The use of deictic versus representational gestures in infancy

Daniel Puccini

Cover picture: Little Jackie Paper by Kendall Rattner

Printed by: Ipskamp Drukkers, Nijmegen, The Netherlands

ISBN: 978-90-76203-46-1

The use of deictic versus representational gestures in infancy

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 vrijdag 11 Januari, 2013
om 11.00 uur precies

door

Daniel Puccini
geboren op 4 maart 1982
Yrisarri, New Mexico, USA

Promotor:	Prof. dr. P. Fikkert
Copromotor:	Dr. U. Liszkowski
Manuscriptcommissie:	Prof. dr. P. Muysken
	Prof. dr. A. Ozyurek
	Dr. R. Pfau (Universiteit van Amsterdam)

Contents

Chapter 1

Introduction

"I cannot doubt that language owes its origin to the imitation and modification, aided by signs and gestures, of various natural sounds, the voices of other animals, and man's own instinctive cries."

— *Charles Darwin, 1871. The Descent of Man, and Selection in Relation to Sex*

1.1 General introduction

Babies use their hands to communicate well before they use language. Research on infants' preverbal communication has focused largely on deictic gestures, that is, gestures which indicate or point out entities in the immediate environment. Much less is known about the role of representational gestures in infancy, which, similar to spoken words, stand in for, or represent the objects or actions to which they refer. The studies comprising this thesis investigated both deictic and representational gestures in infancy, with two studies devoted to the use and emergence of deictic gestures, and two focusing on the comprehension and production of representational gestures. Based on the notion that language is grounded in collaborative joint action, I will argue that the emergence of referential communication arises out of experience in acting jointly with other individuals, and that deictic, as opposed to representational gestures provide the foundation upon which language is built.

1.2 Language as joint action

Language is more than a mere system of codes used to express thought. Language is grounded in action and used to directly influence the real world (Austin, 1975; Searle, 1969). Austin, for example, pointed out that, in communicating, the grammatical form of an utterance need not match the communicative intentions of a speaker in order to be understood. There is no one-to-one code for decoding and interpreting what others say. Instead, language is built on humans' unique ability to interpret others' communicative intentions (Tomasello, 2008). Embracing such a view, Herbert Clark (1996, 2006) argued that a shared joint structure encompasses both language and cooperative joint actions and that joint action and language are really one and the same thing. In both, individuals must collaborate cooperatively in order to accomplish particular actions and meet mutual goals. In the left panel of Figure 1.1, in order to complete the puzzle, the four individuals must first share a common goal of piecing together the puzzle and, in addition, all four individuals must coordinate their actions in order to carry out their goal. The right panel of Figure 1.1 depicts a similar situation where the same requisites apply: In order to reach an agreeable conclusion, the individuals must first share a common goal, and, in addition, they must carefully coordinate their communication in order to arrive at their goal. The common element characterizing both language and interaction is jointness. Neither the people in Panel 1 of Figure 1.1, nor those in Panel 2 would be able to meet their goals alone: They must act jointly by taking turns and coordinating their own attention and mental and physical actions with all others involved. Ultimately, the degree of success in both interactions will be determined by the degree to which the individuals collaboratively work together in a joint fashion.

Figure 1.1 Joint action (left panel)[1] and language (right panel) share a common structure, built on shared intentionality and coordinated joint attention

1.3 Gesture as joint action

Less is known about the emergence of communication as a form of joint action in infancy. The past few decades of research, however, have shown that many defining characteristics of human communication emerge first in other domains before language is even in place. Bruner (1975), for example, argued that infants learn fundamental pragmatic components of language through acting jointly with others. Speech acts, he argued, emerge

[1] http://appleboxs.com/wordpress/tag/applebox/

first in ontogeny from action formats. Based on this hypothesis, infants learn about the cooperative nature of human interaction though interacting with others. This basic understanding provides infants with a basis upon which language can later be mapped.

Before 12 months, infants' interactions are primarily dyadic in nature, involving only a caregiver and an infant (Bateson, 1975; Stern, 1977; Trevarthen, 1979). Around 12 months, however, infants begin to engage in triadic interactions. Triadic interactions differ from earlier dyadic interactions in that they entail shared joint attention and manual action on external objects (Bakeman & Adamson, 1984). It is possible that the ability to coordinate attention between an individual and an object of shared interest sets the stage for the emergence of referential communication.

Interestingly, infants also first communicate referentially towards the end of their first year using deictic gestures such as showing and pointing (Leung & Rheingold, 1981; Murphy, 1978; Carpenter, Nagell, & Tomasello, 1998). Deictic gestures highlight or bring attention to particular entities in the immediate environment (originally called 'proto-declaratives;' Bates, Camaioni & Volterra, 1975). Analogous to adults' use of speech acts, infants point for various cooperative motives (Bates et al. 1975). In a longitudinal study of three infants in their first year, Bates et al. showed that infants point, not only to obtain objects, but also because they want to share interest with others. For example, while a 12-month-old infant might point to a bottle of milk, indicating her desire for the milk, she might also point to the moon, not because she wants to have the moon, but because she wants others to also see the moon, or, put differently, because she wants to share attention and interest with others around her.

The view that infants' initially point referentially to influence others' mental and attentional states, however, has been challenged based on the idea that infants point egocentrically to fulfill their own personal motives. Some, for example, have proposed that they initially point because they appreciate the contingent responses of their recipients, because they enjoy being the center of attention, or even to orient their own attention and behavior (Bates et al., 1975, Carpendale & Carpendale, 2010; Moore & D'Entremont, 2001). Research by Liszkowski and colleagues, however, suggests against these lean interpretations of infant pointing, and has confirmed that infants as young as 12 months do point referentially to share attention and influence others' mental states. When pointing, for example, infants consider, not only whether an adult has already seen a particular referent, but also the adult's attitude towards the object, for example, if she had previously emoted positively or negatively about it (Liszkowski, Carpenter, & Tomasello, 2007). Infants also take the recipient's focus of attention into account before they point. They point less, for example, when a recipient cannot see their pointing gesture, as compared with when the recipient can see the gesture (Liszkowski, Albrecht, Carpenter & Tomasello, 2008). Infants also point to locations where interesting objects had previously been, suggesting that their pointing gestures are indeed referential (Liszkowski, Schäfer, Carpenter & Tomasello, 2009). Taken together, the studies by Liskowski and colleagues show that by 12 months, infants point referentially in order to direct and change others' mental states. To date, however, little is known about how pointing emerges in infancy and how the social pragmatic skills which are manifested in infants' pointing gestures emerge in ontogeny.

Based on the findings that infants' early pointing gestures entail the same cognitive underpinnings as language and joint action, one interesting possibility is that infants become sensitive to others' mental states through interacting with others. Based on this hypothesis, it is possible that the very emergence of referential communication is related to how caregivers and infants interact with one another. The relation between infants' early use of pointing and their skills of joint engagement has not been investigated in detail. In Chapter 2, I explore the possibility that infants' ability to communicate referentially arises from experience interacting with others in joint activities.

1.4 From gesture to language

As discussed, deictic gestures are the first form of referential communication to emerge in ontogeny. These gestures entail complex cognitive underpinnings that enable infants to communicate about objects in their immediate environment. Human language, however, is not comprised of deixis alone. Human language is made possible via a conventionalized system for mental representation, and naturally, infants, who communicate initially with deictic gestures, must also learn to communicate representationally, for example, with words. A crucial difference between deictic gestures and words is that deictic gestures are generally used to communicate about entities in the immediate environment, while words are often used to communicate about displaced referents: that is, entities that exist outside the physical and/or temporal proximity of their users. In order to refer with a word, an individual must have a mental representation of the entity itself, as well as of the symbol connecting the word and its referent (DeLoache, 2004). Despite the major shift in how infants communicate over the course of their second year, surprisingly little is known about how children learn to communicate representationally and how this process unfolds developmentally.

First, words and their referents seldom appear in isolation. Human speech consists of a continuous stream of sounds, not distinct individual words. Infants must therefore extract relevant and distinct phonological strings and map them onto their intended referents. To complicate matters even more, entities never exist in isolation. Accordingly, infants must also identify the intended referent of any given particular word by selecting one from an array of objects in sight. Finally, even after infants are able to identify the referent, how do they know whether the word refers to the whole object, a part of the object, to the objects' shape, or even to its color (Tomasello, 2008)? Research in the past few decades has shown that caregivers play a critical role in this process by scaffolding children's language acquisition and word learning.

One way that caregivers scaffold language acquisition is through the use of child-directed speech: a modified speech register catered specifically to infants' developing language capacity. For example, when talking to infants, caregivers tend to use simpler and shorter utterances, a more variable pitch range, and more word repetitions (Fernald & Kuhl, 1987; Gleitman, Newport, & Gleitman, 1984; Newport, 1975). A more recent study also showed that child-directed speech facilitates word segmentation (Thiessen, Hill, & Saffran, 2005). The modified speech register used by caregivers scaffolds infants' referential understanding by highlighting relevant parts of their utterances, which, in turn, makes it easier for infants to interpret referential intention and thus to formulate representations and establish connections and between words and their intended referents. Taken together, these studies show that caregivers play an active role in catering their language input to the needs of their infants.

Another way caregivers scaffold infants' referential understanding is by integrating deictic gestures into their own communication (Gogate, Bahrick, & Watson, 2000; Gogate, Bolzani & Betancourt, 2006). Deictic gestures are helpful in word learning since they help establish joint attention (Baldwin, 1991; Tomasello & Farrar, 1986). Yet, even with the help of infant-directed speech and infant-directed gesture, how do infants learn that spoken words refer to the same objects to which mothers are gesturing? One possibility is that caregivers systematically integrate speech and gesture into particular activities as a form of multimodal motherese. In Chapter 3, I explore the extent to which infants' communicative input is structured by synchronous deictic gesture – language combinations. Through the systematic integration of deictic gestures and language in particular types of joint activities, I argue that infants are provided with a multimodal scaffold which enables them to interpret communicative intention, thus enabling word learning.

Research also shows a relation between infants' own production of pointing gestures and their developing language capacity. There is a correlation, for example, between the onset of pointing and the age at which infants produce their first words (Brooks & Meltzoff, 2008; Carpenter et al., 1998; Harris, Barlow-Brown, & Chasin, 1995). Further, there is a relation between the age at which infants combine pointing gestures with words in a supplementary fashion (words that add semantic content to points; for example pointing to a dog and saying 'black') and the age at which they begin combining words (Iverson & Goldin-Meadow, 2005). These studies show that, in addition to sharing the same cognitive basis as language, deictic gestures are intimately linked with language development.

One interesting possibility, which will comprise the remainder of this chapter and a large part of this thesis, has to do with the possibility that representational gestures also relate to infants' developing language capacity. It is possible, for example, that representational gestures serve as a developmental medium between deictic communication and spoken language. Similar to spoken words, representational gestures stand in for, or replace their referents (Bates 1979; Capirci, Iverson, Pizzuto, & Volterra, 1996; McNeill, 1992). One important distinction has to do with whether a representational gesture is created online or whether it is rote learned through and used as a linguistic convention. Creative representational gestures (also called 'codified' gestures) are not learned in familiar routines, nor are they retrieved from memory (Caldognetto & Poggi, 1995). Creative representational gestures are produced spontaneously in order to convey to another individual enough information so that he or she can accurately identify the intended referent. By virtue of their communicative function, creative representational gestures are iconic, meaning they bear some perceptual resemblance to their intended referents. In using such a gesture, one uses his or her own body to depict the particular action scheme or object representation that he or she is trying to convey. If I was to visit a cafe in India, for example, and I, nor the waiter spoke each other's language, then I might use creative representational gestures to communicate my order to the waiter. I might, for example form a cup shape with my hand and raise it to my mouth, indicating that I would like something to drink.

Through increased usage, however, creative gestures have a tendency to become conventionalized. This occurs as individuals of a particular community begin to recognize a gesture as conveying a particular meaning. A conventionalized gesture is thus culturally learned and stored in memory. As it shifts into a conventionalized gesture, users begin to abbreviate it because the iconicity is no longer needed to recognize its intended meaning (Armstrong & Wilcox, 2007). Accordingly, conventionalized gestures often bear a more arbitrary relationship with their referents than creative gestures. Conventionalized gestures are analogous to many spoken words since both are learned as conventions within a community and bear little perceptual resemblance to the entities to which they refer. In the Netherlands, for example, someone can indicate that something is delicious by waving one's open palm side-to-side by his or her ear. A gesture such as this would be incomprehensible to anyone who is not familiar with Dutch culture. While speakers might have originally established the gesture as a creative, iconic gesture, it likely gradually became conventionalized and shifted towards being arbitrary. Nowadays, most speakers of Dutch no longer recognize any iconicity in the gesture as it has become a convention which does not require iconicity in order to interpret its intended meaning.

Little is known about the role of representational gestures in infants' developing language capacity. Acredolo and Goodwyn (1985) first reported on a case study that documented the use of representational gestures from 12.5 to 17.5 months. In that study, the authors reported that the infant had acquired a small repertoire of representational gestures by 17.5 months. These gestures, the authors reported, were originally produced spontaneously outside of familiar routines, suggesting that the infant had produced them creatively. In a related follow-up study (which used both parental interviews and longitudinal diary entries), Acrodolo and Goodwyn (1988) confirmed their previous case study with a larger sample of infants. The authors reported that infants typically begin to spontaneously produce representational gestures around 15 months. Reported examples included holding one's arms out to represent an airplane or sniffing to represent a flower (Acredolo & Goodwyn, 1988). However, the frequency with which these gestures are used is very low in infants' second year, especially when compared to their use of deictic gestures and spoken words (Capirci, Iverson, Pizzuto, & Volterra, 1996). And, while infants' deictic gestures are very often accompanied by words or vocalizations (e.g., Cochet & Vauclair, 2010; Liszkowski & Tomasello, 2011), their early representational gestures are generally not (Pizzutto & Capobianco, 2005). Such findings bring into question whether young infants actually use representational gestures creatively and communicatively, or whether they might instead be ritualized actions which are learned from social routines. Likewise, it remains unclear whether infants are capable of using creative representational gestures. Without experimental evidence, we simply do not know. The study in Chapter 5 addresses this question, where I investigate whether Dutch infants at 18 and 24 months will spontaneously and creatively produce representational gestures in a situation where deictic gestures do not suffice.

Despite the lack of empirical research on the role of representational gestures in language acquisition, some research suggests that infants map representational gestures as object labels as easily as they do spoken words (Namy & Waxman, 1998, 2000). Research from the acquisition of signed languages provides further support for this claim as several studies have shown that infants exposed to signed languages from birth acquire language at the same rate as infants learning spoken languages (Capirci, Montanari, & Volterra, 1998; Folven & Bonvillian 1993; Meier & Newport, 1990; Petitto, 1987; Schick, Marschark, & Spencer, 2006). These studies suggest that infants initially have no bias towards either modality and will initially accept both spoken words and manual gestures as object labels. One interesting possibility is that the combinatorial use of representational gestures and spoken words facilitates the mapping of a word and its referent. In a training study of infants from 11 to 36 months, Goodwyn, Acredolo, and Brown (2000) compared the language development of a group of typically developing children with a group of children who had been explicitly trained on multimodal labels including both spoken words and representational gestures. The group of infants who had been trained on representational gestures attained higher language scores than the non-intervention control group across a number of language measures, suggesting that the combinatory use of representational gestures and spoken words had facilitated their language development. To date, however, little research has been able to corroborate these findings, and a recent meta-analysis suggests that the reported benefits of representational gestures on language acquisition might not be well founded (Johnston, Durieux-Smith, & Bloom, 2005).

Despite the lack of empirical support for the claim that representational gestures facilitate language development, the general public has, in the past decade, shown a great interest in such a possibility, as evidenced in the creation and recent surge of commercially available books, DVD's, courses, and other baby media products which encourage parents to teach their infants representational gestures before they speak. The eye tracking study reported in Chapter 4 aims to provide insights into how infants interpret utterances containing arbitrary representational gestures. In particular, I explore whcther infants are better at fast mapping spoken words or arbitrary representational gestures when trained on multimodal labels. Together, Chapters 4 and 5 provide insight into how infants interpret representational gestures and how they might relate to infants' developing language capacity.

1.5 Outline of thesis

The studies comprising this thesis investigate gestural communication in infancy. The experiment reported in Chapter 2 explores the possibility that referential communication arises out of experience in joint activities. Based on natural observations in two tasks, one based on joint manual actions, and another based on pointing, I explore the common underlying structure of preverbal communication and early face-to-face triadic interactions. Further, I investigate whether there is a developmental continuity between the reciprocal turn taking structure of infants' early face-to-face interactions and their later language usage.

Based on the same dataset, Chapter 3 explores, in greater detail, the nature of caregivers' and infants' multimodal deictic communication across contexts. In that study, I investigate whether caregivers systematically integrate language with deictic gestures, and whether the communication of both caregivers and infants is shaped by the context within which dyads are interacting. It is possible that caregivers' multimodal integration of deictic gestures and language along with the activity-specific communication provides infants with a multimodal scaffold that enables word learning and language development.

The second half of the thesis shifts from deictic to representational gestures. In Chapter 4, I use an eye tracking paradigm to investigate whether infants better learn spoken words or arbitrary representational gestures when labels are taught multimodally. Chapter 5 looks into infants' own spontaneous production of representational gestures. Here, I ask whether Dutch infants at 18 and 24 months are capable of producing representational gestures when confronted with a situation where such gestures would be effective in describing to an adult, how to operate a toy. In addition to providing important insights into infants' developing capacity for symbolic communication, the results of Chapters 4 and 5 are relevant in assessing particular programs that encourage the explicit training of representational gestures in infancy. Chapter 6 summarizes the results of all four studies and situates them into our current understanding of how language develops in both ontogeny and phylogeny. Finally, I conclude with a brief discussion of the practical implications of the research presented.

Chapter 2

Joint action and joint pointing at 12 months

Abstract

This study investigated the shared structure of pointing and joint acting in infancy. Thirty-nine 12-month-olds and their caregivers were recorded in two semi-natural tasks: one based on joint manual actions, and another which elicited pointing. Correlational analyses revealed a conversational turn-taking structure in caregivers' and infants' preverbal pointing as evidenced by a sequential and correlated use of points between partners. Further, there was a relationship between the frequency with which infants used conversational pointing and the time they spent in joint manual action. Results suggest a common structure of joint action and communication before language has emerged in earnest, and provide further support to the idea that cooperative referential communication originates in joint activities.

2.1 Introduction

Human communication and joint action share a common infrastructure. Both are based on cooperative principles and both rely on mutual responsiveness and commitment to each other's actions (Bratman, 1992; Grice, 1957). Herbert Clark (1996) asserts that "[l]anguage is really a form of joint action." (p. 3). When communicating, interlocutors work together in order to achieve mutual understanding and, when cooperating, individuals work together to fulfill shared goals. Given the similarities and mutual dependencies between communication and cooperation, one hypothesis is that cooperation and communication emerge in a correlated fashion. Tomasello (2008) has recently argued for, and provided evidence to show a phylogenetic link between communication and cooperation, suggesting that the cooperative structure of human communication emerged on the heels of more general cooperative skills for acting together. However, ontogenetically, little is known about the emergence of communication and how it is related to joint action.

Research has provided evidence for a developmental relationship between infants' vocabulary size and how caregivers and infants engage with one another in joint activities (Tomasello & Farrar, 1986; Trautman & Rollins, 2006). Examples of such activities include looking at or manipulating objects together. This line of research, however, did not address the relationship between the ability to act together and the use of language as a shared activity, as, for example, in conversations. Further, infants begin communicating with gestures before they have acquired language (Bruner, 1975). For example, recent research has revealed cooperative communicative skills underlying the pointing gestures of prelinguistic infants (for a recent overview, see Liszkowski, 2010). Interestingly, the ability to actively participate in joint activities emerges around the same time as pointing, towards the end of the first year of life (Bakeman & Adamson, 1984; Carpenter, Nagell, & Tomasello, 1998). However, to date, few studies have explicitly addressed the ontogenetic relationship between pointing and joint engagement (but see Leung & Rheingold, 1981, and Murphy, 1978). Further, studies on joint engagement have distinguished different degrees to which infants' interactions are really 'joint' as opposed to being scaffolded only by caregivers ("active" vs. "passive" joint engagement; see Bakeman & Adamson, 1984; Carpenter, 2009; Carpenter, et al., 1998). In contrast, little is known about the degree to which infants' early pointing is a joint activity. For example, it remains unknown whether caregivers and infants use pointing as a form of joint activity, possibly constituting a form of gestural proto-conversation.

The emergence of proto-conversations has been traced back to infants as young as two to three months of age when infants begin to engage in face-to-face interactions with adults. These early interactions often involve a reciprocal turn-taking structure whereby a caregiver and infant exchange smiles, vocalizations, and/or positive affect (Bateson, 1975; Stern 1977; Trevarthen, 1979). Proto-conversational exchanges may serve to provide infants with the turn-taking structure of later linguistic, topic-based conversations. Turn-taking is a fundamental part of conversation not only because it allows for an organized, cooperative flow of information, but also because it ensures mutual engagement and joint reference between interlocutors (Clark & Schaefer, 1989; Sacks, Schegloff, & Jefferson, 1974). To date, research on the transition from dyadic, non-referential proto-conversations to linguistic turn-taking conversations has largely overlooked the possibility of a turn-taking structure in prelinguistic gestural communication. One relevant study reported that while reading books with their infants, caregivers sometimes pointed during or in response to infants' pointing, possibly to expand in more detail on certain aspects of mutual interest (Murphy, 1978). One possibility is that infants' gestural communication already entails the beginnings of a topic-related, turn-taking conversation structure via pointing, whereby both caregivers and infants point in response to each other's points, and establish joint reference by pointing to the same referent.

In the current study, we chose a correlational approach to explore the cooperative structure of prelinguistic communication and its relation to joint acting in infancy. Although correlations do not provide evidence for causality, they are an important first step in establishing a hypothesized relation. We observed caregiver-infant dyads interact in two semi-natural situations: one Free Play task based on joint manual action and one Decorated Room task which elicited pointing. In our first approach, we investigated whether caregiver-infant pointing entails a form of preverbal gestural conversation. Here, we looked for a positively correlated usage of caregivers' and infants' points which were contingent upon each other's points, that is, in close temporal proximity. In addition, we investigated whether such conversational points would be used as a means to establish joint reference. For this, we analyzed whether caregivers and infants would sometimes point to the same referents as each other and whether the frequency of this type of pointing would be positively correlated between caregivers and infants. In our second line of inquiry, we investigated whether caregiver-infant joint action would be positively related to infants' gestural communication. Here, we correlated infants' conversational pointing in the Decorated Room task with the time they spent acting jointly in the Free Play task.

2.2 Method

Participants

Thirty-nine 12-month-old infants (20 females and 19 males; mean age = 12;1) and their primary caregivers (35 mothers and 4 fathers) participated in the study. Nineteen infants were index-finger pointers, 17 infants did not yet point with the index finger and indicated only occasionally with the whole-hand, and 2 infants never pointed. All participants lived in a mid-sized city in Germany, were recruited via a mailing list and received a small gift for their participation.

Procedure

Caregivers spent approximately five minutes interacting with their infants in two contexts, each of which resembled a different type of shared activity in the daily lives of caregivers and their infants. The first context was the Decorated Room task (See Figure 2.1). The Decorated Room task was meant to represent a situation where infants and caregivers spend time together, looking at objects. In order to replicate this social interactional context, a decorated room was designed in which a number of interesting items were placed on the wall and around the room. Some examples of objects in the room included stuffed animals, toys, and several pictures of animals. Caregivers were instructed to hold their infants on their hips and explore the items in the room with their infants. Participants were all kept blind to the aims of the study and no mention of language or gesture was given in the instructions. In order to assure that the context remained one of regard, caregivers were asked to avoid touching the objects. Each dyad spent approximately five minutes in the room. The entire session was recorded via four video cameras mounted in each of the four corners in the room.

