Action perception in development: The role of experience

Der Fakultät für Biowissenschaften und Psychologie

der Universität Leipzig

eingereichte

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

Dr. rer. nat.

vorgelegt von

Anne Keitel (geb. Melzer)

geboren am 12.07.1980 in Berlin

Leipzig, den 21.10.2013

Dekanin

Prof. Dr. Andrea Robitzki

Gutachter

Prof. Dr. Ute Kunzmann

Prof. Dr. Moritz M. Daum

Tag der Verteidigung: 30. Januar 2014

Danksagung

An dieser Stelle möchte ich all denen danken, die mich auf unterschiedlichste Art und Weise bei der Entstehung dieser Dissertation unterstützt haben. Ohne sie wäre die vorliegende Arbeit nicht entstanden.

An erster Stelle möchte ich Moritz Daum für die hervorragende Betreuung danken. Du hast mich zu jedem Zeitpunkt dieser Arbeit voll und ganz unterstützt! Durch deine Denk- und Vorgehensweise habe ich viel gelernt. Außerdem gilt mein Dank Wolfgang Prinz, der mit kritischem Hinterfragen und wertvollen Anregungen mein Promotionsvorhaben begleitet hat und stets für ein optimales, freundliches und kollegiales Arbeitsumfeld sorgte. Weiterhin danke ich besonders Professor Ute Kunzmann für ihre Bereitschaft die vorliegende Arbeit zu begutachten.

Vielen Dank auch an alle Kollegen und Mitdoktoranden, die mir mit ihrer Hilfsbereitschaft, anregenden Diskussionen und willkommener Ablenkung zur Seite standen. Ganz besonders möchte ich an dieser Stelle Silke Atmaca, Barbara Vogt, Anja Gampe, Andrea Walter, Peggy Sparenberg, Sylvia Hach und Emily Cross danken.

Für zuverlässige Hilfe bei der Probandenrekrutierung und Datenerhebung möchte ich mich bei Caterina Böttcher, Kerstin Träger und all den fleißigen Hiwis unseres Babylabs bedanken. Außerdem gilt mein Dank Susanne Starke und Nancy Muschall, die für eine bestmögliche Organisation sorgten und für alle Belange ein offenes Ohr hatten.

Den Mitgliedern und Betreuern im Graduiertenkolleg danke ich dafür, dass sie meinen Horizont fachlich und methodisch erweitert haben. Es hat viel Spaß gemacht mit euch.

Mein Dank gilt auch meiner Mama und meinem Papa. Ihr habt mich immer bedingungslos unterstützt und mir so den Rückhalt gegeben, den ich für mein Berufs- und Privatleben brauche.

Ganz besonders danken möchte ich meiner kleinen Familie. Meinen beiden Töchtern, Lara und Lise, dafür, dass sie mich auf die schönste denkbare Weise ablenken und mich so oft an das Wichtigste im Leben erinnern. Außerdem danke ich meinem Mann Christian. Danke, dass du mir den Rücken frei hältst, mich anspornst, immer für mich da bist. Ich freue mich auf unser nächstes Abenteuer!

CONTENTS

1. Introduction 1

1.1. Defining Action Perception 3

1.2. Agent and Observer – Theoretical Accounts of Action Perception 5

1.3. Dissociation Between Measures of Action Perception 10

1.4. Action Perception in Children 12

1.4.1. Nonverbal individual action 13

1.4.2. Nonverbal joint action 14

1.4.2. 'Verbal joint action' in conversations 16

1.5. Research Questions 16

1.5.1. The developing link between production and perception of individual action 17

1.5.2. The difference between individual and joint action 17

1.5.3. The role of semantics and intonation in the perception of conversations 18

1.6. Summary of the Main Results 19

1.7. Considerations, Limitations, and Perspectives 19

2. Production and perception of contralateral reaching: A close link by 12 months of age 27

3. Infants' and adults' perception of individual and joint action 53

4. Perception of conversations: The importance of semantics and intonation in children's development 75

References (Chapter 1) 104

List of Figures 115

List of Tables 117

List of Abbreviations 118

Curriculum Vitae & List of Publications..119
Selbstständigkeitserklärung...120
Bibliographic Details & Kurzzusammenfassung 121
Summary / Zusammenfassung 122
Confirmation of Contribution of Co-Authors...133

Chapter 1

Introduction

'During the earliest stages the child perceives things like a solipsist who is unaware of himself as subject and is familiar only with his own actions."

