Infants' understanding of communication as participants and observers

Guðmundur Bjarki Thorgrímsson

Cover picture: De Kabouterboom: Sweet Chestnut (Castanea Sativa) in Beek, Ubbergen, The Netherlands. Photograph by Thorgrímur Guðmundsson.

Printed by: Ipskamp Drukkers, Nijmegen, The Netherlands

Infants' understanding of communication as participants and observers

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,

volgens besluit van het college van decanen

in het openbaar te verdedigen op woensdag 27 augustus 2014

om 12.30 uur precies

door

Guðmundur Bjarki Thorgrímsson

geboren op 19 oktober, 1982

in Reykjavík, IJsland

Promotor: Prof. dr. H. Bekkering

Copromotor: Prof. dr. U. Liszkowski (University of Hamburg, Germany)

Manuscriptcommissie:

Prof. dr. P. Fikkert

Prof. dr. A. Ozyurek

Dr. K. Rohlfing (University of Bielefeld, Germany)

Table of Contents

Chapter 1

Introduction

Introduction

General introduction

To interact successfully with others requires understanding what drives their actions and anticipating their future behavior. This is true of joint activities ranging from coordinating bodily movements in dance to debating philosophy. There is a great deal of evidence that the ability to understand and predict others' actions originates in the first year of life. Studies suggest that, in the eyes of the infant, the activities of other people are not mere physical movements with causal effects, but purposeful, goal-directed attempts to act on their environment. This represents the early beginnings of human social understanding that later culminates in an understanding of others in terms of their beliefs, desires, and other mental states. Infants' understanding of goal-directed action allows for both early learning about actions and objects through observing or engaging with others, and prediction of others' ongoing or future actions (see Csibra & Gergely, 2007). In the first year of life, we also see the beginnings of infants' engagement with and understanding of communicative actions. They are responsive to the communicative efforts of caregivers in face-to-face interactions from as early as 2 months, and in the second half of the first year, they begin to show understanding of referential communication conveyed through speech, gaze, and gesture. For example, at around 12 months, infants follow pointing gestures to target objects (Carpenter, Nagell, & Tomasello, 1998), understand that points and words refer to the same object (Gliga & Csibra, 2009), and can infer the motive behind the gesture based on the social context within which it is used (Liebal, Behne, Carpenter, & Tomasello, 2009). The ability to understand and engage in referential communication provides further opportunities for social learning during infancy, and is likely to play an important role in infants' entry into symbolic language (Tomasello, Carpenter, & Liszkowski, 2007; Iverson & Goldin-Meadow, 2005).

Introduction

The large amount and wide range of research in this field has revealed much about different aspects and conditions of infants' understanding of instrumental and communicative actions, but important questions have remained unaddressed. In recent years, a number of interaction-based experiments have shown that infants' understanding of action and communication informs not only their passive processing of observed actions, but also their responses in naturalistic social contexts (e.g. Behne, Carpenter, & Tomasello, 2005; Behne, Liszkowski, Carpenter, & Tomasello, 2012; Liebal et al., 2009; Warneken & Tomasello, 2007). However, any naturalistic interactional setting inevitably presents the observer or addressee with additional information about the potential meaning of the action. Although these studies crucially demonstrate that infants' action understanding is put to use in interactive settings, the context within which the actions are presented typically consists of multiple cues to the actor's motive, which may have partly or exclusively prompted the infants' response. Thus, while processing studies that employ looking time measures have demonstrated infants' detailed understanding of the actions themselves - revealing, for example, that infants selectively attribute goals to a (disembodied) reaching human hand based on its shape, trajectory, and environmental constraints - it has remained unclear whether such understanding is also manifested in infants' active interactional responses. The study described in Chapter 2 addresses this question by examining infants' responses to object-directed actions and gestures in a minimal social setting where the action's meaning could not be inferred through social-pragmatic cues. The study also examined the secondary question of whether infants' interpretation of actions and gestures is guided by the spatial relationship between the actor and the object.

Another important question is the extent to which infants' understanding of actions and gestures extends to others' interactions. The majority of research of infants' action understanding focuses on dyadic, one-on-one settings, while observing and overhearing third-party interactions also forms a large part of infants' early social experience. The social interactions of others present a rich potential source of information about various aspects of the social world, including language, joint actions, and social norms. Observing third-party interactions is likely to be particularly important for social and communicative development in traditional communities where children are often in constant company of several people, are involved in community practices from an early age, and are in some cases reported to be infrequently addressed before they start to speak (Brown, 1998; Heath, 1983; Ochs and Schieffelin, 1984, see also Lieven, 1994). Research has shown that infants understand and anticipate the outcome of common individual object-directed actions and are responsive to others' communication, but we know less about their understanding of and expectations about social communicative actions that are directed to another person. Infants' looking patterns show that they visually monitor speakers in third-party conversations, but it has been unclear whether they understand that speech is directed to, and typically provokes responses from, other people. Chapter 3 describes an eye-tracking study that takes a first step toward addressing this question by examining infants' expectations toward third-party addressees of speech. The level of infants' understanding of instrumental actions and gestures in others' interactions has also remained unclear. Infants can imitate and learn information conveyed through actions directed to a third party addressee, but whether their understanding encompasses the addressee's perspective and allows them to predict their response is an open question. The study described in Chapter 4 employed a similar eye-tracking paradigm to examine whether infants can predict how third-party addressees respond to common actions and communicative gestures.

This introductory chapter is organized into three sections. I begin by giving a brief review of what we currently know about infants' processing and visual monitoring of actions and gestures. The second section focuses on how infants' action understanding is reflected in their social interactions and the extent to which it depends on, or is influenced by, social-pragmatic cues. The third reviews literature on infants' understanding of third-party conversations and communicative interactions and discusses the importance of investigating third-party social interactions.

Processing and predicting action

Infants encounter others' actions in various forms and settings. They observe people act on their environment or act jointly with others, and in their own interactions infants are addressed through speech or gesture. Understanding any action, whether a simple reaching movement, or a culture-specific conventional gesture such as the thumbs-up signal, relies on understanding human action as *goal-directed* behavior. An important line of research has examined infants' understanding of goal-directed actions by measuring their looking responses to live or pre-recorded instrumental actions - typically basic acts of reaching, grasping, or moving small objects. In a seminal looking-time study, Woodward (1998) habituated infants with a hand reaching and grasping one of two distinct objects. In the test trials, the object locations were switched, and the hand either reached and grasped the same object, thus taking a new path (old goal/new path), or the other object, thus taking the same path as before (new goal/old path). The infants looked longer at the test trials where the hand reached for the other object (the new goal), indicating that they expected it to consistently grasp the same object rather than take the same path. Importantly, the infants did not pay such selective attention to the goal object when they observed the same actions performed with a mechanical claw instead of a human hand. A follow-up study revealed that infants show such expectations only when an object is grasped but not when it is touched with the back of the hand, indicating that infants pay attention to the specific hand shape (Woodward, 1999). Employing similar paradigms, subsequent studies showed that by 12 months infants expect people to reach and grasp objects in an efficient manner given spatial constraints (Phillips & Wellman, 2005), and understand reaching in terms of goals also when it falls short of grasping the object (Brandone & Wellman, 2009). The visual habituation or looking-time paradigms used in these studies work by repeatedly showing infants an action event until infants look away (habituate) and then presenting two similar test events but with key changes. The infants' looking time to the test events are compared and longer looking times are thought to indicate a violation of their expectations. As such, these studies demonstrate infants' retrospective judgments of action goals. Recent eye-tracking research shows that infants also attend prospectively to others' actions, revealing an ability to not only process but actively predict others' actions. Thus, key looking-time experiments have been replicated using eye-tracking measures and revealed that infants visually anticipate (i.e. "look ahead to") the goal objects of reaching actions (Cannon & Woodward, 2012; Brandone, Horwitz, Aslin, & Wellman, 2014). Similarly, infants have been found to anticipate to the goal location of other common everyday actions, such as an arm transferring a ball into a bucket (Falck-Ytter, Gredebäck, & von Hofsten, 2006), an actor bringing a cup to her mouth (Hunnius & Bekkering, 2010), or a spoon to another person's mouth (Gredebäck & Melinder, 2010).

Infants' looking behavior has also given important insight into their understanding of others' referential actions. Referential actions, such has pointing or gazing at objects, direct the others' attention to specific aspects of the environment, for any number of possible social goals. Studies show that infants follow the direction of others' gaze from as early as 3 months (e.g.

D'Entremont, Hains, & Muir, 1997), and show understanding of its referential nature by at least 12-months (Moses, Baldwin, Rosicky, & Tidball, 2001; Brooks & Meltzoff, 2002). Similarly, infants follow points to their target objects (e.g. Carpenter, Nagell, & Tomasello, 1998; Deak, Flom, & Pick, 2000), and looking time studies show that they understand the relation between the person pointing and the referent object (Woodward & Guajardo, 2002, Sodian & Thoermer, 2004) and expect points and words to co-refer to the same object (Gliga & Csibra, 2009). However, comprehension of referential actions and other communicative actions obviously entails more than passive processing and looking at referents. In the case of pointing, the addressee needs to understand that the pointer intended to direct his or her attention and infer the social motive behind the point (e.g. a request for an object; see Tomasello et al., 2007). Evidence for comprehension of communicative actions in this sense can thus only come from studies of infants' own action responses within communicative interactions. Similarly, a full picture of infants' understanding of instrumental actions requires examining how this understanding is embodied in interaction. In general, given that the role of the social understanding manifested in infants' processing of others' actions is arguably to enable infants to interact with and learn from others, and that this understanding is in turn shaped by their social interactions (Carpendale & Lewis, 2004; 2006; Liszkowski, 2013), it is crucial to test infants' responses in interactions with others.

Action understanding in action

While there is a comparatively large body of research on infants' passive visual processing of others' actions, we know less about their responses in interactive settings. Regarding instrumental actions, there is much evidence that infants imitate others' goal-directed actions (e.g. Carpenter, Akhtar, & Tomasello, 1998; Carpenter, Call, & Tomasello, 2005; Elsner, 2007; Gergely, Bekkering, & Király, 2002), suggesting that they can infer and reproduce action goals through observation. Less research attention has been devoted to investigating whether infants understand and respond appropriately to instrumental actions, such as reaching, when these are presented in the context of a social interaction. One recent study revealed that infants respond to unsuccessful reaching actions directed toward objects within their reach by handing the object over (Warneken & Tomasello, 2007). However, the actions presented in such studies are embedded within a rich social context that typically contains pragmatic cues to the motive behind the action. Such pragmatic cues may be necessary, or even sufficient, for infants to infer the motive and respond appropriately. As a result, it is unclear whether infants have a pragmatic understanding of the actions themselves. Similar confounds characterize studies of infants' understanding of common communicative gestures, such as the palm-up request gesture. Infants often encounter this gesture in give-and-take routines, in which caregiver and infant repeatedly exchange possession of objects (Bruner, 1977; Messinger & Fogel, 1998). Although one experimental study indicates that infants respond to the gesture by handing an object over (Hay & Murray, 1982), the gesture was presented within the context of the familiar routine and was accompanied by a verbal request. Thus, it is unclear whether infants have understanding of the gesture form itself. Understanding the meaning of action and gesture forms independently of contextual information would evidence a more abstract pragmatic understanding of common actions in infancy than has previously been documented. The study reported in Chapter 2 addresses, among others, the question of whether infants respond to common instrumental and communicative actions in an interactive setting, and employs a design where the motive behind the actions cannot be inferred from social contextual cues.

In addition to reaching and the palm-up request gesture, the study in Chapter 2 examined infants' interpretation of points. Unlike reaching - where the shape of the hand is instrumental in grasping an object, and the palm-up gesture - where the hand shape has a requestive meaning by convention (and can also be instrumental in receiving objects) , the pointing gesture itself carries no information about its underlying motive. The motive has to be inferred from the surrounding context, for example, the topic of an ongoing conversation, or the current activity of the pointer or the addressee. Successful communication requires both interactants to be mutually aware of this shared context, which has been termed their conceptual common ground (Clark, 1996). Infants take into account the social context in their interpretation and response to others' points by 12 months (Behne, Carpenter, & Tomasello, 2005; Behne, Liszkowski, Carpenter, & Tomasello, 2012; Liebal, Behne, Carpenter, & Tomasello, 2009). To illustrate, when an adult who has just shared with an infant the activity of placing toys into a basket points at another toy, the infant is likely to respond by placing that toy in, as compared to when the adult did not share this activity prior to the point (Liebal et al., 2009). Similarly, other studies show that infants also understand points that take place within social games (Behne et al., 2005; Behne et al., 2012). However, a shared context of activity is not the only pragmatic cue to the meaning of points available to infants. Points (and other forms of ambiguous communicative actions) also get their meaning from the spatial context in which they take place. The spatial relationship between communicator, the addressee, and the potential referent objects in the environment can help addressees resolve referential or motivational ambiguity. Recent findings suggest that older children use spatial information to disambiguate referentially ambiguous verbal requests. When addressed with an ambiguous verbal request which could refer either to an object within the speaker's reach or an identical one out of the speaker's reach (both within reach for the children), children select and offer the latter (Collins, Graham, & Chambers, 2012). The children's choice likely reflects an expectation that requests should be rational and people should not request assistance for something they can accomplish more easily on their own. It is unclear when sensitivity to pragmatic cues in the form of spatial information emerges in development. In attributing goals to instrumental actions, infants are sensitive to the apparent rationality of the actions. Their attributions are guided by spatial information indicating action efficiency, such as whether the action follows the shortest pathway to a goal object in the presence or absence of a physical obstacle (e.g. Gergely, Nádasdy, Csibra, & Bíró, 1995; Csibra, Bíró, Koós, & Gergely, 2003; Woodward & Sommerville, 2000; Phillips & Wellman, 2005). This suggests that spatial information may also inform infants' interpretation of communicative actions. The study reported in Chapter 2 addresses, among others, the question of whether infants interpretation of the motive of communicative actions are guided by available spatial information, specifically the spatial distance between a communicative partner and a potential referent object. In the study, an experimenter directed communicative and instrumental actions toward an object within the infants' reach, where the object's location was manipulated such that it was either within reach for the experimenter, or out of his reach.

Third-party interactions

As reviewed above, infants demonstrate action understanding both as observers of others acting on objects and as addressees of others' communicative actions. Although the vast majority of research on action understanding focuses on these settings, infants are also surrounded every day by social interactions that take place between other people and they frequently observe and overhear actions that are directed to a third party. Infants' understanding of third-party conversations and other communicative interactions is interesting for two main reasons. First, third-party interactions present infants with opportunities for social learning without direct participation. These interactions are a potential source of information about many aspects of social life, including words, actions, gestures, tools, and social norms. Learning from third-party interactions is likely to be a particularly useful skill in traditional communities around the world where ethnographic reports suggest adults rarely address infants (e.g. Brown, 1998; Heath, 1983; Ochs & Schieffelin, 1984, see also Lieven, 1994). Second, infants' understanding of the structure of others' interactions can give insight into whether they possess a more abstract understanding of communicative actions than is manifested in their own interactions. Infants respond appropriately to others' speech, actions, or gestures in their own interactions (and initiate interactions themselves) but it is unclear whether this reflects an understanding of their underlying function. For example, while infants' looking responses in word comprehension studies show that they associate specific words with their depicted referents, such paradigms cannot tell us whether infants understand the communicative function of speech. Whether or how infants expect other addressees to respond to speech is a better indicator of understanding of the communicative role of the speech signal. Similarly, if infants have expectations about how third-party addressees respond to actions or gestures, it would show that their action understanding is not limited to their own interactions, and suggest that infants may be guided by more abstract understanding of actions than revealed in their own social interactions. Chapters 3 and 4 respectively describe eye-tracking studies that examined infants' understanding of speech and actions in third-party interactions, by measuring their expectations toward the addressees. The following sections present a brief background for each of the studies.

Third-party speech

Infants in their first year understand and predict the outcome of several common everyday actions, such as placing objects in containers (Falck-Ytter et al., 2006), drinking from a cup (Hunnius & Bekkering, 2010), and feeding others (Gredebäck & Melinder, 2010). This raises the question whether infants also understand and anticipate the typical outcome of the act of speaking: a response from the addressee. An interesting possibility is that before learning to interact through language, infants may already understand speech as a communicative action, performed to influence others. Given previous findings that have established links between infants' own action experience and their understanding of others' actions (Sommerville & Woodward, 2005; Sommerville, Woodward, & Needham, 2005; Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012), and what appear to be failures of preverbal infants to predict others' conversational turns (von Hofsten, Uhlig, Adell, & Kochukhova, 2009), it is possible that such an understanding of speech emerges only with infants' own language experience. However, infants may also come to understand the communicative nature of speech through their rich pre-linguistic social and observational experiences. They are frequently addressed with speech, typically in face-to-face interactions with caregivers, and studies show that infants respond through vocalizations, facial expressions, and other bodily movements. Moreover, the rudiments of speech comprehension are in place by as early as 6 months, when infants have started to associate socially significant words such as mommy and daddy with their depicted referents (Tincoff & Jusczyk, 1999; Bergelson & Swingley, 2012). Finally, infants are typically surrounded with conversing adults from birth and frequently observe and overhear speech in others' interactions.

One indication of infants' understanding of the communicative function of speech is whether they expect other people to respond to speech when addressed. A series of recent eye-tracking studies has examined infants' looking patterns while viewing staged conversations between two adults. At 6 and 11 months, infants look at speakers in face-to-face conversations during about half of the conversational turns, and more often than when interlocutors are oriented away from each other in a back-to-back configuration (Augusti, Melinder, & Gredebäck, 2010). Similar studies suggest that by 3 years of age, children monitor conversations with more flexibility and disengage their attention from the current speaker to the addressee, possibly in anticipation of the upcoming conversational turn (von Hofsten, Uhlig, Adell, & Kochukhova, 2009; Bakker, Kochukhova, & von Hofsten, 2011). Further, these studies show that 3-year-olds, but not 1-year-olds, make such gaze shifts more often between human speakers than between temporally matched objects that take turns moving and emitting sound, or between humans who emit mechanical sounds. These studies suggest that children may not understand speech to be a communicative act directed to another person until they are experienced speakers themselves. However, it is not clear whether the shifts of attention between interlocutors seen in these studies reflect expectations about speech. Instead, they may reflect infants' ability to learn to anticipate events through observing the contingent conversations. In the study described in Chapter 3, the eye movements of 12 - and 24-month-old infants were tracked as they watched a person repeatedly address another person with short utterances, to measure their expectations of a response. Crucially, the addressee never replied to the utterances, preventing infants from learning to anticipate a response. Infants' expectations were gauged by measuring the speed with which they shifted their gaze from the speaker to the addressee, and the duration of their subsequent looking at the latter. In a control condition, the speaker emitted natural non-speech sounds, such as a cough. The study also examined the effect of bodily orientation on infants' expectations. The pair was either facing each other in a typical communicative context, or facing away from each other in a back-to-back configuration. Mutual gaze between others signals affiliation for older children (Nurmsoo, Einav, & Hood, 2012) and adults (Thayer & Schiff, 1974), and already at 6 months, infants monitor speakers in conversations more when they are face-to-face than when they are back-to-back (Augusti et al., 2010). Thus, infants may understand face-to-face orientation as a typical communicative context, and expect speech to be more likely to elicit a response in this setting.

Third-party actions

Infants' social learning is facilitated by action demonstrations that take place in a communicative context, where people directly address infants using ostensive cues such as direct eye gaze or speech with high and variable pitch ("motherese"; e.g. Gergely, Egyed, & Király, 2007; Yoon, Johnson, & Csibra, 2008). However, recent research shows that infants can also learn from observing and overhearing actions that are not addressed to them, but to another person. Such scenarios are a part of most infants' everyday social experience. In some traditional cultures, where infants tend not be addressed by caregivers (e.g. Brown, 1998; Lieven, 1994), they may even be the predominant setting in which infants hear language and encounter others' actions.