Immediately following the Decorated Room task, the participants were asked to participate in the second task: the Free Play task (See Figure 2.1). The Free Play task was an unstructured free play situation where caregivers spent approximately five minutes interacting with their infants and a box of toys. Caregivers were instructed to interact "as they would do at home." Again, no mention of language or gesture was given in the instructions. The box of toys included a variety of different toys. Some examples include: a ring-stacking game, two toy cars, several stuffed animals, and a set of toy telephones. Four video cameras recorded dyads' interactions.

Figure 2.1. Setup of the two shared activities: Decorated Room task (left) and Free Play task (right)

Coding and analysis

All coding was done with ELAN software (free, downloadable software created by the Max Planck Institute for Psycholinguistics at http://www.lat-mpi.eu/tools/elan/; Sloetjes & Wittenburg, 2008), which allows for time-locked coding of the videotaped interactions. Pointing was coded in the Decorated Room task. A point was coded when an individual extended the arm partially or fully towards a discernable object, either with the hand or the extended index-finger. Infants' reaching attempts to obtain an object were not coded as points. A point which occurred within 10 seconds of another's point was coded as a '*joint point*.' Joint points were further coded as being either *same referent points*, in which the caregiver and the infant pointed to the same objects, or *different referent points*, in which the caregiver and the infant pointed to different objects.

Joint engagement was coded in the Free Play task, following Bakeman and Adamson (1984). An engagement state had to last at least three seconds. We coded for *Active joint engagement* and *Supported joint engagement*. Active joint engagement was coded when the interaction between the infant and the caregiver fulfilled two conditions: First, the caregiver and the infant were acting on a common or related object. Second, both the caregiver and the infant provided a cue that they were aware of each other's engagement. Awareness on behalf of the caregiver was automatically assumed unless there was clear evidence that the caregiver's attention was not focused on the infant or a common object. For the infant, we defined three possible cues to indicate awareness: The first cue was alternating eye gaze between the object and the caregiver. For this cue, we also required that the infant either produced emotive affect during the gaze alteration or altered gaze between her parent and the object at least two times. The second cue was when an infant clearly reacted to the caregiver's verbal and/or gestural instructions (e.g., to put a toy in a particular location). The third cue was a communicative gesture, such as a point, a show, a request, or a communicative reach. Active joint engagement was terminated after 10 seconds without a further cue indicating mutual awareness, or after one of the individuals clearly disengaged from the interaction, for example, by crawling away, or playing with a different toy on one's own (see Carpenter et al., 1998). An example of Active joint engagement would be when a caregiver stacks rings on a tower with her infant and the infant responds to her vocal and or gestural cues to stack the rings.

Supported joint engagement was coded when the caregiver and the infant were acting together on a common object without an observable cue for mutual awareness. For example, supported joint engagement would be coded if an infant was playing with the ring stacking game but never acknowledged his mother's involvement while playing with the game. Note that in order to obtain a more adequate measure of joint acting, the joint engagement coding referred only to object-directed actions, and not to attentional states alone (e.g. Joint Attention).

A trained assistant recoded data from nine of the dyads. Reliability was assessed using Spearman Rho correlations. We obtained positive, significant correlations for the time that individuals spent in Active joint engagement and Supported joint engagement (respectively, $\rho = .879$, $p = .002$, $\rho = .967$, $p < .001$). Further, none of the means differed significantly between the reliability coder and main coder.

2.3 Results

Joint pointing

We investigated the extent to which caregivers and infants pointed together, as a form of joint activity. Three caregivers and two infants never produced a point. Our analyses on joint pointing therefore focused on 34 dyads. Twenty-one infants (62%) used joint points. The significant majority of these infants were index-finger pointers (16 of 21 infants), while most of the infants who did not point with the index-finger (and only occasionally pointed with the open hand) did not engage in joint pointing (N= 10), Fisher's exact, $p = .004$. When looking only at fully-fledged index-finger pointers, the number of infants who engaged in joint pointing increased from 62% to 84%.

There was indeed a correlated use of pointing by caregivers and infants. Namely, there was a relationship indicating that caregivers who used joint points (N = 20) had infants who also used joint points (N = 23), *phi* = .443, $p = .010$. Further, there was a high positive correlation between the frequencies of joint points by caregivers and infants ($r(34) = .907$, $p < .001$), indicating that the more one followed into the other's points, the more the other followed into one's own points (Figure 2.2). Interestingly, this correlation was specific to joint points, as there was no significant correlation between the overall amount of pointing by caregivers and infants, $r(34) = .194$, p = .271. The mean numbers of joint points did not differ significantly between infants (2.7; range 0-15) and caregivers (2.50; range = 0-16), $t(33) = .754$ $p = .456$. When looking only at index-finger pointers the frequency of infants' joint points increased substantially (M = 4.16).

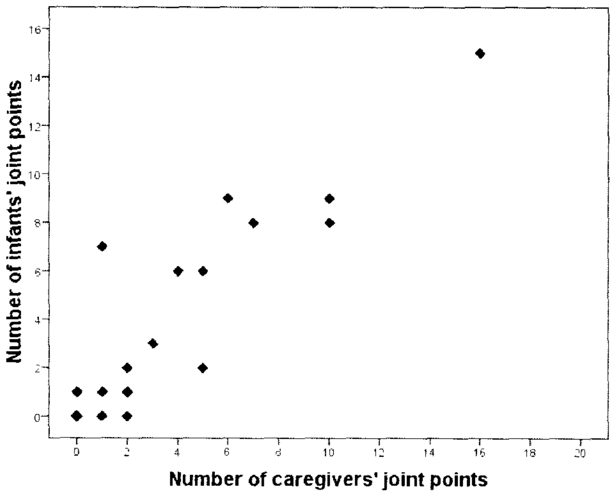


Figure 2.2. Relation between caregivers' and infants' joint points

Next we investigated the microstructure of joint pointing by looking into the actual referents of caregivers' and infants' joint points. On average, 62% of caregivers' joint points and 51% of infants' joint points were to the same referent. There was no difference in the frequency with which caregivers and infants used these same-referent points, $t(34) = 1.20$ $p = .239$, however, there was a substantial relation. On the individual level, caregivers who used same-referent points also had infants who used same-referent points, $phi = .589$, $p = .001$. With regard to the frequencies, the number of same-referent points by caregivers and infants was also positively correlated r $(34) = .610$, $p < .001$. As a whole, the pointing analyses reveal a joint, correlated usage of pointing by both infants and caregivers.

Joint acting and joint pointing

Dyads spent significantly more time in Supported joint engagement (mean proportion = .24) than in Active joint engagement (mean proportion = .10), t $(38)=4.23$, $p < .001$. The vast majority of infants engaged in active joint engagement (82%), however only about half of the sample was index finger pointers (49%), suggesting a developmental primacy of joint acting over index finger pointing.

In order to test for a relation between joint pointing and joint acting, we computed correlations across the two tasks. As Figure 2.3 shows, the proportion of time that dyads spent in active joint engagement correlated positively with the proportion of infants' joint points to the same referent, $r(37) = .327$; $p = .048$. Thus, the more time dyads spent actively engaged during the Free Play, the more infants shared reference in the Decorated Room. There were no significant correlations between overall points or different-referent points and active or supported joint engagement, indicating a specific relation between same-referent joint pointing and active joint engagement.

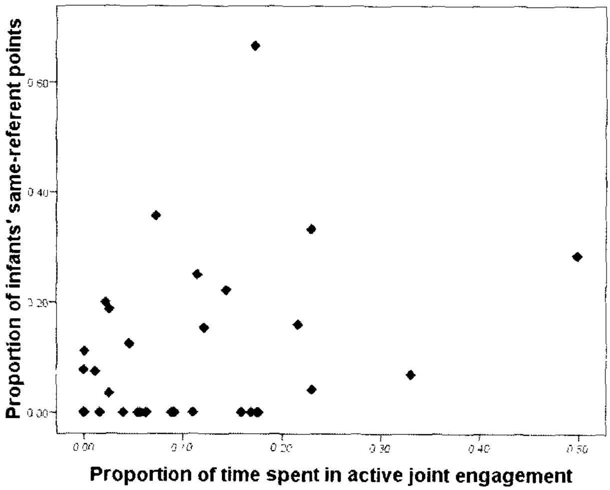


Figure 2.3. Relation between time spent in active joint engagement and proportion of infants' same-referent joint points

2.4 Discussion

In the current study, we investigated the common structure of communication and joint action at 12 months. First, we found evidence for a conversational usage of pointing in caregivers' and infants' communication. Caregivers and infants pointed in temporal proximity with each another, and to the same referent as one another. The frequency of these pointing acts was correlated within dyads, suggesting that pointing is a joint activity between caregivers and infants. Second, we found a relation between infants' and caregivers' gestural conversations and joint engagement. Namely, the frequency with which infants pointed to the same referent as their caregivers correlated positively with the time dyads spent actively engaged in joint action, indicating a relation between prelinguistic communication and joint action. Findings suggest that referential communication is a cooperative activity from the very beginning, and that communication and joint action are built on a common infrastructure from early in ontogeny.

It is unlikely that infants' joint points were simply acts of behavioral mimicry. First, if infants were simply mimicking their caregivers, then the overall amount of points by caregivers and infants should have been highly correlated. However, the high correlations that we found were specific to those points which were in temporal proximity. The current findings reveal that it is not pointing per se which increases the likelihood of infant pointing (see also Leung & Rheingold, 1981), but instead, pointing sequentially by following into another's point. Second, infants' joint referent points correlated with the time they spent actively engaged in joint action, a correlation that would not be expected based on behavioral mimicry alone. The specific set of correlations also suggests against the possible interpretation that the joint points simply occurred by chance. If this was the case, one would not have obtained a correlated usage of conversational pointing between infants and caregivers. The reported results instead support the idea that pointing is used in infancy in a joint fashion, whereby caregivers and infants point to the same referents as one another in order to establish joint reference and to make it explicit that they are indeed sharing attention.

Joint action and joint pointing

The frequencies of joint pointing and active joint engagement were rather low at 12 months of age. It is important to recall, however, that for about half of the infants in the sample, pointing was not yet fully-fledged (see Liszkowski & Tomasello, 2011). The correlation between index-finger pointers and joint pointers suggests that the conversational usage of pointing emerges with index-finger pointing. Indeed, when looking only at index-finger pointers, the frequency of joint pointing was substantially higher. Regarding the frequency of joint engagement, our results correspond to those of previous studies (Bakeman & Adamson, 1984; Carpenter et al., 1998) which have shown that infants' active participation in joint action is an emerging skill that increases over the second year of life. One possibility that warrants further investigation is that both behaviors increase with age in a mutually dependent manner.

Triadic, object-related joint actions emerge around nine months. Such interactions are initially scaffolded by caregivers who engage their infants in so-called supported joint engagement, bringing relevant objects into infants' focus of attention (Bakeman & Adamson, 1984, Trautman & Rollins, 2006). Around 12 months, based on parents' earlier scaffolding, infants begin to engage in active joint engagement, where they too, become active participants in the interactions. Not coincidentally, it is at this very age when infants also begin pointing. The relation between joint acting and joint pointing, and the developmental primacy of joint acting over index-finger pointing, suggest that referential communication arises out of experience in joint actions. Longitudinal evidence would still be needed to establish, concretely, the directionality of the correlation. In support of the hypothesis, however, recent cross-cultural evidence reveals that the amount of triadic joint engagement is related to how early and how frequently infants point (Salomo & Liszkowski, 2010).

Regarding the correlation between parents' and infants' joint pointing,here, too, a longitudinal design is needed to establish the directionality. However, one possibility is that caregivers scaffold infants' communicative development by pointing in temporal proximity to objects of mutual interest. Word learning studies have shown that labeling is especially efficient when the labels follow into infants' attentional focus (Tomasello & Farrar, 1986). Similarly, it is possible that, when pointing follows into infants' attentional focus, the point is immediately relevant for the infant as directing and manifesting attention to a shared referent. In turn, infants attempt to establish shared reference by following into caregivers' reference with pointing. This creates a communicative dialogue based on joint gestural reference. It is likely that these exchanges facilitate infants' understanding of topic based conversation, and helps establish the format for future conversations involving language.

To date, research on pointing has focused primarily on pointing gestures as single communicative acts and has largely overlooked the possible meta-structure of preverbal pointing as a joint activity. The current findings reinforce initial findings by Murphy (1978) and show that preverbal pointing encompasses a turn-taking meta-structure similar to that used in conversation, suggesting that topic-based turn-taking emerges in the gestural modality before language is even in place. Valloton (2011) recently showed that sign-trained babies and their caregivers also engage in gestural conversations. Our results show that gestural conversations develop naturally in infants who follow a typical course of development without sign intervention. Seeing that this turn-taking structure is used in infants' early, face-to-face exchanges and also in their later language, the current findings suggest continuity between infants' early non-referential interactions and their subsequent language usage.

The findings presented here reveal that pointing is a cooperative joint activity from its very emergence, as evidenced by its conversational usage and its concurrent relation to manual joint actions. Caregivers scaffold infants' social and communicative development by actively engaging them in joint activities, involving both manual actions and gestural conversations. Referential communication thus emerges first in the visual modality, most likely through experience in manual joint activities, and the resulting pointing conversations serve as a basis for the use of language as a form of joint action shortly thereafter.

In the following chapter, I explore caregiver – infant interactions in greater detail. In addition to linguistic input, infants are provided with extralinlguistic cues which enable language learning. In Chapter 3, I investigate two possible extralinguistc cues: First, I explore the possibility that infants' gestural and linguistic input is structured by the type of activity within which caregivers and infants are engaged and second, I look into the extent to which caregivers systematically integrate language and gesture when interacting with their infants. Extra-linguistic cues such as these, I argue, provide infants with a multimodal scaffold which fosters the communicative development of infants.

Chapter 3

The type of shared activity shapes caregiver and infant communication

A version of the study presented in this chapter was originally presented in:

Puccini, D., Hassemer, M., Salomo, D., & Liszkowski, U. (2010). The type of shared activity shapes caregiver and infant communication. *Gesture, 10*(2/3), 279-297. doi:10.1075/gest.10.2-3.08puc

Abstract

For the beginning language learner, communicative input is not based on linguistic codes alone. This study investigated two extralinguistic factors which are important for infants' language development: the type of ongoing shared activity and non-verbal, deictic gestures. The natural interactions of 39 caregivers and their 12-month-old infants were recorded in two semi-natural contexts: a free play situation based on action and manipulation of objects, and a situation based on regard of objects, broadly analogous to an exhibit. Results show that the type of shared activity structures both caregivers' language usage and caregivers' and infants' gesture usage. Further, there is a specific pattern with regard to how caregivers integrate speech with particular deictic gesture types. The findings demonstrate a pervasive influence of shared activities on human communication, even before language has emerged. The type of shared activity and caregivers' systematic integration of specific forms of deictic gestures with language provide infants with a multimodal scaffold for a usage-based acquisition of language.

3.1 Introduction

Research investigating children's language input typically focuses on linguistic factors such as the semantics, syntax, or phonology of caregivers' speech. However, according to social-pragmatic theories, children's language acquisition is also heavily dependent on other, non-linguistic factors of communication (Baldwin, 1995; Bruner, 1983, 1981 1975; Tomasello, 2003). Two extralinguistic factors that play a major role in children's acquisition of language are: (i) mutual engagement in shared activities and (ii) non-verbal, deictic gestures.

Shared activities facilitate children's language learning on a micro and a macro level (Tomasello & Farrar, 1986). On the micro level, shared activities enable infants to determine the referent of a word through a shared focus of attention (Tomasello & Farrar 1986; Baldwin, 1991). On the macro level, shared activities enable infants to learn about the different functions for which language is used (Nelson, 1981). Bruner (1983) argued that familiar, conventionalized routines in children's daily lives serve as 'formats' within which language becomes meaningful. According to Bruner, conventionalized routines allow infants to limit the amount of possible interpretations of an utterance since communication is directly relevant to the current activity. One prediction is thus, that the type of shared activity should influence the communication of caregivers and infants.

Several studies have compared caregivers' speech in different types of activities and have revealed equivocal findings. For example, when comparing activities that involved free play and book reading, Snow et al. (1976) did not find a difference between syntactic aspects in caregivers' language such as the relative frequency of noun phrases and verb phrases. However, Tardif, Gelman, and Xu (1999) found that English-speaking mothers used more verb types than noun types in play situations while the opposite was true in a book reading task. On the macro level, other studies have suggested that caregivers' speech acts differ according to which parent is interacting, the social class of the caregivers, and according to the type of activity (Leaper & Gleason, 1996; Ryckenbusch & Marcos, 2004; Snow et al. 1976). For example, caregivers use more behavior directives in free play contexts as opposed to book-reading contexts (Hoff Ginsberg, 1991; Jones & Adamson, 1987). Little research, however, has addressed the structural differences of the types of shared activities themselves. While the main focus has been on free play tasks that are based on joint manual activities, few studies have considered activities based on mutual regard of objects: an activity argued by Werner and Kaplan (1963) to serve an important role in the understanding of symbols and language learning. Although book-reading may share some of the features of a context of regard, books themselves are objects which infants manipulate, and infants sometimes even attempt to manipulate the objects they depict (DeLoache, Pierroutsakos, Uttal, Rosengren, & Gottlieb, 1998). Further, in real life situations, the majority of children in non-western cultures are not frequently confronted with book-reading practices and do not learn language usage through book reading but must rely on other formats of mutual object regard.

Deictic gestures are another pivotal aspect in the communication of caregivers and infants. Caregivers frequently use deictic gestures to reinforce the message conveyed in their speech (Iverson, Capirci, Longobardi, and Caselli, 1999) and infants already use deictic gestures to communicate in meaningful ways before they use language (Bates, Camaioni, & Volterra 1975; Liszkowski, 2010 for an overview). It is currently unknown whether the type of shared activity influences deictic gesture usage. With regard to caregivers, one study suggests that caregivers do not use deictic gestures differently across a free play task (mostly involving book reading) and a counting task (requiring counting several toys together; O'Neill, Bard, Linnell, & Fluck, 2005). However, in that study, the types of shared activities were not very different from each other, as both involved looking at objects or pictures of objects. With regard to infants, no study to date has investigated the influence of the type of shared activity on infants' deictic gesture use. By extending social-pragmatic theories of language acquisition to the gestural modality, we would expect that the usage of non-verbal deictic gestures also varies according to the type of shared activity, for example, if one activity focused on joint manual actions and the other, on joint visual regard of objects. If the type of shared activity influenced infants' use of deictic gestures, this would demonstrate a pervasive influence of shared activities on human communication in the gestural modality from the beginning, even before the emergence of language.

A final aspect relevant to caregiver-infant communication concerns caregivers' multimodal integration of speech and gesture. It is well-established that caregivers integrate gestures with speech (Iverson et al., 1999; O'Neill et al., 2005; Gogate, Bahrick, & Watson, 2000). For example, Gogate et al. (2000) showed that caregivers' '*show*' gestures tended to include more movement when caregivers were teaching infants novel verbs as opposed to novel nouns. O'Neill et al. (2005) showed that caregivers' deictic gestures most often disambiguate referents in their speech as opposed to emphasize or add information to the speech. Kalagher and Yu (2006) further showed that word learning is facilitated when caregivers point to the target referent while naming the object. However, to date, most studies have either focused on a single deictic gesture type (e.g., pointing) or combined all deictic gestures into one category (e.g., deictic gestures). Few studies have investigated whether caregivers systematically combine specific linguistic features with specific forms of deictic gestures, and to what extent different types of shared activities affect how speech and gesture are combined. If the ongoing shared activity influences the use of language and gesture, and if gestures indeed play a facilitative role in the acquisition of language and its usage, we expected that speech and gesture combinations should vary systematically according to the type of shared activity. Caregivers' systematic integration of particular deictic gestures and linguistic features in specific shared activities would provide the beginning language learner with valuable regularities in a multimodal scaffold.

In the current study, we investigated how the type of shared activity shapes caregiver and infant communication and whether caregivers combine particular deictic gestures with specific aspects of speech. We used a semi-natural context of regard in which caregivers and their twelve-month-old infants spent five minutes exploring items displayed around a decorated room. We compared language use (on behalf of caregivers) and deictic gesture use (on behalf of both caregivers and infants) in this new task with language and gesture use in a free play task, in which dyads acted together on objects. Our first goal was to investigate whether different types of activities influenced caregivers' language use with regard to specific speech acts and linguistic references. Our second goal was to determine whether the type of shared activity influenced caregivers' non-verbal, gestural communication and if it might even influence the gestural communication of prelinguistic infants. Our third goal was to determine whether there was a specific pattern with regard to how caregivers combine deictic gestures with language, and to investigate the extent to which these combinations would be affected by the different types of shared activities. Based on social-pragmatic theories of language acquisition and on previous findings, we expected that the type of shared activity would structure caregiver-infant communication such that caregivers would use both speech and gesture - and infants their gestures - differently across the two contexts. Further, if caregivers combined specific deictic gesture types with specific language features, then we predicted that these combinations too, would differ according to the type of shared activity.

3.2 Method

Participants

This study included the same participants as those in Chapter 2, consisting of thirty-nine 12-month-old infants and their primary caregivers. For more details on these participants, see the *Participants* section in the previous chapter.

Procedure

The same video recorded interactions between caregivers and infants from Chapter 2 were also used in this study. Accordingly, the procedure was identical in both studies. In short, each dyad spent five minutes interacting in two semi-natural tasks: the Context of Regard (referred to in Chapter 2 as the Decorated Room task) and the Context of Action (referred to in Chapter 2 as the Free Play task). In the Context of Regard, caregivers were asked to hold their infants on their hips and explore a variety of objects that were displayed around a room. In the Context of Action, caregivers were given a basket of toys and asked to interact with their infants as they normally would. These two situations elicited spontaneous interactions between infants and caregivers.

Coding

All coding was done in ELAN, a free software program developed by the Max Planck Institute for Psycholinguistics, which allows for coding that is time locked with the video data. The coding focused on three aspects in the interaction: Language, Gestures, and Language and Gesture combinations.

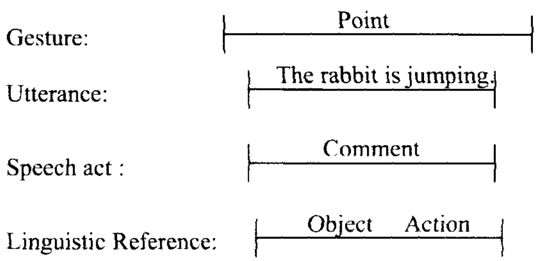
Language. All caregiver speech was first transcribed orthographically and time locked with the video data. Since infants produce so little language at 12 months, we only coded the language of the caregivers. Specifically, we coded linguistic references and speech acts. For linguistic references, our goal was to code referential words. Thus, we coded all words that were used to refer to objects (e.g. It's a *goat*), dynamic actions (e.g. He's *jumping*), and features (e.g. The rabbit is *pink*). Words that were unclear or incomprehensible were not included in the analysis. Words that were repeated in the same utterance were counted as a single reference. Utterances were defined as any unit of speech preceded and followed by silence (Crystal, 1991). For speech acts, we coded all utterances as *Comments* (e.g. *That's a frog.*), *Questions* (e.g. *Where's the fish?*), behavioral *Directives* (e.g. *Put it there.*), and *Invitations* (e.g. *Look here!*). All speech act categories except invitations were adopted from Tomasello and Farrar (1986). Invitations are often used with infants as a summons to draw infants' attention towards objects, after which, caregivers will speak further about the object (Estigarribia & Clark, 2007). The four speech act categories encompassed all speech such that every utterance was assigned to at least one of the four aforementioned categories. It should also be noted that some utterances were used for more than one speech act. For example, a caregiver might say: "Look, it's a frog." in which case the utterance serves as both an invitation and a comment.

