

Words are not enough: how preschoolers' integration of perspective and emotion informs their referential understanding*

SUSAN A. GRAHAM, VALERIE SAN JUAN AND
MELANIE KHU
University of Calgary

*(Received 30 June 2016 – Revised 25 August 2016 – Accepted 1 September 2016 –
First published online 7 November 2016)*

ABSTRACT

When linguistic information alone does not clarify a speaker's intended meaning, skilled communicators can draw on a variety of cues to infer communicative intent. In this paper, we review research examining the developmental emergence of preschoolers' sensitivity to a communicative partner's perspective. We focus particularly on preschoolers' tendency to use cues both within the communicative context (i.e. a speaker's visual access to information) and within the speech signal itself (i.e. emotional prosody) to make on-line inferences about communicative intent. Our review demonstrates that preschoolers' ability to use visual and emotional cues of perspective to guide language interpretation is not uniform across tasks, is sometimes related to theory of mind and executive function skills, and, at certain points of development, is only revealed by implicit measures of language processing.

INTRODUCTION

“When I use a word,” Humpty Dumpty said, in rather a scornful tone, “it means just what I choose it to mean—neither more nor less.” “The question is,” said Alice, “whether you can make words mean so many

* This work was supported by funds from the Canada Foundation for Innovation, the Canada Research Chairs program, and the University of Calgary, and by an operating grant from the Social Sciences and Humanities Research Council of Canada awarded to Susan Graham. Valerie San Juan was supported by a postdoctoral fellowship from SSHRC and an Eyes High Fellowship from the University of Calgary. We are very grateful to our collaborators on the research reviewed in this paper: Jared Berman, Craig Chambers, and Elizabeth Nilsen. We also thank Nina Anderson for her assistance with the preparation of the manuscript. Address for correspondence: S. Graham, Dept. of Psychology, University of Calgary, Calgary AB, T2N 1N4, Canada; e-mail: susan.graham@ucalgary.ca

different things.” “The question is,” said Humpty Dumpty, “which is to be master—that’s all.”

(Lewis Carroll, *Through the Looking Glass*)

As so cleverly illustrated by this exchange between Humpty Dumpty and Alice, inferring a speaker’s intended meaning cannot always be accomplished through words alone. Consider, for example, the following situation: a child looks at her bookshelf and says to her parent “Can you get the book?” Given that there are multiple possible referents (i.e. books) available, how does the parent infer the child’s intended meaning? In the face of this indeterminacy, listeners can use a variety of cues to infer the child’s intended meaning. For example, the parent may consider whether the child has a favourite book she always wants to read; whether there is a particular book the parent, but not the child, can reach; whether there is a book that is not visible to the child and thus can be excluded from consideration; or whether the child sounds happy because a brand new book is on the shelf. As demonstrated by this example, skilled listeners can draw upon information about a speaker’s perspectives to gauge that speaker’s communicative intent. This ability to use information about a speaker’s perspective to make inferences about that speaker’s intended meaning is known as communicative perspective taking.

Communicative situations like the one described in the example above are likely frequently encountered in everyday interactions. Thus, core questions arise around children’s abilities to attend to and integrate other’s perspectives during communicative interactions and whether these perspectives can be integrated rapidly enough to guide language processing in the moment. In this paper, we review research examining the developmental emergence of preschoolers’ sensitivity to a communicative partner’s perspective. We focus particularly on preschoolers’ tendency to use cues both within the communicative context (i.e. a speaker’s visual access to information) and within the speech signal itself (i.e. emotional prosody) to make on-line inferences about communicative intent. First, we review research examining the emergence of communicative perspective taking during the first two years of development, with particular focus on children’s attention towards others’ visual perspectives. Next, we introduce the visual world paradigm as a means of examining HOW cues of perspective become integrated with on-line spoken language processing. We then review research examining children’s sensitivity to a speaker’s visual perspective and emotional prosody in referential communication, addressing current issues in these research areas. We conclude with empirical challenges and future directions.

THE EMERGENCE OF VISUAL PERSPECTIVE-TAKING AND COMMUNICATIVE ABILITIES

Visual perspective taking involves tracking what another person can see in order to form inferences about their knowledge and intentional actions (Moll & Meltzoff, 2011a). For example, knowing that a person cannot see a toy that is hidden by a barrier may lead one to infer that she is unaware of the toy's presence. Around the same time that infants begin to engage in verbal communicative interactions, they also begin to track and reason about the perspectives of others. That is, studies using looking-time measures have found evidence of perspective taking emerging just after infants reach their first birthdays (Caron, Kiel, Dayton & Butler, 2002; Dunphy-Lelii & Wellman, 2004; Luo & Baillargeon, 2007). For example, 14-month-olds will selectively follow the gaze of another person whose visual access to items is not occluded by either a physical barrier (Caron *et al.*, 2002; Dunphy-Lelii & Wellman, 2004) or a blindfold (Brooks Meltzoff, 2002). Similarly, 12.5-month-old infants will track an agent's visual access to a desired item and use the information to interpret the agent's subsequent actions (Luo & Baillargeon, 2007). When assessed explicitly via verbal or behavioural selection responses, visual perspective-taking abilities become evident around two years of age (Moll & Meltzoff, 2011b). For example, 24-month-olds, but not 18-month-olds, will correctly respond to an adult who is searching for a toy ("Where is it? I cannot find it") by selecting an item hidden from the adult (Moll & Tomasello, 2006).

Given the early development of visual perspective taking, when do children first begin to consider the visual perspectives of others in communicative interactions? The first studies to examine this question suggested that before children reach school-age, they are largely egocentric in their referential communication and fail to integrate feedback from their communicative partner (e.g. Glucksberg & Krauss, 1967; Krauss & Glucksberg, 1969). However, advancements in both methods and technology have led to more sensitive means of assessing children's visual perspective taking. We now know that the ability to integrate perspective-taking and communication abilities emerges during infancy and shows marked improvement throughout the preschool years.

Between 12 and 18 months of age, infants begin to differentially adapt their pointing gestures to communicate object location to both knowledgeable and unknowledgeable agents (Liskowski, Carpenter & Tomasello, 2008). During this same period, infants will also vary their interpretation of communicative behaviours (e.g. eye-gaze and emotional reactions towards an object) depending on the visual perspective of their communicative partner (Moll & Tomasello, 2004; Moses, Baldwin,

Rosicky & Tidball, 2001). By the end of their second year, infants begin to use the perspectives of others to disambiguate spoken language. Specifically, in word learning studies, researchers have shown that infants as young as 18 months will attend to where a speaker is looking to correctly infer the referent of a novel label (e.g. Baldwin, 1991, 1993; Tomasello, Strosberg & Akhtar, 1996). By two years of age, children will monitor what a person has or has not seen and will adapt their verbal requests for items to match the knowledge state of their listener (Nayer & Graham, 2006; O'Neill, 1996). Overall, these findings suggest that as soon as infants begin to reason about the visual perspectives of others, they begin to also use this information to inform their interpretation and production of both non-verbal and verbal communicative behaviours.

In summary, the ability to integrate visual perspective taking in receptive and productive communication begins to emerge during the second year of life. In the next section, we shift our focus to research that has begun to examine HOW children develop the ability to integrate perspective-taking abilities with on-line language processing. We begin with a brief overview of the visual world paradigm as used in referential communication experiments.

