, R. (1993). Following the child's lead when teaching nouns to preschoolers with mental retardation. *Journal of Speech, Language, and Hearing Research*, *36*(1), 158–167.

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*, *26*(9), 1235–1240.

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.