Studies show that 18-month-olds can learn object labels through observing the labels being used in others' interactions (Floor & Akhtar, 2006; Gampe, Liebal, & Tomasello, 2012) and learn novel actions demonstrated to a third party addressee (Herold & Akhtar, 2008; Matheson, Moore, & Akhtar, 2013). Infants have also been shown to understand and make use of the information conveyed through gestures in third-party interactions. After observing a person indicate the location of a hidden toy to a third party addressee, infants successfully retrieved the object themselves – but only if the point was communicative in a clearly ostensive fashion (Gräfenhain et al., 2009). These studies reveal infants' ability to imitate and learn the information conveyed through actions and gestures directed to a third person, but it is not clear to what extent they understand the structure of the interaction itself. Learning from third-party interactions requires at a minimum to attend to and understand the actions of the model, but it is an open question whether it is necessary for infants to observe the action modeled in an interactive context and attend to the addressee. More generally, it is unclear whether or when infants understand third-party interactions from the perspective of the interactants, or whether they simply make use of the information conveyed. Recent eye-tracking and looking-time studies suggest that 12-month-old infants anticipate and predict responses from third-party addressees of speech (Thorgrimsson, Fawcett, & Liszkowski, 2012; Martin, Onishi & Vouloumanos, 2012). Thus, infants may have a similar understanding of actions and gestures in third-party interactions, possibly encompassing the perspective of the interactants and enabling them to predict the outcome of the interaction. To address this question, the study reported in Chapter 4 used eye-tracking to investigate infants' response expectations toward a third party recipient of common object-directed actions and gestures. The infants watched a person (the Actor) reach, point, or direct a palm-up request gesture toward an object located between the two persons, and the other person (the Recipient) respond by either taking the object for himself, or giving the object to the Actor. Infants' understanding of the actions was gauged by measuring their ability to anticipate the outcome of the Recipient giving the object to the Actor.

Outline of thesis

The study reported in Chapter 2 examined 14-month-old infants' pragmatic understanding of object-directed actions and gestures. In an interactive setting, an Experimenter pointed, reached, or directed a request gesture (open hand, palm up) toward an object located next to the infant. To ensure that infants' responses were to the actions themselves, the behavior of the Experimenter prior to and during the action did not provide infants with social-pragmatic cues to the motive behind his actions. To examine infants' sensitivity to pragmatic cues provided by spatial information, the location of the object was manipulated, such that it was either within easy reach for the Experimenter or out of his reach.

Chapter 3 reports a study where infants' expectations toward third-party addressees of speech were measured in two eye-tracking experiments. The first experiment presented 12- and 24-month-old infants with videos of two people in profile view, where one person either produced a short utterance or a natural non-speech sound (e.g. a cough). Infants' expectations were gauged by comparing how quickly and for how long they looked at the addressee following the two types of sounds. To further examine the expectations of the younger age group, the second experiment presented 12-month-olds with a similar paradigm, but more engaging and easier to follow, and employed a less demanding experimental design.

The study reported in Chapter 4 examined infants' understanding of actions and gestures in third-party interactions. Infants observed two people in profile view, where one person (the *Gesturer*) pointed, reached, or directed a request gesture to an object located between the pair. The other person (the *Addressee*) then responded by grasping the object and dispensing it through one of two tubes which led to each person, thereby either giving the object to the Gesturer, or taking it for himself. As the Addressee's hand and the tube entrances were concealed by an occluder, the infants could not see into which tube he placed the ball, and had to look at the tube exit to see the outcome. The measure of interest was whether infants would anticipate an outcome, and if so, whether they would anticipate the Addressee to give or take the object. Importantly, the Gesturer did not verbally communicate any motive, and as both actors were visible only from the neck down, no information could be gleaned from their facial expressions or their gaze direction.

Chapter 2

12-month-olds show pragmatic understanding of object-directed actions based on hand shape and spatial context

Abstract

We investigated 12-month-old infants' pragmatic understanding of object-directed instrumental and communicative actions in a minimal social setting where the action's meaning could not be inferred through social-pragmatic cues. An Experimenter pointed, reached, or directed a palm-up request gesture toward an object located in front of the infant, and infants' propensity to obtain and hand over the object was measured. A secondary question was whether infants' action interpretation is guided by the potential pragmatic cue inherent in the spatial relations between an actor and object. To examine this question, the object was located either within easy reach, or out of reach for the Experimenter. Results reveal that infants responded to all three actions by handing the object over and did this more often than in a control group where no action was performed. The results indicate that 12-month-old infants understand and respond to these actions based on their hand shapes, suggesting that infants' understanding of actions and gestures is more abstract than previously documented. In response to pointing, infants offered the object more often when it was out of reach for the Experimenter, revealing an early emerging ability to make use of spatial information as a pragmatic cue to resolve ambiguity in communication.

Introduction

Infants routinely experience other people's instrumental and communicative actions, such as reaching and pointing, both as passive observers and as active participants in social interactions. There is evidence that infants who observe others acting on objects, process their actions as goal-directed behavior. For example, infants' visual processing of actions reveals that they expect a person to consistently reach for or point toward the same object. These studies focus on infants' perception and processing of actions in the context of passive observation, using measures such as looking time (Woodward, 1998; 1999; Thoermer & Sodian, 2001; Woodward & Guajardo, 2002; Sodian & Thoermer, 2004; Phillips & Wellman, 2005; Brandone & Wellman, 2009); visual anticipation (Gredebäck, Stasiewicz, Falck-Ytter, Rosander, & von Hofsten, 2009; Cannon & Woodward, 2012), or brain activity (Southgate, Johnson, El Karoui, Csibra, 2010). As the focus is on how infants understand the action event itself, for example its trajectory or the hand shape, the actions presented in these studies are typically devoid of other bodily, facial, vocal, or verbal cues. Instead, the manipulation of interest is often spatial or physical cues, such as obstacles preventing direct access to a goal object. This research has revealed that infants expect people to direct actions toward goal objects in an efficient manner given the spatial constraints of their environment (Csibra, 2003; Csibra & Gergely, in press).

However, it is not clear whether the action understanding revealed in processing studies also enables infants to respond appropriately to others' actions in interactive settings. This would be important to know since it has been argued that social understanding emerges in the service of social acting (Carpendale & Lewis, 2006; Reddy, 2010; Liszkowski, 2013). There is evidence that infants respond appropriately to others' instrumental and communicative actions when these are embedded in shared social activities. For example, infants initiate and respond to a pointing gesture depending on the type of joint activity they were previously engaged in (Liebal, Behne, Carpenter, & Tomasello, 2009; Liebal, Behne, & Tomasello, 2010). In contrast to processing studies, interaction-based studies embed the actions in a natural social context, preceded by social interactions and accompanied by facial, vocal, bodily, or linguistic cues. As a result, it is currently unclear which of these many information sources are necessary or sufficient for infants to understand and respond appropriately. In fact, it is possible that infants' action responses are based on information from the preceding shared social context ("common ground", Clark, 1996) alone, and not on the characteristics of the actions themselves.

For example, regarding reaching, Warneken & Tomasello (2007) showed that when 14-month-olds witness a person accidentally drop an object and reach for it unsuccessfully, they will help him by handing it over. However, the infants' responses may not have been prompted by the reach itself, but by one or more additional social contextual cues. For example, the person was engaged in activity with the object before dropping it, indicating his desire to use it, and the drop was marked as an accident by a vocalization. Recent findings indeed show that given the right social context infants help others without any solicitation of a response (Knudsen & Liszkowski, 2012; Warneken, 2012). In fact, it is unclear whether infants respond to reaching because they understand it as a communicative act; because they simply adopt the other's instrumental goal, or whether their reaction is based on other contextual cues. One interaction-based study has examined infants' responses outside a rich social context. In this study, 7-month-olds responded to a grasping action or an unsuccessful reach to one of two objects within their reach by touching the same object (Hamlin, Hallinan, & Woodward, 2008). However, while touching the same object another person has tried to reach for could be an attempt to reproduce the other's goal (or simply a continuation of the familiar action sequence of reaching and grasping an object, see also

Brandone & Wellman, 2009), it does not indicate pragmatic understanding of the action.

Regarding communicative actions, interaction-based studies show that infants understand different social motives underlying pointing. From 12 months of age infants can infer the social motive of a point based on the joint activities within which it is embedded. Thus, they respond appropriately to points that take place within games (Behne et al., 2012), or follow joint activities on objects (Liebal et al., 2009). Again, however, it is unclear how infants would interpret and react to others' pointing when no additional social-contextual information is provided. Another communicative action that shares goal aspects of both reaching and pointing is the conventional request gesture, where the hand is open and the palm faces up. The goal of the gesture is to obtain an object, but as a communicative gesture it requires a compliant social partner to fulfill this goal, and the hand shape is not an efficient individual means towards the goal. Further, observational data suggests that infants themselves virtually never use the request gesture to request objects (Bakeman & Adamson, 1986; Hay & Murray, 1982; Puccini, Hassemer, Salomo, & Liszkowski, 2010). Caregivers use this gesture, typically along with a verbal request, to initiate "give-and-take" routines in infancy (Bruner, 1977; Messinger & Fogel, 1998) and experimental findings show that infants respond to the gesture within such routines by handing over objects in their possession (Hay & Murray, 1982). However, it is unclear whether infants understand the gesture as representing a request or whether they react based on the socially shared context and routine.

The question is thus whether infants have abstracted pragmatic understanding of the form of others' instrumental and communicative actions that enable them to understand and respond appropriately, independently of pragmatic cues such as a shared activity, prosody, or body posture. Since infant communication often lacks extensive common ground, the action form itself is an important source of information which could aid in building up a common interaction-based context. Another question is whether infants can use non-social, spatial information to guide their pragmatic understanding and acting when there are no disambiguating pragmatic cues. Non-social spatial cues are commonly used to gauge rational action understanding in processing studies. For example, to examine whether infants expect people to consistently reach and grasp the same object, the object's location is changed (e.g. Woodward, 1998; 1999), and to examine their understanding of action efficiency, the actor either has unconstrained access to the object, or has to circumvent a barrier to reach it (Phillips & Wellman, 2005; Brandone & Wellman, 2009).

Regarding rational action understanding in infants' interactions, imitation studies suggest that infants preferably imitate actions that appear to be freely chosen (and thus rational) compared to those that are forced on the actor through environmental constraints (e.g. Gergely, Bekkering & Király, 2002). A recent follow-up study however suggests that this effect may be explained by an unrelated perceptual confound (Beisert, Zmyj, Liepelt, Jung, Prinz & Daum, 2012). Regarding communicative actions, it is even less clear if or when infants' understanding of communicative actions is informed by spatial cues and expectations about rational actions. In fact, it has been argued that infants' instrumental and communicative action understanding are initially guided by separate systems that follow different processing constraints (Csibra, 2003). Attending to the spatial relationships between oneself, one's communicative partner, and the potential referent objects in the environment, however, would seem crucial for an addressee to understand a message and resolve referential or motivational ambiguity. In support, recent findings with slightly older children show that when addressed with an ambiguous verbal or gestural palm-up request which could refer either to an object within the speaker's reach or an identical one out of the speaker's reach (both equidistant from the child), children select and offer the latter, presumably because they assume that her request is rational and refers to the object out of her reach rather than the one she could more easily obtain herself (Collins, Graham, Chambers, 2012; Grosse, Moll, & Tomasello, 2010).

In the current study we crossed research on infants' action processing and pragmatic understanding. We tested, first, whether infants have an abstracted pragmatic understanding of the form of actions, and second, whether they use non-social spatial information to guide their pragmatic interpretations of instrumental and communicative actions, perhaps in accordance with rationality principles. As in pragmatic understanding studies, we measured infants' action responses to the actions of an experimenter in an interactive context. However, as in processing studies, these actions were presented without the rich social context in which they are usually embedded and without additional verbal, vocal, facial, or other cues. Further, we manipulated the distance of the action to the target object to investigate the role of non-social spatial cues on infants' pragmatic understanding. The objects were equidistant from the infant but either within reach or out of reach for the actor. We tested 12-month-olds on their pragmatic understanding of three actions: instrumental (non-communicative) reaching, communicative index-finger pointing, and conventional palm-up requests. In addition, we employed a no action control condition. Given that infants' action processing reveals understanding of actions that are presented mostly without social cues, and that their interactive responses suggests some pragmatic understanding, we expected that infants would respond appropriately to the three actions and offer the object more often than infants in a No Gesture control group, revealing an abstracted pragmatic understanding of the form of the actions. We also expected infants' responses to be modulated by spatial information. Regarding non-communicative reaching, we expected that infants would understand the action as object-directed and, based on a cooperative motivation, offer the object to the Experimenter. Given that infants' processing of instrumental actions depends on the actions' apparent rationality, we expected them to offer the object more often when it was out of reach for the Experimenter than when it was close and he could obtain it by himself (in the latter case infants might be puzzled and refrain from offering it). Regarding the palm-up gesture, we expected that infants understand its canonical form as representing a communicative request for an object and, based on the cooperative nature of communication, comply with the request. Since the gesture represents an object request, and since it is commonly used in a context of the give-and-take routine where the object is within reach for the caregiver, inferences about rationality may not be relevant and infants should offer the object irrespective of the distance. Regarding pointing, we expected that infants would have different pragmatic interpretations available (e.g., an invitation to look at it, a request to hand it over, or an offer for the infants to take it; Tomasello et al., 2007). The spatial relation between the pointer and the referent may enable infants to disambiguate the motive behind the point if their interpretation is guided by what they perceive to be rational for the pointer in each case. In this case, infants should understand the point to the out-of-reach object as a request and offer it to the experimenter, but the point to the within reach object rather as an invitation to attend to it or take it and offer it less often.

Method

Participants

Sixty-four 12-month-olds (31 boys, 33 girls; mean age = 12;18, range = 11;27 to 12;29) participated in the experiment. An additional 28 infants participated but were not included in the final sample: eighteen due to not finishing the experiment as a result of fussiness, eight due to not yielding enough trial data points (see below), and two due to experimenter error. Infants were recruited from a database of families who expressed interest in participating in research. Infants were primarily white and from middle-class backgrounds, living in a medium-sized European city. Parents received a small gift for participating.

Materials and design

Each participant was tested for her tendency to offer objects in response to an action or gesture. Figure 1 provides a schematic overview of the setup. Room dividers (200cm x 50cm) were located behind and to each side of one of the experimenters (the Placer). The toys designated locations were demarcated by cardboard frames which were attached to the underside of the tablecloth and formed an inconspicuous bump in the cloth surface. The frames allowed the Placer to find the designated place through touch and without overtly attending to it.

Sixteen infants (mean = 7.75 boys and 8.25 girls) were randomly assigned to each of four groups. Each infant received a total of eight trials, of which four had the toy within reach of the Actor (proximal trials) and four out of his reach (distal trials). The trial types were presented in alternating pairs and the order of presentation was counterbalanced across infants. The position of the Actor relative to the infant was varied such that he sat on the left side of the infant in four consecutive trials, and on the right side in the remaining four, with order similarly counterbalanced across infants. The toys presented in the distal and proximal trials were selected to be of comparable shape, size, or material, such that a given distal trial toy had a matching proximal trial toy (with the Actor on the same side). The toys were a white plastic dog and a yellow plastic alligator; a red rubber ball and a green wicker ball; a stuffed bear and a stuffed donkey; a green plastic car and a blue rubber car. The order of toys was counterbalanced in blocks of four.

Procedure

Before the experiment, each infant had a short warm-up session where they played with both experimenters for approximately 10-15 minutes, or until the infant seemed at ease with them. The role of the Placer was played by three different female experimenters, whereas the Actor was always the same male experimenter. The experiment took place with the infant sitting in a hook-on chair attached to a table, the Placer sitting opposite to her, and the Actor perpendicular to her and slightly further away (see Figure 1). To test whether the infant would be able to reach the toys, a small toy was placed on the table before the experiment began, positioned in one of the two locations which would later hold the toys. The infants were encouraged to take the toy, but if they failed to reach it, the Placer placed the subsequent toys slightly closer to the infants than the designated locations.

Familiarization and placing. In all four conditions, the infants were familiarized with the toys at the start of each trial. The Placer introduced the toy to the infant and placed it on the table in front of the Actor, at a distance equal to that of the toy's eventual location during the action. To demonstrate that the location was within the Actor's reach, he picked up the object, made a brief positive comment on it, and put it back on the table. The Placer then pushed the toy over to the infant and encouraged her to take it and explore it. If the infant was reluctant to take the toy, the Placer tried to get the infant interested by playing with it briefly. While the infant explored the toy, the Placer responded naturally to the infants' communicative attempts, but refrained from initiating interactions. To ensure that the Actor was seen as a potential communicative partner, he made a brief positive comment on the toy while the infant was exploring it ("that's nice"). The Actor did not otherwise initiate interactions, but responded nonverbally to any communicative attempts. If the infants showed or offered the toy to the Actor during this time, he acknowledged the action with a brief ambiguous comment ("yes"), after which the Placer took the toy from the infant. This response did not reveal whether or not the Actor wanted to possess the toy, and was thus unlikely to influence infants' interpretation of his subsequent actions. When the infant had explored the object for 45 seconds, the Placer retrieved the toy again. The Placer then looked behind the room divider behind her and expressed surprise as if she had noticed something. As she turned and began to move behind the room divider, she placed the toy distractedly in its designated location, looking neither at the toy nor the infant, giving the impression that her placing was incidental. In the *proximal trials*, the Placer placed the toy in the location corresponding to the Actors' side, such that it was within easy reach for him, and in the *distal trials* she placed it on the side opposite to him and out of his reach. The designated locations were at a distance of 15 cm diagonally from the arm of the infants' chair. If the infant took the toy immediately after the Placer had placed it and before the Actor had a chance to play his part, or if the infant was unable to reach the toy, the trial was terminated.

Point. In the Point condition, the Actor then greeted the infant ("Hey [Name]") and once the infant was attending to him, pointed to the designated location. The Actor pointed for 5 seconds, during which he alternated gaze between the infant and the toy, and repeated the gaze alternation if the infant did not attend to the toy. This was followed by a 5 second interval, during which the Actor's hand was retracted. If the infant did not respond by taking the object, the gesture was repeated, and when another 15 seconds had elapsed without a response, the trial was terminated. If the infants touched the object or took it into their possession at any point, the Actor retracted his hand and looked at the infant. If the infant then offered the object to the Actor, he received it without taking it into his possession, but instead placed it down on the table and made a short ambiguous remark (yes, that's nice"), thereby terminating the trial. If there was no offer, the trial was terminated when 10 seconds had elapsed, with the same remark. As the Actor made the same ambiguous remark in response to the infants taking and offering the object, and did not take the toy into his possession, his response did not reinforce any particular interpretation of his motive.

Request. In the Request condition, the Actor directed his empty upturned hand toward the toy (request gesture). The procedure was otherwise identical to that of the Point condition.

Reach. In the Reach condition, unlike the Point and Request conditions, the Actor did not greet the infant. Instead he reached for the toy as soon as it had been placed, leaning slightly forward in his seat in the distal trials.

No gesture. In the No Gesture condition, the Actor did not direct an action toward the toy. However, to control for the attention directed to the Actor by his greeting and action in the experimental conditions, the Actor in the No Gesture condition drew attention to himself in a non-communicative manner. When the toy was placed, the Actor sighed audibly, glanced briefly at the infant, and then gazed distractedly ahead. Mirroring the 5 second gesture duration and 5 second interval in the experimental groups, the sigh was repeated after 10 seconds if the infant did not respond by taking or touching the object, and when another 20 seconds had elapsed without a response, the trial was terminated. As in the other conditions, the Actor responded to offers by receiving the toy, placing it on the table, and making a short remark. If there was no offer, the trial was terminated when 10 seconds had elapsed with the same remark.

Coding and analyses.

The infant's response was coded as an offer when she took the toy and moved it closer to the Actor, with the apparent intention of giving it to him. Instances where infants held the object and extended their arm toward the Actor with the apparent intent to offer, but then retracted when the Actor tried to receive it, were considered as offers. Instances where infants moved the toy closer to the Actor but it was unclear whether the act was intentional (e.g. when the infant threw the toy in the direction of the experimenter) were not coded as offers. The primary experimenter coded the infants' behavior. To assess interrater reliability, a research assistant blind to the hypotheses coded 16 randomly selected infants (25% of the total sample). The interrater reliability was Kappa = 0.912.

Our dependent measures were the number of infants who offered the object at least once, and the proportion of trials with an offer. First, we addressed our hypothesis that infants have an abstracted pragmatic understanding of the form of actions. Following our predictions and experimental design, we used three planned comparisons to test whether there were more offers in the experimental conditions than in the control condition. Second, we addressed our hypothesis that spatial information would influence the pragmatic understanding of instrumental and communicative actions. We conducted three planned comparisons to test whether infants offered the distal object more often than the proximal one. Since data was not normally distributed and comparisons were based on unequal variances. all comparisons were conducted with non-parametric tests. The results, however, did not change when employing parametric t-tests.

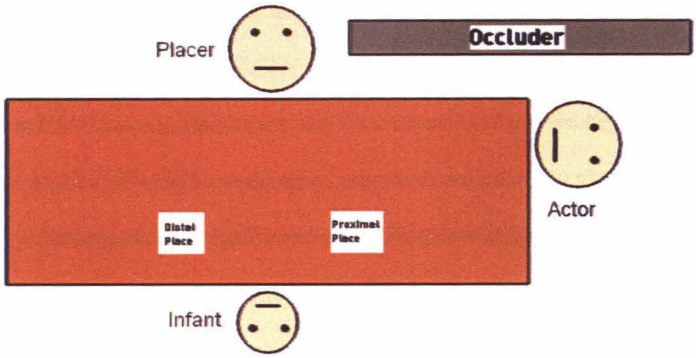

Figure 1. A schematic overview of the experimental set-up during the familiarization and placement.