Gestures. We coded the following deictic gestures of caregivers and infants: <u>Action Demonstration</u>: An individual performs an action with an object, with the intent for the other individual to duplicate the action; <u>Object Demonstration</u>: An individual draws attention to an object by moving and animating it for the other to see; <u>Show</u>: An individual draws attention to an object by holding it in view of the other; <u>Give</u>: An individual transfers possession of an object from self to another; <u>Place</u>: An individual transfers an object to the ground to draw the other's attention to the object; <u>Point</u>: An individual uses the hand or a part of the hand to indicate an external referent to another; <u>Request</u>: An individual requests an object held by another by reaching out one's own hand, palm up; <u>Reach</u>: An individual indicates desire of an object by reaching for it, without the sole intent of retrieving the object by oneself. The gesture categories were based on previous literature and on observation of caregiver-infant interactions (see e.g. Bates et al., 1975; Blake, O'Rourke, & Borzellino, 1994; Clark, 2003; Trautman and Rollins, 2006).

Language & Gesture combinations. Language & Gesture combinations were based on temporal synchrony between the gesture and the speech with which the gesture occurred (Iverson et al. 1999). For linguistic references, the temporal synchrony had to be between the actual referential word and the gesture. For speech acts, the temporal synchrony had to be between the utterance in which the speech act occurred and the gesture. An example of a Language & Gesture combination is provided in (1).

In (1), the point spans across the entire duration of the utterance. The point spans across one speech act (a comment) and two linguistic references (an object reference (rabbit) and an action reference (jumping).

(1)



Reliability

Nine randomly selected dyads were re-coded for language and gesture occurrences by a second, trained assistant. Inter-rater reliability revealed significant correlations for each language category, both with respect to speech acts (all ρ's > .86, *p*'s ≤ .001), and linguistic references (all ρ's > .93, *p*'s ≤ .001). For gestures, inter-rater reliability revealed significant correlations for the frequencies of each gesture type (ρ's > .69 *p*'s < .05).[1]

3.3 Results

Coding revealed that dyads spent, on average, slightly more time in the Context of Action (mean = 312 seconds) than in the Context of Regard (mean = 301 seconds), $t(38) = 2.409$, $p = .021$. We therefore calculated language and gesture frequencies per minute, which allowed for comparison between the two activities.

Language

A 2 (Shared activities) x 3 (Linguistic reference types) repeated measures ANOVA on the mean frequency of references per minute revealed that caregivers used overall significantly more linguistic references in the Context of Regard than in the Context of Action $F(1, 38) = 23.00$, $p < .001$. Further, caregivers used linguistic reference types with different frequencies, $F(1.64, 62.27) = 55.09$, $p < .001$ (adjusted for Greenhouse-Geisser correction). These two effects interacted significantly, $F\ (1.63, 61.82) = 85.84$, $p < .001$ (adjusted for Greenhouse-Geisser correction). On account of the different amounts of speech in the two activities, we resolved the interaction with t-tests on the proportions of linguistic reference types relative to the total amount of linguistic references in each activity. Figure 3.1 shows that caregivers used significantly more object references and more feature references in the Context of Regard than in the Context of Action, $t(37) = 9.36$, $p < .001$, $t(37) = 2.77$, $p = .009$, respectively. Further, they used significantly more action references in the Context of Action than in the Context of Regard, $t(37)= 9.21$, $p < .001$.

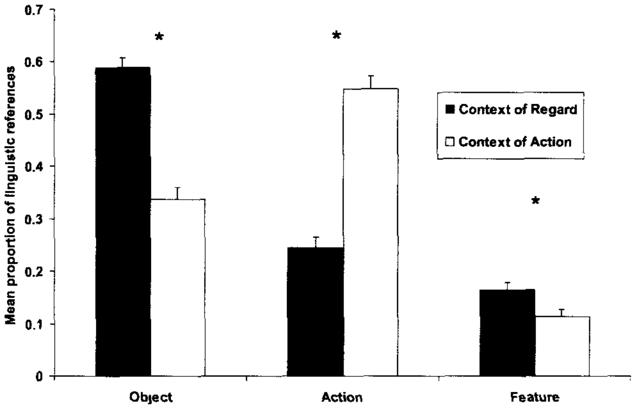


Figure 3.1. Caregivers' use of linguistic references across contexts

A 2 (Shared activities) x 4 (Speech act types) repeated measures ANOVA on the mean frequency of speech acts revealed that caregivers used significantly more speech acts in the context of Context of Regard than in the Context of Action, $F(1,38) = 23.78, p < .001$. Further, caregivers used speech acts with different frequencies, $F(1.41, 53.49) = 177.80, p < .001$ (adjusted for Greenhouse-Geisser correction). These two effects interacted significantly, $F(2.06, 78.19) = 18.35, p < .001$ (adjusted for Greenhouse-Geisser correction). On account of the different amounts of speech in the two activities, we resolved the interaction with *t*-tests on the proportions of speech act types relative to the total amount of speech acts in each activity. Figure 3.2 shows caregivers' use of speech acts in the two activities, whereby caregivers used significantly more directives in the Context of Action than in the Context of Regard, $t(37) = 6.58, p < .001$), and more invitations in the Context Regard than in the Context of Action, $t(37) = 2.00, p = .052$. There was no significant difference between caregivers' use of comments in the two contexts, $t(37) = .374, p = .71$, nor was there a significant difference in their use of questions, $t(37) = 1.70, p = .098$.

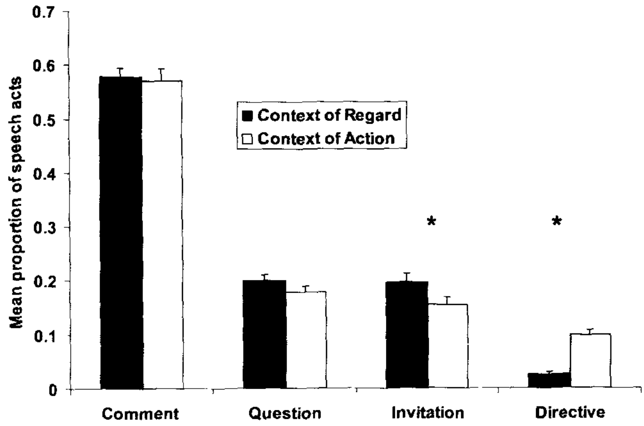


Figure 3.2. Caregivers' use of speech acts across contexts

Gestures

Figure 3.3 shows the frequencies of each gesture type for both caregivers and infants in the Context of Action. As is evident from Figure 3.3, caregivers and infants used a variety of deictic gestures in the Context of Action. Figure 3.4 shows that both caregivers and infants pointed significantly more in the Context of Regard than in the Context of Action, $t(38) = 6.17, p < .001$; $t(38) = 6.60$ (38) $p < .001$, respectively. For caregivers, 92% (N = 36) pointed at least once in the Context of Regard, whereas only 59% (N = 23) pointed at least once in the Context of Action (McNemar, $p < .001$). For infants, 95% (N = 37) pointed at least once in the Context of Regard, whereas only 23% (N = 9) pointed at least once in the Context of Action (McNemar, $p < .001$).

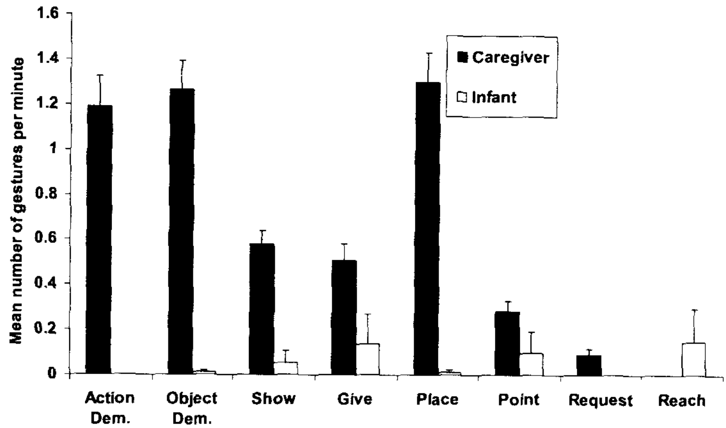


Figure 3.3. Caregiver and infant gestures in Context of Action

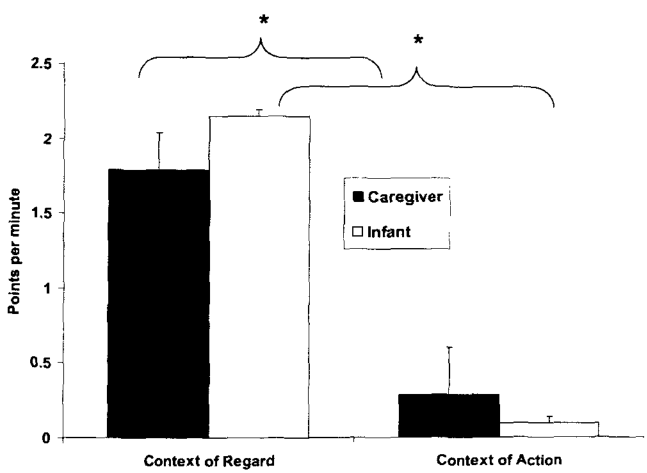


Figure 3.4. Caregiver and infant pointing across contexts

Language & gesture combinations

Figure 3.5 shows that in both contexts, caregivers' deictic gestures were most often accompanied by speech. In the Context of Action, 63% of all deictic gestures were accompanied by speech, although there was also variation with respect to each individual gesture type (see Figure 3.5). In the Context of Regard, 94% of caregivers' pointing gestures were accompanied by speech.

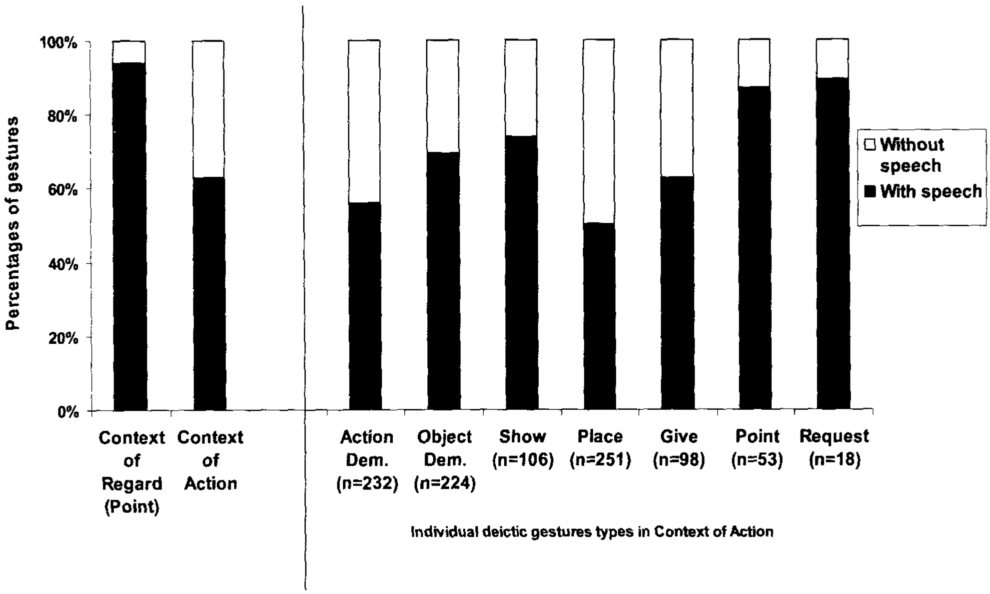


Figure 3.5. Percentages of caregivers' deictic gestures accompanied by speech across contexts

Context of Action. Since many of the gestures were used infrequently in the Context of Action and since not all gestures were used by all caregivers, parametric tests were not appropriate. We therefore combined data of all caregivers into one single set and conducted chi-square analyses. The chi-square tests made it possible to discern whether certain gestures co-occurred with particular speech act types and linguistic reference types more than would have been expected by chance.

A chi-square test revealed that gesture types and linguistic references did not co-occur randomly: $\chi^2(12) = 66.61, p < .001$. In order to investigate this association further, adjusted standardized residuals (ASRs) were calculated for each individual chi-square cell. ASRs indicate which cells contribute to the significant chi-square value and they allow for the comparison between cells (Agersti, 2007; Sheskin, 2004). These calculations allowed us to determine which particular Gesture & Speech act and Gesture & Linguistic Reference combinations occurred more or less than would be expected by chance. ASRs above 2.0 reflect a significant association between two variables ($p < .05$) while ASRs below -2.0 reflect a significant dissociation between two variables ($p < .05$; Agersti, 2007; Sheskin, 2004). ASRs for the entire chi-square are displayed in Table 3.1. The residual analysis revealed significant associations between the following gestures and linguistic references: Action demonstration & Action reference, Object demonstration & Object reference, Show & Object reference, and Give & Action reference, indicating that these particular Gesture & Linguistic reference combinations occurred more often than would be expected by chance. In the example in (2), a caregiver used an action demonstration (by ostensively stacking the rings) with an action reference (*stack*) as well as an object reference (*ring*) while commenting about an action to her infant.

(2)

Gesture:	Action Demonstration
Utterance:	You stack the ring here, like this
Speech act :	Comment. Invitation
Linguistic Reference:	Action Object

There were significant dissociations between the following: Action demonstration & Object reference, Object demonstration & Action reference, Object demonstration & Feature reference, Show & Action demonstration, and Give & Action reference, indicating that these particular combinations occurred less than would be expected by chance.

Table 3.1. *Adjusted standardized residuals for Gesture & Linguistic reference combinations in Context of Action*

	Linguistic Reference		
Gesture	Object	Action	Feature
Action Demonstration	**-4.5**	**4.8**	-0.6
Object Demonstration	**5.2**	**-3.4**	**-2.4**
Show	**3.3**	**-2.9**	-0.5
Place	-1.1	-0.3	2
Give	**-3.5**	**2.1**	1.9
Point	-0.6	0.2	0.6
Request	-1.2	1.1	0.1

Note. Statistically significant associations (positive value) and dissociations (negative value) are in boldface; $p < .05$.

Another chi-square analysis was computed on Gesture & Speech Act combinations in the Context of Action. This revealed that deictic gesture types and speech acts did not co-occur randomly, χ^2 (18) = 82.78, $p < .001$. ASRs were calculated to further break down the association (see Table 3.2). These revealed significant associations between: Object demonstration & Comment; Place & Invitation; Give & Directive; Point & Directive, and Request & Question, and significant disassociations between: Give & Invitation; Point & Question; and Request & Invitation. In the example in (3), a caregiver uses a Give gesture to transfer a ring to her infant and a Directive speech act to direct the infant to stack it on the tower.

3)

Gesture: Give

Utterance: Put this ring on the stack.

Speech act : Directive

Linguistic Reference: Action Object Object

Table 3.2. *Adjusted standardized residuals for Gesture & Speech act combinations in Context of Action:*

	Speech Act			
Gesture	Comment	Directive	Invitation	Question
Action				
Demonstration	0.6	-0.7	0.9	-1.3
Object Demonstration	**2.4**	-1.1	-1.7	-0.5
Show	-0.5	-1.3	0.8	0.8
Place	-1.5	-1.1	**3.1**	-0.5
Give	0	**4.1**	**-4.4**	1.7
Point	-1.2	**2.3**	1.8	**-2.1**
Request	-1.3	-0.6	**-2.3**	**4.8**

Note. Statistically significant associations (positive value) and dissociations (negative value) are in boldface; $p < .05$

Context of Regard. In the Context of Regard, we analyzed only the language that accompanied caregivers' points since all of the other gestures involve direct contact with the objects and the participants had been asked not to touch the objects. Figure 3.6 shows caregivers' use of Point & Linguistic reference combinations in the Context of Regard. A one-way repeated measures ANOVA on the mean frequencies of combination types revealed a significant difference between the various Point & Linguistic reference type combinations in the Context of Regard , $F(1.05) = 21.802$, $p < .001$ (adjusted for Greenhouse Geisser correction). Paired *t*-tests revealed that caregivers used Point & Object reference combinations significantly more often than Point & Feature reference combinations, $t\ (38) = 4.815$, $p < .001$ and Point & Action reference combinations $t\ (38) = 4.639$, $p < .001$. They also significantly used more Point & Feature references than Point & Action references, $t(38) = 2.037$, $p = .049$.

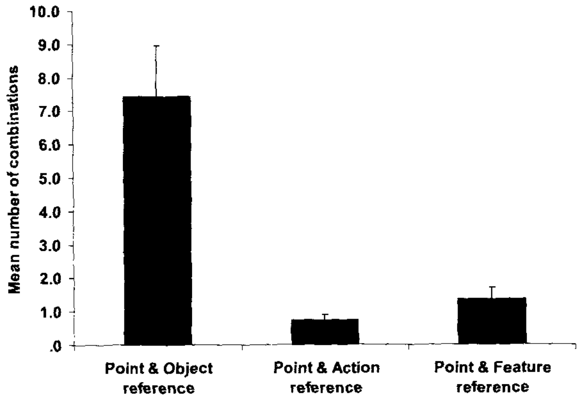


Figure 3.6. Caregivers' Point & linguistic reference combinations in the Context of Regard

Figure 3.7 shows caregivers' use of Point & Speech act combinations in the Context of Regard. A one-way repeated measures ANOVA revealed a significant difference between the frequencies of the various combinations, F (1.37) = 26.94 $p < .001$ (adjusted for Greenhouse-Geisser correction; see Figure 3.7). To break down this effect, paired t-tests were computed. These showed that caregivers used Point & Comment combinations more than all other combinations (all p's $< .01$), followed by more Point & Invitation combinations (all p's $< .01$). Point & Directive combinations were used least often (all p's $< .001$).

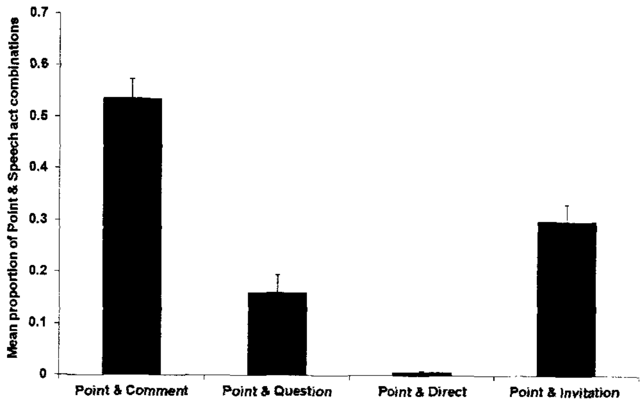


Figure 3.7. Caregivers' Point & Speech act combinations in the Context of Regard

3.4 Discussion

The current study shows how different types of shared activities structure caregivers' and infants' communication. Depending on the type of shared activity, caregivers exposed infants to different types of linguistic input, to different types of deictic gestures, and to specific gesture-language combinations. The type of shared activity also influenced the way prelinguistic infants communicated with their deictic gestures. The findings support social-pragmatic theories of language usage and acquisition and show how shared activities structure human communication from the beginning. The type of shared activity and caregivers' systematic integration of specific forms of deictic gestures with language likely provides infants with a multimodal scaffold for a usage-based acquisition of language.

The type of shared activity had a pervasive influence on the verbal and non-verbal communication of caregivers and even on the communication of prelinguistic infants, which reveals that shared activities structure communication irrespective of both the modality and the presence of language. In a context where action and manipulation of objects was possible, both caregivers and their prelinguistic infants used a variety of proximal deictic gestures such as show, place, give, or object demonstration. In this context, they rarely used points. However, when the interaction was focused on regard of objects, both caregivers and infants pointed frequently. It is worth emphasizing that the total amount of objects was approximately equal between the two contexts, thus, infants had equal opportunities to point in each. The high prevalence of the pointing gesture in contexts of regard, and its relative absence in contexts of action suggests that pointing is used in infancy primarily as a means to share interest in distal, non-manipulable objects (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004). In support of this, our speech act analysis revealed that caregivers pointed in the Context of Regard most often to comment, while they pointed in the Context of Action most often to direct infants' behavior. This illustrates how the inherent ambiguity of the pointing gesture (see also Quine, 1960) is disambiguated through the shared activity within which it is used (see also Liebal, Behne, Carpenter, & Tomasello, 2009). With regard to the ontogenetic origins of pointing, one intriguing possibility is that distal reference via pointing may build on an understanding of reference underlying proximal gestures, which may first emerge in the context of shared manual activities.

The influence of the type of shared activity pertained unequivocally to caregivers' language usage. We found more object references when caregivers and infants were in a context based on regard of objects than when they were in a context based on acting on objects. Conversely, there were more action references when dyads were in a context of acting on objects than when they were looking together at objects. This illustrates, on the micro level, how different types of shared activities help narrow the referential interpretation of words. The use of speech acts was also dependent on the type of shared activity: There were more invitations to look at objects when dyads were in the context based on regard of objects than when they were in the context focused on acting on and manipulating objects. Conversely, there were more behavior directives in the Context of Action. On the macro level, this demonstrates how the usage of language is shaped by the type of shared activity.

With regard to the integration of different kinds of deictic gestures and speech, as measured in the Context of Action, we found that caregivers integrated language and gesture systematically: Specific gestures were used in combination with particular types of language. For example, action references systematically accompanied action demonstrations, thus highlighting action, and object references systematically accompanied show gestures, thus highlighting objects. Further, the speech acts that accompany particular gestures offer insight into the various uses of the gestures themselves. Request gestures, for example occurred frequently with questions, e.g. 'Can I have the yellow block?' where the speech and the gesture were both used to request an object from the infant. Although the place gesture and the give gesture are morphologically fairly similar, they were combined with different speech act types, revealing that they are actually used differently from one another. The integration of the two modalities constitutes an advantageous form of reinforcing multimodal communication through which caregivers maintain infants' attention and scaffold their communicative development. It is likely that infants use speech accompanying gestures to narrow down the possible referential interpretations of words. Thus, with the help of activity-dependent, gesture-language combinations, infants themselves can learn to refer by making associations between utterances they hear, the gestures they see, and the objects and actions that are relevant to the ongoing shared activity.

Our results also have important methodological implications. First, the results accentuate the need to consider language acquisition in context: Researchers should be cautious in assuming that a language or gesture sample from a single type of activity (e.g. "free play" in the sense of joint manual action) accurately portrays one's communication. They should also be cautious when combining data from multiple settings. Results show that the type of shared activity and caregivers' accompanying gestures play an integral role in shaping caregiver-infant interactions. Researchers studying infants' natural language development should therefore strive for a comprehensive portrayal of infants' language input by including factors other than the linguistic content itself. Second, our findings are also relevant for cross-cultural research. Cultures differ greatly in the amount of time they spend in various types of shared activities, and in how frequently various gesture types are used (Salomo & Liszkowski, 2010). It is thus likely that infants' social and linguistic development depends on how prevalent certain types of interactions are within the culture.

Taken together, the finding that caregiver and infant communication is shaped by the type of shared activity, and the finding that caregivers integrate language and gesture systematically, lend support to the notion that infants' communicative input is structured by factors other than language itself. These factors help infants infer communicative intent and extract relevance from the input (Bruner, 1975; Tomasello, 2003). Following usage based theories of language acquisition, this enables infants to both disambiguate the intended referent in caregivers' speech, and to learn about the functions for which speech is used. As a whole, our study supports a rich, socio-pragmatic view of language acquisition whereby human communication is structured by non-linguistic gestures and by the activities within which it is used.

The findings reported in Chapters 2 and 3 suggest that infants' communicative input is highly structured. Infants' own early communicative exchanges are embedded in social activities within which infants are exposed to many deictic gestures. Likewise, infants' own gestural production is a social activity, in which caregivers' and infants exchange gestures as a type of early protoconversation. Deictic gestures thus play a major role in the early communicative exchanges of infants and it is likely that these gestures provide a foundation upon which language is built. In Chapters 4 and 5, I look into how infants interpret representational gestures. It is possible, for example, that representational gestures, too, play a major role in early language development. If this is the case, one would predict that infants could interpret representational gestures as object labels and that they, themselves would use representational gestures creatively to communicate with adults when their deictic gestures and spoken words are insufficient. These questions are addressed in the following chapters.