(J. Piaget)

Our own actions are a vital part of our life from birth. The above quotation demonstrates Jean Piaget's view that young children aged 18 months and under are 'solipsistic' (extremely egocentric) and only familiar with their own actions (1954, p. 355). Yet, children also observe actions performed by others from early on. This dissertation aimed to investigate this side of the story: How children develop the ability to understand actions performed by others and how the increasing 'familiarity' with their own actions shapes this understanding.

Imagine that you enter a busy restaurant and observe the following scene: You see many different people sitting at their tables. Some of them will be in the midst of eating, or reaching out to grasp their glass of wine. Others will raise their glass to toast their friends. And others again will talk to each other animatedly. For adults, it will be no difficulty to understand the actions these people are performing, because adults are able to use their knowledge to infer others' goals and intentions from observed movements and the occurring context. Moreover, adults are able to *predict* others' action goals and thus prepare quick and suitable responses when they interact with others.

Now imagine a 1-year-old child observing the same scene. Will this child understand the actions in the same way as adults do, and predict the action goals as easily as adults do? It is very likely that there will be differences. Some actions, such as someone grasping a glass, will be apparent for even a small child. But the meaning of two people raising a toast to each other might not be instantly comprehensible for infants. And to follow a conversation seems inconceivable without some semantic skills and the basic knowledge about how conversations work.

The development from child to adult is associated with achieving manifold experience. Throughout the ages, experience has been proposed as the main influence that shapes cognition. For example, more than 2000 years ago, Julius Caesar wrote 'Experience is the teacher of all things'. In the beginning of the last century, Albert Einstein maintained this view by saying 'The only source of knowledge is experience'. Rather recently (i.e., in the last 20 years), scientific interest in how experience influences action perception has intensified. More precisely, experience in performing actions (i.e. active experience, or the ability to produce an action) and to a lesser extent experience in observing actions (i.e., passive, mainly visual experience). One reason for this interest was certainly the notion of a close link between perception and action, or more precisely, a common representational domain between planned and observed actions (Prinz, 1990, 1997; Hommel, Müsseler, Aschersleben, & Prinz, 2001). This link has since been accepted in adults and infants, but research continues on how exactly experience, or action, shapes perception.

The present dissertation aimed to investigate how different levels of experience influence the perception of other's actions in distinctive areas. The restaurant scene described above illustrates that actions can be performed by one individual or jointly (for example, one person reaching for a glass, or two people raising their glasses to toast each other, respectively), and that they can be nonverbal (as the above) or verbal (for example, two people having a conversation). In order to succeed in the performance of those nonverbal and verbal actions and interactions, different skills (and therefore different kinds of active experience) are necessary. Infants learn to perform various manual actions during their first year of life (e.g., Bourgeois, Khawar, Neal, & Lockman, 2005). During their second year of life, they learn to coordinate those individual actions with others in joint action (Brownell, 2011). During their third year of life, children master to verbally interact with others (Clark, 2003). Thus, each of the actions in the different areas requires new skills that are learned successively during the first 3 years of life. Accordingly, the research included in this dissertation concerned the perception of an individually performed manual action, a jointly performed manual action, and a verbal interaction, by children with more or less experience in the respective areas and by adults, who are typically very experienced in these everyday actions.