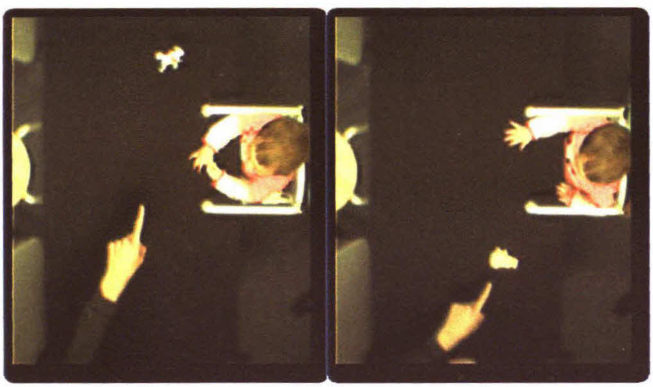

Figure 2. A frame from the ceiling camera showing the Actor pointing in a Distal trial (left) and in a Proximal trial (right)

Results

Our first question was whether infants in the Request, Reach, and Point groups would respond to the Actor's action with an offer, and whether they would do so more often than in the No Gesture control. Chi-square tests revealed that more infants offered the toy at least once in the Request (n=10), Reach (n=9), and Point (n=11) conditions, than in the No Gesture condition (n=2) (Request vs. No Gesture: $\chi2$ (1) = 8.533, p = .003; Reach vs. No Gesture: $\chi2$ (1) = 10.494, *p* = .001; Point vs. No Gesture: $\chi2$ (1) = 6.788, *p* = .009; see Figure 3). Our main dependent measure was the proportion of trials in which infants offered the object to the Actor. The overall means for each group are shown in Figure 4. Planned contrasts revealed that infants in each of the experimental groups offered the object more often than infants in the No gesture control group (Mann-Whitney tests: Request vs. No Gesture: U = 63.00, z = -2.819, *p* = .003; Reach vs. No Gesture: U = 73.00, z = -2.449, *p* = .007; Point vs. No Gesture: U = 57.500, z = -2.991, *p* = .002; all *p*'s one-tailed).

To examine whether higher rates of offering in the experimental groups than in the No Gesture control were due to the infants taking the toy more often in the former, additional analyses were performed on the proportion of trials in which the infants took the toy into possession. Planned contrasts revealed that infants in the Request (*M* = .72, *SD* = .32) and Point (*M* = .62, *SD* = .31) groups were no more likely to take the object than infants in the No Gesture group (*M* = .69, *SD* = .34) (Request vs. No Gesture: U = 121.500, z = -.254, p = .800; Point vs. No Gesture: U = 104.000, z = -.914, *p* = .361; all *p*'s two-tailed). The difference between the Reach (*M* = .52, *SD* = 0.32) and the No Gesture group approached significance (U = 82.000, z = -1.747, *p* = .081), but in the opposite direction, such that the infants in the Reach group tended to take the object less often. Thus, the infants did not take the object more often in the experimental groups than in the No Gesture control group.

The second question was whether infants would be more likely to offer the toy when it was out of reach for the Actor than when it was within his reach. The means are displayed in Figure 5. As predicted, infants in the Point group offered the toy proportionally more often in the distal than in the proximal trials. T = 7.5, *p* (one tailed) = .036 (Wilcoxon Signed Ranks). No differences were found between distances for the other three groups.

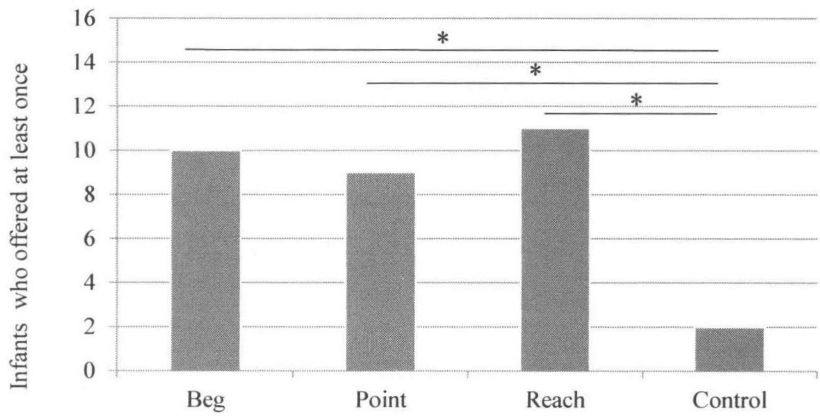

Figure 3. Number of infants in each group who offered the object in one or more trials. Asterisks indicate a significant difference ($p < .05$)

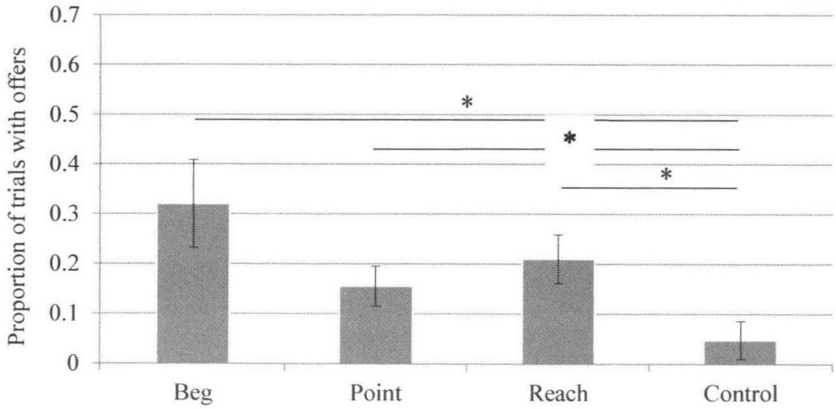

Figure 4. Average proportion of trials in which infants in each group offered the object. Asterisks indicate a significant difference ($p < .05$).

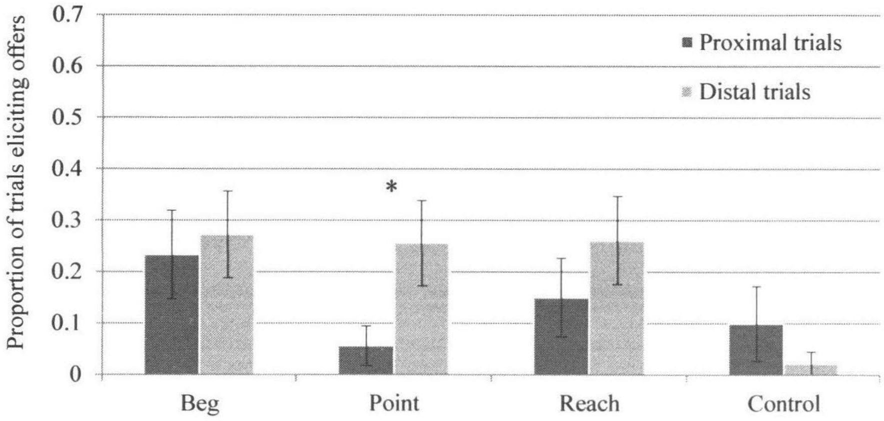

Figure 5. Average proportion of trials in which infants in each group offered the object in proximal trials and distal trials. Asterisks indicate a significant difference (*p* (one-tailed) < .05).

Discussion

Our first question was whether 12-month-old infants have abstracted pragmatic expectations about others' instrumental and communicative actions. We found that infants offer objects in response to a non-communicative instrumental reach, an ostensive pointing gesture, and a conventional request gesture. Infants in all groups offered the objects more often than in a control group where no action was performed but where the procedure otherwise closely mirrored that of the experimental groups. This indicates that the infants' offers were not unrelated attempts to interact or initiate interactions with the Actor, but were made in response to the actions. Importantly, the social context surrounding the actions in the current study did not give infants any cues indicating the actor's desire to obtain the object, and the actor did not express his desire through speech or vocalization. As a result, the infants could only have inferred the motive underlying the actions based on the hand shape of the actor. Thus, the current study shows that 12-month-olds understand and respond to actions and gestures even in the absence of social-contextual, verbal, or vocal cues, suggesting that infants at this age have abstract pragmatic understanding of common actions and gestures.

It is important to note that the differences found in proportion of offers between groups were not simply due to infants taking the object more often into possession in the experimental groups. Infants in the experimental groups were equally likely to take the object into possession as those in the control group. Although overall rates of offering were low, this is not unusual for infants at this age: even when they are explicitly requested through both verbal and gestural means to hand over objects in their possession, they comply with around 40% of requests (Messinger & Fogel, 1998; Hay & Murray, 1982).

Although the two gestures were not accompanied by any vocal, verbal, or contextual cues, the infant was greeted by name and the Actor alternated gaze between the infant and the object during the gesture. The address is likely to have helped infants recognize the subsequent event as a form of communication (e.g. Senju & Csibra, 2008), and the gaze alternation probably helped establish the object as the referent of the gestures. These behaviors did not, however, provide infants with information regarding the motive of the actor. Our pattern of results also suggests that infants' interpretation could not have been driven by these behaviors. First, the infants offered the objects in response to the reach, where the Actor neither looked at nor addressed the infant. Second, the proximity of the objects influenced infants' responses to the two gestures in different ways, suggesting that infants' interpretation was based on the specific hand shape of the gestures.

Our second question was whether infants are able to use spatial information to guide their pragmatic understanding of instrumental and communicative actions. We found that when a person pointed to an object, infants offered the object more often when it was out of reach for the person than when it was within his reach, suggesting that they were more likely to ascribe a motive of request to the gesture when the distance to the object was too great for the person to easily reach the object on his own. When the point was directed to an object within reach, the infants may have instead interpreted it as an invitation for them to attend to the object or to take it. These findings show that infants can use spatial information as a pragmatic cue to infer the motive of ambiguous communicative actions. Previous studies have shown that preschoolers take a speaker's spatial distance from a potential referent into account in inferring the referent of ambiguous referential expressions (Hanna & Tanenhaus, 2004; Collins, Graham, Chambers, 2012), and there is some indication that 21-month-old infants can do this as well (Grosse, Moll, & Tomasello, 2010). The current study shows that infants are sensitive to spatial information in communication from a very early stage of language development. This sensitivity is likely to aid infants' communicative development by helping them resolve referential and motivational ambiguity. Further, it is likely that the infants in the current study construed the point as a request for the more distant object because they understood that requesting assistance to obtain the object was rational only when the Actor could not easily obtain it by himself. Thus, assumptions about rationality appear to guide not only infants' interpretation of observed instrumental actions (Csibra, Bíró, Koós, Gergely, 2003; Sodian, Schoeppner, & Metz, 2004; Csibra, 2008; Gredebäck & Melinder, 2009), but also their interpretation and response to communicative actions.

The finding that infants offer objects in response to an unsuccessful reach suggests that, in addition to processing observed reaching actions as a goal-directed activity (Woodward, 1998; Woodward, 1999; Cannon & Woodward, 2012), infants understand their specific goal to be obtaining the target object, and actively respond to the action by helping the actor obtain it.

Pragmatic understanding of object-directed actions

Previous studies had shown that infants respond to unsuccessful object-directed reaches by touching the object (Hamlin, Hallinan, & Woodward, 2008), and by handing the object over after the reaching person had previously used and dropped it by accident (Warneken & Tomasello, 2007). In the current study, infants did not stop at touching or taking the object into possession, but went on to offer it to the Actor, suggesting that they understood his goal to be to obtain it. In addition, infants responded in the absence of any preceding social context and accompanying cues that would enable them to infer the motive, suggesting that their response was based on the action itself. The finding that infants helped the Actor obtain the object is consistent with previous findings of infants' prosocial motivation (Liszkowski, Carpenter, Striano, & Tomasello, 2006; Warneken & Tomasello, 2007).

In light of studies indicating that infants determine the goal of an action based on its perceived efficiency (Gergely, Nádasdy, Csibra, & Bíró, 1995; Csibra , Bíró, Koós, Gergely, 2003; Sodian, Schoeppner, & Metz, 2004; Csibra, 2008; Gredebäck & Melinder, 2009), it is surprising that infants in the reach group did not offer the object significantly more often in the distal trials - where the distance between the actor and the object constrained the Actor and rendered his action efficient as an attempt to obtain the object - than in the proximal trials - where the proximity to the object rendered his action inefficient as an attempt to obtain the object. It is unlikely that the infants were not aware that the object in the proximal trials was within the Actor's reach, both because his arm was not fully extended, and because the infants had previously seen him reach for and pick up the object from the same distance during the familiarization. A speculative explanation is that infants disregarded the apparent rationality of the reaching action because they understood it not (only) as an instrumental attempt to obtain the object, but as communicating to them a request to hand it over. On this account, the infants offered the object in both distances because they interpret any reaching action that takes place in an interactive setting (where the infant, actor, and object are at a close distance) as a communicative request, despite the lack of overt communicative cues.

The finding that infants offer objects in response to the palm-up request gesture indicates that they understand it to represent a request for an object. In the current study, infants offered the object even though they were not in possession of the object when the Actor gestured, and despite having heard no verbal request from the Actor. In addition, as for the other actions, the Actor's motive could not be gleaned from the social context or any accompanying cues. Unlike the point group, infants in the Request group offered the object equally often in the proximal and the distal trials. This is likely because infants understand the palm-up gesture to have a fixed conventional meaning of request, whereas their interpretation of points depends on the social or spatial context.

Taken together, the results indicate that infants understand and respond to instrumental and communicative actions independently of an informative social context and accompanying verbal, vocal, or bodily cues. The ability to understand the meaning of the hand shape of common actions and communicative gestures is likely to help infants learn from and communicate successfully in their social interactions. The finding that spatial context and rationality influence infants' interpretation of point runs contrary to the predictions of a recent theory of action understanding in infancy which posits that infants' understanding of instrumental and referential actions is governed by separate and independent non-mentalistic action interpretation systems (Csibra, 2003). Briefly, referential action understanding in the first year is proposed to be guided by a principle of directionality, whereas goal-directed (instrumental) action understanding is guided by a principle of efficiency. As these systems are posited to integrate only in the second year of life, the theory predicts that younger infants should not be able to use referential actions to infer goals or motives, nor use goal-directed acts to infer referents. However, the current study shows that 12-month-olds interpret the social motive of a referential action based on spatial-contextual information. This indicates that a more integrated action understanding is already in place in the first year, where infants' communicative experiences may instead be guided by their understanding of people as goal-directed, intentional agents (see Tomasello, Carpenter, Call, Behne, & Moll, 2005).

Chapter 3

Infants' expectations about speech in third-party interactions

Thorgrímsson, G.B., Fawcett, C., Liszkowski, U. (submitted). Infants' expectations about speech in third-party interaction

Abstract

Infants expect people to direct actions toward objects, and they respond to actions directed to themselves, but do they have expectations about actions directed to third parties? Two experiments used eye tracking to investigate 1- and 2-year-olds' expectations about communicative actions addressed to a third party. Experiment 1 presented infants with videos where an adult (the Emitter) either uttered a sentence or produced non-speech sounds. The Emitter was either face-to-face with another adult (the Recipient) or the two were back-to-back.

The Recipient did not respond to any of the sounds. We found that 2-year-olds, but not 1-year-olds looked quicker and longer at the Recipient following speech than non-speech, suggesting that they expected her to respond to speech. These effects were specific to the face-to-face context. Experiment 2 presented 1-year-olds with similar face-to-face exchanges but modified to engage infants and minimize task demands. The infants looked quicker to the Recipient following speech than non-speech, suggesting that they expected a response to speech. The study shows that by 1 year of age infants expect communicative actions to provoke a response from a third-party listener.

Introduction

The ability to understand and anticipate the actions of other people is an essential part of human development. The rudiments of this ability are found in infancy, enabling infants to learn from and communicate with others before they develop the means to participate in linguistic exchanges. Infants monitor others' actions with interest and understand much about what goes on when people act on their environment. They understand that actions can be directed at objects (Woodward, 1998; Behne, Carpenter, & Tomasello, 2005), and they infer unseen goals of ongoing actions (Nielsen, 2009; Southgate & Csibra, 2009; Southgate, Johnson, El Karoui, & Csibra, 2010). Recent eye tracking studies show that infants also make visual anticipations to the goal location of actions, such as an arm transferring a ball into a bucket (Falck-Ytter, Gredebäck, & von Hofsten, 2006), an actor bringing a cup to her mouth (Hunnius & Bekkering, 2010), or a spoon to another person's mouth (Gredebäck & Melinder, 2010).

Of course, not all human actions involve an observable physical movement toward a goal object. People act on each other remotely through speech, gesture, and other forms of communicative actions. Infants are surrounded with the communicative interactions of others and in some non-Western communities observation of third-party interactions may even be the most common form of communication they experience in the first year. How much do infants understand about others' interactions? Do they understand communicative actions as attempts to influence other people (i.e. communication as person-directed action)? Several studies indicate that infants in their second year can learn through observing object-directed actions that communicate information to a third party. For example, at 14 months, infants can infer the location of a hidden object from a pointing gesture addressed to a third party (Gräfenhain, Behne, Carpenter, & Tomasello, 2009). At 18 months they can learn novel object labels from third-party labeling (Floor & Akhtar, 2006), and learn from object-directed emotions displayed to a third party (Repacholi & Meltzoff, 2007). These studies reveal that infants understand and make use of information conveyed through communicative actions that are not directed to them, demonstrating that infants' understanding extends beyond their own interactions. It is less clear, however, what infants understand of the structure of the interaction itself, that is, whether they expect communicative actions to be directed at interactants. Attending to and anticipating how people respond to communication in third-party interactions would facilitate social-observational learning, and provide evidence that infants understand communicative exchanges more abstractly beyond their own interactions.

Recent evidence suggests that infants monitor and understand third-party gestural interactions and can anticipate their outcome. 14-month-old infants viewing a person point or direct a request gesture (open hand, palm up) to an object, successfully anticipated that the addressee would transfer the object to the gesturer (Thorgrimsson, Fawcett, & Liszkowski, 2014). Infants also appear to have some expectations toward third-party addressees of speech. Looking times measures suggest that infants expect a person addressed with a nonsense word to select the same object the speaker had previously expressed a preference for (unbeknownst to the addressee; Martin, Onishi & Vouloumanos, 2012). These findings suggest that infants may understand speech as an action directed to another person to provoke a response, enabling them to actively anticipate how third-party verbal interactions unfold. Such an ability would likely be driven not only by infants' observational experience, but also by their own social interactions, as expectations about and engagement with others' communicative actions are evident from as early as two months of age. In these face-to-face "proto-conversations", infants actively respond to speech and expect and prefer contingent responses from caregivers (Feldstein et al., 1993; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001; Murray & Trevarthen, 1985; Nadel, Carchon, Kervella, Marcelli, & Réserbat-Plantey, 1999 ; Toda & Fogel, 1993).

A recent line of eye-tracking studies have examined infants' looking patterns while viewing third-party conversations. Infants begin to visually orient toward third-party speakers in the second half of their first year and orienting to speakers or other sources of sound becomes more flexible over the first three years. Thus, 6- and 11-month-olds viewing a third-party conversation shift their gaze from the previous to the current speaker during roughly half of the conversational turns, exhibiting rudimentary abilities to follow a third-party turn-taking conversation (Augusti, Melinder, & Gredebäck, 2010). By 3 years of age, children also disengage from the current speaker and orient their gaze to the prospective speaker during the majority of turns (von Hofsten, Uhlig, Adell, & Kochukhova, 2009). In the latter study, 3-year-olds, but not 1-year-olds, were also found to make more such gaze shifts between human interlocutors than between turn-taking objects presented in a matched control condition. A recent follow-up study (Bakker, Kochukhova, & von Hofsten, 2011) similarly found that for 3-year-olds, but not 1-year-olds, such gaze-shifts were more frequent between people than between objects. Further, the older age group made more shifts between people or objects when these emitted speech than when they emitted mechanical sounds. Together, these eye tracking studies suggest that before 12 months of age, infants tend to fixate speakers or other sound sources during their turn, and that by 3 years, children are more flexible in monitoring others' interactions in that they disengage attention from human speakers to addressees.