[1] Not all correlations for infant gestures were significant. ρ for infants' Reach gestures was .61, $p = .08$, and for infants' Show gesture, $\rho = .55$, $p = .129$. These correlations are non-significant because infants rarely gestured in this context which greatly reduces the statistical power of the test.

Chapter 4

15-month-old infants fast map words but not representational gestures of multimodal labels

A version of the study presented in this chapter was originally presented in:

Puccini, D., & Liszkowski, U. (2012). 15-month-old infants fast map words but not representational gestures of multimodal labels. Frontiers in Developmental Psychology, 3, 101. doi:10.3389/fpsyg.2012.00101.

Abstract

This study investigated whether 15-month-old infants fast map multimodal labels, and, when given the choice of two modalities, whether they preferentially fast map one better than the other. Sixty 15-month-old infants watched films where an actress repeatedly and ostensively labeled two novel objects using a spoken word along with a representational gesture. In the test phase, infants were assigned to one of three conditions: Word, Word + Gesture, or Gesture. The objects appeared in a shelf next to the experimenter and, depending on the condition, infants were prompted with either a word, a gesture, or a multimodal word-gesture combination. Using an infant eye tracker, we determined whether infants made the correct mappings. Results revealed that only infants in the Word condition had learned the novel object labels. When the representational gesture was presented alone or when the verbal label was accompanied by an arbitrary representational gesture, infants did not succeed in making the correct mappings. Results reveal that 15-month-old infants do not benefit from multimodal labeling and that they prefer words over arbitrary representational gestures as object labels in multimodal utterances. Findings put into question the role of multimodal labeling in early language development.

4.1 Introduction

Multimodal speech-gesture combinations are an integral part of language development. Caregivers, for example, often provide labels for infants in temporal synchrony with gestures such as pointing and showing (Gogate, Bahrick, & Watson, 2000; Gogate, Bolzani, & Betancourt, 2006; Masur, 1997; Ninio, 1980). These multimodal speech-gesture combinations scaffold infants' referential understanding (Iverson, Capirci, Longobardi, & Caselli, 1999), since deictic gestures help establish joint attentional episodes which are crucial to the process of word learning (Baldwin, 1991; Tomasello & Farrar, 1986). Further, infants combine their own deictic gestures with words, and these speech-gesture combinations are predictive of the two-word stage (Iverson & Goldin-Meadow, 2005). As a whole, these studies show that deictic gestures are an integral part of multimodal gesture-speech combinations and are intimately connected with language learning, laying the grounds for first language acquisition.

Much less is known about the role of representational gestures in infants' word learning. Whereas deictic gestures direct attention to objects in the immediate environment, representational gestures stand in for the entities to which they refer (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Capirci, Iverson, Pizzuto, & Volterra, 1996; Iverson, Capirci, Caselli, 1994; McNeil 1992). Representational gestures can be either iconic, which have some perceptual resemblance to their referent, or arbitrary, which have no perceptual resemblance to their referent. Interestingly, infants younger than two years learn arbitrary gestures as easily as iconic gestures, suggesting that they are not sensitive to gestural iconicity until later in development (Bates et al., 1979; Namy, 2008; Namy, Campbell, & Tomasello, 2004).

One interesting possibility is that, similar to deictic gestures, representational gestures, also facilitate language development. A recent study suggests a correlation between parents' multimodal labeling with iconic gestures and infants' acquisition of the labels (Zammit & Schafer, 2011). However, in that study, parents' multimodal labeling with deictic gestures as well as their labeling without gestures also correlated with infants' acquisition of the labels. Accordingly, the study does not reveal a direct relation between iconic gestures and word learning. It has also been suggested that infants' production of representational gestures may be related to early vocabulary development, as indicated by a concurrent correlation (Acredolo and Goodwyn, 1988). The idea of a relation between representational gestures and language has recently received a great deal of attention in the public. In particular, it has led to a mini-industry offering "baby-sign" courses to parents and their babies, guided by the claim that babies can be taught representational gestures to communicate before they can talk. In an experimentally controlled training study (Goodwyn, Acredolo, & Brown, 2000), one group of parents was encouraged to provide their infants with multimodal labels for a number of words that are commonly learned around this age range. Infants who received multimodal training outperformed infants in another group who received no explicit training in nearly all of the receptive and productive language measures at nearly all ages investigated (from 15 to 36 months), suggesting that multimodal labels may facilitate language development. However, there are several criticisms of this study (see Johnston, Durieux-Smith, & Bloom, 2005). The authors, for example, did not report on how subjects were recruited, nor did they report on how infants were assigned to groups. Therefore, one cannot rule out the possibility that infants in the multimodal group outperformed their peers because their parents were more motivated to begin with. And, although the study included an additional control group exposed to increased verbal labeling, no comparisons between the multimodal group and the verbal group were reported. Accordingly, it remains unclear whether prelinguistic infants benefit from multimodal utterances involving representational gestures more so than from verbal utterances without representational gestures.

Other studies have used fast mapping paradigms to investigate the role of representational gestures in word learning. On the basis of the finding that young infants can associate labels with novel objects with very little exposure (Houston-Price, Plunkett, & Harris, 2005; Schafer & Plunkett, 1998), Namy & Waxman (1998, 2000) and Namy (2001) investigated whether infants can also fast map arbitrary representational gestures. They found that 17- and 18-month-olds fast map both spoken words and representational gestures, suggesting that infants' early symbolic capacity is not specific to a single modality. Other fast mapping studies show that infants as young as 13 months can also associate other types of stimuli, including beeps and tones as labels for novel objects (Namy, 2001; Namy & Waxman, 1998; Woodward & Hoyne, 1999). In contrast, infants as young as 6 months expect object labels in the form of spoken words (Fulkerson & Waxman, 2007). As Fulkerson and Waxman acknowledge, the conflicting findings in these studies might be due to the methodologies employed. Namely, Fulkerson and Waxman trained infants on object labels using disembodied sound-object pairings which involved no joint attention, while Namy (2001) used an interactive paradigm that assured infants were jointly attending to the intended referent while hearing the object labels.

Few fast mapping studies have directly compared the learning of gestural labels versus spoken labels when presented simultaneously in multimodal utterances. One recent experimental study suggests that multimodal labeling facilitates 3-year-olds' comprehension of novel verbs (Goodrich & Hudson Kam, 2009). In that study, participants were first shown a distinct action for each of two objects. They were then exposed to two multimodal labels containing a spoken novel verb and an iconic action depiction corresponding to each of the objects. At test, participants heard the spoken novel verb without the iconic action depiction and had to choose to which object the verb referred. Results showed that 3-year-old children benefited from iconic gestural information when learning novel verbs. Another study (Wilbourn & Sims, in press) investigated whether 26-month-olds can learn multimodal labels including spoken words and arbitrary manual gestures. They reported that 26-month-olds can correctly identify the referent of a gesture-word combination when trained using a multimodal label, but not when trained with a gestural label alone. The authors, however did not include a group of infants who were exposed solely to spoken words, thus, it remains unknown whether the gestural label actually facilitated word learning. Further, it is still unclear whether younger infants at the cusp of acquiring language benefit from gestural labels in multimodal utterances. One question is thus whether preverbal infants fast map multimodal labels, and, when given the choice of two modalities, whether infants preferentially fast map one better than the other.

In the current study, we investigated how infants interpret utterances containing multimodal object labels. Using an infant eye tracker, we taught infants multimodal labels for two novel objects: Each time a label was produced, infants heard a spoken label coupled with an arbitrary representational gesture. At test, infants were presented either with the verbal label, a verbal label accompanied by a gestural label, or only a gestural label. Measuring their looks to the objects, we assessed whether infants had made the correct object-label associations. Following Schafer and Plunkett (1998), we used a two-label procedure which is a more ridged method of establishing word learning, and which has been shown to be more reliable for looking measures than reaching measures (Gurteen, Horne, & Erjavec, 2011). We expected that if infants would be able to map a word to a referent, they would look above chance to the referent. If representational gestures reinforce the association between the label and its referent, then infants should fast map better when both the representational gesture and the word are available at test. However, if infants do not know how to interpret the representational gestures, then infants should perform better when the word is decoupled from the gesture.

4.2. Methods

Participants

Sixty 15-month-old infants participated in this study (mean age: 15;14; 30 males and 30 females). Infants were randomly assigned to one of three conditions: Word (n = 20), Word + Gesture (n = 20), and Gesture (n = 20). An additional 15 infants were tested but excluded from analysis (three in the Word condition, five in the Word + Gesture condition, and seven in Gesture condition; see *test phase* section). Five infants were excluded for fussiness; seven because less than 50% valid gaze points were recorded during the training (see results), two due to caregiver interference, and one because his mother reported that he was familiar with one of the novel objects. All infants were recruited through a database, and all received a small gift for their participation. Written consent was obtained from all of the legal caregivers of all participants.

Experimental setup and procedure

Infants were seated on the laps of their caregivers approximately 50cm in front of a Tobii 1750 remote eye tracker, equipped with an infant add-on. The eye tracker records gaze data at 50hz and has an average accuracy of 0.5° visual angle and a spatial resolution of 0.25° visual angle. Stimuli were presented on a 17- inch flat screen monitor. The visual area was 1,280 x 1,024 pixels, and extended over the entire area of the screen. Infants' eye-gaze was calibrated using a nine-point calibration. If fewer than seven points were calibrated successfully, the calibration was repeated. Infants' behavior during the experiment was also recorded with a digital video camera mounted on a tripod below the monitor.

Table 4.1. Spoken words, manual gestures, and novel objects used in this study

Novel words	Novel manual gestures	Novel objects
/ʃIm/	 ASL YES	 A
/ni:p/	 ASL NO	 B

Note: Novel objects were counterbalanced so that for half of the infants, /ʃIm/ and ASL YES were paired with Object A, and for the other half, /ʃIm/ and ASL YES were paired with Object B

Stimuli

Stimuli were recorded with a Canon HV 30 camera and edited with Adobe Premiere Pro CS4. The video clips consisted of a training phase and a test phase. During the training phase, an actress sat behind a table with two tall, narrow shelves standing on either side of the table: one to the left and one to the right. Each shelf contained two compartments, within which the objects could sit (See Figure 4.1). The experimenter gazed into the camera, and, after a brief greeting, she proceeded to teach two novel labels for two novel objects, as if the camera was an infant. Each time the actress labeled the objects, she produced both an arbitrary word and an arbitrary manual gesture, presented simultaneously. The words were both one syllable CVC words and had no phonemes in common: /ʃIm/ and /ni:p/. The gestures were meant to be within the motor repertoire of a typically developing 15-month-old infant and were based on real contrastive signs in a natural signed language. They were thus easily discernable from one another (roughly YES and NO in American Sign Language). Table 4.1 provides a summary of the spoken words, manual gestures, and novel objects used in the study. The experiment was conducted by adhering to the guidelines for good scientific practice by the Max-Planck-Society.

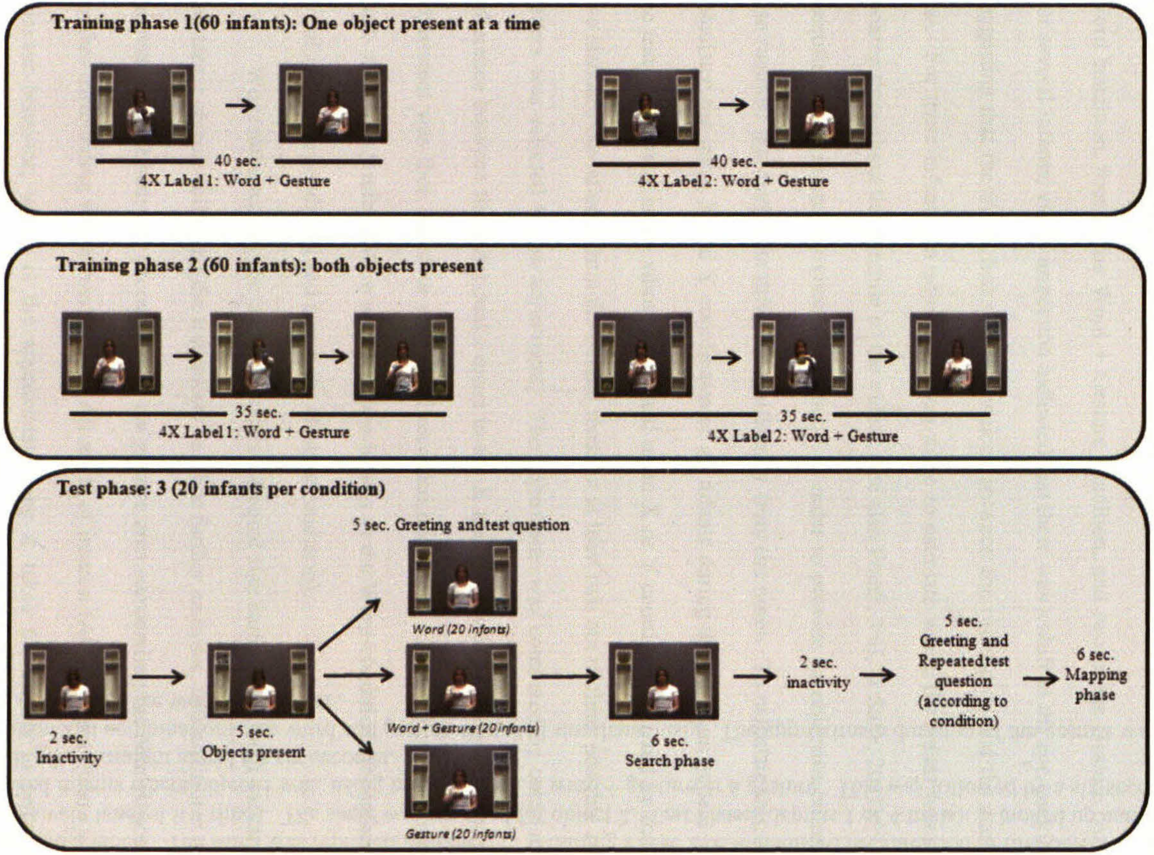
In the training phase, infants were exposed to a total of 16 word + gesture labels, eight for one object and eight for the other object. The training phase was divided into two parts, each of which presented infants with four word + gesture object labels for one object and four word + gesture labels for the other object. In the first part, only one novel object was visible to infants at a time. Using infant directed speech, the actress labeled the first object four times, using both the spoken label and the gesture simultaneously. The labels were embedded in familiar naming contexts in order to facilitate learning (Namy & Waxman, 2000). During the training, the actress ostensively drew attention to the objects by shifting her gaze between the object and the camera. She said:

"Look!" (gaze to object and back), "A [word plus gesture]! Look here," (gaze to object and back to camera), "A [word plus gesture]" (gaze to object and back to infant). "This is a [word plus gesture], Wow!" (gaze to object and back to infant), "A [word plus gesture]." The actress then placed the first object out of sight and presented the second object. Using the same script, she proceeded to teach the label for the second object using the other word and gesture.

The second part of the training was meant to familiarize infants further with the objects' multimodal labels and to introduce them to the task that would later be used in the test phase. Each of the two novel objects sat in a separate compartment and the actress stated that she wanted to find one of the objects saying, "Hmm? Where is the [word plus gesture]?" She then leaned over and gazed into the compartments to search for the object. When she found the appropriate object, she emoted positively, and said: "There it is!" She then took the object out from the compartment, showed it to the camera, and again, using ostensive language and alternating gaze between the camera and the object, labeled it an additional three times with the word and accompanying gesture. She said: "The [word plus gesture]! Look! The [word plus gesture]. Wow, the [word plus gesture]." The same procedure was repeated for the second object, using the other word and sign. In total, infants heard and saw the multimodal object labels eight times for each object during the training phase.

Immediately following the training phase, infants viewed the test phase. Each trial began with the actress looking downward. After 1 second, the two objects each appeared in a separate compartment. Using film editing software, objects were superimposed so that they appeared as if they had actually been present. After 5 seconds, the actress raised her head, gazed into the camera as if addressing an infant and said: "Hello, where is the [target label]?" In the Word condition, the target label was the spoken word. In the Word + Gesture condition, the target label was the spoken word along with the gesture, and in the Gesture condition, the target label was only the gesture. Following the question, the actress continued to gaze directly into the camera for 6 seconds, which served as the search phase, in which infants were expected to locate the target object. After the 6 seconds had elapsed, the same question was repeated. After a further 6 seconds, an attention getter was displayed in the center of a black screen to bring infants' attention to the middle of the screen (approximately two seconds). Then the second trial started, within which the other label served as the target. In total, there were four trials, each consisting of two questions with the same label thus totaling eight questions, with the target label alternating between the trials, and the order being counterbalanced across participants.

We counterbalanced the positions of the objects based on the two following criteria: First, over the course of all trials, each object appeared in each compartment once. Second, the target object never appeared in the same box for two consecutive trials. All infants received a fixed trial order whereby the positions of the target- and non-target objects across trials were held constant for all infants. We also counterbalanced which object was last seen before the test phase as well as to which object each set of labels was paired. Infants received no feedback with regard to whether they had made the correct mappings.



Training phase 1(60 infants): One object present at a time
40 sec.
4X Label 1: Word + Gesture
40 sec.
4X Label 2: Word + Gesture
Training phase 2 (60 infants): both objects present
35 sec.
4X Label 1: Word + Gesture
35 sec.
4X Label 2: Word + Gesture
Test phase: 3 (20 infants per condition)
2 sec.
Inactivity
5 sec.
Objects present
5 sec. Greeting and test question
Word (20 infants)
Word + Gesture (20 infants)
Gesture (20 infants)
6 sec.
Search phase
2 sec.
inactivity
5 sec.
Greeting and
Repeated test
question
(according to
condition)
6 sec.
Mapping
phase

Figure 4.1 (above) Schematic overview of the procedure. Training Phase 1: E ostensively drew attention to object 1 and then labeled it four times using multimodal labels. The same was repeated for object 2. Training Phase 2: E announced her intention to find object 1. She then found it and ostensively labeled it 4 times. The same was repeated for object 2. Test Phase (depicts 1 of 4 trials): E looked up and, according to the condition, asked infants where object 1 was, using either a word, a word + gesture or a gesture. This was followed by a six second test phase, in which E gazed straight ahead for six seconds.
Note: For all multimodal combinations, the word and gesture occurred simultaneously. The approximate duration of the gesture was 1sec. and the approximate duration of the word was 500ms.

Analysis and data reduction

Infants who had less than 50% valid gaze points recorded during the training were excluded from analysis. In total, this resulted in the exclusion of seven infants (one in the Word condition, four in the Word + Gesture condition, and two in the Gesture condition). For several infants, visual inspection indicated that there was a shift in the eye tracking data, suggesting that the infant had shifted positions between the calibration and the start of the test. For these infants, an adjustment was made to estimate where the infant was actually looking throughout the duration of the videos (see also Frank, Vul, & Saxe, 2011). We first identified these infants by visual inspection. In order to provide an objective criterion, we also required that their data differed significantly from the norm. To calculate this, we first calculated the mean X and Y coordinates of all infants during the attention getters between the test trials. Any infant whose individual mean X or Y coordinate differed by more than two standard deviations from the overall mean in at least two out of three possible attention getters was selected for an adjustment. The adjustment was computed by calculating the difference between the individual's mean in the X and Y axes and the overall mean. This adjustment was then added to each valid recorded coordinate for that particular infant. In total, data from 5 infants were adjusted (two infants in the Word condition, one infant in the Word + Gesture condition, and two in the Gesture condition).

We created rectangular areas of interest of equal size surrounding both the target and non-target objects following the test questions. For further analyses, two additional areas of interest were created: one encompassing the gesture area surrounding the actress' torso, and another surrounding the actress in full. All areas of interest (AOIs) began 200ms after the relevant behavior, which is the approximate time it takes for infants to program eye movements (Canfield, Smith, Brezsnyak, & Snow, 1997). Our primary dependent measures were (1) the proportion of gaze points to the target object relative to the non-target object, and (2) whether infants' first look was to the target or the non-target object. We used Analyses of Variance to compare performance between conditions, and one-tailed, one-sample t-tests to test whether infants performed above chance.

4.3. Results

Figure 4.2 shows the mean proportion of gaze points that infants looked to the target object, relative to the non-target object during the search period following the test questions. A one-way ANOVA on the mean proportion of gaze points to the target object relative to the non-target object revealed a significant difference between conditions, $F(2,57) = 7.83$, $p = .001$, $p\ \eta^2 = .22$. Post hoc tests (two-tailed) using Least Significant Difference comparisons showed that infants in the Word condition ($M = 1.8$ms.; $SD = 1.1$) looked significantly longer to the target object than infants in both the Word + Gesture condition (Mean difference = -.16, $p = .003$; $d = .45$; $M = 1.7$ms, $SD = 1.9$) and in the Gesture condition (Mean difference = -.18, $p = .001$, $d = .64$; $M = 1.1$ms., $SD = 1.5$). Infants in the Word + Gesture and Gesture conditions did not differ significantly from one other (Mean difference = .03, $p = .58$, ns.). One-sample t-tests revealed that only infants in the Word condition looked significantly longer to the target object than would be expected by chance (Word condition: $t\ (19) = 3.37$, $p = .002$; Word + Gesture condition: $t(19)$.= -1.03, $p = .841$; Gesture condition: $t(19) = -2.33$, $p = .985$; all one-tailed probabilities for the upper tails).

Figure 4.3 shows the mean proportion of infants' first looks to the target object relative to the non-target object across the eight test questions. A one-way ANOVA did not reveal significant differences between conditions ($F(2,56)= 2.098$, $p = .13$), however the overall pattern followed that of the looking time measure. Infants in the Word condition looked significantly more often first to the target object than would be expected by chance ($t(19) = 2.008$, $p = .030$), while infants in the Word + Gesture condition and Gesture condition did not (respectively, $t(19) = .333$, $p = .372$; $t(19) = .920$, $p = .816$ one-tailed probabilities for the upper tails).

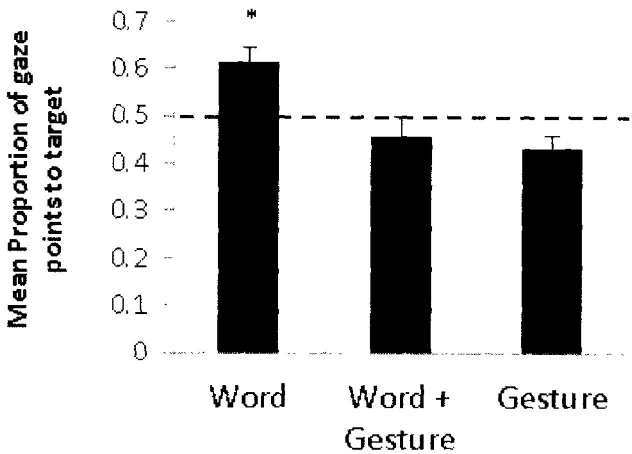


Figure 4.2. Mean proportion of gaze points to the target object relative to the non-target object; Note: Asterisk indicates significant performance above chance, $p < .05$

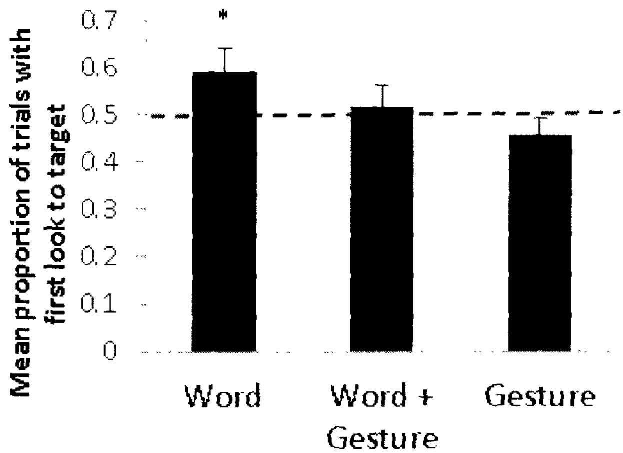


Figure 4.3. Proportion of trials with the first look to the target object; Note: Asterisk indicates significant performance above chance, $p < .05$

Finally, we analyzed infants' performance on the first question of the first trial since it avoids any possible biases stemming from switching labels, objects moving positions, repeated questioning, or fatigue. A one-way ANOVA on the mean proportion of gaze points to the target object relative to the non-target object on the first question of the first trial revealed a marginally significant difference between conditions, $F(2, 42) = 2.73$ $p = .077$ (See Figure 4.4). One-tailed *t*-tests again confirmed that only infants in the Word condition performed above chance (Word: $t(16) = 2.698$, $p = .008$; Word + Gesture condition: $t(15) = 1.582$, $p = .068$; Gesture $t(12) = -.869$, $p = .799$; all one-tailed probabilities for the upper tail).