Before describing the specific research questions in detail, I will first define the concept of action perception. Second, the distinction between the perception and the performance of actions will be highlighted, which includes different theoretical accounts of action perception. Next, I will focus on different measures of action perception and how they are related, in order to facilitate the understanding of the following empirical findings. Empirical findings are organised according to the different areas mentioned above (perception of nonverbal and verbal action and interaction). Based on this overview, the resulting research questions will be outlined, followed by a summary of their results. The chapter will conclude by highlighting some limitations and future perspectives.

1.1. Defining Action Perception

In order to define action perception, I will disentangle the terms action and perception first. In a nutshell, human *action* consists of two main components: a movement and a goal (Prinz, 1997; Elsner, 2007; Csibra & Gergely, 2007). A more scientific phrasing establishes that ‘actions and reactions need to be regarded as segments of body movements that are individuated on the basis of goals’ (Hommel et al., 2001). Thus, when using the term ‘action’, we always mean goal-directed actions. In its simplest form, this goal can be immediately visible and achieved in a single movement: I reach out to grasp an object that lies in front of me. But a goal can also be abstract (I speak to inform my conversation partner about my thoughts) and lie in the far future (I study to get good grades in order to get a good job). These examples also demonstrate that goals are organised hierarchically (Jeannerod, 1994), for example, from overarching goals to sub-goals that can be achieved by ‘elementary motor acts’ (Csibra, 2007). An even more precise description of the hierarchy of actions characterises four different levels (Kilner, Friston, & Frith, 2007; see also Hamilton & Grafton, 2008; Kilner, 2011):

‘(1) The intention level that defines the long-term goal. (2) The goal level that describes short-term goals that are necessary to achieve the long-term intention. (3) The kinematic level that describes the shape of the hand and the movement of the arm in space and time. (4) The muscle level that describes the pattern of muscle activity required to execute the action.’

Importantly, language is also considered an action in the present dissertation (Pickering & Garrod, 2013; Glenberg, 2007). Because the above hierarchy of actions was presumably not intended for the case of language, it needs to be slightly adapted. One example for an intention of ‘producing language’ is to convey thoughts; an example for a goal is to express a word; on the kinematic level, the movements of the mouth and tongue can be described; and the muscle level is similar to that of other actions.

The second part of action perception, the term *perception*, is generally defined as ‘The process of becoming aware or conscious of a thing or things in general; the state of being aware; consciousness; † [spiritual] understanding [obs.].’ (‘perception, n.’, *The Oxford English Dictionary*, 2005). In psychological sciences, perception describes an immensely broad research field, which ranges for example from basic physiological processes, such as the incidence of light on the cornea, to the other end of the continuum, for example, acting on a perceived stimulus. For the present work, the term perception needs to be narrowed down to an operational definition in the context of visually presented verbal and nonverbal actions. The general definition of perception already suggests that action perception is much more than a simple sensation of an action. 'Action understanding' is a term that is often used alternatively to action perception (including this work), and that is intuitively comprehensible, but it is difficult to define. Some authors try to describe action understanding as something like 'grasping of the *sense* of the actions performed by others' (Gallese, 2006), which replaces one abstract term with others. For this work I specify action understanding as 'extracting the immediate or further goal' (Csibra, 2007), which matches the idea that goal detection forms the 'core ability of action understanding' (Gallese, 2009). Goal detection involves obtaining a mental representation of the goal. Regarding the previously described hierarchy, this targets the second level (goal level) of actions. Such a goal could be the glass in reaching actions, or the words in spoken language. In summary, my operational definition of action perception is as follows: Action perception is the observation of actions performed by others and the obtainment of a mental representation of this action, including the action goal. The first level of Kilner's hierarchy, the intention level, will only play a minor role for this dissertation.