THE VISUAL WORLD PARADIGM

The visual world paradigm is the basic method used to study spoken language comprehension in real time, drawing upon the systematic relation between eye-movements and language processing (Allopenna, Magnuson & Tanenhaus, 1998; Sedivy, Tanenhaus, Chambers & Carlson, 1999; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). In this paradigm, researchers track participants' eye-movements as they respond to spoken instructions in the context of a visual display (see Huettig, Rommers & Meyer, 2011; Snedeker & Huang, 2016, for recent reviews of the paradigm). Using this paradigm, research has demonstrated that spoken language is processed incrementally – that is, both child and adult listeners interpret words and sentences as they unfold over time, rather than waiting to hear an entire sentence before making inferences about a speaker's intended meaning (e.g. Allopenna *et al.*, 1998; Swingley, Pinto & Fernald, 1999; Tanenhaus *et al.*, 1995; Trueswell, Sekerina, Hill & Logrip, 1999). Furthermore, this incremental interpretation occurs in real time, with listeners launching eye-movements to intended referents within the first few hundred milliseconds of hearing a target word (e.g. Tanenhaus *et al.*, 1995; Trueswell *et al.*, 1999).

Research using the visual world paradigm led to fundamental insights into the interactive nature of the language processing system – that is, adult and child listeners integrate linguistic, paralinguistic, and non-linguistic

information in real time to guide their interpretations of utterances (e.g. Chambers, Tanenhaus & Magnuson, 2004; Collins, Graham & Chambers, 2012; Graham, Sedivy & Khu, 2014; Sedivy, 2003; Snedeker & Trueswell, 2004; Trueswell *et al.*, 1999). To illustrate, a seminal study by Chambers, Tanenhaus, Eberhard, Filip, and Carlson (2002) examined how adult listeners coordinate linguistic and non-linguistic information when listening to referential statements. In this study, adult participants were instructed to manipulate physical objects (e.g. “Put the cube inside the can”), in the context of displays where there were two possible candidate referents (e.g. a large can and a small can). The size of the theme object (e.g. the can) was varied across conditions such that it could either fit in both containers or only in one container. From the earliest moments of processing, listeners’ visual attention was restricted to only those containers large enough to accommodate the object, indicating that they were rapidly integrating contextual information and knowledge of the possible actions with the unfolding utterance.

The visual world paradigm has also been used to examine the timing and integration of visual perspective taking during on-line referential communication (e.g. Brown-Schmidt & Heller, 2014; Hanna, Tanenhaus & Trueswell, 2003). In this variation, a discrepancy of perspective is established between a listener and a speaker by varying the physical co-presence of objects available for reference on a visual display. For example, a listener may hear an instruction to manipulate a target referent (e.g. “Pick up the duck”) on a display where only one of two candidate referents is mutually available to both themselves and the speaker (e.g. one of two ducks is occluded from the speaker’s view). If listeners use the perspective of their speaker to constrain their interpretations of reference, then they should ignore items on the display that their speaker cannot see – i.e. PRIVILEGED GROUND INFORMATION – in favour of items that are mutually visible to both themselves and the speaker – i.e. COMMON GROUND INFORMATION. The type of information a listener considers (i.e. privileged ground vs. common ground) during referential interpretation can be measured via their eye-gaze towards display items as a critical sentence unfolds on-line. In this way, the visual world paradigm offers a valuable means of assessing both the types of perspective cues that listeners consider as they interpret reference on-line as well as the timing with which perspective information becomes integrated with linguistic input. In the next section, we review developmental research that has used the visual world paradigm to examine visual perspective taking during the preschool years, with particular focus on research examining how preschoolers interactively coordinate visual perspective information with the linguistic properties of unfolding referential statements.

VISUAL PERSPECTIVE TAKING DURING ON-LINE COMMUNICATION

To date, the majority of experimental studies that have examined children's perspective taking using a visual world paradigm have focused on one type of perspective reasoning – namely, reasoning about information that is visually shared or not shared between themselves and a speaker (e.g. Nadig & Sedivy, 2002; Epley, Morewedge & Keysar, 2004). This research has yielded valuable insights into two key issues: (i) when preschoolers begin to use visual perspective taking to guide their on-line comprehension and production of referential utterances; and (ii) the timecourse and efficiency with which preschoolers recruit perspective information during on-line language processing.

Preschoolers' use of visual perspective taking to guide referential communication

During the preschool years, children undergo significant improvements in their ability to integrate visual perspective with both the comprehension and production of referential statements (e.g. Matthews, Lieven, Theakston & Tomasello, 2006; Nadig & Sedivy, 2002). In a series of studies in our lab, we have assessed preschoolers' sensitivity to others' visual perspectives in both productive and receptive language, examining the emergence of these abilities during the preschool years.

In one of our first studies (Nilsen & Graham, 2009), we examined three- to five-year-olds' integration of visual perspective taking in a comprehension task, where children had to follow a speaker's instructions to retrieve objects on a display. We also examined four- to five-year-olds' ability to use a listener's visual perspective in a production task, where children had to instruct an experimenter to retrieve objects on a display. In both tasks, we examined whether children's explicit responses varied with the visual perspective of their communicative partner. On the comprehension task, we also examined whether children's implicit eye-gaze towards display items would be constrained by visual perspective cues. Results of the comprehension task indicated that three- to five-year-olds accurately tracked what a speaker could see in order to correctly interpret the referent of an ambiguous utterance. That is, when interpreting an ambiguous instruction (e.g. "Pick up the duck" in a display with two ducks), children were more likely to constrain their visual attention towards items that were mutually visible than to items that were exclusively visible to themselves. Across both experiments, children were also more likely to select items that were visible to themselves and the speaker.

The results of the production task showed that four- to five-year-olds considered their listener's visual perspective, when forming their own instructions. That is, children used more adjectives to request a target referent (e.g. "Pick up the *big* duck") when their communicative partner

had visual access to two competing referents rather than one unambiguous referent. Other research has shown that children as young as three years of age will similarly adapt their productions to fit a listener's perspective (Matthews *et al.*, 2006), but not in contexts where the child and listener's perspective are simultaneously competing for the child's attention or where cues of visual perspective change on a trial by trial basis. Our findings, using a visual world paradigm, therefore demonstrate that preschoolers are able to selectively and flexibly track the visual perspective of their listener in order to adapt their production of referential utterances (see also Nadig & Sedivy, 2002).

Thus, around three to four years of age, children can use a speaker's perspective to guide their referential interpretations and begin to adapt the clarity of their own messages to match the visual perspective of their listener. In the next set of studies, we asked whether preschoolers can take this understanding one step further and use visual perspective information to evaluate the clarity of an utterance from the perspective of another person (Nilsen & Graham, 2012; Nilsen, Graham, Smith & Chambers, 2008). Message evaluation is a critical component of referential comprehension, as detection of sentence ambiguity could highlight to the listener the need to rely on non-linguistic cues of reference such as visual perspective. In this third-party paradigm, a sticker is hidden in a location and children either share the speaker's perspective (i.e. see where the sticker was placed), or share the message recipient's perspective (i.e. do not see the sticker's location). The message recipient is provided with a statement about the sticker location that is either ambiguous (e.g. "it's under the rubber duck" in the presence of two rubber ducks) or unambiguous (e.g. "it's under the big duck" in the presence of a big and a small rubber duck). After hearing the statement, children are asked to evaluate the message recipient's knowledge of the sticker location and the quality of message (e.g. "Was that a good clue or a tricky clue"; see also Robinson & Robinson, 1982; Sodian, 1988). Thus, in this paradigm, children must ignore their own perspective in order to interpret the quality of a message from the perspective of another person.