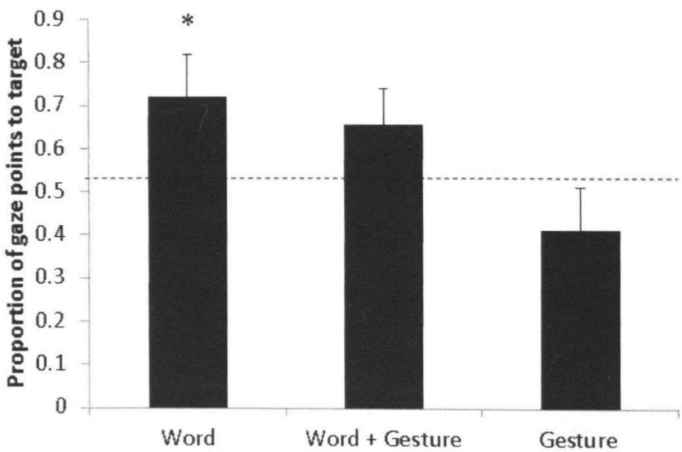


Figure 4.4. Mean proportions of gaze points to target object relative to the non-target object on Trial 1, Question 1; Note: Asterisk indicates significant performance above chance, $p < .05$; Some infants did not look to the target or the non-target on this question. Thus, the numbers of infants varies according to condition (word: $n = 17$ Word; Word+Gesture: $n = 16$; Gesture: $n = 12$)

To look at infants' individual performance, we computed binomial tests. These compared the number of infants who first looked to the target with the number of infants who first looked to the non-target on the first question of the first trial. Only in the Word condition, was there a significant difference, indicating that only infants in the Word condition had made the correct mappings (13 of 17 infants; p = .049, two-tailed; see Figure 4.5).

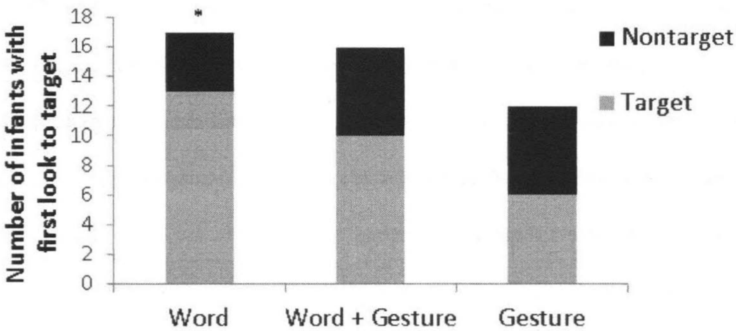


Figure 4.5. Number of infants with a first look to the target on the first question of Trial 1; Note: Asterisk indicates significant difference from chance, $p < .05$

Additional analyses

It is possible that infants did not make the correct mappings in the Gesture condition because they did not attend to the gestures, either in the training phase and/or during the test phase. During the training phase, infants focused their attention on an average of 11.2 of 16 possible gestures (70%), with no significant difference between the gestures (5.55 for ASL YES and 5.65 for gesture ASL NO; $F(1, 57) = 1.56$, $p = .21$), and no significant difference between conditions, $F(2, 57) = 1.56$, $p = .21$. There were, however, no significant correlations between how many signs infants attended to in the training phase and their performance in the test phase. When analyzing each condition separately, the correlations remained non-significant, Word: $r = .108$, $p = .650$; Gesture: $r = .167$, $p = .482$; Word + Gesture $r = -.043$, $p = .856$.

During the test phase, infants in the Gesture condition looked to the gesture area for almost all of the questions ($M = 7.67$ questions; $SD = .594$). A significant one-way ANOVA on the mean number of test questions in which infants attended the sign space ($F(2,52) = 16.65, p < .001, p\eta^2 = .39$) revealed that infants in the Gesture condition looked to the gesture area in significantly more test questions than infants in the Word condition ($M = 4.58$, $SD = 2.17$; Mean difference = 3.09, $p < .001$) and marginally more than those in the Word + Gesture condition ($M = 6.61$, $SD = 1.75$; Mean Difference = 1.06, $p = .062$). Infants in the Word + Gesture condition also looked at the gesture area in significantly more trials than infants in the Word condition (Mean difference = 2.03, $p <.001$). Thus, the gestures clearly elicited infants' attention in the test phase.

A further analysis also confirmed that infants, in the test phase, attended more to the gesture area in the conditions with gestures than in the Word condition. A significant one-way ANOVA on the mean proportion of gaze points per question to the gesture area ($F(2,52) = 17.83, p <.001, p\eta^2 = .406$) relative to the total amount of gaze points revealed that infants looked to the gesture area significantly more in the Gesture condition ($M = .23$ $SD = .12$), than in both the Word + Gesture condition (where the face and the gesture were competing for visual attention; $M = .14$ $SD = .11$; Mean difference = .09, $p = .007$), and the Word condition (where there was no gesture; $M = .04$ $M = .05$; Mean difference = .19, $p < .001$), and significantly more in the Word + Gesture condition than in the Word condition (Mean difference = .10 $p = .003$). Taken together, the additional analyses indicate that (1) in the training phase, infants in the Gesture condition attended to fewer gestures, while they had presumably heard all the auditory labels, and (2) in the test phase, infants allocated more attention to the gesture area in the two conditions where a gesture was present.

4.4. Discussion

In the current study, infants were taught multimodal labels for objects. Results from the test phase showed that infants were able to map the verbal label to the referent. However, when the verbal label was accompanied by a representational gesture, or when the representational gesture was presented alone, infants did not succeed in identifying the correct referents. These findings suggest that when mapping labels to objects, infants initially rely more on the verbal reference of multimodal utterances than on representational gestures, and further, that accompanying representational gestures may interfere with making the correct mappings.

Why did infants in the Gesture condition fail to correctly identify the referents at test? It is unlikely that infants lack the general capacity to integrate information from two modalities. For example, 12-month-olds appropriately process multimodal speech-gesture utterances that include *deictic* gestures (Gliga & Csibra, 2009). Further, when deictic gestures and spoken words are put in competition with one another, infants sometimes even prefer gestures over spoken words as referential cues (Grassman & Tomasello, 2010). One possible explanation for why infants failed in the Gesture condition is that infants do not sufficiently attend to *representational* gestures in multimodal utterances. Indeed, results from our additional analyses showed that infants, in the training phase, attended to only 70% of the gestures. Representational gestures, by the very nature of the visual modality, require that infants divide their visual attention between the referent and the gesture. This, of course, is not the case for spoken words as infants can visually attend to a referent and hear its spoken label simultaneously. Neither is this the case for deictic gestures since they direct attention away from the gesture, and to the referent, thus facilitating the mapping of an auditory label to the attended object. In the current paradigm, infants' visual attention was explicitly directed at the referents with ostensive referential gaze and showing gestures. It is likely that these deictic cues directed infants' attention to the novel objects and thus away from the gestural labels. Since the labels were already provided in the auditory modality, the gestural labels were presumably irrelevant for establishing a mapping.

However, infants did not simply ignore the gestures altogether. First, if infants had ignored the gestures altogether, then they should have performed equally well in both conditions where a word was present at test. This was not the case. To the contrary, infants performed worse when a spoken word was paired with a gesture. Second, our additional analyses show that during the test phase, infants in the Gesture condition indeed attended to the gestures, yet still failed to identify the correct referents. These findings suggest that during the test phase, the gestures actually interfered with infants' ability to identify the referents. Because there were no deictic cues to guide infants' visual attention, it is likely that infants paid more attention to the gestures in the test phase, which ultimately impeded their ability to map the words to the objects. Typically, word learning entails a two-way association between a spoken word and an object. However, when representational gestures are paired with spoken words, as was the case in this study, there is a three-way association since both a spoken word and a representational gesture are mapped onto the object. The three-way association (word-gesture-object) proved to be more difficult for infants to map than the two-way association (word-object). It is possible that older infants are able to form these associations since they have more experience with spoken language. This would explain why the 26-month olds in Wilbourn & Sims (in press) made the mappings but the 15-month-olds in the current study did not.

The claim that infants younger than 18 months have no preference for either modality (Namy, 2001; Namy & Waxman, 1998), or that they even prefer the gestural over the verbal modality (Goodwyn et al., 2000), was not substantiated by the current findings. Instead, results clearly support the notion that by 15 months, when confronted with multimodal utterances containing arbitrary representational gestures, infants rely on the verbal instead of the gestural reference in word learning. The age at which infants form this preference is still a matter of debate. With regard to unimodal utterances, some studies have found a preference for spoken labels arising as early as 6 months of age (Fulkerson & Waxman, 2007) while others suggest that it does not develop until later, for example, around infants' second birthdays (Namy, 2001; Namy & Waxman, 1998; Woodward & Hoyne, 1999). Of course, by 15 months, infants learning spoken languages have already had a great deal of experience with spoken words and may even comprehend some commonly used spoken words as young as 6 months (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999). However, it is important to emphasize that typically developing hearing infants – who receive spoken language input – can indeed fast map gestural labels that are unimodal (i.e. gestural labels that are not combined with spoken words; Namy 2001, Namy & Waxman 1998). Similarly, research on the acquisition of signed languages suggests that infants who are exposed to signs from birth (e.g. infants born to signing parents) acquire language at the same rate as infants acquiring spoken languages (Capirci, Montanari, & Volterra, 1998; Folven & Bonvillian 1993; Meier & Newport, 1990; Petitto, 1987; Schick, et al 2005). These studies show that infants can map representational gestures to objects, and it might well be the case that they also have no a priori preference for either modality. However, the current study shows that when the label is presented in two modalities simultaneously, representational gestures do not facilitate early word learning. This is because the deictic cues necessary for establishing joint reference take visual attention away from the gestural – but not the auditory – label. Representational gestures may even hinder infants' mapping of a spoken word to its referent when deictic cues are lacking, since representational gestures take attention away from the referent.

A wide range of programs currently promote the use of multimodal speech with young infants. Although the proclaimed benefits of these programs are extensive, empirical evidence supporting the claims is lacking (see Johnston et al., 2005 for a meta-analysis). Results from the current study do not support the claim that representational gestures accompanying spoken words facilitate early word learning. It is, of course, still possible that early, increased exposure to multimodal labels could influence infants' attentional skills and their processing of the more complex three-way associations of multimodal utterances. More research needs to be carried out which investigates how much exposure is actually necessary for infants to form expectations about word modality, and when the cognitive requirements emerge that enable infants to use spoken words in combination with representational gestures as object labels.

Another important consideration has to do with the nature of the signs themselves. In the current study, we used arbitrary spoken words and manual gestures, which bear no physical resemblance to the actual objects. Research shows that multimodal utterances containing iconic gestures do facilitate word learning in toddlers and adult second language learners (Kelly, McDevit, & Esch, 2009; Marentette & Nicoladis, 2011). However, infants younger than 26 months routinely fail to recognize gestural iconicity, and presumably would not benefit from multimodal utterances including iconic gestures (Namy, 2008; Namy, et al., 2004; Tolar, Lederberg, Gokhale, & Tomasello, 2008).

Findings from this study show that 15- month-old infants make use of verbal but not gestural references in multimodal utterances containing representational gestures. These findings question whether arbitrary multimodal labels facilitate early word learning at 15 months. It remains speculative whether a facilitative effect of representational gesture-speech combinations on word learning would occur with increased earlier exposure to multimodal representational labels (e.g., through 'baby signing'). More likely, the multimodal combinatory use of representational labels places too much cognitive demand on young infants since it requires the formation of multiple associations. If so, representational gesture-speech combinations are likely not facilitative of early language acquisition. A host of research has shown that gestures play a leading role in the acquisition of language (e.g., Iverson & Goldin-Meadow, 2005), however, few studies have explicitly discussed the fact that these facilitative effects are initially attributed primarily to deictic gestures. Research shows that the iconicity of representational gestures is not realized until around 26 months (Namy, 2008), and suggests that infants' use of iconic gestures is mediated by their vocabulary size (Nicoladis, Mayberry, & Genesee, 1999). Such findings reinforce the argument that representational gestures become relevant only after representational language is acquired (see Liszkowski, 2010). Thus, while deictic gestures and deictic gesture-speech combinations pave the way for language (Iverson & Goldin-Meadow, 2005), representational gestures presumably do not. Notwithstanding the enormous impact of deictic gesture-speech combinations on language acquisition, current findings suggest that representational gestures play a minor role in early multimodal labeling and word learning. Nevertheless, to date, little research has investigated infants' own production of representational gestures. The research presented in Chapter 5 addresses this void by investigating whether, under favorable circumstances, infants will creatively produce representational gestures to communicate with an adult.

Chapter 5

Representational gestures before 24 months?

Abstract

This study investigated Dutch infants' production of representational gestures at 18 and 24 months. In a first study, we asked whether infants would communicatively pantomime actions to inform an ignorant partner about how to operate a toy. In this paradigm, we tested whether infants would use action pantomimes both with and without a tool in their hand. Results showed that even by 24 months, Dutch infants do not use their own bodies to communicatively pantomime actions. Infants at both ages, however, did use action pantomimes when they had an exact replica of the tool in their hands. Infants also frequently used deictic gestures to communicate with the experimenter. At 24 months, however, few of infants' gestures were accompanied by vocalizations, which suggested that they might not have been using them communicatively. Study 2 was designed to test the communicativeness of the pantomimed actions in Study 1.

Study 2 showed that infants continued to pantomime actions even when no experimenter was present, suggesting that even by 24 months, Dutch infants do not use representational gestures to communicate. They instead rely on words and deictic communication. This pattern suggests against theories that propose that representational gestures serve as a medium between deictic communication and spoken language and suggests that deictic - and not representational - gestures give rise to spoken language.

5.1 Study 1: Introduction

Infants have long been known to communicate with manual gestures before they speak. Most experimental research on infants' preverbal gestures has investigated infants' *deictic* gestures, placing particular emphasis on the cognitive skills involved in the use of such gestures and how they relate to infants' developing language. Infants' early deictic communication has been shown to entail complex cognitive underpinnings and has been shown to be fundamentally linked to infants' developing language capacity, with a substantial body of research providing evidence for relations between the production of deictic gestures and infants' developing pragmatic, lexical, and syntactic skills (Bates, Camioni, & Volterra, 1975; Tomasello, Carpenter, & Liszkowski, 2007; Iverson & Goldin-Meadow, 2005)

Much less is known about the use of representational gestures infancy. Unlike deictic gestures, representational gestures entail an understanding of symbolic representation much the same as spoken words. Symbolic representation serves a cornerstone of human language since it enables individuals to communicate about displaced referents which are not located in the here and now (Hockett & Altmann, 1968). Bates (1979) hypothesized that the shift from deictic communication to representational communication served as a "second dawn" in humans' developing capacity for communication. One interesting possibility is that representational gestures, too, are related to infants' developing cognitive skills and language capacity (Zammit & Schafer, 2011). In support of such a hypothesis, several studies have reported that infants spontaneously produce representational gestures as early as 14 months, before language has emerged in earnest (Acredolo & Goodwyn 1988; Iverson, Capirci & Caselli, 1994; McNeil 1992). It is even possible that they serve as a developmental precursor to spoken language (Tomasello, 2008). In a longitudinal study of four children from 10 to 24 months, Camaioni, Aureli, Bellagamba, and Fogel (2003) reported that the frequency of representational gestures in infants' second year progressed through an inverted U-shaped trajectory peaking around 18-19 months and subsiding shortly thereafter, presumably as spoken words become more predominant in infants' communication. The authors interpreted this pattern as evidence that representational gestures provide a developmental bridge from infants' deictic communication (with gestures) to symbolic communication (with spoken words).

Studies also suggest that the early training of representational gestures expedites spoken language acquisition (Goodwyn & Acredolo, & Brown, 2000). While such studies are informative, and indeed, provide a useful first step towards understanding how representational gestures relate to infants' linguistic development, they fall short in providing concrete answers regarding the actual use and understanding of representational gestures in infancy. Nearly all existing studies investigating infants' production of representational gestures have been observational. Based on observation alone, it is nearly impossible to know whether infants were producing representational gestures communicatively and creatively, or whether they were simply reproducing particular actions that they had learned in conventionalized formats. As such, it remains speculative whether representational gestures really serve as a necessary bridge towards spoken language.

Another possibility is that representational gestures emerge along with, or even after language is in place. According to DeLoache's dual representation model, the development of a symbolic capacity occurs gradually across infants' second and third years. Based on DeLoache's model, the capacity to understand and use symbols entails the formation of two mental representations: one of the object itself and another of the actual symbol, as well as a connection between the two (DeLoache, 2004). A growing body of research shows that infants do not possess a dual understanding of symbols until well into their second or third year. For example, with regard to comprehension, children as old as 30 months have troubles differentiating symbols from their referents when the two are perceptually similar to one another (DeLoache, Uttal, & Rosengren, 2003). Only by 26 months, do infants understand that replicas of objects can refer to like objects (Tomasello, Striano, & Rochat, 1999), and only by 36 months, do infants understand that drawings can be used to refer to objects (Callaghan, 2000). Further, infants do not seem to realize the iconicity of gestures until after their second birthdays (Namy 2008; Namy, Campbell, & Tomasello, 2004). Some researchers have taken infants' pretense acts as indications of representational communication (Shore, O'Connell, & Bates, 1984). Rakoczy, Striano, and Tomasello (2005) suggest that early acts of pretense arise out of social engagements where infants imitate others' instrumental actions. However, others have cautioned that these acts are initially imitations of others' actions (Caselli, 1990). One early form of pretense is when infants pantomime actions in play situations. Although these early pantomimes are generally not communicative, it is possible that they act as a precursor to symbolic communication. To my knowledge, no study to date has investigated this possibility. Children's apparent difficulties in comprehending iconic symbols in their second and even into their third year brings into question the age at which they can communicate with representational gestures.

Little is known about the actual circumstances within which children will use representational gestures or if they can use them at all before they have acquired language. In the current study, we investigated the production of deictic and representational gestures at 18 and 24 months. Knowing that infants' in their second year are cooperative and willing to help others (Warneken & Tomasello, 2007, Warneken, Chen, & Tomasello, 2006), and that they will also do so through the use of deictic gestures (Knudsen & Liszkowski, 2012, 2011; Liszkowski, Carpenter, Striano, & Tomasello, 2006), we asked whether they would creatively pantomime a relevant action in order to communicate to an ignorant adult about how to operate a toy. In order to further explore the circumstances within which infants might use representational gestures, we also investigated whether infants were capable of using an exact replica of tool to demonstrate from afar how to operate a toy. If representational gestures bridge the transition from deictic communication to speech, then we predicted that 18-month-old infants would creatively produce such gestures to inform an ignorant adult how to operate a toy. Further, if these gestures were indeed communicative, then we expected them to be accompanied by vocalizations. However, if representational gestures emerge along with or even after speech, then we predicted that infants would resort deictic communication, even in a situation in which representational gestures would be the most effective.

5.2 Study 1: Methods

Participants

Sixteen 18-month olds (mean age = 18;18, range = 18;3 – 18;29; 8 males and 8 females) and sixteen 24-month-olds (mean age = 24;20, range = 24;10 – 24;29, 9 males and 7 females) participated in the study. An additional six 18-month-olds were tested but excluded from analysis because they were inattentive and/or fussy. All participants were recruited from a mailing list of babies born in or around Nijmegen, The Netherlands; all infants were learning Dutch as their first language, and all received a small gift for their participation.

Material

There were 10 toys in total, all of which produced a pleasant effect when operated correctly. The actions afforded by the toys were all simple actions within the motor repertoires of 18- and 24-month-old children. There were six toys which could be operated with a tool (*tool toys*; three involving a horizontal action and three involving a vertical action), three which could be operated with the hand (non-tool toys), and one additional tool-toy used in the familiarization trial. The six tool toys required that a specific tool be used to produce the specific effect associated with that particular toy. The actions for the vertical toys were up/down motions, e.g. a hammer was needed to pound balls down a chute, and the actions for the horizontal tool-toys were side-to-side motions, for example, using a Velcro stick to retrieve balls from a bottle lying on its side. For each of the tool toys, there were two identical tools which could both be used to produce the desired effect (e.g. two hammers and two Velcro sticks). The three non-tool toys could all be operated with the use of a particular manual action, for example, shaking a rattle to produce a sound.

Setup

Infants were seated in an infant chair aside an L-shaped table while their caregivers sat directly behind them in another chair. Infants who refused to sit in the infant chair were permitted to sit on their caregiver's lap. In these cases, caregivers were explicitly instructed to make sure that infants remained on their lap throughout the experiment. Two additional chairs (one for E1 and another for E2) sat across the table, opposite the infant, and a stool sat to the right of the table. A framed plexi-glass barrier sat on the table between the infant and E2's chair. To the left of the table was a room partition, behind which, the stimuli were hidden.

Procedure

After a brief warm up period, infants, accompanied by their parents, were led to the experiment room. The entire study consisted of one demonstration trial, six tool trials, and three non-tool trials. Tool trials were always preceded by a warm up trial in which the infant witnessed E1 explain to E2 how to operate a toy, using both language and representational gestures. The procedure consisted of a familiarization phase and a test phase. A stepwise depiction of the procedure is provided in Figure 5.1. The familiarization phase was meant to teach the infants the actions associated with each particular toy. In the familiarization phase, E1 retrieved a toy from behind the room partition, sat down, set the toy down across from the infant and introduced it to the infant, saying "Look what I have!" Meanwhile, E2 sat quietly behind the partition, out of sight from both E1, and the infant. E1 ostensively demonstrated how to use the toy three times. She said, "Look here. You do it like this," and then ostensively performed the intended action for the infant. After E1 had demonstrated how to use the toy, she moved the toy towards the infant and allowed her to also play with the toy. E1 helped the infant operate they toy until she was able to produce the desired effect using the same action as had been demonstrated. After the infant had demonstrated that she was able to reliably reproduce the intended action on the toy, E1 took the two tools and the toy and invited E2 to come play with the toy. She set the toy behind the plexiglass barrier, and E2 emerged from behind the room partition. E1 then gave E2 one of the tools, retained one for herself, and said, "Look what a nice toy we have." and sat down in the stool next to the table facing the wall, disengaged from the play.

The test phase consisted of two phases. The goal of the Phase 1 of the test phase was to test whether infants would use representational gestures to inform an ignorant adult (E2) how to operate the toy. E2 first greeted the infant and then sat down behind the plexiglass barrier, in front of the toy. E2 expressed his interest in the toy and his desire to play with it while making it clear that he did not know how to operate it. He said to the infant, "Wow, this is a great toy, but I don't quite know how it works. Hmm, maybe you can help me?" E2 repeatedly shifted gaze between the infant and the toy. In addition, his language was always accompanied by searching behavior. For example, E2 would raise his arms at his side with his palms up, and have a perplexed looks on his face, which conveyed to the infant that he did not know how to operate the toy.