However, the extent to which infants visually monitor speakers in a conversation does not necessarily reflect their expectations about responses to speech. Since these studies all presented contingent, alternating conversational turns, the infants may have shifted their gaze from speaker to listener (or between turn-taking objects) because they learned to anticipate the contingency of the stimuli, but not because they expected the utterances to provoke a response from the listener. This possibility is especially evident considering that from as early as 2 months of age, infants quickly learn to anticipate side-alternating visual events (Canfield & Haith, 1991). Thus, the fact that 1-year-olds shift their gaze to a contingently reacting object as much as to a human listener in a conversation (von Hofsten et al., 2009), may reflect only their sensitivity to the matched contingencies. Demonstrating that infants switch their gaze to a third-party addressee in the absence of previous or ongoing contingent interaction would give stronger evidence that they expect speech to provoke a response from another person. In addition, since infants are more likely to identify a person as the source of familiar natural human sounds than unfamiliar artificial sounds, assessing infants' expectations about speech will benefit from a comparison with their expectations about common non-communicative human sounds, such as coughs or yawns.

The current study built on the previous research to examine infants' expectations about communicative actions in more detail. We investigated whether infants expect an utterance directed to a third party to be responded to in the absence of previous contingent interactions. To contrast communicative with non-communicative actions in the current study, infants were also presented with identical videos in which the Emitter made a natural non-speech sound (e.g. a cough). We tested 12- and 24-month-olds to investigate whether both older, verbal infants and younger infants with much less linguistic experience would have expectations about speech. In the study, we presented infants with videos of an adult actor in profile (the *Emitter*) uttering a short sentence in the presence of another adult actor (the *Recipient*). Crucially, the Recipient never produced a response, ruling out the possibility that infants could anticipate responses based only on their contingency. We reasoned that if infants expect third parties to respond to speech they should do so in the current situation even without having seen a response occur, and possibly from the first trial on. Previous studies suggest that infants as young as 4 months of age have expectations about event outcomes despite having had no previous experimental exposure to the event (Wang, Baillargeon, & Brueckner, 2003). Moreover, anticipatory eye-tracking studies have shown that despite repeated viewings, infants do not learn to anticipate unusual outcomes of familiar actions (such as a person bringing a spoon to her ear; Hunnius & Bekkering, 2010), suggesting that their action expectations are established outside of the experimental setting.

We also manipulated the orientation of the actors. As mutual gaze between third parties signals affiliation for preschoolers (Nurmsoo, Einav, & Hood, 2012) and adults (Thayer & Schiff, 1974), and speakers typically face their addressees, we expected the body orientation and gaze-direction of the actors to influence infants' expectations. Further, 6-month-olds monitor interlocutors more when they are in mutual than non-mutual gaze (Augusti et al., 2010), and 10-month-olds not only distinguish mutual gaze from non-mutual gaze but also show signs of surprise when an entity that a person looks at and converses with is revealed to be an object, rather than another person (Beier & Spelke, 2012). Studies on infants' understanding of object-directed actions further suggest that gaze direction is a strong cue to goal expectation (Phillips, Wellman, & Spelke, 2002). Thus, it is possible that infants understand speech to be communicative or provoke a response only when the speaker is facing her addressee. In Experiment 1, the actors were either in a typical conversational context with face-to-face body orientation and mutual gaze, or were facing away from each other in a back-to-back configuration (as in Augusti et al., 2010). In Experiment 2 we maintained a face-to-face body orientation but one in which the actors averted gaze and looked down at a paper. In addition, with a between-subject design, Experiment 2 allowed us to analyze how the lack of a contingent response influenced infants' expectations across repeated trials in the same condition. Together, the current study thus allowed an assessment of whether 12- and 24-month-old infants have expectations about third-party communicative actions, and whether these expectations are specific to speech and modulated by body orientation, in the absence of contingent responses.

We were interested in three main comparisons. Our main question was whether infants have stronger expectations of a response to speech than non-speech sounds when interactants are in a typical face-to-face configuration. In addition, we were interested in whether infants' expectations in the speech condition would remain in situations in which the interactants did not look at each other. Finally, we wanted to know whether infants have developed expectations about speech only by the time they are already verbal themselves at 24 months of age, or whether these expectations are already in place at 12-months, when infants have had very limited verbal experience. Our two dependent measures were infants' latency to orient from the Emitter to the Recipient, and the total time that infants looked at the Recipient after the sounds, with the general hypothesis that earlier and longer looking to the Recipient indicates an expectation for her to respond. We tested two specific hypotheses. Our first hypotheses was that when the two actors faced each other, infants would be more likely to expect a response following speech than non-speech, as this is a typical communicative context. Second, given infants' early sensitivity to third-party mutual gaze (Augusti et al., 2010; Beier & Spelke, 2012), and their ability to use gaze to predict actions (Phillips et al., 2002), we predicted that infants would expect a response from speech only when the Emitter is oriented towards the Recipent. Regarding age, one possibility is that infants' understanding of third-party verbal interactions depends on their own verbal communicative experience (von Hofsten et al., 2009; see also Sommerville & Woodward, 2005) and that 12-month-olds – who are only beginning to utter their first words – have not yet developed expectations about third-party speech. Alternatively, 12-month-olds' observations of others' conversations and their participation in proto-conversations may have already led them to form such expectations. To address this developmental question, we examined 12-month-olds' performance more closely in Experiment 2, using an optimized and simplified version of the paradigm used in Experiment 1.

Experiment 1

Method

Participants

Twenty-seven 12-month-olds (11 boys, 16 girls; mean age = 12;20, range = 11;24 to 12;29) and twenty-six 24-month-olds (17 boys, 9 girls; mean age = 24;18, range = 23;29 to 24;29) participated in the experiment. Counterbalancing was based on 24 infants per age group, but in case of dropouts or cancellations, additional infants had been scheduled to ensure the full sample size would be reached. These additional infants were retained to increase power. An additional 33 12-month-olds participated but were not included in the final sample: sixteen due to not meeting the looking criteria (i.e. looking at the Emitter during the period in which she emitted a sound in at least two trials per condition; see Data Reduction), nine due to not finishing the experiment as a result of fussiness, four due to technical error, and four due to unsuccessful calibration. For the 24-month-olds, an additional 9 infants participated but were not included: seven due to not meeting the looking criteria, one due to not finishing the experiment as a result of fussiness, and one due to unsuccessful calibration. Infants were recruited from a database of families who expressed interest in participating in research. Infants were primarily white and from middle-class backgrounds, living in a medium-sized European city. Parents received a small gift for participating.

Apparatus

The infants' eye movements were monitored using a Tobii 1750 remote eye-tracker with an infant add-on. The eye tracker is integrated with a 17" TFT monitor with a native resolution of 1280 x 1024 pixels. The Tobii 1750 has a 50 Hz sample rate, an accuracy of 0.5°, precision of 1°, and allows head movements of up to about 30 cm horizontally, 15 cm vertically, and 20 cm in depth.

Stimuli

Infants were presented with 24 videos (31.37 x 7.66 visual degrees), each with a duration of 6-7 s in total. The videos showed two actors in profile, visible from the shoulders up. Each video began with a period of inactivity (1 or 2 s), followed by the Emitter emitting a sound (mean duration 1.4 s), then by another period of inactivity (4 s). In half of the videos the Emitter uttered speech and in the other half she uttered a non-speech sound. The orientation of the two actors was also manipulated, such that in half of the videos they were face-to-face and in mutual gaze and in the other half they were back-to-back and gazing away from each other (see Figure 1). The conditions were thus *face-to-face/speech*, *face-to-face/non-speech*, *back-to-back/speech*, and *back-to-back/non-speech*. The position and identity of the Emitter was counterbalanced across infants, such that both actors played the role of Emitter and Recipient but each actor was always on the same side of the screen. To avoid too much repetition across trials, three examples of each sound type were presented (see Table 1) and each of them was presented twice, with different durations of the initial inactive period (1 or 2 s). The aim was to make the trials less predictable and more likely to keep the infants engaged. Each condition was presented as a block of 6 trials with both block order and sound example order counterbalanced across infants.

Figure 1. A frame from the face-to-face condition (above) and the back-to-back condition videos (below) of Experiment 1. The rectangles define the two Areas of Interest.

Table 1

Characteristics of the Speech and the Non-speech Sounds in Experiment 1

Speech utterances	Non-speech sounds
"Hoi, hoe gaat het?" (Hi, how are you?)	Throat clearing
"Wil je wat eten?" (Do you want something to eat?)	Vocal sigh
"Ik ben moe vandaag" (I am tired today)	Song (syllable "ta" repeated five times at different pitch levels)
Mean duration = 1256.67 ms, *SD* = 174.32	Mean duration = 1580 ms, *SD* = 133.75

Procedure

Infants were seated in a safety car seat that was placed in their parents' lap so that the infants' eyes were approximately 60 cm from the monitor. Before the experiment the infants' gaze was calibrated using a 9-point calibration during which the experimenter monitored the infants' attention on the screen and showed a short animated video of talking puppets in place of the calibration video to recapture attention when needed. Each infant viewed all four conditions of the experiment (3.5 minutes total). Between trials within a block, a short (1 s) animated video (a small abstract moving picture accompanied by sound) appeared in the middle of the screen in an attempt to ensure that the infants' gaze would be on the screen and centered at the onset of each trial; however, the onset of the trial was not contingent on infants' gaze to these short animations. Between trial blocks, similar but more engaging animated videos (a looming diamond shape accompanied by sound) were presented to center the infants' gaze and to help sustain their attention throughout the study. These longer videos were terminated by an experimenter when the infants' attention was on the screen or once the maximum duration of 9 s was reached. Audio was transmitted through a single desktop speaker connected to the computer and hidden from view behind the monitor.

Data reduction

Rectangular areas of interest (AOI) were created around each of the actors (7.98 x 8.65 visual degrees each; see Figure 1). Infants' unfiltered gaze data points registered within these AOI's at a rate of 50 per second were used to calculate the dependent measures. To account for possible errors in gaze estimation, the AOI's covered an area approximately 30 pixels (0.8 visual degrees) wider and higher than the actors (e.g. Gredebäck & Melinder, 2010). Two critical time phases were selected from the data: the sound phase, during which the Emitter made a sound, and the post-sound phase, during which both actors were silent and immobile. Since the time it takes for infants and adults to initiate a saccade is around 200 ms (Becker, 1972; Canfield, Smith, Brezsnyak, & Snow, 1997), each phase was offset by this amount of time. Thus, the sound phase began 200ms after the onset of the sound and the post-sound phase began 200 ms after the end of the sound, resulting in a duration of 3.8 s for the latter. Two dependent measures - gaze shift latency and looking time - were extracted from the post-sound phase data. To ensure that only trials in which infants were aware of who emitted the utterance were analyzed, ensuring a conservative measure of infants' attention, trials in which no fixation was registered on the Emitter AOI during sound phase were excluded from the analysis of both dependent measures. To measure these fixations to the Emitter during speech, we used the Tobii Clearview fixation filter, and classified gaze data points as fixations when they occurred within a radius of 30 pixels for at least 100 ms (e.g. Hunnius, de Wit, Vrins, & von Hofsten, 2011; Doi, Tagawa, & Shinohara, 2010). The fixation filter was only used for this one measure. All remaining data reduction is based on the unfiltered gaze data points. The number of excluded trials differed slightly between age groups and conditions: a 2 (sound type: speech vs. non-speech) x 2 (orientation: face-to-face vs. back-to-back) x 2 (age, between subjects) mixed analysis of variance (ANOVA) revealed that more trials were excluded in the speech conditions ($M = 1.915$, $SD = .997$) than in the non-speech conditions ($M = 1.475$, $SD = 1.020$), $F(1, 51) = 9.750$, $p = .003$, $\eta^2_p = .160$. This difference is likely due to a slightly shorter duration of the speech sounds, which gave infants less time to orient to and fixate the Emitter. In addition, more trials were excluded per condition for the 12-month-olds ($M = 2.102$, $SD = .868$) than for the 24-month-olds ($M = 1.288$, $SD = .867$), $F(1, 51) = 11.648$, $p = .001$, $\eta^2_p = .186$, possibly because the older age group was more attentive to the display overall. To reduce noise in the data, infants who fixated the Emitter in fewer than two trials of a condition were excluded from our analyses (n = 16). For within-subject designs, two trials per condition are considered the minimum to effectively reduce noise in similar experimental paradigms (see Fernald, Zangl, Portillo, & Marchman, 2006). Nevertheless, to ensure that these exclusions do not bias our results, we conducted an additional analysis on all infants who fixated the Emitter in at least one trial per condition.

The measure of gaze shift latency was used to capture how quickly the infants shifted their gaze from the Emitter to the Recipient following the Emitter's sound. Only trials that had a fixation on the Emitter during the sound could be considered. The measure was calculated by subtracting the start time of the post-sound phase from the time when the first gaze data point was registered within the Recipient AOI. In the main analyses, trials in which infants did not look at the Recipient during the post-sound phase were assigned a score of 3800 ms (the maximum possible latency). The benefits of assigning a maximum score are that these trials represent meaningful behavior (i.e., not looking to the AOI following the sound) and including them yields more data. However, to ensure that the results are not an artifact of these assignments, an additional analysis excluded trials without a look to the Recipient. Our second main measure was a looking time difference score to gauge how long the infants looked to the Recipient compared to the Emitter in the post-sound phase. Difference scores were used to help diminish individual variability in overall looking times. The difference score was calculated by subtracting the looking time to the Emitter from the looking time to the Recipient during the post-sound phase. Looking time was indexed by the number of gaze data points registered within each of the AOI's.

Results and Discussion

Gaze shift latency

Our main question was whether infants looked quicker to and longer at the Recipient in a face-to-face orientated context when she was addressed with speech compared to non-speech. In addition, we were interested in whether the orientation of the actors would have an effect on infants' looking following the sounds, and whether we would find evidence for specific expectations about speech already at 12 months of age. Figure 2 displays the latency scores across conditions and age groups. The scores were square-root transformed to a normal distribution for the analyses. A first omnibus test with a 2 (sound type: speech vs. non-speech) x 2 (orientation: face-to-face vs. back-to-back) x 2 (age: 12 vs. 24 months) mixed analysis of variance (ANOVA) revealed a main effect of orientation ($F(1,51) = 8.881$, $p = .004$, $\eta^2_p = .148$), as well as interactions between orientation and sound ($F(1,51) = 6.276$, $p = .015$, $\eta^2_p = .110$), and between sound and age ($F(1,51) = 4.868$, $p = .032$, $\eta^2_p = .087$). To directly address our main questions, we compared latency for speech versus non-speech in the face-to-face context using an ANOVA including age as a between-subject factor. A 2 (sound type: speech vs. non-speech) x 2 (age) mixed ANOVA revealed a main effect of sound type ($F(1,51) = 6.336$, $p = .015$, $\eta^2_p = .111$), such that infants looked quicker to the Recipient following speech than non-speech, whereas the interaction between sound type and age was not significant ($F(1,51) = 2.322$, $p = .134$, $\eta^2_p = .044$). In the back-to-back context, however, the same analysis did not reveal any significant differences between sound types ($F(1,51) = 1.388$, $p = .244$, $\eta^2_p = .027$), suggesting that the effect of speech in decreasing latency to look at the Recipient was selective to the context where the two actors faced each other. To test directly for the effect of sound type in the face-to-face context for each age group, separate t-tests were performed. These revealed that the 24-month-olds looked quicker to the Recipient following speech ($t(25) = -2966$, $p = .007$), whereas the 12-month-olds did not ($t(26) = -.681$, $p = .502$).

Additional analyses were performed on the latency measure where trials in which the infants did not look at the Recipient were excluded. The measure was square-root transformed to a normal distribution. The omnibus ANOVA revealed a marginal main effect of orientation, with quicker looking to the Recipient in the face-to-face context ($\underline{F}(1, 44) = 3.904$, $p = .054$, $\eta^2_p =$

.082). For the face-to-face context, a 2 (sound type) x 2 (age) mixed ANOVA revealed a marginally significant interaction between age and sound type ($F(1,50) = 3.685$, $p = .061$, $\eta^2_p = .069$). Analyses of simple effects show that the 24-month-olds, but not the 12-month-olds, shifted their gaze significantly quicker to the Recipient following speech than non-speech (respectively, $F(1,50) = 5.91$, $p = .019$, $\eta^2_p = .106$; $F(1,50) = .08$, $p = .778$, $\eta^2_p = .002$). In the back-to-back context, however, the same analysis did not reveal any significant differences between sound type or age (respectively, $F(1,45) = .129$, $p = .721$, $\eta^2_p = .003$; $F(1,45) = .973$ $p = .329$, $\eta^2_p = .021$). Thus, when trials without looks to the Recipient are excluded, the 24-month-olds were found to look quicker to the Recipient following speech in the face-to-face context.

Recall that our main analysis included only infants who fixated the Emitter in two or more trials per condition. We re-analyzed the data using a less conservative exclusion criterion, where all infants who fixated the Emitter in at least a single trial per condition were included. Results were similar to the main analysis with the more conservative inclusion criteria. A 2 (sound type) x 2 (age) mixed ANOVA's for the face-to-face context revealed a main effect of sound type ($F(1,72) = 4.890$, $p = .030$, $\eta^2_p = .064$), but also an interaction between sound type and age, ($F(1,72) = 4.869$, $p = .031$, $\eta^2_p = .063$), with simple effects again revealing that 24-month-olds looked quicker to the Recipient following speech than non-speech ($F(1,72) = 9.03$, $p = .004$, $\eta^2_p = .111$), whereas 12-month-olds did not ($F(1,72) = 2.322$, $p = .997$, $\eta^2_p = .000$). For the back-to-back context, no main effects or interactions were found.

Taken together, the latency analyses all reveal that in the face-to-face context, the 24-month-olds tended to look more quickly to the Recipient following speech than non-speech, whereas the 12-month-olds did not seem to distinguish between the sound types.

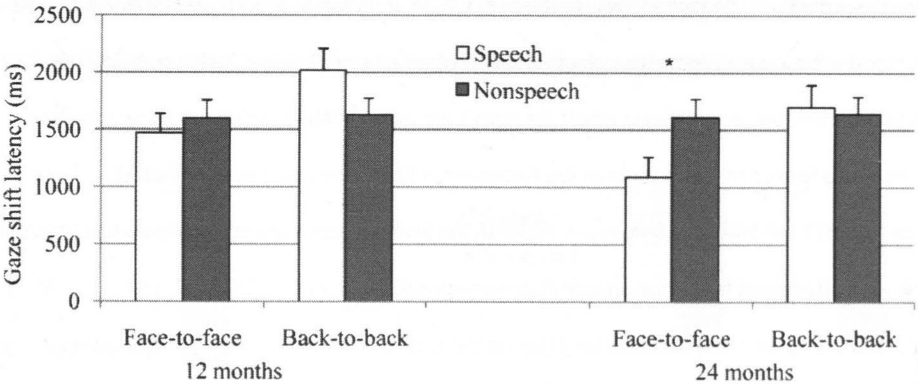

Figure 2. Average gaze shift latency in Experiment 1(untransformed), for all four conditions. Asterisks indicate a significant difference ($p < .05$)

Looking time differences

Our second main measure was looking time differences. We analyzed the difference score for infants' looking time at the Recipient and the Emitter (looking time at Recipient – looking time at Emitter). For the sake of clarity, in the following description and in Experiment 2, higher difference scores (longer looking to the Recipient relative to the Emitter) are described as longer looking to the Recipient. Figure 3 displays the means across conditions and age groups. The initial omnibus 2 (sound type: speech vs. non-speech) x 2 (orientation: face-to-face vs. back-to-back) x 2 (age: 12-months vs. 24-months) mixed ANOVA revealed a main effect of orientation, with longer looking at the Recipient in the face-to-face than in the back-to-back context ($F(1,51) = 7.639$, $p = .008$, $\eta^2_p = .130$), and this effect interacted with sound type ($F(1,51) = 15.059$, $p < .001$, $\eta^2_p = .228$). In addition, orientation interacted with age ($F(1,51) = 5.992$, $p = .018$, $\eta^2_p = .105$). To directly address our main questions, we compared difference scores for speech versus non-speech separately for each orientation as we did for the latency measure. The ANOVA's revealed that infants looked longer at the Recipient following speech than non-speech in the face-to-face context ($F(1,51) = 9.014$, $p = .004$, $\eta^2_p = .150$), whereas the interaction between sound type and age was not significant ($F(1, 51) = 1.101$, $p = .299$). In addition, there was a significant effect for age in the face-to-face context, revealing that the 24-month-olds looked overall longer at the Recipient in the face-to-face context than the 12-month-olds ($F(1,51) = 4.120$, $p = .048$, $\eta^2_p = .075$). In the back-to-back context, however, infants looked less at the Recipient following speech than non-speech ($F(1,51) = 4.258$, $p = .044$, $\eta^2_p = .077$). To test directly for the effect of sound type in the face-to-face context for each age group, separate t-tests were performed. These revealed that the 24-month-olds looked longer at the Recipient following speech than non-speech ($t(25) = 2.425$, $p = .023$), and that the 12-month-olds had a marginally significant tendency to do the same ($t(26) = 1.736$, $p = .094$),

As was done for the latency measure, 2 (sound type) x 2 (age) mixed ANOVA's for the face-to-face and back-to-back contexts were performed on all infants who fixated the Emitter in at least one trial – a lower inclusion criteria than in the main analysis. Again, these results closely mirror our main analysis with the more conservative inclusion criteria. For the face-to-face context, the ANOVA revealed a main effect of sound type ($F(1,72) = 8.055$, $p = .006$, $\eta^2_p = .101$), whereas the interaction between sound type and age was not significant ($F(1,72) = 2.370$, $p = .128$, $\eta^2_p = .032$). For the back-to-back context, no main effects or interactions were found.