Using this third-party paradigm, we conducted a longitudinal study to examine children's implicit and explicit message evaluation between the ages of four and five years (Nilsen & Graham, 2012). Our results demonstrated that, at four years of age, children only demonstrated implicit sensitivity to message ambiguity. That is, even when children were aware of the sticker's location, they gazed equally towards both locations when hearing an instruction that was exclusively ambiguous to the other person. By 4.5 years of age, children began to show evidence of explicit message evaluation: first recognizing when a message was sufficiently clear for the message recipient to interpret reference, and then

later, at five years of age, recognizing when a message was too ambiguous for an agent to infer reference. Implicit sensitivity to message ambiguity at four years of age, however, was not predictive of later developing explicit message evaluation.

In summary, by using variations of the a visual world paradigm, we have found that preschool children integrate visual perspective taking to constrain both implicit and explicit comprehension of referential statements by as early as three years of age. The ability to flexibly use visual perspective taking to inform the explicit evaluation and production of referential statements also begins to emerge between four and five years of age. In the case of message evaluation, however, implicit awareness of message ambiguity may be evident before children are able to explicitly judge the quality of a spoken utterance.

Visual world paradigms, however, are not only useful for developmental trajectories. They also provide a unique means of assessing the timecourse of communicative perspective taking. In the following section we review studies that have begun to examine how rapidly and efficiently children integrate visual perspective taking with on-line language processing.

Timing of preschoolers' recruitment of visual perspective information

Because spoken language is processed incrementally as it unfolds in real time, perspective inferences must be rapidly generated so that this information is coordinated with other cues of reference. The question of when, during sentence processing, perspective cues become integrated with linguistic input has been the subject of a lively debate in the adult literature, with proponents advocating for both early and late integration accounts. Early integration accounts propose that individuals are inherently motivated to track their communicative partner's perspective, and thus perspective constraints are considered from the earliest moments of sentence processing (Brown-Schmidt & Heller, 2014; Heller, Parisien & Stevenson, 2016). According to these accounts, the ability to use perspective information to constrain the interpretation of a spoken utterance depends on the strength of these cues relative to other sources of information (e.g. ambiguity of linguistic input, number of competing referents on a display, etc.). If perspective cues are strongly represented, then evidence of perspective-taking integration should be seen as a sentence is unfolding. Conversely, late integration accounts propose that perspective constraints may not always be available to influence the earliest moments of sentence processing (Apperly, Carroll, Samson, Humphreys, Qureshi & Moffitt, 2010; Keysar, 2007). According to these accounts, individuals do not always track perspective cues automatically, and the cognitive demands associated with generating perspective inferences would make it inefficient for the language processing

system to coordinate these cues with other sources of information during on-line sentence processing. As a result, late integration accounts predict that perspective cues are often not considered until after a spoken utterance has been heard and linguistic input has been processed.

To date, only a few studies have examined the timing of children's perspective taking during on-line sentence processing (Epley *et al.*, 2004; Nadig & Sedivy, 2002). In one of the first studies to address this question, Nadig and Sedivy (2002) examined five- and six-year-olds' ability to interpret referential instructions (e.g. "Pick up the duck") using displays that contained four items, two of which were referential matches for the critical noun (i.e. two similar ducks). Children were significantly faster at identifying the referent on trials where the speaker had visual access to only one of the two candidate referents (i.e. privileged ground trials) vs. trials where the speaker could see both candidate referents (i.e. common ground trials). Eye-gaze data further demonstrated that, on privileged ground trials, children began to constrain their attention towards the target while the instruction was still being heard (approximately 200–760 ms after the onset of the noun). These findings demonstrate that children integrated perspective cues early to constrain their interpretation of a referential statement, as it was still unfolding. A recent study in our lab yielded similar results with younger children (Khu, Chambers & Graham, unpublished observations). That is, we found that four-year-olds selectively used common ground information to guide their interpretation of referential statements within the earliest moments of processing (i.e. as soon as the critical noun began to unfold).

In contrast, Epley and colleagues (2004) found sensitivity to another's visual perspective emerged much later in sentence processing. In this study, the referential comprehension of both children, ranging in age from four to twelve years, and adults was examined using a similar but more challenging procedure than Nadig and Sedivy (2002). Participants followed instructions that contained size or spatial ambiguity (e.g. "Move the small truck" in a display with multiple trucks) on displays that contained nine items, rather than four items. A set of three display items matched the critical noun (e.g. three trucks of ascending sizes), but the strongest referential candidate was always occluded from the view of the speaker (e.g. the smallest truck was hidden behind a screen). The target referent was thus the best referential candidate that could be seen by both the listener and the speaker (i.e. the medium-sized truck). Eye-movement patterns indicated that both children and adults considered privileged ground items first before shifting their focus to target items in common ground. While adults were able to recover their attention towards the target early enough (an average of 1329 ms following the offset of the instruction) to produce an accurate reaching response, children's recovery was

significantly slower (an average of 3647 ms after the offset of the instruction) and often led to inaccurate reaching responses. These results suggest that both children and adults demonstrated late integration of perspective information, although adults were better able to use this information to correct, if not constrain, their interpretation of a referential statement.

Overall, these findings suggest that perspective information is available to listeners early in speech processing; however, when this information is integrated with linguistic information may depend on the complexity of the communicative task (see San Juan, Khu & Graham, 2015 for a discussion). That is, children are less likely to show early integration of perspective information when there is more contextual information to consider (i.e. more display items) and there is greater competition between different sources of information. Thus, in the Epley *et al.* (2004) task, children's representation of the speaker's perspective may have been outweighed by the relative strength of other competing cues of reference (e.g. the fact that the item in privileged ground was a stronger referential match to the linguistic input). Alternatively, children in this task may have had more difficulty generating inferences about their speaker's perspective because there was more direct conflict between their own and their communicative partner's perspective (Moll, Meltzoff, Merzsch & Tomasello, 2013). If children were less efficient at generating perspective inferences in this type of context, then they would not have been able to consider this information until well after the linguistic input had been processed.

Summary of visual perspective taking and communication

Application of the visual world paradigm has provided developmental researchers with a more sensitive means of assessing the implicit and explicit integration of visual perspective taking during spoken language processing. This has led to a more detailed understanding of when communicative abilities emerge during the preschool years. These methods have also expanded the opportunity to examine the timing and efficiency with which children integrate visual perspective taking during communication. However, further research in this area is necessary for understanding the underlying mechanisms of communicative perspective taking. That is, more studies are needed to clarify the contextual and cognitive factors that influence children's integration of visual perspective taking during the earliest moments of sentence processing.

EMOTIONAL PERSPECTIVE TAKING DURING ON-LINE COMMUNICATION

Attending to and integrating another's visual perspective represents only one aspect of communicative perspective taking. Indeed, it is conceivable that the

need to monitor the emotional perspective of a communicative partner could be more socially relevant, with lack of attention to emotion potentially more socially consequential in a communicative interaction than the need to consider a partner's visual perspective. One means through which speakers may signal their emotional state or disposition is through the emotional prosody that accompanies an utterance. Emotional prosody refers to paralinguistic information that signals a speaker's emotional state or disposition, as indexed by variations in pitch contours, speech rate, intensity, and pitch level (Banse & Scherer, 1996; Frick, 1985). Emotional prosody is often consistent with an utterance's linguistic content (i.e. a speaker's sadness is communicated both through her words and her emotional prosody for the statement "I'm having a bad day" spoken in a sad tone of voice). When linguistic information alone does not fully disambiguate meaning, however, emotional prosody alone can provide clarification. For example, statements like "School starts tomorrow" or "I got the reviews on my manuscript", can convey markedly different meanings if spoken with a happy-sounding voice versus a sad-sounding voice.

In this next section, we consider preschoolers' sensitivity to a speaker's emotional perspective, as signalled by their emotional prosody, to guide inferences about communicative intent. Specifically, we review research documenting: (i) the emergence of preschoolers' sensitivity to emotional prosody to resolve communicative ambiguity; (ii) valence and timing differences in preschoolers' sensitivity to emotional prosody; and (iii) sensitivity to emotional prosody and communicative perspective taking.