1.2. Agent and Observer – Theoretical Accounts of Action Perception

Actions can be regarded from two different perspectives: *Agents* produce their own actions first-hand, knowing their own goals and intentions. *Observers* perceive other people's actions from the outside and have to infer their goals and intentions. Before elaborating on different ideas of action perception, I will first describe the characteristics of action production. An agent needs to plan an action ahead, that is, before a movement is initiated. This requires an internal mental representation, so called 'motor programs' (Summers & Anson, 2009; Rosenbaum & Krist, 1996) or 'action plans' (Rotman, Troje, Johansson, & Flanagan, 2006; Flanagan & Johansson, 2003; Prinz, 1997). These internal representations, in the following called action representations, not only include a representation of the movement but also a representation of the action goal (Gallese, 2009; Hommel et al., 2001). Furthermore, once a movement is initiated and agents are in the process of executing an action, forward models are implemented to predict motor states and the sensory consequences of that movement (Miall & Wolpert, 1996). Every performed goal-directed action thus comprises an internal anticipation of a future goal and of future states.

For an observer, action representations are only apparent via open behaviour and the context in which it occurs. This seems a disadvantage if someone seeks to understand others' actions. And yet, most of the time, we effortlessly manage to understand what people around us are doing. There are a number of theoretical accounts that try to explain how action understanding is achieved. An important framework that addresses the general relationship between perception and action has been provided by Prinz (1990, 1997). The *common coding approach* assumes a common representational domain for action perception and production as opposed to strictly separate coding. The exact nature of this representational overlap is unknown, but it concerns high cognitive levels of coding (Prinz, 1997), and might operate in addition to separate coding (Prinz, 2012). A common representation provides a basis for a bidirectional influence between both action perception and production. In adults, it is well established that action perception can facilitate or interfere with action production, which is usually examined in shortened or prolonged reaction times in the execution of actions (Brass, Bekkering, & Prinz, 2001; Brass, Bekkering, Wohlschläger, & Prinz, 2000; Craighero, Bello, Fadiga, & Rizzolatti, 2002; Kilner, Paulignan, & Blakemore, 2003). Similarly, action production can facilitate or interfere with the processing of observed actions (Hamilton, Wolpert, & Frith, 2004; Miall et al., 2006; Wühr & Müsseler, 2001).

One proposed way to achieve action understanding that relates to a common representation is provided by the *simulation theory* (e.g., Gallese, 2009; Jeannerod, 2001; Rizzolatti & Craighero, 2004). Here, an observed action is simulated through mapping the observed action onto one's own motor representations. This motor simulation is defined as an 'internal representation of motor programs without overt movement' (Cross, Hamilton, & Grafton, 2006; Jeannerod, 2001). The simulation theory posits that, through motor simulation, we are not only able to represent others' movements but also others' action goals and intentions (Gallese, 2009; Rizzolatti & Sinigaglia, 2010), or the 'meaning' of others' actions (Gallese, Keysers, & Rizzolatti, 2004). Proponents of the simulation theory thus assert that motor simulation by itself provides an understanding of the observed action. The neurophysiological foundation of simulation processes, which is also called direct matching, is thought to be a parietal-frontal network, consisting of mirror neurons that respond both when a particular action is executed or observed (the mirror neuron system, MNS; e.g. Rizzolatti & Craighero, 2004). Motor simulation is initiated as soon as the observed action begins. Importantly, this enables the anticipation of an action goal in the observer (Rizzolatti & Sinigaglia, 2010). Motor simulation has become an accepted and important aspect of action perception during the last 10 years. However, the particular proposition that simple motor simulation 'directly' provides action understanding has been challenged recently (Csibra, 2007; Hickok, 2013; Jacob, 2013; Kilner, 2011; Kosonogov, 2012).