To summarize, our main analyses revealed that when the sounds were emitted in a face-to-face context, 24-month-olds, but not 12-month-olds, were quicker to shift their gaze to the Recipient following speech than non-speech. Similarly, 24-month-olds looked longer at the Recipient following speech than non-speech in the face-to-face context, and while 12-month-olds also tended to look longer following speech, this did not reach significance. In the back-to-back context, however, infants did not look quicker or longer to the Recipient following speech than non-speech, but instead looked at her for a shorter duration following speech. For both measures, additional analyses using less conservative inclusion criteria revealed highly similar findings.

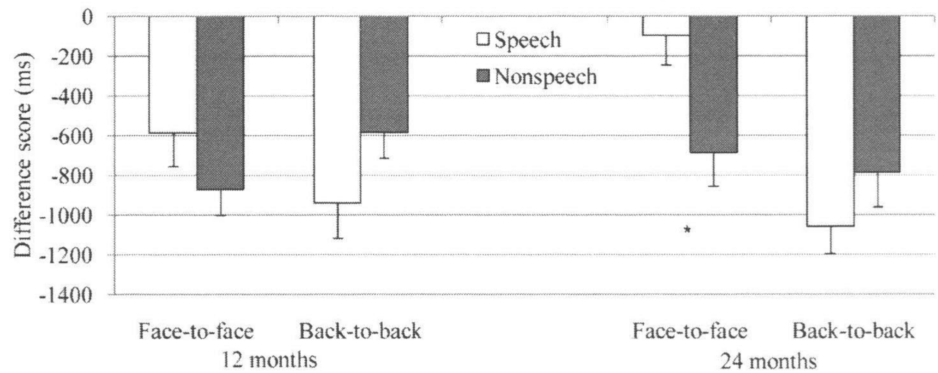

Figure 3. Average difference scores of looking at the Recipient relative to the Emitter in Experiment 1, for all four conditions and both age groups. Asterisks indicate a significant difference ($p < .05$).

Since the measures of gaze shift latency and looking time to the Recipient are influenced by how long infants look at the Emitter following her sound, at first glance it seems possible that the findings reflect not differing expectations toward the Recipient, but differing interests in the Emitter's sound. On this account, infants' longer or quicker looking to the Recipient following speech are due not to a stronger expectation that the Recipient will respond, but to a greater interest in – and thus longer looking to - the Emitter when she emits a non-speech sound. We believe this is an unlikely account for two reasons. First, if infants are more interested in the non-speech sounds, they should look longer at the Emitter following non-speech than speech, irrespective of her orientation. However, the results show that the effect was found only in the face-to-face context. In the back-to-back context this effect either disappeared or reversed, with infants looking longer at the Recipient following non-speech (the equivalent of looking longer at the Emitter following speech). Second, research shows that infants look longer at human faces when they are accompanied by speech sounds than when they are accompanied by human non-speech vocalizations (Vouloumanos, Druhen, Hauser, & Huizink, 2009), suggesting that, if anything, infants would be expected to focus on the Emitter longer after speech as compared to non-speech. Thus, a general preference to look at the Emitter following non-speech is unlikely to account for our findings.

The results reveal that 24-month-old infants show heightened interest in a third-party listener following speech compared to non-speech sounds, suggesting a stronger expectation of a response to speech. Their expectations are restricted to a setting where speech is uttered in a typical face-to-face conversational context. The 12-month-olds, on the other hand, showed little evidence of expecting a response to speech. Possibly, the 24-month-olds' outperformed the younger age group because they have had more experience both as participants in and observers of conversations.

However, given the high attrition rate for the 12-month-olds, and the high number of them who failed to look at the Emitter during her sound, it is possible that the stimuli were simply not engaging enough to sustain their attention, and that they had difficulties realizing which actor had made the sound . Further, the within-subject design with four different conditions and 24 trials may have been too challenging for 12-month-olds' visual attention. Thus, in a second experiment, we made several changes to the paradigm to make the exchanges more realistic and easier to follow than the somewhat artificial exchanges in Experiment 1, and simplified the design to minimize the need to exclude data. The extent of these changes means that Experiment 2 was not a directly comparable variation on Experiment 1, but was designed as a simplified and more natural example of a verbal exchange that younger infants may find more engaging and easier to process. First, to make the social context more natural, the actors smiled and greeted each other before the exchange. The greeting made it clear that both actors were potentially willing to move and communicate. Second, in order to help infants identify the Emitter as the source of the sound, we attracted their attention to the Emitter by having her inhale visibly before emitting her sound. Finally, to reduce the number of trials excluded due to general inattentiveness, we presented fewer trials, using a between-subjects design. The between-subjects design also allowed us to compare infants' performance over trials, a comparison which was not possible in Experiment 1 as potential carryover effects restricted usable trials to the first block, which also led to a low number of infants for the comparison between conditions. Given the possibility that infants learn to anticipate responses through the contingency of conversations, it is important to examine whether they expect a response in the first trials, and how their expectations develop when they repeatedly view utterances that provoke no response.

Apart from improving the paradigm, Experiment 2 was also designed as a more rigorous test of infants' expectations. As gaze direction and body orientation of the actors in Experiment 1 matched, the older infants' looking in the face-to-face context may have been influenced by a tendency to follow the gaze direction of the Emitter after she uttered speech. Infants follow others' gaze direction from early in life (Scaife & Bruner, 1975; D'Entremont, Hains, & Muir, 1997), not only that of their communicative partners, but also that of people depicted in static images (von Hofsten, Dahlström, & Fredriksson, 2005). Although the specific role of speech in gaze-following has not been explored, infants show understanding of the referential nature of both words (Gliga & Csibra, 2009) and gaze (Phillips et al., 2002; Csibra & Volein, 2008; Moll & Tomasello, 2004), and may attend to the object of a speaker's gaze in search of a referent. Thus, it is possible that infants look quicker to and longer at a listener following face-to-face speech (than following face-to-face non-speech) because speech selectively prompts them to follow the speaker's gaze direction to the listener. To address this issue, in Experiment 2 we presented exchanges where the actors were facing each other in terms of body orientation, but were not in mutual gaze during the sound and instead looked at a piece of paper each of them held in their hand. Since the Emitter was not looking at the Recipient but at her piece of paper, infants' looking to the Recipient could not be influenced by the Emitter's gaze direction. Finally, Experiment 2 was specifically designed to probe and corroborate our findings with 12-month-olds, in order to explore further whether infants have expectations about speech already at the early stages of language acquisition.

Experiment 2

Experiment 2 examined 12-month-olds' expectations about third-party interactions further by employing more natural and engaging exchanges in which actors faced each other, but with averted gaze, while they either uttered speech or made a non-speech sound. As in Experiment 1, the Recipient never responded to the sounds, preventing infants from forming expectations based on previous contingent responses in the stimuli. Since we were especially interested in the performance of the younger age group, the second experiment included only 12-month-olds. The between-subjects design allowed us to examine changes in infants' expectations over time, testing two additional hypotheses. If infants' expectations regarding speech are present a priori, they should be seen already in the first half of the experiment, and possibly even in the first trial. Further, if contingent interactions as employed in previous paradigms (von Hofsten et al., 2009; Augusti et al., 2010; Bakker et al., 2011) facilitated anticipations, their absence in the current paradigm might lead to a decrease in expectations.

Method

Participants

Thirty-seven 12-month-olds (22 boys, 15 girls; mean age = 12;22, range = 12;00 to 12;30) participated, with 18 randomly assigned to the speech condition. Counterbalancing was based on 16 infants per condition, but in case of dropouts or cancellations, additional infants had been scheduled and these were included to increase power. Three additional infants participated but were not included in the final sample: Two due to failing to look at the Emitter in at least two trials (see Data Reduction) and one due to not finishing the experiment as a result of fussiness. Infants were recruited as in Experiment 1.

Apparatus

Unlike in Experiment 1, infants' eye movements were measured with a Tobii T120 remote eye-tracker, using a sampling rate of 60Hz. The eye-tracker has an accuracy of 0.5°, precision of 1°, and allows head movements of up to 44 cm horizontally, 22 cm vertically, and 30 cm in depth. The eye tracker is integrated with a 17" TFT display with a native resolution of 1280 x 1024 pixels.

Stimuli

Infants were presented with 8 videos (29.5 x 17.33 visual degrees), which lasted for 1.5 minutes total. Each video began with the actors waving and greeting each other. Importantly, the greetings were completely synchronous in order to avoid the impression of contingent reactivity. Then each picked up a piece of paper and looked at it. The Emitter (shown on the right side in Figure 4) then inhaled visibly and audibly (2 s) before she emitted a sound (mean duration = 1.51 s). After the sound, both actors were inactive (4 s) (see Figure 4). In the speech condition, the Emitter uttered a question or a statement, and in the non-speech condition she either cleared her throat or yawned (see Table 2). For both conditions, the two sound examples were presented in alternation, and each pair was presented twice with the Emitter on the right side of the screen and twice with the Emitter on the left side (video was flipped horizontally), also in alternation. Example type of the first video and Emitter position of the first pair was counterbalanced across infants.

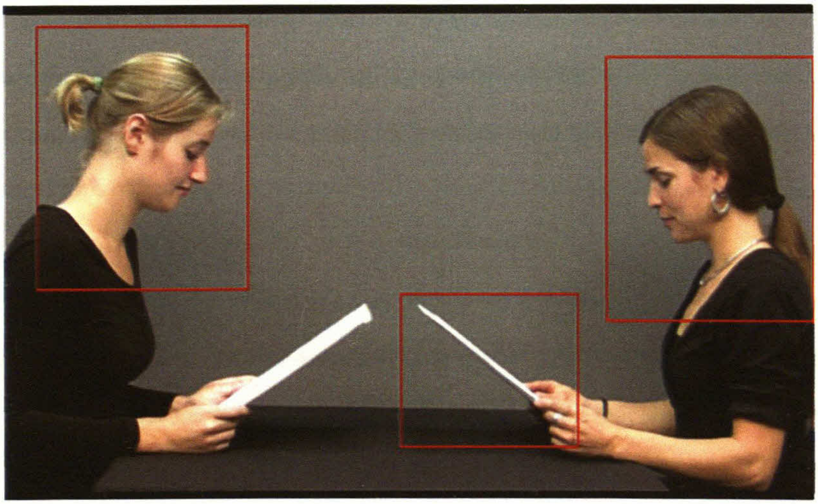

Figure 4. A frame from the post-sound phase of a video from Experiment 2, where the Emitter is on the right side. The rectangles define the Areas of Interest.

Table 2

Characteristics of the Speech and the Non-speech Sounds in Experiment 2

Speech utterances	Non-speech sounds
"Is de kat al buiten?" (Is the cat already outside?)	Throat clearing
"Het is lekker weer" (The weather is nice)	Vocal sigh
Mean duration = 1360 ms, *SD* = 56.57	Mean duration = 1660 ms, *SD* = 84.85

Procedure

Infants were seated in a safety car seat that was placed in their parents' lap so that the infants' eyes were approximately 64 cm from the monitor. They saw a short animation (a colorful looming ball accompanied by sound with a duration of 4s) between trials 4 and 5 to reengage their attention in case it had drifted away from the screen. As the greeting at the start of each video was highly salient, and the Emitter drew attention to herself before she made the sound, no attention getters were displayed between the other trials. In other respects, the procedure was identical to that used in Experiment 1.

Data reduction

As in Experiment 1, rectangular AOI's were created around each of the actors' faces (7.77 x 9.75 visual degrees each; see Figure 4) covering an area approximately 30 pixels (0.71 visual degrees) wider and higher than the actors. Unfiltered gaze points registered in these areas were used to calculate the measures of gaze shift latency and difference score. The sound phase and the post-sound phase were selected as in Experiment 1, and the post-sound phase had the same duration in both conditions (3.8 s). As in Experiment 1, for the latency analysis, trials in which infants did not look at the Recipient were given a maximum score (3800). For both dependent measures, trials in which the infants did not fixate the Emitter during the sound were excluded. On average, 2.27 trials out of 8 per infant were excluded for this reason, but the number of trials excluded did not differ between conditions ($t(35) = .341$, $p = .735$). Infants who did not fixate the Emitter in at least two trials were excluded from the analysis ($n = 2$; both in the non-speech condition). To measure fixations to the Emitter during speech we used the Tobii Studio fixation filter, with the default 35 pixel velocity and distance thresholds. This newer and more accurate fixation algorithm is slightly different than the one used in Experiment 1, with a minimum fixation length of 5 data points (83 ms rather than 100 ms). We re-analyzed the current data with the fixation filter from Experiment 1. When using the filter used in Experiment 1, more trials had to be excluded (36 trials in total), but the main findings remained the same (see Results and Discussion).

Results and Discussion

We conducted ANOVAs on both dependent measures with condition as a between subjects variable and first vs. second trial block of the experiment as a repeated measure to examine performance over trials. The gaze shift latency measure revealed that infants looked quicker to the Recipient following speech than non-speech ($F(1, 34) = 5.896$, $p = .021$, $\eta^2_p = .148$). A marginally significant interaction was found between sound type and block ($F(1, 34) = 3.197$, $p = .083$, $\eta^2_p = .086$). To address our question of learning over trials we conducted a simple effects analysis which revealed that infants looked quicker to the Recipient following speech than non-speech in the first block of the experiment ($F(1, 34) = 8.34$, $p = .007$, $\eta^2_p = .197$), but not in the second block ($F(1,34) = 1.84$, $p = .184$, $\eta^2_p = .051$; see Figure 5). An analysis on the first trial (in which infants fixated the Emitter) similarly revealed that infants in the speech condition were quicker to look to the Recipient than infants in the non-speech condition ($t(35) = -2.770$, $p = .009$).

Several additional analyses were performed on the latency measure. First, a 2 (sound type) x 2 (block) ANOVA was performed where trials without a look to the Recipient were excluded instead of given a maximum latency. This analysis revealed a main effect of sound type ($F(1, 29) = 4.659$, $p = .039$). A first trial analysis on the same data revealed a marginally significant effect of sound type, such that infants in the speech condition looked quicker to the Recipient in the first trial than infants in the non-speech condition, $t(34) = -1.922$, $p = .063$). Finally, another 2 x 2 ANOVA was performed after applying the fixation filter from Experiment 1. This analysis again yielded a main effect of sound type ($F(1,31) = 7.709$, $p = .009$, $\eta^2_p = .199$), such that infants in the speech condition were overall quicker to look at the Recipient.

Expectations about speech in third-party interactions

An ANOVA for the difference score measure, using the same factors as the analysis for latency, revealed a similar pattern of looking as the 24-month-olds showed in Experiment 1, with longer looking to the Recipient than the Emitter. However, these differences were not statistically significant ($F(1,34) = 1.240$, $p = .273$, $\eta^2_p = .035$), nor were there other significant main effects or interactions (p's $> .10$; see Figure 6).

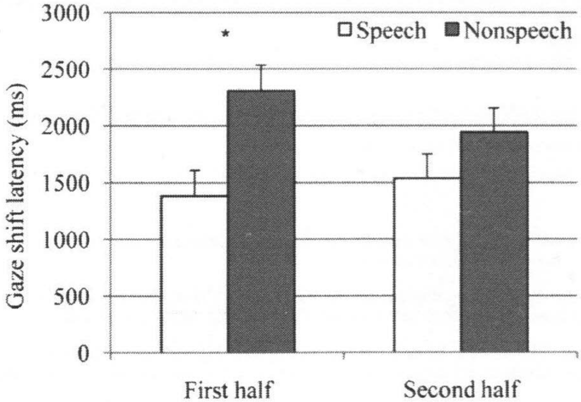

Figure 5. Average gaze shift latency in the speech and non-speech conditions of Experiment 2, for the first and the second half of the trials. Asterisks indicate a significant difference ($p < .01$)

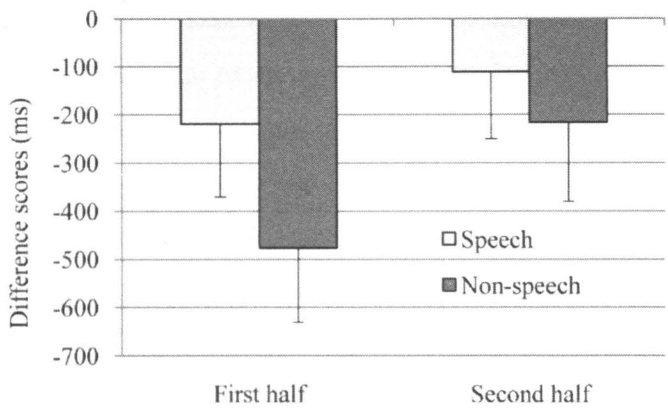

Figure 6. Average difference scores of looking at the Recipient relative to the Emitter in the speech and non-speech conditions of Experiment 2, for the first and the second half of the trials.

It is theoretically possible that infants did not detect the Emitter's downward gaze direction, but instead assumed from the body orientation of the two actors that the Emitter's attention was directed to the Recipient. However, based on the reviewed literature of infants' understanding of gaze direction we find this unlikely. Further, after the Emitter has greeted the Recipient, and before she makes the sound, she shifts her gaze down from the Recipient toward the paper and then picks up the paper and tilts it upward. These events make the Emitter's downward gaze direction highly salient and presumably easy for infants to notice, especially given that infants follow gaze in both dynamic (e.g., Senju & Csibra, 2008) and even static images, where there is no change in gaze direction (von Hofsten, Dahlström, & Fredriksson, 2005).

To summarize, infants in the speech condition were quicker to shift their gaze to the Recipient than infants in the non-speech condition, suggesting that they had stronger expectations for her to respond to speech. As the two actors were not in mutual gaze during the sound, but instead gazed down to their papers while facing each other in body orientation, the differences found between conditions could not have been due to infants' selectively following the gaze-direction of the Emitter after she spoke. Interestingly, the latency measure indicates that infants' differential expectations of a response to speech over non-speech ceased in the second half of the experiment. Unlike Experiment 1, no significant differences were found for the measure of looking time, although the means were in the predicted direction. While the lack of this effect remains unexplained, one possibility is that the higher complexity of the visual scene in Experiment 2 led infants to distribute their attention more evenly. Further, both actors were more active in Experiment 2 (greeting each other and picking up and looking at papers), so infants may have equally expected further actions from both of them.

General Discussion

We found that 12- and 24-month-old infants looked quicker and longer to a third-party listener following speech compared to non-speech sounds, suggesting a stronger expectation of a response to speech. In Experiment 1, only the 24-month-olds showed evidence of expecting a response, whereas 12-month-olds did so when presented with the more natural and animated exchanges in Experiment 2. In addition, the 24-month-olds' were found to have such expectations toward the listener only when the speaker and listener faced each other, but not when they were back-to-back. The findings indicate that infants distinguish speech from other natural human sounds as having the potential to influence another person, at least when the orientation of the speaker and listener is typical for a communicative interaction. In addition, these looking patterns could not have resulted from a tendency to follow the gaze of the Emitter to the Recipient, since in Experiment 2, where the two actors averted their gaze, 12-month-olds showed evidence of having a stronger expectation of a response to speech. Thus, the current findings suggest that infants in the early stages of language acquisition not only monitor ongoing third-party conversations (von Hofsten et al., 2009; Augusti et al., 2010; Bakker et al., 2011), but also expect third-party communicative acts to provoke a response. A recent looking time study showed that 12-month-olds understand that speech, as opposed to a cough or an emotional vocalization, can function to transfer information about an object to an addressee in a third-party interaction (Martin et al., 2012). The current finding compliments and extends this previous study with online-processing measures by showing that infants also actively anticipate that third-party addressees will respond to speech, but not to natural non-speech sounds.

At first glance our findings may seem at odds with those from a previous study that included comparisons between infants' looking at a normal conversation and the same conversation but with the speech replaced by mechanical sounds (Bakker et al., 2011). No differences were found in how often 1-year-olds or 3-year-olds disengage from the current to the prospective speaker, nor how long they looked at the scene. However, since humans do not emit mechanical sounds, infants in that condition may have perceived the back-and-forth sounds between the two people looking at each other as a form of communication (e.g. Johnson, Slaughter, & Carey, 1998). In contrast, the current study used natural non-speech sounds familiar to infants, which may have helped them distinguish non-communicative from communicative sounds. Further, in contrast to previous studies (von Hofsten et al., 2009; Augusti et al., 2010; Bakker et al., 2011), the design of the current experiments did not involve contingent reactions, which, if contingent on infants' own behavior, can lead them to infer communication even from non-human, non-verbal agents (Deligianni, Senju, Gergely, & Csibra, 2011). Instead, both experiments controlled for contingency between the actors so that it could not influence infants' gaze patterns. Even the initial greeting in Experiment 2 was synchronous and thus did not demonstrate contingent turn-taking.