Developmental emergence of sensitivity to emotional prosody in communication

In the first year of life, infants display sensitivity to emotional prosody. Infants as young as 1 month of age show preferences for infant-directed speech, which has distinct prosodic modifications that typically convey positive affect, over adult-directed speech (Cooper & Aslin, 1990; Fernald, 1985; Singh, Morgan & Best, 2002). During this first year, infants begin to discriminate the different intonational patterns used by mothers to convey distinct communicative intent types (i.e. comforting or soothing, affection or approval, and directive affect; e.g. Fernald, 1989, 1992, 1993; Kitamura & Burnham, 2003; Kitamura & Lam, 2009). Furthermore, infants will respond in an appropriate manner to different types of emotional prosody. For example, 5-month-old infants smile more when hearing approval vocalizations produced in infant-directed speech than when hearing prohibition vocalizations, even if these vocalizations are produced in an unfamiliar language (Fernald, 1993). Thus, even before the

onset of productive language, infants detect and respond to emotional prosody.

As children acquire language, they must learn to integrate emotional prosody with linguistic information. From a research standpoint, this issue has been investigated from two directions: first, examining children's relative attention to information conveyed by linguistic content versus that conveyed by emotional prosody when the two information sources are in conflict (e.g. Friend, 2000; Morton & Munakata, 2002; Morton, Trehub & Zelazo, 2003), and second, examining preschoolers' attention to emotional prosody when the linguistic content is indeterminate, rather than in conflict with emotional cues (e.g. Berman, Chambers & Graham, 2010; Berman, Graham & Chambers, 2013; Berman, Graham, Callaway, & Chambers, 2013). Both lines of research have yielded insights into the developmental emergence of children's ability to integrate emotional prosody with linguistic content.

Children's sensitivity to conflicting linguistic and emotional prosody cues. Research examining children's resolution of conflicting linguistic and emotional prosody cues indicates that children's sensitivity to emotional prosody shifts during infancy and the preschool and school-age years (Friend, 2000; Friend & Bryant, 2000). At the early stages of language development, 15-month-olds rely on emotional prosody to guide their behaviour, when emotional prosody and lexical content provide incongruent messages (Friend, 2001). As children reach preschool age, they are more likely to rely on the linguistic content of an utterance over emotional prosody, when the two sources of information conflict. For example, Morton and Trehub (2001) presented four- to ten-year-olds with sentences that described either happy or sad events (e.g. "I got an ice cream for being good", for a happy event), spoken with both positive (happy-sounding) and negative (sad-sounding) emotional prosody. When presented with conflicting information (i.e. a sentence describing a sad event paired with happy emotional prosody), four- to eight-year-olds relied almost exclusively on the content of the sentences to judge the emotional state of the speaker. By nine years of age, children began to decrease their reliance on the linguistic content of the utterances and, like adults, used the speaker's emotional prosody to gauge the speaker's emotional state. In a subsequent study, Morton and Munakata (2002) demonstrated that preschoolers' adherence to linguistic content over emotional prosody persists even when they are explicitly instructed to attend to emotional prosody.

The findings described above suggest that four- to eight-year-olds prioritize lexical content over emotional prosody in these conflict paradigms. More implicit measures, however, suggested that children are not fully disregarding the information signalled by the emotional prosody.

Specifically, in the Morton and Trehub (2001) experiments, children showed longer response latencies on conflict trials, indicating that they recognized the incongruity between the two sources of information. Thus, preschoolers' tendency to privilege linguistic information over emotional prosody in these tasks likely reflects difficulty resolving conflicting sources of information, rather than a failure to recognize the meaning of emotional prosody (Morton *et al.*, 2003; Waxer & Morton, 2011).

Children's use of emotional prosody to resolve linguistic indeterminacy. In a series of studies in our lab, we have approached the question of preschoolers' integration of linguistic content and emotional prosody from a different direction. Rather than using conflict tasks, we asked whether preschoolers might show greater sensitivity to emotional prosody in tasks that are less cognitively demanding—namely, when the linguistic information is indeterminate, rather than in conflict with emotional prosody. We also reasoned that employing a visual world paradigm and measuring both explicit behavioural responses (i.e. pointing) and eye-movements would allow us to gain insight into both preschoolers' real-time processing of, and more conscious and controlled responses to, emotional prosody.

In the first study to address this question, we presented three- and four-year-olds with formally ambiguous referential descriptions (i.e. “Look at the ball”, in the presence of more than one ball) and examined whether they would use emotional prosody to identify the speaker's intended referent (Berman *et al.*, 2010). On each trial, preschoolers saw arrays that contained three photographed objects: two objects of the same category that varied in their physical state (e.g. an intact ball and a deflated ball) and an unrelated object (e.g. a star). Children were instructed to find one of the two objects belonging to the same category using an ambiguous phrase (e.g. “Look at the ball”), spoken using one of three different types of emotional prosody (happy-sounding, sad-sounding, or neutral). Four-year-olds' eye-gaze patterns, but not their pointing responses, demonstrated appropriate sensitivities to emotional prosody. As the ambiguous noun unfolded, children fixated the broken object most often when hearing sad-sounding emotional prosody, less when hearing neutral prosody, and much less when hearing happy-sounding prosody. This effect emerged only during the noun region: during the early part of the utterance (i.e. “Look at the”) there was no influence of emotional prosody on eye-gaze behaviour. Neither three-year-olds' eye-gaze patterns nor their pointing behaviour reflected any sensitivity to emotional prosody.

Results from this study suggest that there is a developmental progression in the use of emotional prosody for language comprehension between three and four years of age. Four-year-olds, however, appear to be in a transitional period in their ability to integrate emotional prosody with

linguistic information, as their sensitivity to emotional prosody was not reflected in their explicit behavioural decisions. In a subsequent study, using the same paradigm, we found that five-year-olds evidenced use of emotional prosody to guide referential understanding, when assessed with both eye-gaze measures and pointing measures (Berman, Graham, & Chambers, 2013, Experiment 1). We further documented this developmental transition between four and five years in another set of studies examining preschoolers' use of emotional prosody to learn new words (Berman, Graham, Callaway, & Chambers, 2013). In these experiments, we presented four- and five-year-olds with two novel objects, first in their original state and second in an altered state (broken or enhanced). Children heard an instruction to find the referent of a novel word, produced with sad-sounding, neutral, or happy-sounding emotional prosody. Both four- and five-year-olds' gaze patterns indicated that they linked the novel word with the object that best matched the speaker's emotional prosody (e.g. the broken object when the instruction was produced with sad-sounding affect, the enhanced object when the instruction was produced with happy-sounding prosody). Only five-year-olds, however, demonstrated their use of emotional prosody in their explicit referential decisions.

Taken together, these findings indicate that, between four and five years of age, preschoolers move from an implicit understanding of emotional prosody in referential communication tasks to a more explicit use of this cue. Three-year-olds, however, did not appear to show any evidence of integrating emotional prosody with linguistic information. What might account for the three-year-olds' apparent lack of success in such tasks? We addressed this question in a recent study, examining specifically whether three-year-olds' difficulties in our earlier study stemmed from an inability to identify the acoustic cues corresponding to different types of emotional prosody (Berman, Chambers & Graham, 2016). Here, we presented three- and five-year-olds with utterances produced with happy-sounding, neutral, or sad-sounding emotional prosody in the presence of faces depicting happy, neutral, or sad facial expressions. Children were instructed to point to the face that reflected how the speaker was feeling when she made a specific utterance. Only five-year-olds pointed to the face that matched the utterance's emotional prosody, providing further evidence of the developmental changes in sensitivity to emotional prosody between three and five years. In contrast, both three-year-olds' and five-year-olds' gaze patterns demonstrated that they could link happy-sounding and sad-sounding emotional prosody to the appropriate emotional face. Matching neutral emotional prosody to neutral faces proved difficult for children of both ages. These results suggest that three-year-olds can recognize happy-sounding and sad-sounding emotional prosody and link it to the

appropriate facial expression. Thus, the difficulties demonstrated by three-year-olds in the Berman *et al.* (2010) study are likely isolated to the process of linking vocal affect with the intent to refer to objects.