A different simulation account that focuses on the anticipatory nature of action perception and provides a detailed description of how action understanding might work is the *predictive coding framework* (Kilner et al., 2007; Kilner, 2011). This framework posits that predictions are made on all hierarchical levels of an action (intentions, goals, motor commands, and kinematics). First, prior expectations about goals or intentions are estimated from the context in which an action occurs. The estimated goal is then used to generate a prediction of the 'sensory consequences' (i.e., kinematics) based on one's own motor system. This predicted kinematics is compared with the actual observed kinematics and a prediction error is generated (i.e., the level of confidence with which a prediction is correct, Cross et al., 2012). Through reciprocal interactions among cortical hierarchies, the predictions on each level are updated until the prediction error is minimised and the most likely cause of an action is inferred. The simulation of actions is accomplished by an action-observation network (AON, see Kilner, 2011), of which the MNS is part of. Goals or intentions, however, are predicted through a 'semantic retrieval' process that functions on a more abstract level and is not part of the AON (Kilner, 2011).

An account that does not ascribe a crucial role of motor simulation to action understanding is the *teleological stance* (Csibra & Gergely, 1998, 2007; Gergely & Csibra, 2003). In this context, ‘teleological’ means ‘relating to ends or final causes‘ (‘teleological, adj.’, *The Oxford English Dictionary*, 2011) and stresses the importance of the goal for action understanding. The teleological stance was developed to specifically explain how infants can understand observed actions without inferring underlying mental states (e.g., intentions, beliefs) in the agent. The mechanism that allows infants to interpret a goal is called teleological reasoning. Teleological reasoning enables infants to relate relevant aspects of reality (action, goal state, and current situational constraints) through the principle of rational action (Gergely & Csibra, 2003). This inferential principle characterises that agents generally achieve their goals in the most efficient manner. A famous example that was used to demonstrate teleological reasoning in infants is one of a computer-animated small circle jumping over an obstacle to reach a larger circle. When the obstacle was removed, infants were surprised when the small circle performed the now unnecessary and inefficient jumping action (Gergely, Nadasdy, Csibra, & Bíró, 1995). This shows that infants expected the small circle to take the most efficient way to reach its goal (straight to the large circle). In general, this account has been used to explain findings showing that infants are able to attribute action goals where simulation is not possible, for example, when observing non-human agents (Gergely et al., 1995; Luo & Baillargeon, 2005; Bíró & Leslie, 2007) or biologically impossible motions (Southgate, Johnson, & Csibra, 2008). Notably, teleological reasoning does not require prior experience with the action, because it considers relevant constraints (e.g., the context) even of novel situations. Furthermore, the teleological stance explicitly states that attributing a goal to an observed ongoing action includes the anticipation of a future state and thus enables ‘action anticipation’ (Csibra & Gergely, 2007). Csibra (2007) recently integrated the mechanism of teleological reasoning into a larger framework where he included motor simulation. Here, action understanding entails what he called ‘motor emulation’. Motor emulation does not cause action understanding, but is the result of action understanding (achieved, for example, through teleological reasoning). According to this 'motor emulation account', the function of this kind of motor activity is to enable us to be engaged in joint action.

Yet another idea combines elements of the simulation theory and the teleological stance (Brass, Schmitt, Spengler, & Gergely, 2007). It posits that motor simulation provides action understanding in situations that are highly familiar to the observer. Only when motor simulation is insufficient because the observer does not possess matching motor representations (for example in novel or unusual situations), action understanding is accomplished via inferential interpretive processes, such as teleological reasoning.

Proponents of all accounts concerning action understanding continue to interpret evidence in favour of their respective hypotheses (e.g., Csibra, 2007; Kilner, 2011; Rizzolatti & Sinigaglia, 2010). The present dissertation does not provide further insights in the underlying mechanisms of action understanding. But the introduced accounts differ in their relevance for this work. The teleological stance is without doubt an important account for infants' action understanding. However, because this work addresses the influence of experience on action perception, it is less relevant that infants are able to infer action goals without active experience under some circumstances, such as when observing non-human agents (e.g., Gergely et al., 1995). For this work, it is more relevant that one's own motor system is involved during action perception, be it as a cause of action understanding (simulation theory, e.g. Rizzolatti & Sinigaglia, 2010) or as a result (predictve coding framework, Kilner, 2011; motor emulation account, Csibra, 2007).