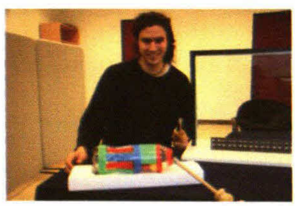
After 30 seconds, Phase 2 began. In phase 2, we were interested in whether infants would use an exact replica of a tool to demonstrate the use of a toy. E1 turned around and offered the replica tool to the infant, saying, "Maybe you can show him how it works." Using facial expressions and searching behavior, E2 continued for 20 seconds, to convey his ignorance about how the toy was operated. If, throughout the entire test phase, the infant pantomimed the appropriate action, E2 tried to elicit the gesture again by asking the infant to repeat the action, saying, "Sorry, once again, how does it work?' If the infant produced the gesture a second time, E2 would thank the infant and perform the action himself, saying "Oh! Now, I see. That's how it works! Thank you." E2 always ended the test phase by inviting the infant to play with the toy for approximately one minute, regardless of whether the infant had gestured in the trial. To end the trial, E2 took the toy and returned behind the room partition. Meanwhile, E1 would retrieve another toy from behind the partition to begin a new trial. In total, there were six trials involving tool toys, each of which included of a different toy (See Table 5.1).

1. E1 teaches infant how to operate the toy

2. E1 sets toy behind window, invites E2 to play with toy, and gives him the tool

3. E2 does not know how to operate toy

4. E1 gives infant an exact replica of the tool



5. E2 allows infant to play with toy

Figure 5.1. Stepwise depiction of the procedure used on the six trials involving tool toys (photographs depict experiment from infants' perspective)

Representational gestures

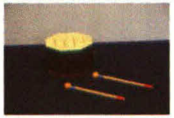
The procedure for the non-tool toys was the same as that of the tool toys, except that, E1 left the scene immediately after having invited E2 to see at the toy. These trials were designed to further test whether infants could pantomime the use of a toy using their own bodies. For these trials a different set of three toys was used (see Table 5.1). The presentation order of the tool- toys and the non-tool toys was blocked. Within the tool-toys, we also blocked the presentation order of the horizontal toys and the vertical toys so that the three toys with horizontal actions always followed one another, as did the three trials with vertical actions. Finally, we counterbalanced the order of presentation of the horizontal, vertical, and non-tool toys.

Coding and data analysis:

Our main question was whether children would communicatively pantomime actions to inform an ignorant partner about how to operate a toy. During the test phase, we coded action pantomimes with a tool in hand and action pantomimes without a tool in hand. An action pantomime without a tool was coded when an infant used his or her body to depict the appropriate use of a tool. For example, an infant might create a fist and shake his hands in the air to depict the use of a rattle. An action pantomime with a tool was coded when an infant reproduced an action which was appropriate for a particular toy while holding a tool in his or her hand. For example, an infant could pick up her hammer and make banging actions in the air. All pantomimed actions had to incorporate the same manual action as was previously demonstrated for that particular toy in the familiarization phase. In addition, in order to check whether infants were generally willing to communicate, we also coded deictic gestures, which were directed towards E2. Deictic gestures included points, shows, offers, and communicative reaches. Gestures and actions which were obviously intended for infants' mothers were excluded from analysis (for example, when an infant turned around and looked at the mother while gesturing). Finally, we coded whether infants' deictic gestures and pantomimed actions were accompanied by vocalizations. A gesture or action was coded as being accompanying by a vocalization when it overlapped temporally with the vocalization.

Reliability

One third of the videos were re-coded by a second trained experimenter to assess coder reliability on pantomimed actions with and without tools, and deictic gestures. Inter rater reliability was nearly perfect in all three categories. Coders agreed on 99% of pantomimed actions without tools (Kappa .85), 92% of pantomimed actions with tools (Kappa .81), and on 91% of deictic gestures (Kappa = .81).

Tool toys (Study 1 and Study 2)		
	Description	Picture
1b.	Drum	
2b.	Hammer used to pound balls down a chute	
3b.	Stick with Velcro used to fish balls from a bottle lying vertically	
4b.	Stick with Velcro used to fish small toys out from a jar	
5b.	Magnet tool used to remove toy cars from covered area	
6b.	Ridged stick used to make trilling sound inside hole in box	

Non-tool toys (Study 1)		
	Description	Picture
1a.	Hand used to ring notes of vertical xylophone	
2a.	Rattle	
3a.	Hand used to move beads back and forth	

Table 5.1. Photographs and descriptions of toys used in the study. Note: Only the tool-toys were used in Study 2

5.3 Study 1: Results

Action pantomimes without tools

We first investigated how frequently infants gestured in the non-tool trials. One infant in each age group became fussy before these trials took place, therefore, the analyses focuses on 15 infants per group. Infants rarely gestured in the non-tool trials. The left panel of Figure 5.2 shows the mean proportion of trials within which infants produced action pantomimes without tools and deictic gestures. Only three out of 15 18-month olds (20% mean proportion of trials with gesture = .09.) and two of 15 the 24-month-olds (13%; mean proportion of trials with gesture = .11) produced action pantomimes without tools. Six out of 16 18-month-olds (33%; mean proportion of trials = .15) and 9 out of 15 24-month-olds (60%; mean proportion of trials = .13) produced a deictic gesture. A mixed ANOVA on the mean proportion of trials with a gesture (2 Gesture types: deictic vs. action pantomime with a tool) with Gesture type as within subject factor and Age as a between subject factor revealed a main effect of Gesture type, ($F(4.87) = 4.86$, $p = .04$), and no interaction, indicating that infants used deictic gestures in significantly more trials than they did action pantomimes without tools. Figure 5.2 shows a screenshot of an infant pointing to a toy through the Plexiglas barrier to indicate to E about a toy. In this particular instance, the infant chose to point, even though a representational gesture would have been more appropriate and indeed, more informative.

Figure 5.2. While trying to elicit representational gestures, an infant points at the toy through the Plexiglas barrier, even though a representational gesture would be more informative

We next investigated infants' production of gestures in Phase 1 of the tool-trials, where no tool was available for the infant (see the right panel of Figure 5.3). In these trials, few infants produced action pantomimes, however, many infants produced deictic gestures. Whereas only 4 out of 16 (25%) of the 18 month olds and 4 out of 16 of the 24 month olds (25%) used action pantomimes, 11 of 16 18-month-olds (69%) and 11 of the 24-month-olds (69%) produced deictic gestures. Figure 5.3 displays the mean proportion of trials with deictic gestures versus action pantomimes without tools by infants at both age groups. A mixed factor ANOVA on the mean proportion of trials with a deictic gesture vs. action pantomime (Gesture as a within subjects factor and Age as between subjects factor) revealed a main effect of Gesture and no interaction, $F(1, 30) = 23.44$, $p <.001$, indicating again, that infants produced deictic gestures on significantly more trials than they did action pantomimes ($p < .001$).

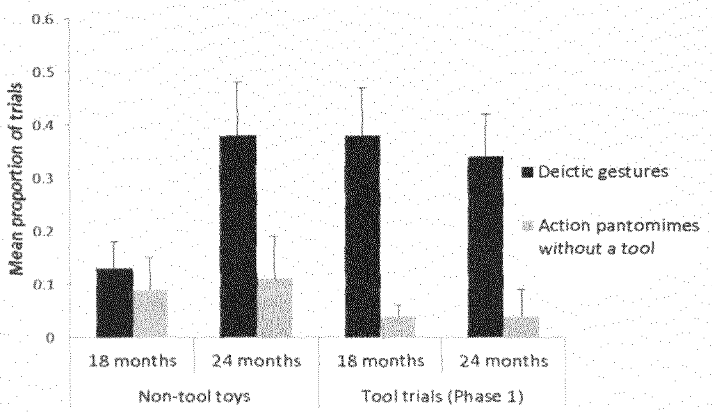


Figure 5.3. Left panel: Infants' use of deictic gestures and action pantomimes in the three non-tool trials; Right panel: Infants' use deictic gestures and action pantomimes in Phase 1 of the tool-trials (when infants did not have a tool available)

Action pantomimes with tools

We next looked into how frequently infants produced action pantomimes with tools. In Phase II, eleven out of 16 18-month-olds (69%) and 14 out of 16 24-month olds (88%) pantomimed the appropriate action when they had an exact replica of the tool available. Within the same trials, 11 of 16 18 month olds (69%) and 11 of the 24 month olds (69%) produced a deictic gesture. Figure 5.4 shows the mean proportion of trials within which infants produced a deictic gesture versus an action pantomime with a tool. A mixed ANOVA (with Age as a between subjects factor and Gesture as a within subject factor) revealed no significant difference in the proportion of trials within which infants produced a deictic gestures versus an action pantomime with a tool.

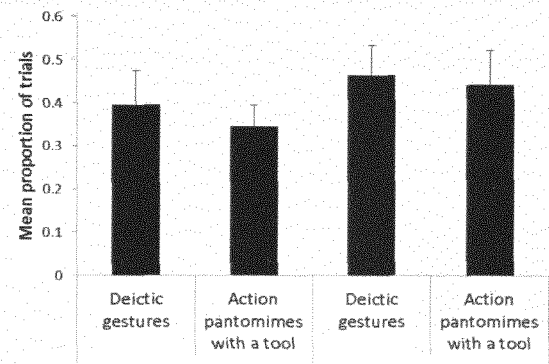


Figure 5.4. Infants' use of deictic gestures and action pantomimes in Phase 2 of the tool-toys (when in infants had a tool available to demonstrate)

The high frequency of action pantomimes with tools produced by infants in both age groups suggests that infants might have been using the tools to demonstrate how to operate the toys. However, based on these analyses alone, it remains unknown whether the actions produced by infants were really demonstrative or whether infants would also produce such actions in the absence of another person. To investigate the degree to which infants' actions and gestures may have been communicative, we performed an additional analysis which looked into how often infants' deictic gestures and action pantomimes with tools were accompanied by vocalizations. If their actions were accompanied by vocalizations, then it is likely that they were using them in a communicative manner (Leung & Rheingold, 1981).

Figure 5.5 shows the mean proportion of deictic gestures and action pantomimes with tools which were accompanied by vocalizations at 18- and 24-months. A mixed ANOVA on the mean proportion of gestures and actions accompanied by vocalizations (with Gesture type: deictic vs. action pantomime, as a within subject factor and Age as a between subject factor) revealed a main effect of Gesture ($F(1,22) = 46.62$, $p = .017$, and an interaction between Gesture and Age (F(1,22) – 10.742, $p = .003$. To break down the interaction, we performed simple effects analyses, which showed that the difference was significant at 18 months ($F(1,22) = 17.11, p < .001$), but not at 24 months, $F(1,22) = .25, p = .623$.

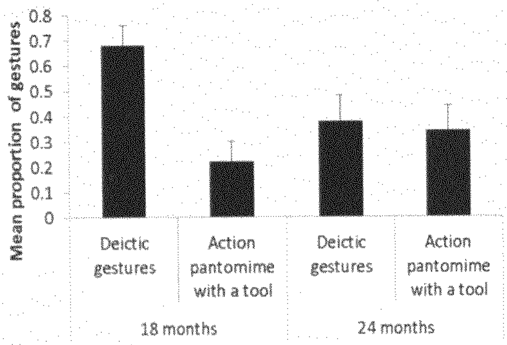


Figure 5.5. Deictic gestures and action pantomimes with a tool accompanied by vocalizations at 18 and 24 months

5.4 Study 1: Discussion

The purpose of Study 1 was to explore, in an experimental setting, infants' creative use of representational gestures at 18- and 24-months. If representational gestures serve as a bridge between deictic gestures and spoken language (Camaioni, et al. 2003), then we expected that infants at both age groups would be able to pantomime actions with their own bodies in order to inform an ignorant adult about how to operate a toy. The current study did not support this claim, as infants at both age groups failed to produce action pantomimes without tools. Nevertheless, we obtained preliminary evidence that infants could communicate with action pantomimes when they had an exact replica of the tool available. This suggests that it might be easier for infants to use tool replicas to demonstrate action schemas than to symbolically represent them using their own bodies.

However, even by 24 months, only about a third of infants' action pantomimes were accompanied by vocalizations. Even the deictic gestures of the 24-month-old infants were rarely accompanied by vocalizations. The lack of communicative cues accompanying their gestures brings into question whether vocalizations really provide an accurate measure of how communicative their gestures were. A better way to investigate the communicativeness of these gestures would be to examine whether infants would continue to produce such gestures in the absence of an experimenter. Based on the results of Study 1, it remains unknown whether infants were actually communicating with action pantomimes. Another possibility is that infants were pantomiming the actions for their own selves. In order tease apart these two possibilities a follow-up study was designed.

5.5 Study 2: Introduction

Study 2 was designed to test further the extent to which the action pantomimes of infants in Study 1 were communicative. To test this, I invited a second group of 18- and 24-month olds to participate in a study which omitted the second experimenter from the task. We know from previous studies that deictic gestures tend to decrease when there is no partner present (Franco, Perucchini, & March, 2009). If the gestures from the infants in Study 1 were communicative, then I expect that infants would produce fewer deictic gestures and action pantomimes when no experimenter was present. However, if the actions from infants in Study 1 were not communicative, then I expected that they would produce fewer deictic gestures, but equally as many action pantomimes.

5.6 Study 2: Methods

Participants

Sixteen 18-month olds (mean age = 18;16, range = 18;6 – 18;25; 8 males & 8 females) and 16 24-month olds (mean age = 24;15, range = 24;0 – 24; 24;27; 7 males & 9 females) participated in Study 2. Two additional infants were tested but excluded from analyses because they were inattentive or fussy (one infant in each age group). All participants were recruited from a mailing list of babies born in or around Nijmegen, a mid-sized city in The Netherlands. All infants were learning Dutch as their first language and all received a small gift for their participation.

Setup, stimuli, and procedure

The setup was identical to that in Study 1. The same six tool-toys from Study 1 were also used as stimuli in Study 2. We dropped the non-tool trials and Phase 1 of the tool-toys since infants almost never produced action pantomimes without tools. The procedure was similar to Phase II in Study 1, however, there was no E2 involved. After teaching an infant how to use a toy (see Study 1), E indicated to the infant that he had to leave for a moment but that he would return shortly. As was the case in Study 1, E placed the toy behind the Plexiglas window, gave one tool to the infant and left the scene behind the room partition so that he was no longer in sight. E remained behind the room partition for 20 seconds (the same length as in Study 1) which served as the test period to investigate whether infants would still perform the gestures and action pantomimes in E's absence. Deictic gestures and action pantomimes with a tool were coded based on the same criteria as reported in Study 1. Reliability was computed on deictic gestures (81% agreement; Kappa =.83) and action pantomimes (88% agreement; Kappa = .71), which revealed high to very high agreement in both categories.

5.7 Study 2: Results and discussion

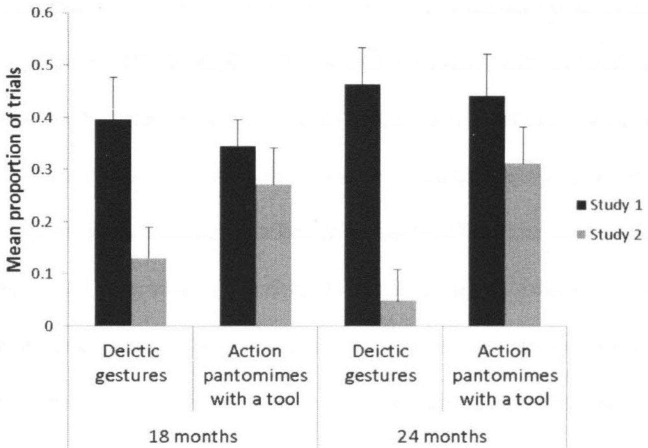


Figure 5.5. Mean proportion of trials within which infants produced deictic gestures and appropriate tool-in-hand-actions.

The gray bars in Figure 5.5 show the mean proportion of trials within which infants at 18 and 24 months produced deictic gestures and action pantomimes with a tool in Study 2, when E was not present. For convenience, the results from Study 1 were duplicated in the same graph (black bars). A mixed ANOVA on the mean proportion of trials with a gesture was computed with Gesture type (Deictic gesture vs. Action pantomime with a tool) as a within-subjects factor with both Age and Study as between subject factors. This revealed a main effect of Gesture ($F(1,60) = 4.15$, $p = .05$), which showed that infants produced, overall, more deictic gestures than action pantomimes, and an interaction between Gesture and Study, $F(1, 60) = 9.24$, $p = .004$. We broke down the interaction using simple effects analyses. These revealed that infants produced significantly fewer deictic gestures in Study 2 than in Study 1, ($F(62,1) = 30.07$, $p < .001$). Infants, however, produced equally as many action pantomimes across studies, $F(62,1) = 2.21$, $p = .142$). These results suggest that the deictic gestures produced by infants at 18 and 24 months were communicative, while their action pantomimes were not.

5.8 General discussion

We investigated whether 18- and 24- month-old infants would creatively and spontaneously produce representational gestures to inform an ignorant adult about how to operate a toy. Infants at both age groups failed to use their own bodies to pantomime relevant actions. Within the same paradigm, however, we found that infants at both ages did pantomime actions when they had an exact tool replica in their hand. It is possible, however, that their action pantomimes were not communicative, since they were rarely accompanied by vocalizations, and, as Study 2 showed, infants in both age groups produced the actions even in the absence of an experimenter. Taken together, the two studies show that, even under favorable circumstances, Dutch infants younger than two years rely on deictic gestures for referential communication. Although 18- and 24-month-old infants did pantomime actions, it is unlikely that did so communicatively. The results suggest against hypotheses of communicative development that advocate representational gestures as a bridge between deictic gestures and spoken language.

Why did infants at both 18- and 24- months fail to produce representational gestures? One possibility is that they were not able to recognize the intentions of the ignorant experimenter. However, infants as young as 12 months have been shown to appropriately and spontaneously use pointing gestures to volunteer information to an ignorant adult who displays relevant behavior (Knudsen & Liszkowski, 2012, 2011; Liszkowski, et al., 2006). Further, E2's behavior was accompanied by infant-directed speech which provided infants with additional cues about his intentions and inability to operate the toys. Although we did not perform systematic analyses on the language that accompanied infants' behaviors, infants' speech was nevertheless useful in showing that they wanted to communicate about how to operate the toys. For example, during the test phase, one 24-month-infant, said "You can drum with this [showing his own drumstick]. Another 24-month-old told E2. "take balls out with this [showing E2 his stick]." Although comments such as these were infrequent, they help to show that infants were capable of understanding that E2, that they did not know how to operate the toy, and that they were willing to provide help. Infants' frequent production of deictic gestures in the same paradigm provides additional evidence that they were motivated to communicate with E2. During Phase 1 of the test phase, some infants even pointed to E1, presumably because they had learned that E1 also had a tool available which could be used to operate the toy. Gestures such as these show that infants were willing to cooperatively volunteer information to the best of their ability, but simply did not know how to use representational gestures to convey the necessary information.

Infants did, however, pantomime appropriate actions when they had an exact replica of the tool in their hand. One intriguing possibility was that infants found it easier to demonstrate how to use an object when they held it in their hands rather than when they had to imagine it and execute the appropriate action schema without anything in their hands. However, as Study 2 showed, infants' continued to pantomime appropriate actions even when the experimenter was not present, suggesting that they were not used in a communicative fashion. If they were not communicative gestures, then what purpose, might these gestures have served? In a dual representation model of symbolic development (DeLoache, 2004), infants must first build mental representations of the objects and their respective referents before they can represent them and communicate with them creatively. One possibility is that infants' action pantomimes were acts of pretense which served to help infants internalize the affordances of the toys. Such acts might help infants formulate mental representations of the objects. Only after infants have construed mental representations and internalized the objects, would they be able to re-produce the action representation in a communicative setting (DeLoache, 2004).

Representational gestures

Interestingly, children have also been shown to think aloud and narrate their own actions with speech, especially in situations where they are confronted with a problem for which they seek a solution (Vygotsky, 1978). It has been argued that so called *private-speech* aids children in solving difficult tasks. It is possible that infants in the present study, who had just learned how to operate several novel toys, were also confronted with a situation which demanded extra cognitive resources. Consequently, it is possible that, by producing private action pantomimes, infants were learning about the affordances of the new toys, even though the toys were not readily accessible. Although admittedly speculative, it is possible that by physically re-presenting the actions for themselves, infants could have been establishing mental representations of the new objects.

If representational gestures do not emerge by 24 months, then when do they emerge and how do they fit in with infants' developing capacity for spoken language? Contrary to the claims of Camaioni, et al. (2003) and Tomasello (2008), it is likely that representational gestures emerge along with or even after spoken language is in place. The current results suggest that infants do not begin using representational gestures until after their second birthdays. They corroborate previous findings which show a spurt in infants' production of representational gestures at 26 months (Ozcaliskan & Goldin-Meadow, 2011). It is likely that this spurt occurs as infants become more proficient with spoken language. One possible explanation for why spoken words are easier than representational gestures has to do with how different types of symbols are related to their referents. As Liszkowski (2010) states, representational gestures are derived from manual actions which are used to act on the world. The use of a representational gesture therefore requires that an infant decouple the actual action from its representation. Spoken words, on the other hand, are, by default, mental tools used to influence the mind. Spoken words, which bear arbitrary relationships with their referents, are possibly understood as mental tools from the very beginning and thus, do not require any decoupling. It is possible that it is more effortful to decouple actions from their representations, which would explain why it is easier to use spoken words than representational gestures. It is also important to note that the creative use of a representational gesture requires that infants invent their own symbols. The learning of spoken words, on the other hand, is scaffolded by caregivers. Infants most likely find it easier to reproduce words that they have learned from interacting with others as opposed to creatively inventing their own.

Taken together, the two studies presented provide evidence that even by 24 months, Dutch infants do not creatively produce representational gestures but instead rely on deictic communication. By 2 years, spoken words are used frequently by infants. The findings suggest against theories which postulate that representational gestures emerge before spoken words. It is more likely that, without sufficient exposure to representational gestures, infants can only creatively produce representational gestures later in ontogeny, i.e. after they have the necessary skills to decouple iconic representations and action depictions from their respective referents. Future research should investigate in more detail, the extent to which this is finding can be generalized to infants learning languages other than Dutch.

Chapter 6

Epilogue

This thesis investigated the comprehension, production, and multimodal use of gestures in infancy. Together, the broad scope of research aimed at furthering our understanding of how infants communicate before language. Chapter 2 investigated gesture as a form of preverbal joint action. In Chapter 3, I looked deeper into the use of deictic gestures on behalf of both caregivers and infants, where I investigated how the type of shared activity structures the communication of caregivers and their infants and how caregivers integrate particular types of *deictic* gestures with particular language in order to scaffold infants' communicative development. Chapter 4 looked into how infants interpret multimodal utterances containing arbitrary representational gestures and Chapter 5 investigated infants' own creative production of representational gestures and demonstrative actions. As a whole, the studies suggest that referential communication arises from collaborative joint action in the form of deictic gestures, and that representational gestures emerge later, along with language.

6.1 Summary

Based on the notion that language is really a form of joint action (Clark, 1996, 2006), that pointing is fundamentally related to language (Iverson & Goldin-Meadow 2005), and that pointing and joint action emerge on the same time schedule in infancy (Bakeman & Adamson, 1984; Carpenter, Nagell, & Tomasello, 1998), Chapter 2 set out to investigate the degree to which infant pointing also constitutes a preverbal form of joint action. Thirty-nine 12-month-old infants were observed in two activities: one free play task which was based on manual action and manipulation of objects, and another decorated room task which elicited pointing. As predicted, I found a relation between manual joint action and pointing. In particular, I found that the amount of time that infants spent jointly engaged in the free play task correlated positively with the frequency that infants shared reference via pointing in the decorated room. Although the direction of the correlation could be fully established based on this finding alone, additional analyses suggested that nearly all infants in the sample could engage in joint action, while not all infants could point, suggesting a developmental primacy of joint action over referential communication. Based on these results, it is likely that referential communication arises out of experience interacting with others. In Chapter 2, I also reported on a turn taking structure which sometimes characterized the pointing gestures of infants and their caregivers: Both caregivers and infants pointed sequentially and in temporally proximity with each others' points. Results from Chapter 2 thus provide evidence for a developmental continuum between the turn-taking structure in infants' early non-referential face-to-face communications and that used in language.