In summary, children's success at integrating emotional prosody with linguistic information varies during the preschool years as a function of communicative task. That is, young children are better able to use emotional prosody to infer communicative intent when linguistic information is indeterminate (Berman *et al.*, 2010; Berman, Graham, & Chambers, 2013) versus when linguistic information is in conflict with emotional prosody (e.g. Morton & Trehub, 2001). Furthermore, there is significant progression in children's abilities to coordinate emotional and linguistic information between three and five years of age, with three-year-olds showing implicit sensitivity to emotional prosody only under some conditions and five-year-olds demonstrating more robust integration of these two sources of information.

Valence effects. In addition to developmental differences, our studies on preschoolers' integration of emotional prosody with lexical content have documented valence differences in children's sensitivity to emotional prosody, both in terms of the types of representation created and in the timecourse of processing prosody in the unfolding speech stream. First, although five-year-olds will use both positive and negative emotional prosody to map a novel word to a novel object, children were only successful at extending and generalizing these newly mapped words when the words were learned using negative vocal affect (Berman, Graham, Callaway, & Chambers, 2013). This finding suggests that negative emotional prosody (versus positive emotional prosody) enabled children to establish a more robust representation of the word in this task.

Second, our studies have demonstrated comparatively greater sensitivity to negative-sounding emotional prosody versus positive-sounding emotional prosody in the earliest moments of speech processing. Specifically, when five-year-olds were presented with unambiguous referential contexts, they used negative emotional prosody early in the utterance to anticipate a particular referential outcome (Berman, Graham, & Chambers, 2013). That is, when presented with an unambiguous referential description produced with negative emotional prosody (e.g. "Look at the ball", in the presence of a ball, a duck, and a cellphone; Experiments 2 & 3), children began to anticipate reference to the one broken object in the scene well before the disambiguating noun. The effect of positive emotional prosody, in contrast, was not observed until after the onset of the noun. Similarly, the gaze patterns of both three- and five-year-olds showed that sad-sounding emotional prosody led children to identify a sad face in the first 800 ms of an unfolding utterance (Berman *et al.*, 2016). In contrast,

children did not identify a happy face on the basis of happy-sounding speech until approximately 800 ms into the utterance.

Our findings are consistent with other research demonstrating an advantage for negative emotional prosody versus positive emotional prosody, in terms of both accuracy and timing of emotion recognition. For example, three- to five-year-olds more accurately identify sadness, on the basis of paralinguistic cues, versus happiness, anger, or fear (Nelson & Russell, 2011). Similarly, adults are more successful at using emotional prosody to identify sadness versus happiness in a speaker's voice, even if utterances are produced in a foreign language (Paulmann & Pell, 2011; Pell, Monetta, Paulmann & Kotz, 2009; Pell, Paulmann, Dara, Allasseri & Kotz, 2009; Scherer, Banse & Wallbott, 2001). Furthermore, adults, like the preschoolers in our studies, are significantly quicker to identify sad vocal emotion versus positive vocal emotion. In one study, for example, it took adults approximately 400 ms longer to recognize happiness compared to sadness, on the basis of emotional paralinguistic cues (Pell & Kotz, 2011). Finally, the timing advantage for negative-sounding emotional prosody observed in our studies and those of others is generally consistent with the proposal that both adults and children are biased towards negative information when processing information (Rozin & Royzman, 2001; Vaish, Grossmann & Woodward, 2008).

Does sensitivity to emotional prosody reflect perspective taking? The research reviewed above documents the critical role of emotional prosody in spoken language comprehension. These studies, however, do not unequivocally demonstrate that children are using emotional prosody to reason about a speaker's emotional perspective. That is, preschoolers' use of emotional prosody to resolve communicative ambiguity could arise from established associative links between vocal patterns and their own emotional reactions (e.g. I would be sad if my beachball was deflated) or object states, rather than inferences about a speaker's perspective.

In a recent study in our lab, we developed an on-line communicative perspective-taking task that more clearly tested whether preschoolers could use emotional prosody to reason about a speaker's emotional perspective (Khu, Chambers & Graham, unpublished observations). In this task, four-year-olds played a competitive game with a speaker, in which a 'loss' for the child meant a 'win' for the speaker, and vice versa. Accordingly, children could not rely on their own emotional reactions or previous associations to infer the speaker's emotional state and communicative intent (e.g. when the speaker sounded sad, it corresponded to a win for the child). Children's eye-gaze was tracked and their responses recorded as they heard ambiguous statements spoken with either happy- or sad-sounding emotional prosody. The implicit gaze measures indicated that preschoolers used the speaker's emotional perspective to influence their

on-line language comprehension. For example, their eye-gaze patterns indicated that they anticipated that they would lose and that the speaker would win when the speaker sounded happy. The influence of emotional prosody on children's interpretations did not occur until after the utterance had ended, suggesting that this information exerted relatively late effects on children's language processing. In addition, evidence of emotional perspective taking was only weakly reflected in children's explicit responses.

Summary. The research reviewed in this section documents preschoolers' sensitivity to emotional prosody in referential communication, highlighting the developmental changes that occur between three and five years of age and the powerful role of negative-sounding emotional prosody. This research also demonstrates that preschoolers can use emotional prosody to generate inferences about a speaker's emotional perspective and integrate these perspectives in on-line language processing. In the next section, we consider the cognitive abilities that might support preschoolers' integration of perspective information in referential communication.

COGNITIVE ABILITIES AND THE INTEGRATION OF PERSPECTIVE AND LINGUISTIC INFORMATION

Theoretical accounts of communicative perspective taking have posited a role for two key sets of cognitive abilities that may support listeners' integration of perspective and linguistic information, namely theory of mind skills and executive function (Nilsen & Fecica, 2011; San Juan *et al.*, 2015). In what follows, we review research that has examined relations between children's abilities in these domains and communicative perspective taking.

Theory of mind

Theory of mind is the ability to represent and form inferences about other people's mental states. It encapsulates both the ability to track what another person can or cannot see and the ability to represent states of knowledge and intention. During the preschool years, there are marked developmental changes in children's theory of mind skills (Gopnik & Slaughter, 1991; Wellman, Cross & Watson, 2001). For example, most three-year-olds make incorrect predictions about the actions of an agent who holds a false belief, whereas most five-year-olds correctly predict that the agent's false belief will guide her behaviour (Wellman & Liu, 2004). During this same developmental period, there are significant developmental changes in children's visual perspective-taking abilities (Flavell, Speer, Green, August & Whitehurst, 1981; Masangkay, McCluskey, McIntyre, Sims-Knight, Vaughn & Flavell, 1974; Moll & Tomasello, 2006; Moll & Meltzoff, 2011a; Moll *et al.*, 2013) and emotional perspective-taking

abilities (Denham & Couchoud, 1990; Hughes & Dunn, 1998). Theory of mind abilities may support communicative perspective taking by providing children with the representational ability to track and form inferences about their communicative partner's perspective and referential intent.