Although both our dependent measures are also influenced by how long infants look at the Emitter, the results from Experiment 1 show that infants' longer and quicker looking at the Recipient following speech is not due to a general preference to look at the Emitter following non-speech. Such a preference would have led to longer looking at the Emitter irrespective of the actors' orientation, whereas the 24-month-olds only showed the effects in the face-to-face context. Another possible explanation for infants' looking patterns is that the infants did not have a stronger expectation of a response to speech, but rather that speech facilitated their gaze-following, leading to shorter latencies and longer looking at the Recipient following speech in the face-to-face context. However, the design of Experiment 2, where the Emitter faced the Recipient but averted her gaze downward during the exchange, ensured that infants' looking to the Recipient was not influenced by the Emitter's gaze direction.

One remaining question is why the 12-month-olds showed evidence of expecting a response to speech in Experiment 2, but not in Experiment 1. We did not design Experiment 2 to make a direct statistical comparison to Experiment 1, but in Experiment 2 we made various modifications to the paradigm from Experiment 1 that we believe succeeded in facilitating infants' ability to process and follow the exchange. For example, infants were much more likely to look at the Emitter during the sound in Experiment 2 - where the Emitter drew attention to herself before the sound by inhaling - than in Experiment 1, where the Emitter made the sound suddenly and unpredictably. The greeting that preceded the sound in Experiment 2 may also have helped infants recognize that the actors were generally social and communicative, or made it clearer that they were socially engaged with each other. However, as these procedures were identical in the speech and the non-speech conditions, they cannot account for the differences we found between conditions.

Infants' increased expectations for the Recipient to respond to speech in Experiment 1 were only present when the actors were facing each other. Previous research shows that infants as young as 6 months are sensitive to whether conversational partners are face-to-face in mutual gaze, or back-to-back (Augusti et al., 2010), and at 10 months they also distinguish between these contexts for static, silent images (Beier et al., 2012). The findings from Experiment 1 indicate that at least by 2 years of age, infants also see the face-to-face orientation with mutual gaze as a setting in which speech is likely to provoke a response. This may be because they understand the face-to-face orientation as a typical communicative context, having learned that speakers typically face their addressees. As the 12-month-olds in Experiment 1 did not have increased expectations in either context, and Experiment 2 did not manipulate the gaze direction or body orientation of the actors, it is less clear how body orientation or mutual gaze affects expectations at this age. Nevertheless, 12-month-olds in Experiment 2 showed evidence of having a stronger expectation of a response to speech when the actors were face-to-face but with their gaze averted, suggesting that mutual gaze during speech is not a prerequisite for them to expect a response from a third-party listener, though face-to-face body orientation may be.

It is unclear how much of the semantic content of the utterances infants understood, and whether this influenced their expectations. It cannot be ruled out that the 24-month-olds outperformed the 12-month-olds in Experiment 1 partly because they understood more language. However, as infants at 12-months expect even unfamiliar (nonsense) words to be able to convey information to a third party (e.g., Martin et al., 2012), it is likely that perceiving the form of speech, rather than understanding its semantic content, may be enough to create an expectation of a response.

Expectations about speech in third-party interactions

A further question is how infants develop expectations about third-party communicative actions. One possibility is that infants build such expectations from their own experience in communication. In contrast to claims relating infants' relevant experience specifically to linguistic skills (von Hofsten et al., 2009), our results with 12-month-olds suggest that minimal experience in linguistic communication may be sufficient, and point to the possibility that infants' expectations are formed through non-verbal vocal interactions in the first year. Another possibility is that infants develop their expectations through observing and overhearing others' conversations, without needing to participate as a communicative partner with others. While our stimuli presented no contingent reactions that infants could learn from, it is likely that their expectations stem at least in part from previous experience with others' contingent turn-taking. It will be important to examine whether infants' understanding of communicative actions addressed to them is prerequisite for their understanding of third-party communication, or when these two integrate and bring infants to a more abstract understanding of the communicative act (see also Moore, 2007).

The current study adds to recent findings on infants' action understanding. Beyond infants' expectations about object-directed actions, the study suggests that infants also have expectations about others' communicative actions. That is, even at the time infants are only beginning to communicate through speech, they appear to have developed the expectation that speech should be responded to by another person. This skill is valuable because attending to the response of a third-party recipient enables social learning not only about individual object-directed goals but also about social goals of acting together (see also Fawcett & Liszkowski, 2012), and may even be a major route of learning in traditional communities where infants and children are afforded fewer opportunities for direct communication with others.

Expectations about speech in third-party interactions

Chapter 4

Infants' expectations about gestures and actions in third-party interactions

Thorgrimsson, G.B., Fawcett, C., & Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Frontiers in Psychology, 5: 321*

Abstract

We investigated 14-month-old infants' expectations toward a third party addressee of communicative gestures and an instrumental action. Infants' eye-movements were tracked as they observed a person (the Gesturer) point, direct a palm-up request gesture, or reach toward an object, and another person (the Addressee) respond by grasping it. Infants' looking patterns indicate that when the Gesturer pointed or used the palm-up request, infants anticipated that the Addressee would give the object to the Gesturer, suggesting that they ascribed a motive of request to the gestures. In contrast, when the Gesturer reached for the object, and in a control condition where no action took place, the infants did not anticipate the Addressee's response. The results demonstrate that infants' recognition of communicative gestures extends to others' interactions, and that infants can anticipate how third-party addressees will respond to others' gestures.

Expectations about gestures in third-party interactions

Introduction

Infants are highly attuned to other's actions. In their first year, they interpret both instrumental actions, such as reaching, and communicative actions, such as pointing, as goal-directed behavior (Woodward & Guajardo, 2002). For example, they expect others to consistently reach for the same object (Woodward, 1998; Cannon & Woodward, 2012) and to reach in an efficient manner given the environment (Brandone & Wellman, 2009). They follow pointing gestures to objects (Deák, Flom, & Pick, 2000) and can infer the social motive of a point based on social contextual cues (Behne, Carpenter, & Tomasello, 2005; Liebal, Behne, Carpenter, & Tomasello, 2009). Understanding actions in terms of goals or motives enables infants to learn about actions and objects through observation and interaction, and to engage in relatively complex nonverbal communicative interactions. However, although the focus of the majority of research on communication and action understanding in infancy is on dyadic settings - where the infant observes or engages with another person in a one-on-one exchange - these represent only a part of infants' early communicative experience. During their first year infants also routinely observe and overhear communicative interactions between other people that provide infants with another source of information about the social world. In traditional cultures where preverbal infants are rarely directly addressed by caregivers (e.g. the Tzeltal Mayans: Brown, 1998; See also Lieven, 1994), such observational experiences could play a particularly important role in infants' social and communicative development.

Recent experimental research indeed shows that infants monitor and learn from actions that are not directed at or addressed to them, but that they observe and overhear being addressed to a third party. The bulk of this research focuses on infants' ability to learn words or actions used in third-party interactions. For example, 18-month-olds can learn object labels through overhearing the labels being used in others' interactions (Floor & Akhtar, 2006), even when the objects are labeled only indirectly (Gampe, Liebal, & Tomasello, 2012). Studies on imitative learning similarly show that infants at this age will imitate a novel action demonstrated to a third party (Herold & Akhtar, 2008; Matheson, Moore, & Akhtar, 2013) and even attempt to imitate the social nature of an action demonstration (Fawcett & Liszkowski, 2012).

It is less clear, however, how much infants understand about the structure and outcome of third-party communicative interactions. Such understanding would entail not only an ability to imitate or learn from actions addressed to third parties, but to anticipate how the addressees respond. A few recent studies suggest that infants have expectations toward addressees of speech in third-party interactions. Specifically, 12- and 24-month-olds are quicker to shift their gaze from a person to a third-party addressee when the person utters speech, than when the person emits natural non-speech sounds, suggesting a stronger expectation of a response to the former (Thorgrimsson, Fawcett, & Liszkowski, 2011). A looking time study further revealed that 12-month-old infants expect addressees to respond to speech in accordance with the speaker's previous object-directed actions, suggesting that infants recognize that speech can transfer information about an object (Martin, Onishi & Vouloumanos, 2012). Infants also seem to have some understanding of the use of gestures in third party communication. When presented with a scene where one person indicates the location of a hidden toy to a third-party addressee by pointing at it in a communicative fashion, infants were able to locate and retrieve the toy themselves (Gräfenhain, Behne, Carpenter & Tomasello, 2009). Together, these studies suggest that infants have expectations regarding the addressee of speech in third-party interactions and that they can pick up on information conveyed through gestures. It remains unknown, however, whether infants can also anticipate the actions of the addressee of gestures in third-party interactions.

The current study used eye tracking to examine infants' expectations toward third-party addressees of gestures and actions. Fourteen-month-olds watched a third-party interaction where one person (the *Gesturer*) directed a gesture (a point or a palm-up "request" gesture) or an instrumental action (a reach) to an object between them. The other person (the *Addressee*) then responded by grasping the object and dispensing it through one of two tubes that led to each person, thereby either giving the object to the Gesturer, or taking it for himself. As the Addressee's hand and the tube entrances were concealed by an occluder, the infants could not see into which tube he placed the ball, and had to look at the tube exit to see the outcome. Importantly, the Gesturer did not verbally communicate any motive, and as both actors were visible only from the neck down, no information could be gleaned from their facial expressions or their gaze direction.

We expected that the nature of each of the three demonstrated actions would lead infants to anticipate different outcomes. A point is a deictic gesture that gets its meaning from the social context within which it is used. In the context of the study, the point can be construed either as communicating a request for the object, or an offer for the Addressee to take it. As infants point for various social motives, including requesting, sharing interest, and informing, and their interpretation of others' points depends on the social context in which they take place (Behne et al., 2005; Behne et al., 2011; Liebal et al, 2009), they may be ready to interpret the pointing gesture in the current third-party context both as communicating a request for the object, and an offer for the Addressee to take it, leading them to anticipate both outcomes. The request gesture, on the other hand, is a conventional gesture communicating a request to hand over an object, which caregivers use from early on in give-and-take exchanges, and which infants recognize and respond to (Hay & Murray, 1982; Thorgrimsson & Liszkowski, 2012). If infants' recognition of the request gesture extends to third-party interactions, they should only anticipate the give outcome. Finally, infants start making reaching attempts at around 4 months of age (von Hofsten & Lindhagen, 1979), and there is a lot of evidence that infants expect others' reaching attempts to indicate goal-directed behavior, both from action-processing studies (e.g. Woodward, 1998; Woodward & Guajardo, 2002; Cannon & Woodward, 2012) and interaction-based studies (e.g. Warneken & Tomasello, 2007). Although infants will likely understand the reaching action in the current study as an instrumental attempt to grasp the object, a reach is typically not used to communicate a request, and thus it is unclear if infants see it as having relevance to the Addressee. However, it is also possible that infants expect people to help others achieve instrumental goals and thus expect the Addressee to give the object. These three actions were compared to a control condition where the Addressee did not gesture but remained silent and immobile for an equal duration of time.

Method

Participants

Eighty 14-month-old infants (41 boys, 39 girls; mean age = 14;17, range = 13;24 to 14;29) participated in the experiment. An additional 19 infants participated but were not included in the final sample: twelve due to not finishing the experiment as a result of fussiness, five due to watching less than 25% of the demonstration phase, and two due to technical issues. Infants were recruited from a database of families who expressed interest in participating in research. Infants were primarily white and from middle-class backgrounds, living in a medium-sized European city. Parents received a small gift for participating.

Apparatus

Infants' eye movements were measured with a Tobii T120 remote eye-tracker, using a sampling rate of 60Hz. The eye-tracker has an accuracy of 0.5°, precision of 1°, and allows head movements of up to 44 cm horizontally, 22 cm vertically, and 30 cm in depth. The eye tracker is integrated with a 17" TFT display with a native resolution of 1280 x 1024 pixels.

Stimuli and design

The study had a between-subject design with four groups of 20 infants each. In the experimental groups the Gesturer pointed, reached, or directed a request gesture toward the object, but in the no-gesture control group no action was performed. Infants were presented with a total of 13 videos (29.50 x 14.32 visual degrees excluding black bars at the top and bottom of the screen), which collectively lasted approximately 5 minutes. The videos showed two people (Gesturer on left and Addressee on the right) in profile sitting at a table, with a small black shelf between them. A transparent tube ran from each side of the shelf to a bowl in front of each person.

As the tubes would be occluded from view during the test trials, infants in all groups first viewed a *demonstration* video without an occluder where a third person (the Demonstrator) stood behind the shelf and distributed balls through the tubes demonstrating how the tubes functioned (see Figure 1). In the demonstration, the Demonstrator picked up a ball from the top of the shelf, looked at the Addressee or Gesturer (whichever would receive the ball that time) with a brief smile, and placed the ball into the tube leading to his or her bowl. The Demonstrator distributed four balls in this way (two blue and two yellow), alternating between the Addressee and the Gesturer. Using video editing software, the ball's movement through the tube was slowed down such that it took 1 second to travel through the tube, and was accompanied by a previously recorded sound of a ball rolling through a long paper tube. To give the impression that the two people were interested in the balls, they responded by picking up the ball, looking at the Demonstrator and vocalizing happily. The demonstration phase had a duration of 69 seconds and ended with the Demonstrator placing an occluder in front of the shelf.

Next, each of the 12 *test trial* videos revealed the Gesturer and the Addressee seated as before and a single ball on top of the shelf. The upper parts of the tubes were concealed by the occluder, such that only the tube exits could be seen protruding from behind it. To direct infants' attention away from the actors' faces and to their actions, and to prevent infants from detecting their gaze direction, the actors were visible only from the neck down. In the experimental groups, the Gesturer directed an action (point, request, or reach) toward the ball and sustained it for 2 seconds before retracting her arm (see Figure 2). As the extension of the Gesturer's arm was identical in all three experimental groups, the only difference between groups was the Gesturer's hand shape. In the no-gesture control group, the Gesturer did not act toward the object, but sat immobile for the same duration of time. In all four groups, the Addressee then responded by reaching for and picking up the ball, and placing it into one of the tubes. As the Addressee's hand and the tube entrances were concealed by the occluder, the infants could not see into which tube he placed the ball, and would have to look at the tube exit or the bowl to see the outcome. As in the demonstration phase, the balls' movement through the tube was artificially slowed down. Four seconds elapsed from the time the Addressee's hand disappeared behind the occluder and until the ball emerged from the tube. To facilitate anticipation, the last 2 seconds of the ball's movement were accompanied by the same sliding sound as in the demonstration phase. For each group, in one half of the test trials the ball emerged on the side of the Gesturer (*Give* trials) and in the other half on the side of the Addressee (*Take* trials). The twelve test trials were presented as blocks of six Give trials and six Take trials, with block order counterbalanced across infants. The first block always featured blue balls and the second yellow balls, so that for each infant, Give and Take trials were also distinguishable by the color of the balls.

Figure 1. A frame from the demonstration phase showing the Demonstrator distributing balls

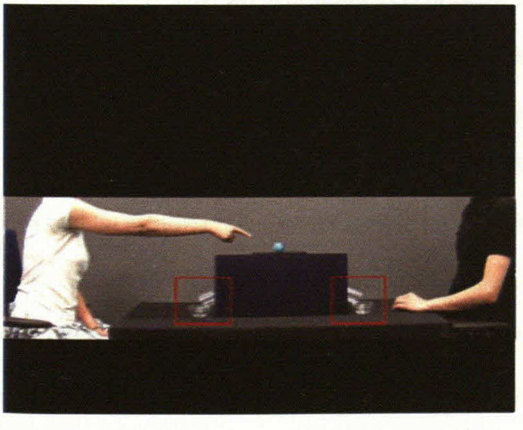

Figure 2. A frame from the test trial in the point group, showing the Gesturer point to the ball. The two red squares delineate the Areas of Interest

Procedure

Infants were seated in a safety car seat that was placed in their parents' lap so that the infants' eyes were approximately 60 cm from the monitor. Before the experiment the infants' gaze was calibrated using a 9-point calibration during which the experimenter monitored the infants' attention on the screen and showed a short animated video of talking puppets in place of the calibration animation to recapture attention when needed. Between trial blocks, a short (4s) animated video (a small shaking cartoon bird accompanied by sound) was played to sustain the infants' attention. Audio was transmitted through a single desktop speaker connected to the computer and hidden from view behind the monitor.

Data reduction

Rectangular areas of interest (AOI's) were created covering the exit of the tube and the bowl on each actor's side (3.30 x 2.83 visual degrees each; see Figure 2). Infants' raw gaze data points registered within these AOI's at a rate of 60 per second were used to calculate the dependent measures. To account for possible errors in gaze estimation, the AOI's covered an area approximately 30 pixels (0.8 visual degree) wider and higher than the tubes and bowls (e.g. Gredebäck & Melinder, 2010). The time window selected for analysis – the anticipatory phase – was calculated by finding the frame when the Addressee touched the ball and the frame when the ball emerged from a tube (5.5 seconds). As the time it takes for infants and actors to initiate a saccade is around 200ms (Becker, 1972; Canfield, Smith, Brezsnyak, & Snow, 1997), the anticipatory phase was shifted forward by 200ms from the onset time of these two frames. Although infants were not expected to anticipate the emergence of the ball from the tube until the Addressee's hand had disappeared behind the occluder, it is possible that they expect the

Addressee to manually transfer the ball into either bowl, and thus the anticipatory phase started as soon as the Addressee grasped it.

Two dependent measures were extracted from the gaze data. The proportional looking time measure was calculated by dividing the looking time to the Gesturer's AOI by the total looking time to both AOI's during the anticipatory phase. Looking time was calculated by summing the gaze data points that fell within the AOI's. The first look measure is a simple binomial measure that specifies which AOI the infant looked at first. It was calculated by subtracting the time at the onset of the anticipatory phase from the time at which the first gaze data point was first registered within each AOI (the latency) and selecting the AOI with the shorter interval. If infants anticipate that the ball will emerge on the Gesturer's side, then their proportional looking time and their first looks to the Gesturer's AOI should be greater than chance (50%).

Results

As the measures tended to be skewed, and tests for normality (Kolmogorov–Smirnov) revealed that the measure of first look deviated from normality for the Give outcome in the point group ($D(20) = .232$, $p = .006$) and marginally so for the Take outcome in the Reach group. ($D(20) = .221$, $p = .011$), non-parametric tests were used. The main analyses are one-sample Wilcoxon signed-rank tests, measured against chance levels (a median of 0.5). The Bonferroni correction was used to correct for multiple comparisons, yielding a two-tailed significance threshold of .0063.

Request

Infants in the request group looked at at least one of the AOI's during the anticipatory phase in 73.75% of the trials (SD =19%). The measure of proportional looking time to the

Gesturer's bowl in the Give and Take trials revealed that the infants tended to look longer at the bowl into which the ball emerged (the *target bowl*) in the Give trials (*Median* = 0.80, W = 162.000, p = .007), but not in the Take trials (*Median* = 0.51, W = 114.500, p = .432), indicating that they only anticipated the give outcome (see Figure 3). The first look measure yielded the same results as the looking time measure: the infants looked first at the target bowl in the Give trials (*Median* = 0.80, W = 124.000, p = .003), but not in the Take trials (*Median* = 0.50, W = 74.000, p = .754; see Figure 4). Thus, although one measure is marginally significant, both indicate that infants in the request group anticipated the give outcome, whereas neither measure indicates that they anticipated the take outcome. To assess learning across the first block of trials, regression analyses were performed on both outcomes for both measures. Proportional looking time was not found to change significantly across trials for the Give outcome (β = 0.019, t(47) = 0.558, p = .58), or for the Take outcome (β = 0.101, t(40) = 2.416, p = .02). Similarly, the first look measure did not indicate learning across trials for the Give outcome (β = 0.148, $\chi 2(1)$ = 0.474, p = .491), or the Take outcome (β = 0.395, $\chi 2(1)$ = 3.719, p = .054). Thus the infants do not seem to be learning to anticipate the outcomes over the course of the trials, but rather have an a priori expectation regarding the request gesture that leads them to anticipate the give outcome.