Chapter 3 reported on a study that looked into how caregivers scaffold infant interactions to facilitate their referential understanding. By observing natural caregiver-infant interactions in two different activities, the results from Chapter 3 showed that infants have a variety of cues available to help facilitate the mapping of a word to its intended referent. First, caregivers systematically integrate certain gesture types with certain types of language.

Second, infants are exposed to different types of communication in different types of shared activities. Caregivers' systematic integration of *deictic* gestures with language creates a multimodal scaffold that infants can use to learn about pragmatic and semantic aspects of reference.

The results reported in Chapter 4, however, suggested that the facilitative effects of multimodal gesturing in infants' early language input might be restricted to deictic gestures alone. When confronted with multimodal labels containing arbitrary representational gestures, infants failed to correctly identify the labels' referents, regardless of whether the cue contained a spoken word plus a representational gesture or whether it contained only a representational gesture. These results call into question whether infants really benefit from the combinatory use of representational gestures and spoken language, as, for example is promoted in commercially available programs such as Baby Signs.

The findings in Chapter 5 lend further support to the claim that deictic – and not representational – gestures give rise to language. In a paradigm specifically geared towards infants' spontaneous production of representational gestures, infants at both 18 and 24 months resorted to deictic gestures when they wanted to inform to an ignorant adult about how to operate a toy. Study 1 in Chapter 5 also showed that infants did produce many action pantomimes with tools, suggesting that they might be able to communicate about the use of a tool when they were given an exact replica of the tool needed to operate the toy. Study 2, however, showed that infants continued to produce such actions, even in the absence of an experimenter. Taken together, the findings suggest that infants gradually formulate mental representations across their second year, and suggest against theories of language development which pose a that representational gestures act as a developmental bridge between deictic communication and language. It is more likely that deictic gestures bridge the developmental transition from deictic communication to language since they can be integrated with language in order to achieve joint reference, which is crucial for establishing mental representations and thus, learning words.

6.2 Gestures in ontogeny

Results from Chapter 2 showed that the more likely infants were to engage in triadic face-to-face interactions, the more likely they were to establish joint reference via pointing. This finding provides initial support for the claim that referential communication arises out of experience in joint engagements (Tomasello, 2008). It is possible that these engagements play a crucial role in the emergence of referential communication because they provide caregivers a platform within which they can bring relevant objects into infants' focus of attention. In doing so, infants could begin to establish triadic joint attention which is a fundamental, defining property of referential communication.

One important consideration, however, has to do with the fact that the reported relationship between social interaction and pointing was correlational, and a correlation, of course, does not imply causality. Accordingly, one cannot say definitively whether the amount of social interaction in infants' daily lives predicts the emergence of referential communication. While Chapter 2 lends support to this idea, a longitudinal study would be necessary to examine the exact nature of this relationship. For example, if social interaction modulates infants' gesture use, then one would predict that pointing would emerge earlier according to the amount of social interaction that infants receive in their daily lives. Cross cultural research investigating the emergence of pointing would help test this hypothesis. Based on the notion that different cultures engage differently with their children, a recent study by Salomo & Liszkowski, (in press) showed that this does seem to be the case, whereby infants point more (and presumably earlier) in cultures that include more triadic social interactions in their daily lives. This study lends further support to the claim that social interaction influences the emergence of pointing.

Another limitation of the study in Chapter 2 has to do with the question of whether five-minute lab interactions accurately represent the day-to-day social interactions of infants and their caregivers. While the lab sessions allowed for a more controlled investigation into how gestures relate with social interactions at 12 months, longitudinal or cross-sectional data including at-home recordings of infants' and their caregivers would indeed be useful in establishing a more natural depiction of infants' day-to-day social interactions.

Chapter 2 also showed that infants and caregivers point in a turn-taking fashion. Within these early conversational exchanges, it is possible that infants learn that their actions as well as their communicative gestures can invoke action on behalf of their caregivers, and further, that communication is really another type of joint action (Clark, 1996, 2006). It is interesting, however, that the correlation was specific to infants' joint points, i.e. those that were temporally aligned with their caregivers' points. Again, a longitudinal study investigating the emergence of pointing would provide insight into whether infants' first points are social in nature (Liszkowski & Tomasello, 2011) or whether they are egocentric (Carpendale & Carpendale, 2010). On the basis of the correlative results in Chapter 2, I would hypothesize that pointing is a social gesture from its very emergence and that joint action gives rise to deictic communication, which in turn paves the way for the emergence of spoken language.

Chapter 3 highlighted the fact that infants' communicative input was characterized by an abundance of integrated deictic gesture – language combinations. With the help of activity-specific multimodal communication, caregivers provide infants with a developmental scaffold within which deictic gestures and language input are highly structured. It is likely that caregivers' contextualized and systematic integration of deictic gestures and language facilitates language development in several ways. First, deictic gesture – language combinations help infants build mental connections between words and their referents. Deictic gestures are useful in establishing these connections because they divert attention away from the gesturer and to the respective referent, thus establishing joint attention. In doing so, infants are able to map the sounds they hear to the objects to which they are attending. This is a critical step towards understanding reference since it provides infants with a rudimentary understanding of how words can refer to objects. Second, children learn words best in situations when they are able to interpret an adult's communicative intention (Tomasello, 2003). Infants can use both the type of shared activity as well as co-speech gestures as valuable cues to interpret the communicative intentions of their caregivers, thus enabling infants to realize the intended referent of their caregivers' utterances and facilitating word learning.

One interesting question has to do with the degree to which adult-infant speech differs from adult-adult speech. It is possible, for example, that caregivers are not actually modifying their speech for their infants, but that the type of ongoing activity modulates speech and gesture use. This could be the case regardless of whether one is interacting with an adult or with an infant. If this was indeed the case, then the observed communication of caregivers would not be specific to interactions with infants, but instead a more broad characteristic of human interaction. Without adult data to serve as a comparison, we simply do not know the degree to which the two differ. Although this is an interesting opportunity for future research, it should be mentioned, that infants can presumably still benefit from the systematic pairing of language and gesture regardless of whether this tendency is specific to infant interactions. In sum, caregivers scaffold infants' early referential understanding by integrating deictic gestures with language input. Taken together, Chapters 2 and 3 show that deictic gestures are ubiquitous in the early interactions between caregivers and infants and it is likely that these gestures provide infants with a necessary foundation upon which language can be built.

One overarching question of this research was whether representational gestures too, facilitate infants' developing language capacity. Chapter 4 showed, within a fast mapping paradigm, that infants failed to map arbitrary representational gestures to their intended referents after having been trained on multimodal labels. The finding that they also failed to make the correct mappings when spoken words were accompanied by arbitrary representational gestures suggests that arbitrary representational gestures might have hindered infants' ability to make the correct mappings. Although the studies did not directly address how deictic gestures facilitate word learning, Chapter 3 showed that caregivers often integrate deictic gestures with spoken words. Since joint attention is achieved via deictic gestures, it is likely that they do in fact facilitate word learning. In contrast to deictic gestures, the results from Chapter 4 suggest that arbitrary representational gestures do not promote joint attention, and , consequently, do not facilitate word learning.

One interesting avenue for future research would be to investigate whether infants could map non-arbitrary (i.e. iconic) representational gestures to their referents and if so, whether these combinations would facilitate word learning. Non-arbitrary, iconic gestures are more natural than arbitrary gestures and are likely more prevalent in infants' natural input. As a result, it is feasible that infants might capitalize from the iconicity of such gestures in word learning when they are paired with spoken words. For example, it is possible that infants would find it easier to learn the gesture 'ball' when the hand-shape actually depicts the shape of a ball. The research in Chapter 4 on arbitrary gestures does not address this possibility.

However, with regard to infants' own production of iconic gestures, and corroborating results from previous observational studies (Ozcaliskan & Goldin-Meadow, 2011), the results presented in Chapter 5 suggest that Dutch infants younger than 2 years rarely, if ever, produce representational gestures in a creative and communicative manner. It is possible that earlier reported instances of representational gestures were actually ritualized actions which had been learned from caregivers in social routines. Based on observational evidence alone, it is impossible to determine exactly when a gesture was first produced and within which context it was learned. The experimental study in Chapter 5 avoided this pitfall since all the target gestures had to be appropriate for the particular toys used in each trial. Under these circumstances, infants almost never produced representational gestures. Even in situations which were specifically geared towards the use of such gestures, infants communicated deictically, even though deictic gestures were insufficient to accurately convey the use of the toys.

In one respect, it is not so surprising that the Dutch infants did not creatively produce representational gestures. A major tenet put forth in this thesis is that caregivers scaffold infants' communicative development through the use of multimodal communication. If so, then there is really no need for infants to creatively construct their own means with which to communicate before they have language. Caregivers, who frequently integrate deictic gestures and spoken words, rarely integrate representational gestures with speech when communicating with their infants, especially when compared to spoken words (Acredolo & Goodwyn, 1993). Based on the disproportionate frequency of words versus representational gestures in Dutch infants' input, one would expect spoken words to emerge before representational gestures. Research on the acquisition of signed languages supports this claim. For example, infants exposed to natural signed languages from birth reach linguistic milestones at or around the same time as those learning spoken languages, most likely because they have sufficient exposure to a natural signed language (Capirci, Montanari, & Volterra, 1998; Folven, & Bonvillian 1993; Meier & Newport, 1990; Petitto, 1987; Schick et al., 2005). Pointing also emerges in infancy at around the same rate, regardless of whether infants are learning a spoken or signed language (Petitto, 1987). The common course of development among signed and spoken languages suggests that, regardless of the modality, language emerges in a similar fashion with deictic gestures emerging through social interaction and paving the way for representational communication shortly thereafter.

Taken together, the results of the studies presented in this thesis suggest against hypotheses claiming that representational gestures emerge before spoken language (Acredolo & Goodwyn, 1988; Camaioni, Aureli, Bellagamba, and Fogel, 2003; Tomasello 2008). Infants, by 12 months communicated often with deictic gestures, did not benefit from arbitrary representational gestures in word learning, and failed to communicate creatively with iconic representational gestures, even at 24 months. One hypothesis that is in line with these findings holds that representational gestures emerge along with or even after the capacity for spoken language is in place. First, both representational gestures and spoken words require skills of mental representation. It is possible that these skills emerge gradually over infants' second and third years, as evidenced by the developing understanding of other symbol types such as drawings and pictures (Rochat & Callaghan, 2005), as well as their developing skills of pretend play (Rakoczy, Striano, & Tomasello, 2005). Second, in contrast to spoken words, it is possible that representational gestures emerge from action schemas, possibly from imitating others' relevant actions (Rakoczy, Striano, & Tomasello, 2005; Tomasello, 2008) or from language- specific constructions (Furman, Ozyurek, & Küntay, 2010). As discussed in Chapter 5, the decoupling of an action schema from its referent may be more difficult than imitating an arbitrary spoken word (Liszkowski, 2010). This could even provide an advantage to spoken words over creative representational gestures, resulting in their earlier emergence in hearing children.

One interesting question that has yet to be addressed has to do with whether representational gestures would play a different role in language acquisition in culture other than Germanic languages such as English, Dutch, and German. For example, what role might they play when where representational gestures are more prevalent in infants' input. Furman et al. (2010) have shown that, in contrast with English and Dutch speaking children, Turkish speaking children from 18 months to three years use as many representational gestures as they do deictic gestures. As the authors argue, the high frequency of representational gestures might be explained by the high frequency of caused motion verbs in Turkish. Based on this explanation, infants' use of representational gestures is mediated by their native language. Such findings are relevant because they suggest against hypotheses claiming that representational gestures are cognitively more demanding than spoken words (see for example, Liszkowski, 2010). In combination with Furman et al.'s findings, the results presented in Chapters 4 and 5 suggest that representational gestures do not emerge independently of language, but are instead, strongly linked with spoken words and language structure. Thus, it is possible that the role of representational gestures in language acquisition differs greatly across language communities.

In sum, the studies presented in this thesis support the idea that referential deictic communication arises from experience in joint activities. Deictic gestures are commonly produced by infants before spoken words, and provide a foundation upon which language can be built. In contrast, creative representational gestures emerge later in ontogeny along with or even after language is in place (Furman et al., 2010; Nicoladis et al., 1999; Ozcaliskan & Goldin-Meadow, 2011). Dutch infants younger than 2 years resort to the use of deictic gestures, even in situations where representational gestures would be more appropriate. As such, representational gestures likely come into play during or after language has already been established. In combination with previous studies, the current findings lend support to the idea that while deictic gestures give rise to language, representational gestures do not.

6.3 Gestures in phylogeny

The studies comprising this thesis also provide insights into the phylogenetic emergence of language. Many researchers have hypothesized that human language emerged out of a proto-language involving gestures (Hewes, 1978, 1992; Armstrong & Wilcox, 2007; Tomasello, 2008). Armstrong & Wilcox (2007: 30) state, "The common ancestor of chimpanzees and humans probably had a limited, but highly significant vocal repertoire and a more substantial capacity for communicating involving visible gesture, including iconic and indexic gestures." The results in Chapter 2 suggest that it was *deictic* gestures, in particular, that arose from experience in joint collaborative activities. It is plausible that deictic communication was likely the most effective form of cooperative communication in early face-to-face interactions because it allowed for individuals to establish joint attention with referents that were already located in the immediate environment.

A primary difference between deictic and symbolic communication is that deictic communication is generally restricted to the here and now, while symbolic communication can be used to refer to displaced referents. The most widely accepted hypothesis supporting a gestural origin of language holds that action schemas gave rise to iconic gestures and that iconic representational gestures gave rise to spoken language (Armstrong & Wilcox, 2007). Intuitively, this is a reasonable assumption. Unlike spoken words, iconic gestures bear some perceivable resemblance to their referents and, as proficient symbol users, it is logical to believe that iconic gestures are cognitively simpler, and thus more easily recognizable than spoken words. This could lead to the assumption that iconic gestures are be better candidates than words to refer to displaced referents. An alternative possibility, however, is that representational gestures are cognitively, more demanding than spoken words.

First, research suggests that the non-arbitrary relationship between representational gestures and their referents is more difficult to represent than the seemingly arbitrary relationship between spoken words and their referents (DeLoache, 2004). If this is the case, then it is possible that spoken words are easier to represent than iconic gestures. Second, as the results from Chapter 2 suggest, the ability to refer deictically arises out of collaborative interactions in proximal joint activities. A likely possibility regarding the phylogenetic emergence of language is that, similar to the caregiver-infant interactions, early symbolic communication arose from communicative exchanges where individuals referred deictically to proximal objects. In this case, humans' earliest rudimentary word forms would have been vocalizations that accompanied deictic gestures. Crucially, if this were the case, then these combinations were most likely produced in the presence of the entities to which individuals were referring. Consequently, the need to represent something symbolically would be unnecessary since deictic gestures served to establish joint attention. An apparent benefit of combining deictic gestures with spoken words (as opposed to representational gestures) is that it could have allowed individuals to share reference. This would have enabled individuals to establish concrete connections between the vocalizations and particular entities which were identifiable in the immediate environment. Only after such connections were established, would it have been possible to communicate about displaced referents using the conventionalized spoken word forms. Based on this rationale, the combinatory use of deictic gestures and spoken language could have enabled both infants and our early hominid ancestors to interpret communicative intentions and establish joint attention.

Eventually, frequent vocalizations would have become conventionalized, after which they could be used to refer to entities outside of the proximity of their referents. Although admittedly speculative, this process would support Werner and Kaplan's (1964) theory of symbolic development which holds that associations between words and their referents are initially established locally, and, only after these associations are established, are individuals able to cognitively distance themselves from the referent and communicate about entities which are temporally and physically displaced.

A related issue has to do with the emergence of grammar in human communication. Armstrong and Wilcox (2007) argue that grammar and syntax evolved from combining iconic manual gestures with one another. Another possibility, however, is that grammar and syntax arose via the integration of deictic gestures with spoken language. In ontogeny, there is a relationship between the age at which infants combine deictic gestures with supplementary spoken words and the age at which they combine words with other words (Iverson & Goldin-Meadow, 2005). This shows that the beginnings of syntactic structure could emerge first in deictic communication involving pointing gestures. Again, deictic gestures could be advantageous over representational gestures for early hominids because they would enable individuals to establish joint attention with other individuals, and subsequently communicate with each other about entities in the immediate environment by predicating their pointing gestures with spoken words. The combination of multiple symbolic representations, e.g. two words, likely arose later, after the conventionalized forms and mental representations had been established within a community.

Proponents of gestural origins of language often advocate that representational gestures were used to communicate before spoken language (Hewes, 1978; Armstrong & Wilcox, 2007, Tomasello, 2008). Based on the infant research presented in this thesis, however, I would hypothesize that representational gestures arose after or along with spoken language. One possibility is that representational gestures piggy-backed off of spoken words. Unlike Dutch infants, adults often integrate representational gestures with their speech

(McNeill, 1992; Kendon 2004) and recent research shows that representational co-speech gestures are often language specific: That is, they reflect grammatical structures which are embedded into the grammars of the spoken languages of their speakers (Brown & Gullberg, 2010; Furman et al., 2010; Hickmann, Hendriks, & Gullberg, 2011). This suggests that speech and gesture are a tightly interwoven representational system. Deictic gestures, on the other hand constitute a prelinguistic universal (Liszkowski, Brown, Callaghan, Takada, & De Vos, 2012). Infants point before they use representational communication, regardless of the language they are learning. This would suggest that while deictic gestures emerge universally, and independently of language, representational gestures emerge as a consequence of language, possibly as complementary co-speech gestures. I hypothesize a similar course of development in phylogeny whereby representational communication arose on the heels of deictic gestures.

6.3 Practical implications

The studies comprising this thesis also have practical implications concerning the use and training of representational gestures in infancy. In the past decade, millions of parents have invested substantial amounts of time and money into training their hearing infants to use representational gestures. Some nurseries and daycare centers have even adopted baby signing programs into their own curriculums. Without doubt, the rising popularity of this practice is fueled largely on claims made by the creators of baby signing programs. Importantly, however, there exists very little empirical support in favor of the programs' purported benefits. The studies in Chapters 4 and 5 aimed to contribute towards a better understanding of how representational gestures are used and understood in infancy. The finding from Chapter 4 - that 15-month-old infants did not make the correct mappings between arbitrary representational gestures and their referents after having been trained on multimodal labels– in combination with results from Chapter 5 – that infants do not creatively produce representational gestures before their second birthdays – suggest shat infants might not have a natural predisposition to learn representational gestures without sufficient exposure to a natural signed or spoken language. The fact that infants did make the correct mappings when the labels were decoupled from the representational gestures suggests that by 15 months, infants realize that spoken words are typically used to label objects.

It is also important for parents to consider their own individual goals in teaching representational gestures to their infants. One widespread claim of these programs is that representational gestures emerge before spoken words, enabling better communication between infants and their caregivers. Although infants can presumably learn to use representational gestures when enough training is involved, (for example when they are learning natural signed languages), it is unlikely that have any inherent advantage over spoken words. As discussed in Chapter 5, spoken words are possibly easier for hearing infants to use than representational gestures because they do not require an infant to decouple the mental representation from the action or object itself.

Another point that is rarely mentioned has to do with the actual words that first emerge in infants with and without sign intervention. Baby signing programs usually encourage parents to initially train infants on words which are most commonly used in infancy. Examples include words such as *mama, papa, more, please,* and *thank you.* Typically, these are also some of the first words to emerge in the spoken modality without any sign intervention. The reported benefit of signed words over spoken words could simply be due to increased training in the gestural modality. And, in fact, research has shown that increased input in the spoken modality can also lead to faster acquisition of target words when the words are presented in everyday activities (DeLoache et al., 2010).

Finally, it is important that companies and individuals strive to endorse products which are based on soundly established claims. Over the past decade, baby media has become increasingly popular in the United States and around the world, with an enormous range of products available to parents which are said to benefit babies' cognitive and communicative development. A recent empirical study, however, suggests that these baby media programs do not actually live up to their claims. Deloache et al. (2010) compared word learning in several groups of children, aged 12 to 18 months. A group of infants who had been instructed to watch a commercially available DVD program several times a week for four weeks did not learn words any better than the control group. Interestingly, infants learned words best when they did not watch any videos but whose parents had repeatedly exposed them to several pre-specified target words while engaged in everyday activities. While baby-media products are most-often promoted in good spirit, they sometimes fail to live up to their claims. As DeLoache et al. (2010) showed, natural interactions between caregivers and infants are likely more facilitative to cognitive and communicative development than baby media products. More research is needed to determine definitively whether programs such as Baby Signs truly facilitate language acquisition. Nevertheless, sound empirical research on such products should be conducted and conclusive evidence should be established before, rather than after such products are made readily available to the public.

References

Acredolo, L. P., & Goodwyn, S. W. (1985). Symbolic Gesturing in Language Development. *Human Development, 28*(1), 40–49.

Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development, 59*(2), 450-466.

Agresti, A. (2007). An introduction to categorical data analysis, 2nd ed. Hoboken, New Jersey: Wiley.

Armstrong, D. F., & Wilcox, S. (2007). *The gestural origin of language*. USA, Oxford University Press.

Austin, J. L. (1975). *How to Do Things With Words*. Oxford, Oxford University Press.

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

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*(5), 875-890.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In C. Moore, P. J. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 131–158) Hillsdale, NJ: Erlbaum.

Bates, E. (1979). The emergence of Symbols: Ontogeny and Phylogeny. In: W.A. Collins (Ed.), *Children's language and communication.* Hillsdale, NJ: Lawrence Elbraum.

Bates, E., L. Benigni, I. Bretherton, L. Camaioni, & V. Volterra. (1979). *The emergence of symbols: Cognition and communication in infancy.* New York: Academic Press.

Bates, E., Camaioni, L. & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly, 21*(3), 205-226.

Bateson, M. C. (1975). Mother – infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences, 263*(1), 101-113.

References

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA*.

Blake, J. O'Rourke, P. & Borzellino, G. (1994). Form and function in the development of pointing and reaching gestures. *Infant Behavior and Development 17*(2), 195-203. doi:10.1016/0163-6383(94)90055-8

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, *101*(2), 327-341.

Brown, A., & Gullberg, M. (2010). Changes in encoding of path of motion in a first language during acquisition of a second language. *Cognitive Linguistics*, *21*(2), 263–286. doi:10.1515/COGL.2010.010

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*, 207-220.

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, *2*(01), 1-19.

Bruner, J. (1981). The social context of language acquisition. *Language & Communication*, *1*(2-3), 155-178. doi:10.1016/0271-5309(81)90010-0

Bruner, J. (1983). The acquisition of pragmatic commitments. In R. Golinkoff (Ed.), *The transition from prelinguistic to linguistic communication* (pp. 27-42). Hillsdale NJ: Erlbaum.

Caldognetto, E.M., & Poggi, I. (1995) Creative Iconic Gestures: Some evidence from Aphasics. Raffaele Simone (Ed). *Iconicity in Language.* Amsterdam:.John Benjamins,.

Callaghan, T.C. (2000). Factors affecting children's graphic symbol use in the third year: Language, similarity, and iconicity. *Cognitive Development*, *15*(2), 185–214.

Camaioni, L., Aureli, T., Bellagamba, F., & Fogel, A. (2003). A longitudinal examination of the transition to symbolic communication in the second year of life. *Infant and Child Development, 12*(1), 1-26. doi:10.1002/icd.333

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. L. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development, 62* (2).

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language, 23*(03), 645-673. doi:10.1017/S0305000900008989

Capirci, O., Montanari, S. & Volterra, V. (1998). Gestures, signs, and words in early language development. In J. M. Iverson & S. E. Goldin-Meadow (Eds.), *The Nature and Functions of Gesture in Children's Communication* (pp. 45–59). Rome: Italian National Council of Research, Institute of Psychology.