Although many researchers have proposed a theoretical link between children's mentalizing abilities and communicative perspective taking (Achim, Fossard, Couture & Achim, 2015; Nilsen & Fecica, 2011; Sperber & Wilson, 2002), only a handful of studies have directly examined the relation between these two capacities. This research has shown that three- to six-year-old children's accurate production and repair of referential statements is positively related to their performance on both visual perspective-taking tasks (Roberts & Patterson, 1983) as well as standard measures of false-belief understanding (Resches & Pereira, 2007). False-belief understanding has also been shown to predict children's comprehension of spoken instructions and detection of referential ambiguity (Maridaki-Kassotaki & Antonopoulou, 2011; Resches & Pereira, 2007).

Two recent studies in our lab have specifically examined whether children's theory of mind abilities are related to their on-line integration of perspective information in referential communication. Our results have demonstrated that theory of mind skills predicted four-year-olds' communicative perspective taking in both visual perspective-taking and emotional perspective-taking referential communication tasks (Khu *et al.*, unpublished observations). Importantly, the relations between theory of mind and communicative perspective taking were specific to the relevant domain. That is, visual perspective taking measured using an off-line task was related to four-year-olds' successful integration of a speaker's visual perspective in an on-line referential communication task (Khu *et al.*, unpublished observations). Likewise, off-line emotional perspective taking was associated with the real-time integration of the speaker's emotional perspective in a referential communication task (Khu *et al.*, unpublished observations).

Although research has demonstrated links between theory of mind and communicative perspective taking, the nature of this relation is underspecified. That is, it is unclear if the relation is necessarily causal and unidirectional in nature. For example, one longitudinal study has shown that children's ability to track perspectives in conversation (e.g. being able to infer the correct recipient of a spoken utterance) predicts later development of false-belief understanding (Bernard & Deleau, 2007). Thus, the relation between theory of mind and communicative development may be bidirectional, as children who engage in more communicative exchanges may experience greater opportunities to represent and reason about differing perspectives (Harris, de Rosnay & Pons, 2005).

Executive function

Beyond theory of mind abilities, executive function has also been proposed as a critical component of communicative perspective taking (Brown-Schmidt, 2009; Lin, Keysar & Epley, 2010). Executive function may facilitate spoken language processing by providing individuals with the cognitive control needed to (i) inhibit their own perspective in favour of their communicative partner's perspective, (ii) simultaneously consider and integrate multiple cues of reference, including perspective information, and (iii) select a response that appropriately matches their communicative partner's state of knowledge or emotional state. To date, studies have demonstrated that individual differences in executive function significantly predict preschool children's comprehension of referential statements (Gillis & Nilsen, 2014; Nilsen & Graham, 2009, 2012). In one study, we examined the relation between three- to five-year-olds' communicative perspective taking and their performance on various measures of executive function (e.g. inhibitory control, working memory, and cognitive flexibility; Nilsen & Graham, 2009). Although individual differences in executive function did not predict children's performance on production measures, a positive correlation was found between children's inhibitory control and their ability to consider a speaker's visual perspective while interpreting a referential statement. Similar relations have been found in studies examining children's message evaluation. Specifically, both cognitive flexibility (Gillis & Nilsen, 2014) and inhibitory control (Nilsen & Graham, 2012) have been shown to predict preschool children's emerging detection of message ambiguity. Thus, executive function appears to assist children with integration of perspective information during spoken language comprehension.

Not all studies that have examined children's comprehension of spoken utterances, however, have found significant correlations with executive function measures (Khu *et al.*, unpublished observations; Nilsen, Mangal & MacDonald, 2013). For example, Nilsen and colleagues (2013) found that inhibitory control measures did not correlate with the performance of typically developing children and children with Attention Deficit Hyperactivity Disorder on a complex comprehension task (similar in procedure to Epley *et al.*, 2004). Similarly, Khu *et al.* (unpublished observations) failed to find relations between children's working memory, conflict inhibitory control, or delay inhibitory control, and performance on communication tasks that involved taking a speaker's visual or emotional perspective. Further research is thus needed to clarify how the contributions of executive function vary across different communicative tasks. It also remains to be seen whether a similar relation exists between executive function and children's ability to produce statements that are tailored to their listener's perspective.

CONCLUSIONS AND FUTURE DIRECTIONS

As we have reviewed, preschoolers are remarkably skilled at integrating perspective information with on-line language comprehension, with significant development occurring between three and five years of age. Our review has highlighted research demonstrating that preschoolers' ability to use visual and emotional perspective information to guide language interpretation is not uniform in character, is sometimes related to theory of mind and executive function skills, and, at certain ages, is revealed only by implicit measures of language processing. Together, the research reviewed here helps to broaden theoretical models of communicative perspective taking, underscoring the importance of examining how different types of perspective inferences shape children's referential understanding.

Although research has significantly advanced our understanding of preschoolers' communicative perspective taking, there remain a number of key issues for further empirical consideration. We discuss three such considerations below.

What types of perspective representations are needed to guide communication?

Communicative perspective taking encompasses a broad range of abilities, including, but not limited to, the ability to track the visual perspective of a communicative partner and/or the emotional prosody of spoken utterance to form inferences about referential intent. What remains unclear is the type of perspective representations that are necessary to influence both implicit constraints on visual attention as well as explicit interpretation of spoken utterances. As a number of studies reviewed in this paper have shown, discrepancies sometimes exist between children's eye-gaze patterns and elicited responses (e.g. Berman *et al.*, 2010; Nilsen *et al.*, 2008). That is, implicit awareness of a communicative partner's perspective does not always influence the explicit comprehension of spoken utterances.

At present, it is unclear whether these discrepancies are due to underdeveloped cognitive abilities, such as executive function, that would assist children in selecting an explicit social response. Alternatively, these findings could also be indicative of the types of perspective representations that are necessary to influence explicit referential interpretation. That is, implicit awareness of a speaker's perspective may be sufficient to constrain visual attention but not sufficiently robust to outweigh competing cues of reference. The few studies that have examined a relation between children's mentalizing abilities and communicative perspective taking have suggested that explicit awareness of a partner's perspective may be important, if not necessary, for communicative perspective taking to

develop (e.g. Khu *et al.*, unpublished observations). As it currently stands, however, it is unclear if discrepancies between implicit and explicit responses are indicative of children's inability to (i) rapidly generate robust, if not explicit, representations of perspective, and/or (ii) integrate perspective cues with other sources of information.

Related to the issue of perspective representations, recent accounts have also suggested that there may be limits on the types of perspective inferences that individuals can generate efficiently enough to influence real-time social responses (Butterfill & Apperly, 2013; Low, Apperly, Butterfill & Rakoczy, 2016). In support of these accounts, researchers have found that both children and adults are able to rapidly form inferences about Level I perspective taking (e.g. understanding WHAT another person sees from a different perspective) but show significant delays in reasoning about Level II perspective taking (e.g. understanding HOW another person sees the same item from a conflicting perspective) (Low & Watts, 2013; Surtees, Butterfill & Apperly, 2012). Examining whether similar limits exist in communicative perspective taking may provide insight into whether children's ability to integrate perspective taking with spoken language processing is dependent on the complexity of perspective inferences being formed.

How might social experience influence the development of communicative perspective taking?