Point

Infants in the point group looked at at least one of the AOI's during the anticipatory phase in 64.58% of the trials (*SD* = 17.70%). The measure of proportional looking time to the Gesturer's bowl revealed that they looked longer at the target bowl in the Give trials (*Median* = 0.79, W = 158.000, p = .002), but not in the Take trials (*Median* = 0.55, W = 117.500, p = .162), indicating that they only anticipated the give outcome (see Figure 3). The first look measure similarly revealed that infants had a tendency to make more target first looks only in the Give trials (*Median* = 0.75, W = 105.000, p = .009), but not in the Take trials (*Median* = 0.50, W = 86.000, p = .650; see Figure 4). Thus, even though one measure was only marginally significant, both indicate that infants in the point group anticipated the give outcome, but not the take outcome. To assess learning across the first block of trials, regression analyses were performed on both outcomes for both measures. Proportional looking time was not found to change significantly across trials for the Give outcome (β = 0.037, t(32) = 0.93, p = .359), or for the Take outcome (β = 0.042, t(43) = 1.014, p = .316). Similarly, the first look measure did not indicate learning across trials for the Give outcome (β = 0.208, $\chi 2(1)$ = 0.505, p = .477), or the Take outcome (β = 0.076, $\chi 2(1)$ = 0.146, p = .702). Again, this indicates that infants did not learn to anticipate the outcomes during the experiment.

Reach

Infants in the reach group looked at at least one of the AOI's during the anticipatory phase in 63.30% of the trials (*SD* =19.48%). The measure of proportional looking time to the Gesturer's bowl revealed that infants did not looked longer at the target bowl in the Give trials (*Median* = 0.62, W = 95.000, p = .162), or in the Take trials (*Median* = 0.50, W = 93.000, p = .194; see Figure 3). The first look measure similarly revealed that infants did not look first at target bowl in the Give trials (*Median* = 0.60, W = 77.500, p = .621), or in the Take trials (*Median* = 0.50, W = 43.500, p = .715; see Figure 4). Thus, both measures indicate that infants in the reach group failed to anticipate either outcome. To assess learning across the first block of trials, regression analyses were performed on both outcomes for both measures. Proportional looking time was not found to change significantly across trials for the Give outcome (β = 0.082, t(30) = 1.643, p = .111), or for the Take outcome (β = 0.019, t(36) = 0.54, p = .593). Similarly, the first look measure did not indicate learning across trials for the Give outcome ($\beta = 0.29$, $\chi2(1) = 1.407$, $p = .236$), or the Take outcome ($\beta = -0.162$, $\chi2(1) = 0.421$, $p = .516$). As for the previous gestures, no learning over trials was apparent for infants.

No gesture

Infants in the no gesture group looked at at least one of the AOI's during the anticipatory phase in 52.94% of the trials (*SD* = 22.40%). The measure of proportional looking time to the Gesturer's bowl revealed that infants did not look significantly longer at the target bowl in the Give trials (*Median* = 0.59, $W = 107.000$, $p = .144$), or in Take trials (*Median* = 0.71, $W = 106.000$, $p = .049$), indicating that they did not anticipate either outcome (see Figure 3). The first look measure revealed that infants did not look first at the target bowl at higher than chance levels in the Give trials (*Median* = 0.600, $W = 92.500$, $p = .192$), or in the Take trials (*Median* = 0.73, $W = 68.000$, $p = .019$; see Figure 4). Thus, both measures indicate that infants in the no gesture group failed to anticipate either outcome. To assess learning across the first block of trials, regression analyses were performed on both outcomes for both measures. Proportional looking time was not found to change significantly across trials for the Give outcome ($\beta = -0.041$, $t(29) = -1.086$, $p = .287$), or for the Take outcome ($\beta = 0.029$, $t(37) = 0.798$, $p = .430$). Similarly, the first look measure did not indicate learning across trials for the Give outcome ($\beta = -0.272$, $\chi2(1) = 0.941$, $p = .332$), or the Take outcome ($\beta = 0.086$, $\chi2(1) = 0.144$, $p = .704$). With no gesture present, infants also did not show a change in response over trials to indicate that they were learning from the observed outcomes.

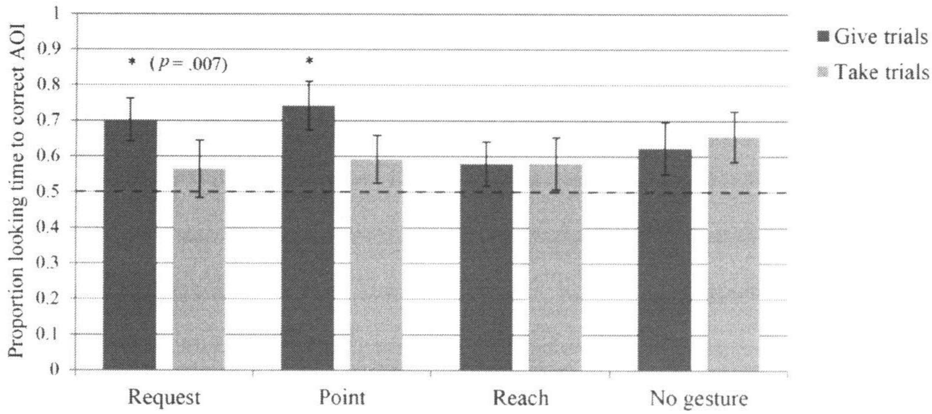

Figure 3. Average proportion of looking time to the target bowl AOI during the anticipatory phase for each of the four groups, separated into give trials – where the Addressee gave the object to the Gesturer – and take trials – where the Addressee took the object for himself. Dashed line indicates chance level and asterisks indicate a significant difference from chance ($p < .0063$, unless otherwise specified).

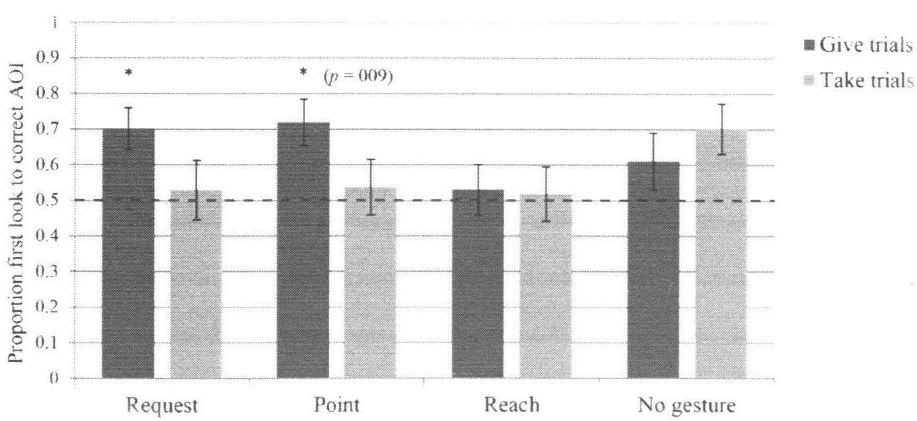

Figure 4. Average proportion of first looks to the target bowl AOI during the anticipatory phase of each of the four groups, separated into give trials – where the Addressee gave the object to the Gesturer – and take trials – where the Addressee took the object for himself. Dashed line indicates chance level and asterisks indicate a significant difference from chance ($p < .0063$, unless otherwise specified).

Discussion

We found that 14-month-old infants expected a third party addressee to produce a specific action in response to another person's communicative gestures. As neither the motive behind the person's gesture, nor the intention of the addressee could be inferred from their gaze direction or facial expressions, the infants' expectations toward the addressee were based only on the hand shape of the person's gesture. When the Gesturer pointed or directed a palm-up request gesture toward an object located between them, the infants expected the Addressee to give the object to her. Infants' expectations are reflected in their ability to anticipate that the Addressee will respond to the gestures by transferring the object to the Gesturer through a tube, and their inability to anticipate the opposite response of transferring the object to himself. Importantly, infants in a control group where no action was produced did not show evidence of anticipating the Addressee's response, ruling out the possibility that infants can anticipate the Addressee's actions without recognition of the gesture. In the reach group, where the Gesturer reached toward the object, infants also failed to anticipate the actions of the Addressee, possibly because they did not understand the reaching as a communicative action.

The current study complements previous findings that infants' understanding of communicative gestures is not restricted to dyadic settings, but extends to observed third-party interactions (Gräfenhain et al., 2009). The current findings further reveal that infants monitor and show understanding of nonverbal third-party interactions not only when the interaction has relevance to their own desired goal (as in Gräfenhain et al., 2009), but also when they are passive observers. Moreover, mirroring findings on infants' expectations about speech directed to third parties (Martin et al., 2013; Thorgrimsson et al., 2011), the study indicates that infants have specific expectations about the actions of third-party addressees of communicative gestures.

Expectations about gestures in third-party interactions

The findings indicate that infants' expectations about the interactions were in place before they first observed the outcome. Given that infants' anticipatory looking did not show a consistent change across the first block of trials, their expectations do not seem to have developed through repeated viewings of the outcome (in which case they would be based on infants' readiness to learn to associate certain actions with a specific outcome). It should also be noted, however, that although infants anticipated the give outcome in the give trials, their performance was at chance for the take trials, suggesting that their expectations are not robust enough to hold up to repeated evidence that the Addressee will take the object for himself.

From these findings, we cannot yet determine how infants develop expectations regarding these gestures. Their experience observing third-party interactions and their own dyadic interactions may both be contributing factors. The novelty of the interaction presented in the current study makes it unlikely, however, that their expectations are the product only of associative learning from observing the sequence of events in previous third-party interactions. The interaction we presented featured a novel and indirect means to transfer the object and neither participant presented any ostensive cues, yet infants anticipated the addressee's response. This suggests that infants' understanding of the function of these gestures is abstract enough for their expectations to be generalized across different contexts. This is in line with results from a recent interaction-based study showing that infants respond appropriately to the point and the palm-up request even in the absence of accompanying social-contextual cues (Thorgrimsson & Liszkowski, 2012).

It is interesting to speculate how infants' third-party expectations may relate to their own frequent experience with these gestures in communicative interactions. Despite the fact that infants' own pointing, and their interpretation of others' points shows that they are aware of the gesture's potential to communicate a variety of social motives, the infants were prepared to construe the point in the current third-party context as an object request, but not as an offer for the gesturer to take the object. However, although we know of no relevant empirical data, it is highly likely that infants point more frequently to request objects than to offer them, which may shape their interpretation of the function of points in others' interactions. Regarding the palm-up request, infants' anticipation in the request group suggests that infants expect its hand shape it to indicate an object request, even when they encounter the gesture outside a dyadic interactive context. As infants are very rarely observed to use this gesture to request (Hay & Murray, 1982; Bakeman & Adamson, 1986; Puccini et al., 2010), their expectations more likely stem from their experience as addressees in the give-and-take routine. Regarding the reach, infants' failure to anticipate the outcome in the reach group suggests that they did not expect the action to communicate a request to the Addressee, or to provoke the Addressee to help. However, the nature of the reach leave the results open to an alternative explanation. By coming very close to reaching the object but then retracting her arm, the Gesturer may have given the impression that she could have obtained the object but chose to leave it in its place. It is important to note, however, that this procedure was identical in the other groups, yet the infants were able to recognize the gestures as communicating a request for the object.

Together with recent research on infants' expectations about speech directed to a third party (Martin et al., 2013; Thorgrimsson et al., 2011), the current findings indicate that infants monitor third-party interactions and have specific expectations regarding addressees already at 14 months of age and possibly earlier. Attending not only to people's actions, but to the reactions they provoke from addressees, is likely to expand infants' opportunities for social-observational learning. For example, attending to the contribution of both interactants in social interactions should help infants learn from others' joint cooperative activities. The findings also raise the question of whether the arguably more demanding skill of learning novel object labels and novel arbitrary actions from third-party interactions may also emerge by this age. The fact that infants in the current study predicted the addressee's response to the gestures suggests that they may have taken his perspective into account when interpreting the actions. One theory of social-cognitive development posits that the ability to understand others' interactions from the perspective of the interactants is precisely what enables infants to learn from third-party interactions (Moore, 2007). Although there is some indication that learning novel actions is still an emerging skill at 18 months (Floor & Akhtar, 2006; Herold & Akhtar, 2008), a recent study shows word learning at 18 months even under difficult conditions (Gampe et al., 2013). Future research should examine younger infants' ability to learn words and actions from others' interactions.

The current findings demonstrate that infants have expectations about how third-party addressees respond to common gestures in an interaction. Further, they show that infants' recognition and use of gestural information is not limited to interactions that have direct relevance to infants' own goals. The ability to understand and predict others' social interactions is valuable as it offers infants the opportunity to learn through observing and overhearing the many interactions that surround them every day.

Chapter 5

Epilogue

Epilogue

Understanding and anticipating the actions of others has its beginnings in infancy. Infants' action understanding enables them to learn from others' actions and engage in preverbal communicative interactions and is likely to play a crucial role in the development of language and social understanding. Research on action understanding has revealed that infants have a basic understanding of actions as goal-directed behavior; they can predict the course and outcome of various common actions; they follow the direction of referential actions such as pointing and gazing, and they show pragmatic understanding of communicative actions. This research has established, through the use of a number of different experimental measures, that infants' action understanding encompasses a wide variety of actions. However, these studies tend to test various aspects of infants' action understanding in broadly similar settings, resulting in gaps in our knowledge about the level and breadth of action understanding in infancy. For one, the focus is mostly on how infants process or respond to actions in one-on-one dyadic settings, whereas observing actions in the context of third-party interactions is a common and potentially informative social experience in infancy, measures of which can reveal whether infants' action understanding extends to others' interactions. Further, while infants' own action responses to others' actions demonstrate pragmatic action understanding, the settings in which these interactions are tested are replete with additional social-pragmatic cues to the action's meaning, leaving open the question whether infants have understanding of the actions themselves. To get a fuller picture of action understanding in infancy I have examined in this thesis on the one hand whether infants monitor and understand communicative actions in third-party interactions, and on the other the level of their pragmatic understanding of actions in interactive settings.

Summary of findings

The study reported in Chapter 2 investigated infants' pragmatic understanding of object-directed actions and gestures. An Experimenter pointed, reached, or directed a palm-up request gesture toward an object within the infants' reach. Infants responded to all actions by handing the object over and did this more often than in a control group where no action was performed. As neither the social context nor the Experimenter's vocal or verbal behavior conveyed information about the motive behind his actions, the results indicate that 12-month-old infants can understand and respond to these common actions based on their hand shapes. Previous studies had shown that infants respond appropriately to reaching actions and request gestures that take place within a rich social context with several cues to the actor's motive. The results of the study in Chapter 2 show that while infants' pragmatic understanding of actions and gestures may be facilitated by such accompanying pragmatic cues, these are not critical. The study also examined whether the spatial relationship between an actor and a target object informs infants' action interpretation. In response to pointing, infants were found to offer the object more often when the object was out of reach for the Experimenter, revealing that infants make use of spatial information as a pragmatic cue to the motive behind ambiguous pointing gestures.

The study reported in Chapter 3 investigated infants' understanding of the role of speech in others' interactions. Two eye-tracking experiments measured infants' expectation of a response from a person addressed with speech. Infants at 12 and 24 months were quicker to shift their gaze to the addressee following speech from another person than following nonspeech sounds, and the 24-month-olds also looked at the addressee longer. In addition, the expectations of the older age group were restricted to a typical face-to-face communicative context. The results indicate that 12-month-old infants, who are only beginning to speak themselves, expect speech to provoke a response from a third party addressee. These expectations suggest that infants have developed an understanding of the structure of verbal interactions, possibly through observing others' communication, or their own preverbal communicative experiences. In addition, the results suggest that a face-to-face orientation between others signals a communicative situation for infants by at least 2 years of age.

The study reported in Chapter 4 investigated infants' understanding of actions and gestures in third party interactions. To this end, the study used eye tracking to measure infants' ability to anticipate how a third party addressee responds to another person's actions and gestures. Infants' anticipations indicate that they expected the point and the request gesture to result in the addressee giving the object to the other person, whereas infants failed to anticipate the outcome when the person reached for the object. In a control condition where no gesture was presented, infants did not anticipate the give outcome. As the three actions were identical in terms of trajectory and timing, and neither the motive behind the person's action, nor the intention of the addressee could be inferred from their gaze direction or facial expressions, the infants' differential expectations toward the addressee were based only on the hand shape used in the actions. Thus, the results from Chapter 4 indicate that infants understand the function of gestures in others' interactions, in that they have specific expectations toward third-party addressees of communicative gestures.

Pragmatic action understanding

The results from Chapter 2 show that infants respond appropriately to common action and gestures not only when these are embedded within shared social activities, or accompanied by other social-pragmatic cues, but also when actions are presented within a social setting stripped of such cues. While previous findings have been ambiguous with respect to whether infants understand the actions themselves, or if their responses depend on accompanying cues (e.g. Warneken & Tomasello, 2007; Hay & Murray, 1982), the results from Chapter 2 indicate that the hand shapes of the actions themselves are meaningful for infants. In the case of the palm-up request gesture, it is likely that the gesture acquired meaning for infants through the give-and-take routine that infants and caregivers often engage in from early on (Bruner, 1977). Note that the infants offered the object despite the gesture being presented outside the context of this routine, and in a setting where the object was not in possession of the infant, suggesting that infants understand the gesture form to signal an object request. Similarly, infants' responses to an object-directed reach suggest that they understand it as an attempt to obtain an object, and are motivated to help. A pragmatic understanding of actions based on the action forms themselves allows for more flexibility in infants' early social interactions, as infants do not have to rely on social-pragmatic cues.

The study also revealed that infants' interpretation of communicative pointing is informed by the spatial relation of the communicator and the referent. Older children have been previously demonstrated to similarly take a person's ease of access to objects into account when interpreting referentially ambiguous object requests but the current results indicate that sensitivity to spatial information in communication emerges as early as 12 months of age. Infants' early pragmatic understanding of referential actions is thus guided both by the shared social activity within which the action takes place (Behne et al., 2005; Liebal et al., 2009; Behne et al., 2012) and the actor's spatial distance from the referent.

Finally, complimenting other interaction-based studies (Behne, Carpeneter, & Tomasello, 2005; Behne, Liszkowski, Carpenter, & Tomasello, 2012; Buttelmann, Carpenter, & Tomasello, 2009; Knudsen, & Liszkowski, 2012; Liebal, Behne, Carpenter, & Tomasello, 2009; Liszkowski, Carpenter, & Tomasello, 2008), the study provides further evidence that infants' social understanding is manifested not only in their passive processing of observed actions, but in their communicative interactions.

Third-party interaction understanding

The two eye-tracking studies suggest that infants expect speech to provoke a response in others' interactions and that they have specific expectation about others' responses to communicative gestures. While previous findings have shown that infants monitor and follow third-party conversations, and can learn from third-party action demonstrations and object labeling, the two studies indicate that infants also pay specific attention to the addressees in such interactions and have expectations about their responses. Attending not only to people's actions, but to the reactions they provoke from third parties, is likely to expand infants' opportunities for social-observational learning. For example, attending to an addressee's object-directed actions in response to others' utterances may enable infants to identify the intended referent and learn novel object labels. Similarly, attending to the contribution of both interactants in social interactions should facilitate infants' ability to learn from observed joint cooperative activities. Future research should examine infants' abilities to learn from addressee responses in third-party interactions. Although they do not speak directly to this issue, the results from Chapter 3 raise the possibility that infants may understand speech as a communicative signal, even before having had substantial linguistic experience. A recent study (published after these findings) lends support to this notion, by showing that infants expect addressees of (nonsense) speech to hand over an object to the speaker that the speaker previously demonstrated a preference for (Martin et al., 2012). Recognizing that the speech form communicates information about the environment at an age where language development is just beginning suggests that this ability may contribute to early language acquisition.

Epilogue

In the study of third-party action understanding reported in Chapter 4, the infants showed not only a general expectation of a response from the addressee, but expected a specific outcome based on the action the addressee witnessed. Findings have previously shown that older infants can learn from instrumental actions and words directed to third party addressees (Floor & Akhtar, 2006; Herold & Akhtar, 2008; Gampe, Liebal, & Tomasello, 2012), and can make use of information conveyed to addressees through pointing and gazing (Gräfenhain et al., 2009), but the study reported in Chapter 4 is the first to show that infants have expectations about the behavior of third-party addressees of gestures. The fact that infants not only respond to communicative gestures in their own interactions, but expect other addressees to do the same, shows that infants' communicative understanding extends beyond their own interactions. Infants' action understanding may thus be guided by an abstract representation of communicative gestures already at the beginning of the second year

The infants had these expectations despite having experienced the interaction as passive observers for whose own goals or desires the interaction had no practical relevance. The finding indicates that infants monitor and understand third-party interactions not from an egocentric perspective but from the perspective of the interactants themselves, from at least 14 months of age and possibly earlier. Given the claim, partly supported by experimental evidence, that infants' ability to learn from third-party interactions depends on them understanding actions in terms of the agent's subjective experience (Moore, 2007; Herold & Akhtar, 2008), the findings raise the possibility that infants may be able to learn from third-party interactions at a considerably earlier age than suggested by previous studies. In light of these findings, it will be important to examine further the development of infants' understanding of and ability to learn from third-party interactions.