Carpendale, J. I. M., & Carpendale, A. B. (2010). The Development of Pointing: From Personal Directedness to Interpersonal Direction. *Human Development, 53*(3), 110–126. doi:10.1159/000315168

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science, 1*(2), 380-392. doi:10.1111/j.1756-8765.2009.01026.x

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development, 63*(4).

Caselli, M. (1990). Communicative gestures and first words. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 56-67). New York: Gallaudet University Press.

Clark, H. H. (2006). Social actions, social commitments. In: Enfield, N. Levinson, S. *The roots of human sociality: Culture, cognition, and interaction*, 126–150. Oxford: Berg Press.

Clark, H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 243–268). Mahwah, NJ: Erlbaum.

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

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science, 13*(2), 259-294. doi:10.1016/0364-0213(89)90008-6

Cochet, H., & Vauclair, J. (2010). Pointing gestures produced by toddlers from 15 to 30 months: Different functions, hand shapes and laterality patterns. *Infant Behavior and Development*, *33*(4), 431–441. doi:10.1016/j.infbeh.2010.04.009

Crystal, D. (1991). *A dictionary of linguistics and phonetics. 3rd edition*. Cambridge, MA: Basil Blackwell.

DeLoache, J. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*(2), 66–70. doi:10.1016/j.tics.2003.12.004

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., Strouse, G. A., et al. (2010). Do babies learn from baby media? *Psychological Science*, *21*(11), 1570–1574.

DeLoache, J. Pierroutsakos, S., Uttal, D., Rosengren, K., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science*, *9*(3), 205-210. doi:10.1111/1467-9280.00039

DeLoache, J. S., Uttal, D. H., & Rosengren, K. S. (2004). Scale Errors Offer Evidence for a Perception-Action Dissociation Early in Life. *Science*, *304*(5673), 1027–1029. doi:10.1126/science.1093567

Estigarribia, B. & Clark, E. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language, 34*(04), 799-814. doi: 10.1017/S0305000907008161

Fernald, A. & Kuhl, P. Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development 10*:3, 279-293.

Franco, F., Perucchini, P., & March, Barbara. (2009). Is Infant Initiation of Joint Attention by Pointing Affected by Type of Interaction? *Social Development 18*:1, 51-76.

Folven, R. J., & Bonvillian, J. D. (1991). The transition from non-referential to referential language in children acquiring American Sign Language. *Developmental Psychology, 27*(5), 806-816. doi:10.1037/0012-1649.27.5.806

Folven, R.J., & Bonvillian, J. D. (1993) Sign language acquisition: Developmental aspects. In: Marschark, M., & Clark, M. D. (Eds.), *Psychological perspectives on deafness.* Hillsdale NJ: Psychology Press.

Frank M. C., Vul, E., & Saxe, R. (2011). Measuring the development of social attention using free viewing. *Infancy. 17*: 4. Doi: 10.1111/j.1532-7078.2011.00086.x

Fulkerson, A., & Waxman, S. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition 105*,(1), 218-228.

Furman, R., Ozyurek, A., & Küntay, A. C. (2010). Early language-specificity in Turkish children's caused motion event expressions in speech and gesture. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *Proceedings of the 34th Boston University Conference on Language Development. Vol. 1* (pp. 126-137). Somerville, MA: Cascadilla Press.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language, 11*(01), 43–79. doi:10.1017/S0305000900005584

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*(3), 347 -353. doi:10.1111/j.1467-9280.2009.02295.x

Gogate, L.B., Bahrick, L.E., & Watson, J.D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development, 1,* 876- 892.

Gogate, L.B., Bolzani, L.H., & Betancourt, E. (2006). Attention to maternal multimodal Naming by 6- to 8-month-old infants and learning of word-object relations. *Infancy, 9,* 259-288. doi:10.1207/s15327078in0903_1

Goodrich, W., & Hudson Kam, C. L. (2009). Co - speech gesture as input in verb learning. *Developmental Science, 12*(1), 81-87. doi:10.1111/j.1467-7687.2008.00735.x

Goodwyn, S. W., & Acredolo, L. P. (1993). Symbolic gesture versus word: Is there a modality advantage for onset of symbol use? *Child Development*, *64*(3), 688–701.

Goodwyn, S.W., Acredolo, L. P. & Brown, C. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior*, *24*(2) 81-103. doi: 10.1023/A:1006653828895 xxx

Grassman, S., & Tomasello, M. (2010). Young children follow pointing over words in interpreting acts of reference. *Developmental Science, 13*(1), 252-263. doi:10.1111/j.1467-7687.2009.00871.x

Grice, (1957). Meaning. *The Philosophical Review 66*, 377-88.

Gurteen, P., Horne, J., & Erjavec, M. (2011) Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure: Looking versus reaching measures. *Journal of Experimental Child Psychology. 109,*(2) 201–217. doi:10.1016/j.jecp.2010.12.001

Harris, M., Barlow-Brown, F. & Chasin, J. (1995).The emergence of referential understanding: pointing and the comprehension of object names. *First Language, 15*, 19-34.

Hewes, G. W. (1978). The phylogeny of sign language. In Schlesinger & Namir, (Eds.) *Sign language of the deaf: Psychological, linguistic, and sociological perspectives.* New York: Academic Press pp. 11–56.

Hewes, G. W. (1992). Primate communication and the gestural origin of language. *Current Anthropology*, 65–84.

Hickmann, M., Hendriks, H. & Gullberg, M. (2011). Developmental perspectives on the expression of motion in speech and gesture. A comparison of French and English. *Language, Interaction and Acquisition, 2*, 129-156. John Benjamins.

Hockett, C.F. & Altmann, S. 1968. A note on design features. In Sebeok, Thomas A., (ed), *Animal communication; techniques of study and results of research*, pp. 61–72. Bloomington, Indiana: University Press.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development*, *62*(4), 782-796. doi: 10.1111/j.1467-8624.1991.tb01569.x

Houston-Price, C., Plunkett, K. & Harris, P. (2005). Word learning wizardry at 1;6. *Journal of Child Language*, *32*(01), 175-189. doi:10.1017/S0305000904006610

Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, *9*(1), 23–43. doi:10.1016/0885-2014(94)90018-3

Iverson, J.M., Capirci, O., Longobardi, E., & Caselli, M.C. (1999). Gesturing in mother-child interactions. *Cognitive Development*, *14*, 57-75.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*(5), 367 -371. doi:10.1111/j.0956-7976.2005.01542.x

Johnston, J. C., Durieux-Smith, A., & Bloom, K. (2005). Teaching gestural signs to infants to advance child development: A review of the evidence. *First Language, 25*(2), 235 - 251. doi:10.1177/0142723705050340

Jones, C. & Adamson, L. (1987). Language use in mother-child and mother-child-sibling interactions. *Child Development, 58*(2), 356-366.

Kalagher, H. & Yu, C (2006). The effects of deictic pointing in word learning. In *Proceedings of the 5th International Conference of Development and Learning.* Bloomington, IN.

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*, 313-334. doi:10.1080/01690960802365567

Kendon, A. (2004). *Gesture. Visible action as utterance.* Cambridge: Cambridge University Press.

Knudsen, B., & Liszkowski, U. (2012). Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science, 15*, 113-122. doi:10.1111/j.1467-7687.2011.01098.x.

Knudsen, B., & Liszkowski, U. (2011). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy.* Advance online publication. doi:10.1111/j.1532-7078.2011.00105.x

Leaper, C., & Gleason, J. (1996). The relationship of play activity and gender to parent and child sex-typed communication. *International Journal of Behavioral Development, 19*(4), 689. doi:10.1080/016502596385523

Liebal, K., Behne, T., Carpenter, M. & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science, 12*, 264-71._ doi: 10.1111/j.1467-7687.2008.00758.x

Leung E., & Rheingold H. (1981) Development of pointing as a social gesture. *Developmental Psychology 17*(2), 215–220. doi: 10.1037/0012-1649.17.2.215

Liszkowski, U. (2010). Deictic and other gestures in infancy. *Acción psicológica, 7*(2), 21-33.

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & de Vos, C. (2012). A Prelinguistic Gestural Universal of Human Communication. *Cognitive Science 36*(4), 698-713. doi:10.1111/j.1551-6709.2011.01228.x

Liszkowski, U., Albrecht, K., Carpenter, M., & Tomasello, M. (2008). Infants' visual and auditory communication when a partner is or is not visually attending. *Infant Behavior and Development, 31*(2), 157-167. doi:10.1016/j.infbeh.2007.10.011

Liszkowski, U, Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*(3), 297-307. doi: 10.1111/j.1467-7687.2004.00349.x

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Reference and attitude in infant pointing. *Journal of Child Language, 34*(1), 1-20. doi:10.1017/S0305000906007689.

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

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science,* 20(5), 654-660

Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of infant pointing. *Cognitive Development*, *26*(1), 16-29. doi:10.1016/j.cogdev.2010.10.001.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought.* Chicago: The University of Chicago Press.

Marentette, P., & Nicoladis, E. (2011). Preschoolers' interpretations of gesture: Label or action associate? *Cognition*, *121*(3), 386-399. doi:10.1016/j.cognition.2011.08.012

Masur, E. F. (1997). Maternal Labeling of Novel and Familiar Objects: Implications for Children's Development of Lexical Constraints. *Journal of Child Language*, *24*(02), 427-439.

Meier, R. P., & Newport, E. L. (1990). Out of the hands of babes: On a possible sign advantage in language acquisition. *Language*, *66*(1), 1-23. doi:10.2307/415277

Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, *2*(2), 109–129. doi:10.1207/S15327647JCD0202_1

Murphy, C. M. (1978). Pointing in the context of a shared activity. *Child Development*, *49*(2), 371-380. doi:10.2307/1128700

Namy, L. L. (2008). Recognition of iconicity doesn't come for free. *Developmental Science*, *11*(6), 841-846. doi:10.1111 j.1467-7687.2008.00732.x

Namy, L. L. (2001). What's in a name when it isn't a word? 17-month-olds' mapping of nonverbal symbols to object categories. *Infancy*, *2*(1), 73-86. doi:10.1207 S15327078IN0201_5

Namy, L. L., Campbell, A. L., & Tomasello, M. (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, *5*(1), 37-57. doi:10.1207/s15327647jcd0501_3

Namy, L L. & Waxman S.R. (2000) Naming and exclaiming: Infants' sensitivity to naming contexts. *Journal of Cognition and Development*, *1*(4), 405–428. doi: 10.1207/S15327647JCD0104_03

Namy, L. L., & Waxman, S. R. (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, *69*(2), 295-308. doi: 10.1111/j.1467-8624.1998.tb06189.x

Nelson, K. (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology*, *17*(2), 170-187. doi:10.1037/0012-1649.17.2.170

Newport, E. L. (1975). *Motherese: The speech of mothers to young children.* In N. Castillan, D. Pisoni, and G. Potts (Eds.), *Cognitive Theory* vol. II. Hillsdale, New Jersey: Lawrence Eribaum Associates.

Nicoladis, E., Mayberry, R. I., & Genesee, F. (1999). Gesture and early bilingual development. *Developmental Psychology*, *35*(2), 514-526. doi:10.1037/0012-1649.35.2.514

Ninio, A. (1980). Ostensive definition in vocabulary teaching. *Journal of Child Language, 7, 565- 573.* doi: 10.1017/S0305000900002853

Ozcaliskan, S., & Goldin-Meadow, S. (2011) Is there an iconic gesture spurt at 26 months? In G. Stam and M Ishino (Eds.) *Integrating Gestures: The Interdisciplinary Nature of Gesture*, 163-174. Amsterdam: John Benjamins.

Petitto, L. A. (1987). On the autonomy of language and gesture: Evidence from the acquisition of personal pronouns in American Sign Language. *Cognition*, *27*(1), 1-52. doi:10.1016/0010-0277(87)90034-5

Pizzuto, E., & Capobianco, M. (2005). The link (and differences) between deixis and symbols in children's early gestural-vocal system. *Gesture*, *5*(1-2), 179–199. doi:10.1075/gest.5.1.13piz

O'Neill, M., Bard, K., Linnell, M., & Fluck, M. (2005). Maternal gestures with 20-month-old infants in two contexts. *Developmental Science*, *8*(4), 352-359. doi: 10.1111/j.1467-7687.2005.00423.x

Puccini, D., Hassemer, M., Salomo, D., & Liszkowski, U. (2010). The type of shared activity shapes caregiver and infant communication. *Gesture, 10*(2/3), 279-297. doi:10.1075/gest.10.2-3.08puc

Quine, W. V. O. (1960). *Word and object (Vol. 4).* Cambridge: The MIT Press.

Rakoczy, H., Tomasello, M., & Striano, T. (2005). How children turn objects into symbols: A cultural learning account. In *Symbol use and symbolic understanding*, L. Namy. (Ed.) 69–97, New York Elbrum.

Rochat, P., & Callaghan, T. (2005). What drives symbolic development? In Namy. L (Ed) The case of pictorial comprehension and production. In: *Symbol use and understanding (pp. 25-46. Mahwah, NJ: Lawrence Erlbaum.*

Ryckenbusch, C. & Marcos, H. (2004). Speech acts, social context and parent-toddler play between the ages of 1;5 and 2;3. *Journal of Pragmatics*, *36*(5), 883-897. doi:10.1016/j.pragma.2003.10.019

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735. doi:10.2307/412243

Salomo, D., & Liszkowski, U. (in press). Socio-cultural settings influence the emergence of prelinguistic deictic gestures. *Child Development.*

Salomo, D., & Liszkowski, U. (2010). Caretaker–infant social interaction and gestures across cultures. *Poster presented at the International Conference on Infant Studies*, Baltimore, USA.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen - month - olds under tightly controlled conditions. *Child Development, 69*(2), 309-320. doi:10.1111/j.1467-8624.1998.tb06190.x

Schick, B.B., Marschark, M., & Spencer, P.E. (eds.) (2005) *Advances in the Sign Language Development of Deaf Children*. New York: Oxford University Press.

Searle, J. (1969). *An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Sheskin, D. (2004). *Handbook of parametric and nonparametric statistical procedures*. USA: CRC Press.

Shore, C., O'Connell, B. & Bates, E. (1988). First sentences in language and symbolic play. *Developmental Psychology, 20*(5), 872-880. doi: 10.1037/0012-1649.20.5.872

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. In: Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008).

Snow, C, Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., & Vorster, J. (1976). Mothers' speech in three social classes. *Journal of Psycholinguistic Research, 5*(1), 1-20. doi:10.1007/BF01067944

Stern, D. (1977). *The first relationship: Infant and mother*. London: Fontana/Open Books..

Tardif, T., Gelman, S., & Xu, F. (1999). Putting the "noun bias" in context: a comparison of English and Mandarin. *Child Development, 70*(3), 620-635. doi: 10.1111/1467-8624.00045

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71. doi:10.1207/s15327078in0701_5

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172 -175. doi:10.1111/1467-9280.00127

Tolar, T. D., Lederberg, A. R., Gokhale, S., & Tomasello, M. (2008). The development of the ability to recognize the meaning of iconic signs. *Journal of Deaf Studies and Deaf Education*, *13*, 225–240

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge MA: MIT Press.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*(3), 705-722. doi:10.1111/j.1467-8624.2007.01025.x

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

Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols? *British Journal of Developmental Psychology*, *17*(4), 563–584. doi:10.1348/026151099165483

Trautman, C.H. and Rollins, P.R. (2006). Child-centered behaviors of caregivers with 12-month-old infants: Associations with passive joint engagement and later language. *Journal of Applied Pycholinguistics*. *27*, 447-463. doi: 10.1017/S0142716406060358

Trevarthen (1979). Communication and cooperation in early infancy: a description of primary intersubjectivity. In: M Bullowa, (ed.), *Before speech: the beginning of interpersonal communication*. (pp. 321-348) Binghamton, New York: Cambridge University Press.

Trevarthen (1979). Communication and cooperation in early infancy: a description of primary intersubjectivity. In: M Bullowa, (ed.) *Before speech: the beginning of interpersonal communication,* (pp. 321-348). Binghamton, New York: Cambridge University Press.

Vallotton, C.D. (2011). Sentences and conversations before speech? Gestures of preverbal children reveal cognitive and social skills that do not wait for words. In G. Stam and M. Ishino (Eds.), *Integrating Gestures: The Interdisciplinary Nature of Gesture* (pp. 105-120). Amsterdam, The Netherlands: John Benjamins.

Vygotsky, L. (1978). *Mind in society.* Cambridge MA, Harvard University Press.

Warneken, F., & Tomasello, M. (2007). Helping and Cooperation at 14 Months of Age. *Infancy, 11*(3), 271-294. doi:10.1080/15250000701310389

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development, 77*(3), 640-663. doi:10.1111/j.1467-8624.2006.00895.x

Werner, H. & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and expression of thought.* London: John Wiley & Sons.

Wilbourn, M.P., & Sims, J.P. (in press). Get by with a little help from a word: Multimodal input facilitates 26-month-olds' ability to map and generalize arbitrary gestural labels. *Journal of Cognition and Development.*

Woodward, A., & Hoyne, K. (1999). Infants' learning about words and sounds in relation to objects. *Child Development, 70*(1), 65-77. doi:10.1111/1467-8624.00006

Zammit, M., & Schafer, G. (2011). Maternal label and gesture use affects acquisition of specific object names. *Journal of Child Language, 38*(01), 201–221. doi:10.1017/S0305000909990328

Samenvatting

Een grote hoeveelheid studies laat zien dat jonge kinderen gebaren gebruiken om te communiceren nog voordat ze leren spreken. De studies van dit proefschrift onderzoeken het ontstaan en het gebruik van deictieke en representationele gebaren in de vroege kindertijd. In hoofdstuk 2 wordt onderzocht in hoeverre wijsgedrag in de vroege kindertijd begrepen kan worden als een preverbale vorm van gedeelde aandacht. Hiervoor werden negenendertig 12 maanden oude kinderen tijdens twee verschillende activiteiten geobserveerd: (1) in een vrij spel situatie met een ouder, waarin het manipuleren van en spelen met objecten centraal stond, en (2) de 'decorated room task' waarin kinderen samen met de ouder naar objecten kijken en zodoende wijsgedrag wordt uitgelokt. Zoals voorspeld, heb ik een samenhang gevonden tussen de mate van gedeelde aandacht tijdens het manipuleren van objecten bij activiteit 1 en het wijsgedrag van het kind tijdens activiteit 2. De tijd die kinderen besteden aan gedeelde aandacht met hun ouder tijdens de vrij spel situatie correleert positief met de mate waarin kinderen tijdens de 'decorated room task' de aandacht van hun ouder op een object richten door te wijzen. Ook al is de richting van deze samenhang reeds duidelijk, verdere analyses laten zien dat bijna alle kinderen in staat waren om aandacht te delen in de vrij spel situatie, maar dat tegelijkertijd niet alle kinderen al referentieel naar dingen wezen. Dit suggereert dat kinderen eerst de vaardigheid ontwikkelen om hun aandacht met iemand te delen en dan pas leren om referentieel te communiceren. Deze bevindingen doen vermoeden dat referentiële communicatie ontstaat door de ervaring die kinderen opdoen tijdens het interacteren met anderen. In hoofdstuk 2 ga ik ook in op de structuur van het communicatieritme ('turn taking') dat soms het wijsgedrag van ouder en kind karakteriseert: ouder en kind wezen sequentieel en vrij snel op elkaar volgend. De resultaten van hoofdstuk 2 leveren bewijs voor de stelling dat er een continuüm bestaat tussen de structuur van het communicatieritme tijdens nonverbale face-to-face communicatie en de structuur van het communicatieritme dat tijdens verbale communicatie wordt gebruikt.

In hoofdstuk 3 wordt een studie beschreven waarin gekeken werd naar de manier waarop ouders het interacteren van hun kind aanmoedigen om zodoende het referentiebegrip van hun kind te bevorderen. Door ouder-kind-interacties tijdens twee verschillende activiteiten te observeren, kon worden vastgesteld dat kinderen over verschillende aanwijzigingen beschikken die hen helpen om een gehoord woord te verbinden aan het door de spreker bedoelde object. Ten eerste integreren ouders specifieke gebaren en specifiek taalgebruik. Ten tweede zijn kinderen blootgesteld aan verschillende manieren van communiceren en verschillende soorten gedeelde activiteiten. De systematische integratie van deictieke gebaren en taal van ouders creëert een vruchtbare bodem van waaruit kinderen pragmatische en semantieke aspecten van referenten kunnen leren.

De bevindingen die in hoofdstuk 4 besproken worden, doen echter vermoeden dat de bevorderende effecten van het multimodaal gesticuleren waarmee kinderen tijdens de vroege spraakontwikkeling geconfronteerd worden, beperkt blijven tot de deictieke gebaren. Als kinderen geconfronteerd worden met multimodale labels die arbitraire representationele gebaren bevatten, zijn ze niet in staat om de juiste referenten te lokaliseren, ongeacht of ze een gesproken woord samen met een representationeel gebaar gepresenteerd kregen of dat ze enkel een representationeel gebaar te zien kregen. Deze resultaten leiden tot de vraag of kinderen werkelijk profiteren van een gecombineerd gebruik van representationele gebaren en gesproken taal, zoals dit bijvoorbeeld wordt gepromoot in commercieel beschikbare programma's als Baby Signs.

De bevindingen van hoofdstuk 5 ondersteunen de stelling dat deictieke - en niet representationele - gebaren tot taalontwikkeling leiden. In hoofdstuk 5 werd met een speciaal ontworpen paradigma om spontane, representationele gebaren van kinderen te onderzoeken vastgesteld dat 18 en 24 maanden oude kinderen deictieke gebaren gebruikten om aan een onwetende onderzoeker uit te leggen hoe een bepaald speelgoed gebruikt diende te worden. Studie 1 in hoofdstuk 5 laat tevens zien dat kinderen veel actiegebaren met hulpmiddelen maken. Dit suggereert dat zij wellicht in staat zouden zijn om over het gebruik van een hulpmiddel te communiceren als zij een exacte replica zouden krijgen van het hulpmiddel dat nodig is om het speelgoed te bedienen. De tweede studie liet echter zien dat kinderen deze gebaren bleven maken, zelfs bij afwezigheid van de onderzoeker. Samengevat doen deze twee studies vermoeden dat kinderen tijdens hun tweede levensjaar geleidelijk mentale representaties ontwikkelen. De bevindingen spreken daarnaast taalontwikkelingstheorieën tegen die stellen dat representationele gebaren de ontwikkelingsbrug vormen tussen deictieke communicatie en taal. Het is waarschijnlijker dat deictieke gebaren de ontwikkeling van deictieke communicatie naar taal overbruggen omdat zij geïntegreerd kunnen worden met taal met het doel gezamelijke referentie te bereiken. Dit is cruciaal voor het vormen van mentale representaties en dus voor het leren van woorden.

Curriculum Vitae

Daniel Puccini was born in Yrisarri, New Mexico, in the United States on March 4th, 1982. He studied linguistics at the University of New Mexico, where he obtained his Bachelor of Arts in linguistics in 2006. From 2007 to 2009, Daniel pursued his Master's degree in linguistes at the University of Amsterdam, where he graduated *cum lauda* in 2009. He began his Phd research in April, 2009 in the Communication Before Language Group of the Max Planck Institute for Psycholinguistics in Nijmegen, where his research focused on communicative gestures and social interaction in infancy. Currently, Daniel has prospects of becoming a farmer, although he spends most of his time picking up rocks and sticks.

Publications

Puccini, D., & Liszkowski, U. (2012). 15-month-old infants fast map words but not representational gestures of multimodal labels. *Frontiers in Developmental Psychology*, *3*, 101. doi:10.3389/fpsyg.2012.00101.

Puccini, D., Hassemer, M., Salomo, D., & Liszkowski, U. (2010). The type of shared activity shapes caregiver and infant communication. *Gesture*, *10*(2/3), 279-297. doi:10.1075/gest.10.2-3.08puc

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*