Nilsen and Fecica (2011) have proposed that the development of cognitive abilities associated with communicative perspective taking may, in turn, be dependent on the quality and degree of children's social experience. To date, most research examining the relation between social experience and communicative perspective taking has focused on children's production of spoken utterances. For example, several studies have now shown that corrective feedback from a listener (e.g. requests for clarification) can lead to improvements in preschoolers' production of referential statements (Matthews, Butcher, Lieven & Tomasello, 2012; Matthews, Lieven & Tomasello, 2007; Nilsen & Mangal, 2012) and better detection of referential ambiguity (Robinson & Robinson, 1981, 1985; Sonnenschein, 1984). Incentives (i.e. stickers) have similarly been shown to lead to improvements in the accuracy of preschoolers' communicative production, suggesting that experience can influence children's motivation to track and form inferences about a communicative partner's perspective (Varghese & Nilsen, 2013). More experience engaging in communicative interactions, perhaps through pretend play, may also contribute to children use of perspective in communication. For example, Roby and Kidd (2008) found that, relative to children without imaginary companions, children with

imaginary companions were more likely to produce descriptions that would help their listener identify a target image and request clarification when interpreting ambiguous descriptions. Thus, social experience appears to influence children's production and, possibly, detection of miscommunicated messages. It remains to be seen, however, whether similar experience and feedback would impact their comprehension of referential statements.

Does communicative perspective taking vary across different social contexts?

Related to the question of how social experience may influence the development of communicative perspective taking, it also remains an open question whether children's communicative perspective taking may vary across different social contexts. For example, Moll, Carpenter, and Tomasello (2011) found that toddlers were more likely to conflate their own perspective with that of a co-present adult when engaged in a collaborative social interaction. It is possible that efficiency and accuracy of communicative perspective taking may vary with contextual factors (e.g. cooperative vs. competitive contexts) that could affect both children's motivation and ability to track differences in perspective.

In closing, addressing the considerations described above will further clarify the cognitive and social factors contributing to the development and efficiency of communicative perspective taking, leading to a more comprehensive account of communicative development.

REFERENCES

- Achim, A. M., Fossard, M., Couture, S. & Achim, A. (2015). Adjustment of speaker's referential expressions to an addressee's likely knowledge and link with theory of mind abilities. *Frontiers in Psychology* **6**, 823.
- Alloppenna, 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), 419–39.
- Apperly, I. A., Carroll, D. J., Samson, D., Humphreys, G. W., Qureshi, A. & Moffitt, G. (2010). Why are there limits on theory of mind use? Evidence from adults' ability to follow instructions from an ignorant speaker. *Quarterly Journal of Experimental Psychology* **63**(6), 1201–17.
- Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development* **62**(5), 874–90.
- Baldwin, D. A. (1993). Early referential understanding: infants' ability to recognize referential acts for what they are. *Developmental Psychology* **29**(5), 832–43.
- Banse, R. & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology* **70**(3), 614–36.
- Berman, J. M., Chambers, C. G. & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology* **107**(2), 87–99.
- Berman, J. M., Chambers, C. G. & Graham, S. A. (2016). Preschoolers' real-time coordination of vocal and facial emotional information. *Journal of Experimental Child Psychology* **142**, 391–9.

- Berman, J. M., Graham, S. A., Callaway, D. & Chambers, C. G. (2013). Preschoolers use emotion in speech to learn new words. *Child Development* **84**(5), 1791–805.
- Berman, J. M., Graham, S. A. & Chambers, C. G. (2013). Contextual influences on children's use of vocal affect cues during referential interpretation. *Quarterly Journal of Experimental Psychology* **66**(4), 705–26.
- Bernard, S. & Deleau, M. (2007). Conversational perspective-taking and false belief attribution: a longitudinal study. *British Journal of Developmental Psychology* **25**(3), 443–60.
- Brooks, R. & Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Developmental Psychology* **38**(6), 958–66.
- Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review* **16**(5), 893–900.
- Brown-Schmidt, S. & Heller, D. (2014). What language processing can tell us about perspective taking: a reply to Bezuidenhout (2013). *Journal of Pragmatics* **60**, 279–84.
- Butterfill, S. A. & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language* **28**(5), 606–37.
- Caron, A. J., Kiel, E. J., Dayton, M. & Butler, S. C. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition and Development* **3**(4), 445–64.
- Chambers, C. G., Tanenhaus, M. K. & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **30**(3), 687–96.
- Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H. & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language* **47**(1), 30–49.
- Collins, S. J., Graham, S. A. & Chambers, C. G. (2012). Preschoolers' sensitivity to speaker action constraints to infer referential intent. *Journal of Experimental Child Psychology* **112**(4), 389–402.
- Cooper, R. P. & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development* **61**(5), 1584–95.
- Denham, S. A. & Couchoud, E. A. (1990). Young preschoolers' understanding of emotions. *Child Study Journal* **20**(3), 171–92.
- Dunphy-Lelii, S. & Wellman, H. M. (2004). Infants' understanding of occlusion of others' line-of-sight: implications for an emerging theory of mind. *European Journal of Developmental Psychology* **1**(1), 49–66.
- Epley, N., Morewedge, C. K. & Keysar, B. (2004). Perspective taking in children and adults: equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology* **40**(6), 760–8.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development* **8**(2), 181–95.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development* **60**(6), 1497–510.
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective. In J. H. Barkow, L. Cosmides & J. Tooby (eds), *The adapted mind*. (pp. 262–282). New York: Oxford University Press.
- Fernald, A. (1993). Approval and disapproval: infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development* **64**(3), 657–74.
- Flavell, J. H., Speer, J. R., Green, F. L., August, D. L. & Whitehurst, G. J. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development* **46**(5), 1–65.
- Frick, R. W. (1985). Communicating emotion: the role of prosodic features. *Psychological Bulletin* **97**(3), 412–29.
- Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental Science* **3**(2), 148–62.
- Friend, M. (2001). The transition from affective to linguistic meaning. *First Language* **21**(63), 219–43.

- Friend, M. & Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly* **46**(2), 342–69.
- Gillis, R. & Nilsen, E. S. (2014). Cognitive flexibility supports preschoolers' detection of communicative ambiguity. *First Language* **34**(1), 58–71.
- Glucksberg, S. & Krauss, R. M. (1967). What do people say after they have learned how to talk? Studies of the development of referential communication. *Merrill-Palmer Quarterly of Behavior and Development* **13**(4), 309–16.
- Graham, S. A., Sedivy, J. & Khu, M. (2014). That's not what you said earlier: preschoolers expect partners to be referentially consistent. *Journal of Child Language* **41**(1), 34–50.
- Gopnik, A. & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development* **62**(1), 98–110.
- Hanna, J. E., Tanenhaus, M. K. & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language* **49**(1), 43–61.
- Harris, P. L., de Rosnay, M. & Pons, F. (2005). Language and children's understanding of mental states. *Current Directions in Psychological Science* **14**(2), 69–73.
- Heller, D., Parisien, C. & Stevenson, S. (2016). Perspective-taking behavior as the probabilistic weighing of multiple domains. *Cognition* **149**, 104–20.
- Huettig, F., Rommers, J. & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychologica* **137**(2), 151–71.
- Hughes, C. & Dunn, J. (1998). Understanding mind and emotion: longitudinal associations with mental-state talk between young friends. *Developmental Psychology* **34**(5), 1026–37.
- Keysar, B. (2007). Communication and miscommunication: the role of egocentric processes. *Intercultural Pragmatics* **4**(1), 71–84.
- Kitamura, C. & Burnham, D. (2003). Pitch and communicative intent in mother's speech: adjustments for age and sex in the first year. *Infancy* **4**(1), 85–110.
- Kitamura, C. & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy* **14**(1), 77–100.
- Krauss, R. M. & Glucksberg, S. (1969). The development of communication: competence as a function of age. *Child Development* **40**(1), 255–66.
- Lin, S., Keysar, B. & Epley, N. (2010). Reflexively mindblind: using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology* **46**(3), 551–6.
- Liszkowski, U., Carpenter, M. & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition* **108**(3), 732–9.
- Low, J., Apperly, I. A., Butterfill, S. A. & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: a primer on the two-systems account. *Child Development Perspectives*.
- Low, J. & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science* **24**(3), 305–11.
- Luo, Y. & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition* **105**(3), 489–512.
- Maridaki-Kassotaki, K. & Antonopoulou, K. (2011). Examination of the relationship between false-belief understanding and referential communication skills. *European Journal of Psychology of Education* **26**(1), 75–84.
- Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E. & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development* **45**(2), 357–66.
- Matthews, D., Butcher, J., Lieven, E. & Tomasello, M. (2012). Two- and four-year-olds learn to adapt referring expressions to context: effects of distracters and feedback on referential communication. *Topics in Cognitive Science* **4**(2), 184–210.
- Matthews, D., Lieven, E., Theakston, A. & Tomasello, M. (2006). The effect of perceptual availability and prior discourse on young children's use of referring expressions. *Applied Psycholinguistics* **27**(3), 403–22.