Epilogue

Conclusions

The results reported in this thesis have revealed novel aspects of infants' understanding of actions and gestures. The study in Chapter 2 indicates that infants' pragmatic action understanding is informed by the form of the actions themselves. Infants were found to understand and respond to common object-directed actions and gestures independently of the pragmatic cues by which they are often accompanied. The study also gave new insight into which aspects of a communicative setting infants can take into account to resolve ambiguity. Infants were found to be guided by spatial relations of actors and referents in their interpretation of ambiguous communicative actions. The second part of the thesis, comprised of two eye-tracking studies, suggests that infants have understanding of the role of speech and gestures in interactions between others. The first revealed evidence that infants at 12 and 24 months of age expect third-party addressees to respond to others' speech, over and above non-speech sounds, even in the absence of a previous contingent interaction. Thus, infants appear to understand the function of the speech signal at an age where they are only beginning to speak themselves - an interpretation given additional support by recent findings that infants understand speech to transfer information in third-party interactions. The study reported in Chapter 3 also showed that by at least 24 months, infants' expectations are restricted to settings where the interactants face each other, suggesting that infants associate verbal communication with a face-to-face orientation. The second study revealed evidence that infants have specific expectations about how third-party addressees respond to communicative gestures, even as passive observers. Infants anticipated that an addressee would respond to a communicative point and a conventional request gesture directed at an object by giving the object to the gesturer. Infants' understanding of communicative gestures thus extends to others' interactions by at least 14 months of age. The results further suggest an early ability to take the perspective of third party interactants, raising the possibility that infants can learn from third-party interactions from an earlier age than has previously been documented.

References

Augusti, E. M., Melinder, A., & Gredebäck, G. (2010). Look who's talking: pre-verbal infants' perception of face-to-face and back-to-back social interactions. *Frontiers in Psychology, 1*, 161.

Bakeman, R., & Adamson, L. B. (1986). Infants' conventionalized acts: Gestures and words with mothers and peers. *Infant Behavior and Development, 9*(2), 215-230.

Bakker, M., Kochukhova, O., & von Hofsten, C. (2011). Development of social perception: A conversation study of 6-, 12-and 36-month-old children. *Infant Behavior and Development, 34*(2), 363-370.

Becker, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca Ophthalmologica, 82*, 233-243.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492-499.

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology, 30*(3), 359-375.

Beier, J. S. & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development, 83*(2), 486-496

Beisert, M., Zmyj, N., Liepelt, R., Jung, F., Prinz, W., & Daum, M. M. (2012). Rethinking 'Rational Imitation 'in 14-Month-Old Infants: A Perceptual Distraction Approach. *PloS*

Bergelson, E., & Swingley, D. (2012). At 6 9 months, human infants know the meanings of

References

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42.

Carpendale, J. I., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27*(01), 79-96.

Carpendale, J., & Lewis, C. (2006). *How children develop social understanding*. Blackwell Publishing.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen-through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development, 21*(2), 315-330.

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve-and 18-month-olds copy actions in terms of goals. *Developmental Science, 8*(1), F13-F20.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

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.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705-717.

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica, 124*(1), 60-78.

Csibra, G., & Gergely, G. (2013). Teleological understanding of actions. In Banaji, M.R., & Gelman, S.A. (Eds.), *Navigating the Social World: What Infants, Children, and Other*

Species Can Teach Us, Oxford University Press, 2013.

Csibra, G., & Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*, 1-11.

Csibra, G., Bıró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*(1), 111-133. *Developmental psychology, 45*(4), 1190.

Csibra, G., Bıró, S., Koós. O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*(1). 111-133.

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, 20, 569-572.

Deak, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12-and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental psychology, 36*(4), 511.

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12-and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology, 36*(4), 511.

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology, 47*(6), 1499-1503.

D'Entremont, Barbara, S. M. J. Hains, and D. W. Muir. "A demonstration of gaze following in 3- to 6-month-olds." Infant Behavior and Development 20.4 (1997): 569-572.

Doi, H., Tagawa, M., Shinohara, K. (2010). Gaze Direction Modulates the Disengagement of Attention From Facial Expression in 10-Month-Olds. *Emotion, 10(2)*, 278-282

References

Elsner, B., Hauf, P., & Aschersleben, G. (2007). Imitating step by step: A detailed analysis of 9- to 15-month-olds' reproduction of a three-step action sequence. *Infant Behavior and Development, 30*(2), 325-335.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*, 878-879.

Fawcett, C., & Liszkowski, U. (2012). Observation and initiation of joint action in infants. *Child Development, 83*(2), 434-441.

Feldstein, S., Jaffe, J., Beebe, B., Crown, C. L., Jasnow, M., Fox, H., & Gordon, S. (1993). Coordinated interpersonal timing in adult infant vocal interactions - a cross site replication. *Infant Behavior & Development, 16*(4), 455-470.

Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Sekerina, I.A., Fernandez, E.M., & Clahsen, H. (Eds.), *Developmental Psycholinguistics: On-line Methods in Children's Language Processing* (pp. 97–135). Amsterdam: John Benjamins.

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy, 9*(3), 327-339.

Gampe, A., Liebal, K., & Tomasello, M. (2012). Eighteen-month-olds learn novel words through overhearing. *First Language, 32*(3), 385-397.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287-292.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: rational imitation in preverbal infants. *Nature, 415*(6873), 755-755.

References

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. Developmental science, 10(1), 139-146.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165-193.

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science, 20*, 347–53.

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). One-year-olds' understanding of nonverbal gestures directed to a third person. *Cognitive Development, 24*(1), 23-33.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition, 114*(2), 197-206.

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & von Hofsten, C. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology, 45*(4), 1190.

Hamlin, J. K., Hallinan, E. V., & Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Developmental Science, 11*(4), 487-494.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science, 28*(1), 105-115.

Hay, D. F., & Murray, P. (1982). Giving and requesting: Social facilitation of infants' offers to adults. *Infant Behavior and Development, 5*(2), 301-310.

Herold, K. H., & Akhtar, N. (2008). Imitative learning from a third-party interaction: Relations with self-recognition and perspective taking. *Journal of Experimental Child Psychology, 101*(2), 114-123.

References

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: a study of infants' visual anticipations during action observation. *Developmental Psychology, 46*(2), 446-454.

Hunnius, S., de Wit, T. C. J., Vrins, S., & von Hofsten, C. (2011). Facing threat: Infants' andadults' visual scanning of faces with neutral, happy, sad, angry, and fearful emotional expressions. *Cognition and Emotion, 25,* 193-205.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological science, 16*(5), 367-371.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development, 66*(2), i-viii, 1-132.

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1,* 233-238.

Knudsen, B., & Liszkowski, U. (2012). Eighteen-and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science, 15*(1), 113-122.

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science, 12*(2), 264-271.

Liebal, K., Carpenter, M., & Tomasello, M. (2010). Infants' use of shared experience in declarative pointing. *Infancy, 15(5),* 545-556.

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), *Input and Interaction in Language Acquisition* (pp. 56 73). Cambridge, U.K.: Cambridge University Press.

Liszkowski, U. (2013). Using Theory of Mind. *Child Development Perspectives.*

References

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month olds communicate helpfully and appropriately for knowledgeable and ignorant partners. Cognition, 108, 732–739.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*(2), 173-187.

Martin, A., Onishi, K. H., & Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12months. *Cognition*, *123*(1), 50-60.

Matheson, H., Moore, C., & Akhtar, N. (2012). The development of social learning in interactive and observational contexts. *Journal of Experimental Child Psychology*.

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental science*, *10*(1), 126-134.

Messinger, D. S., & Fogel, A. (1998). Give and take: The development of conventional infant gestures. *Merrill-Palmer Quarterly (1982-)*, 566-590.*one*. *7*(3), e32563.

Moll, H. & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*, F1–F9.

Moore, C. (2007). Understanding self and others in the second year. In C. Brownell & C.B. Kopp (Eds.), *Socioemotional development in the toddler years: Transitions and Transformations*, (pp. 43-65). New York: Guilford Press.

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), 718-735.

References

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interaction between two month-olds and their mothers. In T. Field & N. Fox (Eds), *Social Perception in Infants* (pp. 101-125). Norwood, NJ: Ablex.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science, 2*(2), 164-173.

Nielsen, M. (2009). 12-month-olds produce others' intended but unfulfilled acts. *Infancy, 14*(3), 377-389.

Nurmsoo, E., Einav, S., Hood, B.M. (2012). Best friends: children use mutual gaze to identify friendships in others. *Developmental Science, 15*(3), 417-425.

Ochs, E., & Schieffelin, B. (1984). Language Acquisition and Socialization: Three Developmental Stories and Their Implications. In *Culture Theory: Essays on Mind, Self, and Emotion*. Richard Schweder and Robert A. Levine, eds. Pp. 276-320. Cambridge: Cambridge University Press.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition, 98*(2), 137-155.

Phillips, A.T., Wellman, H.M., & Spelke, E.S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition, 85*, 53-78.

Puccini, D., Hassemer, M., Salomo, D., & Liszkowski, U. (2010). The type of shared activity shapes caregiver and infant communication. *Gesture, 10*(2-3), 2-3.

Reddy, V. (2010). Engaging minds in the first year: The developing awareness of attention and intention. *Wiley-Blackwell Handbook of Infant Development, The, Volume 1, Second Edition*, 365-393.

References

Repacholi, B. M., & Meltzoff, A. N. (2007). Emotional eavesdropping: infants selectively respond to indirect emotional signals. *Child development, 78*(2), 503-521.

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

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*(9), 668-671.

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing, and reaching as cues to goal-directed action. *Journal of Cognition and Development, 5*(3), 289-316.

Sodian, B., Schoeppner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents?. *Infant Behavior and Development, 27*(1), 31-41.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*(1), B1–B11.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental psychology, 45*(6), 1794-1798.

Southgate, V., Johnson, M. H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychological science, 21*(3), 355-359.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. Psychological Science, *18*, 587–592.

Thayer, S., & Schiff, W. (1974). Observer judgment of social interaction: eye contact and relationship inferences. Journal of Personality and Social Psychology, *30*, 110–114

Thoermer, C., & Sodian, B. (2001). Preverbal infants' understanding of referential gestures. *FirstLanguage, 21*(63), 245-264.

References

Thorgrimsson, G. Fawcett, C., & Liszkowski, U (2011). *12-month-olds expect speech to provoke a response in others' interactions*. Poster presented at the biennial meeting of the Society for Research in Child Development, Montreal, Canada.

Thorgrimsson, G., & Liszkowski, U (2012). *1-year-olds' responses to object-directed actions and gestures*. Poster presented at the annual meeting of the Jean Piaget Society, Toronto, Canada.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172-175.

Toda, S., & Fogel, A. (1993). Infant Response to the Still-Face Situation at 3 and 6 Months. *Developmental Psychology*, *29*(3), 532-538.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child development*, *78*(3), 705-722.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675-690.

von Hofsten, C., & Lindhagen, K. (1979). Observations on the development of reaching for moving objects. *Journal of Experimental Child Psychology*, *28*(1), 158-173.

von Hofsten, C., Dahlström, E., & Fredriksson, Y. (2005). 12-month old infants' perception of attention direction in static video images. Infancy, *8*, 217–231.

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in autism spectrum disorders*, *3*(2), 556–569.

References

Vouloumanos, A., Druhen, M. J., Hauser, M. D., & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences of the United States of America, 106(44)*, 18867–18872.

Wang, S., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: evidence from violation-of- expectation tasks with test trials only. *Cognition*, 93, 167–198.

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271-294.

Warneken, F., & Tomasello, M. (2012). Parental presence and encouragement do not influencehelping in young children. *Infancy*.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1-34.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful 3behaviors. *Infant Behavior and Development*, *22*(2), 145-160.

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*(1), 1061-1084.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, *11*(1), 73-77.

Yoon, J. M., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. Proceedings of the National Academy of Sciences, 105(36), 13690-13695.

Nederlandse samenvatting

De studies in dit proefschrift onderzoeken hoe baby's communicatieve handelingen en gebaren begrijpen, zowel als deelnemers in interacties als waarnemers van de interacties van andere mensen.

Onderzoek heeft aangetoond dat jonge kinderen reageren op gebaren en andere handelingen wanneer die in een sociale context voorkomen, zoals spelen of andere activiteiten. Omdat de sociale context vaak veel informatie over de mogelijke betekenis van gebaren bevat, is het niet duidelijk of jonge kinderen de handvormen van gebaren begrijpen of dat ze slechts reageren op de context waarin deze gebaren voorkomen. In hoofdstuk 2 wordt onderzocht hoe baby's van twaalf maanden reageren op gebaren en handelingen wanneer die zonder een typische sociale context voorkomen. In dit onderzoek bewoog een proefleider zijn arm in de richting van objecten die voor kinderen op tafel lagen. Deze beweging kon of een handeling voorstelden (naar het object reiken) of een gebaar (waarbij onderscheid werd gemaakt tussen wijzen naar het object of een 'verzoek' om het object – met de arm uitgestrekt en handpalm naar boven gericht). De primaire vraag was of de kinderen de gebaren zou interpreteren als verzoeken en de objecten zou geven aan de proefleider. De reacties van de kinderen werden vergeleken met een controlegroep die geen gebaren kregen voorgeschoteld. Het secondaire doel van de studie was om te onderzoeken of de interpretatie van gebaren in de vroege kindertijd werd beïnvloed door de ruimtelijke context. De objecten lagen ofwel binnen ofwel buiten bereik van de proefleider. De resultaten lieten zien dat kinderen de objecten vaker aan de proefleider gaven in de experimentele groepen dan in de controlegroep, wat erop wijst dat ze de gebaren als verzoeken geïnterpreteerd hebben. Kennelijk kunnen jonge kinderen gebaren interpreteren buiten een typische sociale context en kunnen ze de handvormen van deze gebaren begrijpen in hun sociale interacties. We ontdekten ook dat de interpretatie van de kinderen werd beïnvloed door het ruimtelijke context: Ze gaven het object vaker aan de proefleider als het object buiten bereik lag en hij ernaar wees dan wanneer het object binnen bereik van de proefleider lag en hij in die richting wees.. Jonge kinderen kunnen dus gebruik maken voor de pragmatische informatie in ruimtelijke relaties om ambigue communicatieve signalen te interpreteren. Dit vermogen kan belangrijk zijn voor sociaal- en taalontwikkeling in de vroege kindertijd.

Het tweede deel van het proefschrift gaat over wat kinderen begrijpen als waarnemers van de communicatieve interacties van andere mensen, zogenaamde "derde-partij interacties". Derde-partij interacties kunnen een belangrijke bron van sociale informatie voor jonge kinderen zijn, van waaruit ze bijvoorbeeld over woorden, gezamenlijke handelingen, sociale normen en conventies kunnen leren. Wat kinderen verwachten over derde-partij interacties kan ook meer duidelijkheid verschaffen over hoe abstract hun begrip van communicatie is. In de eye-tracking studie beschreven in hoofdstuk 3 wordt onderzocht wat kinderen van 12 en 24 maanden verwachten over spraak in derde-partij interacties. De kinderen keken naar filmpjes waarin twee mensen tegenover elkaar zaten. De ene persoon uitte verschillende 'zinnen' en de andere persoon (de geadresseerde) zei niks terug/reageerde daar niet op. Vervolgens werden de oogbewegingen van de kinderen gemeten om te zien hoe snel zij naar de geadresseerde persoon keken, en hoe lang zij naar die persoon keken. In een controleconditie maakte de acteur natuurlijke menselijke geluiden, zoals een hoest of een geeuw. Daarnaast zaten de twee acteurs ofwel tegenover elkaar ofwel met de ruggen naar elkaar toe gekeerd, om te zien of de wederzijdse oriëntatie van de acteurs de verwachtingen van de kinderen zou beïnvloeden. Resultaten laten zien dat kinderen van 12 en 24 maanden sneller naar de geadresseerde keken na het uitspreken van zinnen dan nade geluiden, wat erop wijst datde kinderen een reactie van de geaddresseerde verwachtten. Kinderen van 24 maanden deed dat bovendien alleen wanner het paar tegenover elkaar zat, waarschijnlijk omdat ze begrijpen dat deze oriëntatie van lichaam en blik typerend is voor communicatieve situaties. De resultaten laten kortom zien dat jonge kinderen verwachtingen hebben over spraak in derde-partij interacties. Deze verwachtingen zijn zelfs al aanwezig op het moment dat ze net beginnen met spreken en wijst er op dat ze de functie van spraak begrijpen.

In de eye-tracking studie beschreven in hoofdstuk 4 wordt naderonderzocht hoe kinderen handelingen en gebaren in derde-partij interacties begrijpen. Kinderen van 14 maanden zagen filmpjes met twee acteurs, waar een van de acteurs (de *Agent*) en handeling danwel gebaar richting een object in het beeld maakte. De handeling bestond uit reiken naar het object, en een gebaar bestond of uit wijzen, of verzoeken (wederom hand uitgestoken met palm omhoog). Het object was tussen de andere twee acteurs gelegen. In de helft van de trials reageerde de geadresseerde door het object aan de Agent geven ("give trials"); in de andere helft pakte de geadresseerde het object zelf ("take trials"). De gedachtegang was dat de handeling van reiken en het gebaar van verzoeken een verlangen weergeven om het object te krijgen, terwijl wijzen ambigu is ten aanzien is van de intentie van de Agent. Het filmpje voor de controlegroep was hetzelfde, alleen waren hier geen gebaren of handeling te zien. Oogbewegingen werden gemeten om te onderzoeken of kinderen kunnen anticiperen op het resultaat van de interactie. De resultaten laten zien dat voor wijzen en het verzoek gebaar de kinderen verwachtten dat de geadresseerde het object aan de Agent zou geven. Kinderen die naar de handeling van het reiken keken en kinderen in de controle groep hadden geen verwachtingen ten opzicht van de geadresseerde. Dit suggereert dt kinderen van 14 maanden kunnen veelvoorkomende gebaren herkennen, niet alleen in hun eigen interacties, maar ook in de interacties van anderen. Dat de kinderen kunnen anticiperen op de respons van het 'verzoek' gebaar wijst erop dat ze de functie van dit gebaar begrijpen. De resultaten suggereren verder dat kinderen van deze leeftijd ook zouden kunnen leren van derde-partij interacties. De twee eye-tracking studies in hoofdstukken 3 en 4 laten zien dat jonge kinderen interesse tonen in het verloop en de uitkomst van derde-partij interacties, ondanks dat de interacties niet direct relevant voor hen zelf zijn. Het besteden van aandacht aan deelnemers in derde-partij interacties kan belangrijk zijn voor sociaal leren in de jonge kindertijd.

Curriculum Vitae

Guðmundur B. Thorgrímsson was born in 1982 in Reykjavík, Iceland. He studied psychology at the University of Iceland, where he obtained his bachelor's degree in 2006. From 2007 - 2009, Gudmundur was enrolled in a Research Master's program in psychology at the University of Amsterdam, where he graduated *cum laude* in 2009. He began his PhD research in the Communication Before Language Group at the Max Planck Institute for Psycholinguistics in 2010, under the supervision of Ulf Liszkowski. His research is described in this thesis.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology. English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences *Andrea Weber*
17. Moving eyes and naming objects *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing Computational and cross-linguistic experimental studies *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production. Effects of syllable preparation and syllable frequency *Joana Cholin*
27. Producing complex spoken numerals for time and space *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction *Rachèl J. J. K. Kemps*
29. At the same time... The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach *Marlies Wassenaar*

32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words. An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*

66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*

67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*

68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*

69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*

70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*

71. Gaze behavior in face-to-face interaction. *Federico Rossano*

72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Conny de Vos*

73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*

74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*

75. Deictic versus representational gestures in infancy. *Daniel Puccini*

76. Territories of knowledge in Japanese conversation. *Kaoru Hayano.*

77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder.*

78. Contributions of executive control to individual differences in word productions. *Zeshu Shao*

79. Hearing and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*

80. High pitches and thick voices. The role of language in space-pitch association. *Sarah Dolscheid.*

81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers.*

82. Mental representations and processing of reduced words in casual speech. *Iris Hanique*

83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann.*

84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85 Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*

87. The acquisition of morphophonological alternations across languages. *Helen Buckler*

88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*

89 Rediscovering a forgotten language. *Jiyoun Choi*

90 The road to native listening. Language-general perception, language-specific input. *Sho Tsuji*

91. Infants' understanding of communication as participants and observers. *Gudmundur B. Thorgrimsson*