- Matthews, D., Lieven, E. & Tomasello, M. (2007). How toddlers and preschoolers learn to uniquely identify referents for others: a training study. *Child Development* **78**(6), 1744–59.
- Moll, H., Carpenter, M. & Tomasello, M. (2011). Social engagement leads 2-year-olds to overestimate others' knowledge. *Infancy* **16**(3), 248–65.
- Moll, H. & Meltzoff, A. N. (2011a). Perspective-taking and its foundation in joint attention. In N. Eilan, H. Lerman, & J. Roessler (Eds.), *Perception, Causation, and Objectivity: Issues in Philosophy and Psychology* (pp. 286–304). Oxford: Oxford University Press.
- Moll, H. & Meltzoff, A. N. (2011b). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development* **82**(2), 661–73.
- Moll, H., Meltzoff, A. N., Merzsch, K. & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology* **49**(4), 646–54.
- Moll, H. & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science* **7**(1), F1–9.
- Moll, H. & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology* **24**(3), 603–13.
- Morton, J. B. & Munakata, Y. (2002). Are you listening? Exploring a developmental knowledge–action dissociation in a speech interpretation task. *Developmental Science* **5**(4), 435–40.
- Morton, J. B. & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development* **72**(3), 834–43.
- Morton, J. B., Trehub, S. E. & Zelazo, P. D. (2003). Sources of inflexibility in 6-year-olds' understanding of emotion in speech. *Child Development* **74**(6), 1857–68.
- Moses, L. J., Baldwin, D. A., Rosicky, J. G. & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development* **72**(3), 718–35.
- Nadig, A. S. & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science* **13**(4), 329–36.
- Nayer, S. L. & Graham, S. A. (2006). Children's communicative strategies in novel and familiar word situations. *First Language* **26**(4), 403–20.
- Nelson, N. L. & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology* **110**(1), 52–61.
- Nilsen, E. S. & Fecica, A. M. (2011). A model of communicative perspective-taking for typical and atypical populations of children. *Developmental Review* **31**(1), 55–78.
- Nilsen, E. S. & Graham, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology* **58**(2), 220–49.
- Nilsen, E. S. & Graham, S. A. (2012). The development of preschoolers' appreciation of communicative ambiguity. *Child Development* **83**(4), 1400–15.
- Nilsen, E. S., Graham, S. A., Smith, S. & Chambers, C. G. (2008). Preschoolers' sensitivity to referential ambiguity: evidence for a dissociation between implicit understanding and explicit behavior. *Developmental Science* **11**(4), 556–62.
- Nilsen, E. S. & Mangal, L. (2012). Which is important for preschoolers' production and repair of statements: what the listener knows or what the listener says? *Journal of Child Language* **39**(5), 1121–34.
- Nilsen, E. S., Mangal, L. & MacDonald, K. (2013). Referential communication in children with ADHD: challenges in the role of a listener. *Journal of Speech, Language, and Hearing Research* **56**(2), 590–603.
- O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development* **67**(2), 659–77.
- Paulmann, S. & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion* **35**(2), 192–201.
- Pell, M. D. & Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS One* **6**(11), e27256.
- Pell, M. D., Monetta, L., Paulmann, S. & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior* **33**(2), 107–20.

- Pell, M. D., Paulmann, S., Dara, C., Allasseri, A. & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: a comparison of four languages. *Journal of Phonetics* **37**(4), 417–35.
- Resches, M. & Pereira, M. P. (2007). Referential communication abilities and theory of mind development in preschool children. *Journal of Child Language* **34**(1), 21–52.
- Roberts, R. J. Jr & Patterson, C. J. (1983). Perspective taking and referential communication: the question of correspondence reconsidered. *Child Development* **54**(4), 1005–14.
- Robinson, E. J. & Robinson, W. P. (1981). Ways of reacting to communication failure in relation to the development of the child's understanding about verbal communication. *European Journal of Social Psychology* **11**(2), 189–208.
- Robinson, E. J. & Robinson, W. P. (1982). Knowing when you don't know enough: children's judgements about ambiguous information. *Cognition* **12**(3), 267–80.
- Robinson, E. J. & Robinson, W. P. (1985). Teaching children about verbal referential communication. *International Journal of Behavioral Development* **8**(3), 285–99.
- Roby, A. C. & Kidd, E. (2008). The referential communication skills of children with imaginary companions. *Developmental Science* **11**(4), 531–40.
- Rozin, P. & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review* **5**(4), 296–320.
- San Juan, V., Khu, M. & Graham, S. A. (2015). A new perspective on children's communicative perspective taking: When and how do children use perspective inferences to inform their comprehension of spoken language? *Child Development Perspectives* **9**(4), 245–9.
- Scherer, K. R., Banse, R. & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology* **32**(1), 76–92.
- Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of Psycholinguistic Research* **32**(1), 3–23.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G. & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition* **71**(2), 109–47.
- Singh, L., Morgan, J. L. & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy* **3**(3), 365–94.
- Snedeker, J. & Huang, Y. (2016). Sentence Processing. In E. L. Bavin & L. Naigles (eds), *The handbook of child language*, 2nd ed. (pp. 409–437). Cambridge: Cambridge University Press.
- Snedeker, J. & Trueswell, J. C. (2004). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* **49**(3), 238–99.
- Sodian, B. (1988). Children's attributions of knowledge to the listener in a referential communication task. *Child Development* **59**(2), 378–85.
- Sonnenschein, S. (1984). How feedback from a listener affects children's referential communication skills. *Developmental Psychology* **20**(2), 287–92.
- Sperber, D. & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language* **17**(1/2), 3–23.
- Surtees, A. D., Butterfill, S. A. & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology* **30**(1), 75–86.
- Swingle, D., Pinto, J. P. & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition* **71**(2), 73–108.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M. & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* **268**(5217), 1632–4.
- Tomasello, M., Strosberg, R. & Akhtar, N. (1996). Eighteen-month-old children learn words in non-ostensive contexts. *Journal of Child Language* **23**(1), 157–76.

- Trueswell, J. C., Sekerina, I., Hill, N. M. & Logrip, M. L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition* **73**(2), 89–134.
- Vaish, A., Grossmann, T. & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological Bulletin* **134**(3), 383–403.
- Varghese, A. L. & Nilsen, E. (2013). Incentives improve the clarity of school-age children's referential statements. *Cognitive Development* **28**(4), 364–73.
- Waxer, M. & Morton, J. B. (2011). The development of future-oriented control: an electrophysiological investigation. *NeuroImage* **56**(3), 1648–54.
- Wellman, H. M., Cross, D. & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development* **72**(3), 655–84.
- Wellman, H. M. & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development* **75**(2), 523–41.