Taken together, the described approaches provide two characteristics of action perception that are significant to this work. First, action perception and action production are linked. This is stated explicitly in the common coding approach, and implied in all other approaches that assume motor simulation in response to a perceived action. An action can only be simulated by the motor system, if observers have obtained a matching representation in their own motor repertoire. And second, anticipatory processes are involved in action perception. It was established in the beginning of this section that, on the one hand, agents need to plan their actions ahead in order to perform them. The anticipation of movements and action goals is an inherent part of action production. On the other hand, observers are able to predict the goals of ongoing actions as well – regardless of the mechanisms that are assumed to play a role in action perception. The simulation theory, as well as the predictive coding framework and the teleological stance, explicitly include goal anticipation. Thus, it is generally accepted that agents and observers both make use of anticipatory processes during the production and the observation of actions, respectively.

One manifestation of anticipatory processes that is particularly important for the present work is anticipatory gaze. When we perform actions, such as playing table tennis or driving a car, we show anticipatory gaze shifts, which means that our gaze precedes important steps (goals and sub-goals) of the task (Land & Furneaux, 1997). It was suggested that the control program for a particular action (i.e., before mentioned action representations) also contains information for the oculomotor system (Land & Furneaux, 1997). Importantly, not only agents show anticipatory gaze shifts but observers do as well. In a seminal study, Flanagan and Johansson (2003) recorded eye movements of participants during the execution of a block-stacking task and during their perception of the same task. They found highly similar gaze patterns during the production of actions and the perception of others' actions. By measuring gaze behaviour, it is thus possible to get a grasp on anticipatory processes in the agent as well as the observer. The development and implications of anticipatory gaze shifts will be discussed in the next section.

1.3. Dissociation Between Measures of Action Perception

Measuring an observer's gaze behaviour is a particularly important approach for the investigation of infants' action perception, because other typical measures, such as manual reaction times, are impossible to adopt with infants. There are two different approaches to use gaze behaviour that will be contrasted in this section: anticipatory and post-hoc measures. The implications of studies that used either approach differ to some extent, which is why I will explain both measures in more detail before moving on to empirical findings of action perception.

Anticipatory gaze behaviour (or gaze latency) has been utilised increasingly since it has been discovered that adults (Flanagan & Johansson, 2003) and 12-month-old infants (Falck-Ytter, Gredebäck, & von Hofsten, 2006) are able to predict the goal of an observed ongoing manual action. If an observer performs anticipatory gaze shifts towards the goal of an action, this suggests that the observer is able to encode the action goal (i.e. obtain a mental representation of it) while the action is still in progress. Besides this on-line *anticipatory measure*, there are 'reactive' measures of gaze behaviour for the study of action perception, which can be summarised as *post-hoc measures*. They have in common that gaze behaviour is analysed after an action is fully presented to the observer (i.e. at least until the goal has been achieved). (Gredebäck & Melinder, 2010). The most common post-hoc measure is looking time. Infants tend to look longer at unexpected actions or action goals, which makes it possible to infer the expectations that infants have built during the observation. Looking time is used, for example, in habituation paradigms, in violation-of-expectation paradigms, or in preferential-looking paradigms. Generally, anticipatory and post-hoc measures can both be used to explore infants' expectations about an action goal (Falck-Ytter et al., 2006; Woodward, 1998) or of the means with which an action was completed (Gergely et al., 1995; Gredebäck & Melinder, 2010). In any case, infants are required to encode an agent's goal (Gredebäck & Melinder, 2010), in order to form expectations about the goal itself or rational means to achieve it. However, anticipatory and post-hoc measures differ in the time and information available to encode an action goal. Whereas the time to encode a goal is strictly limited for its anticipation, more time and more information is available to do so after the action is completed.

Apart from different time limits, there seems to be a functional dissociation between anticipatory and post-hoc measures (for studies contrasting both measures, see Daum, Attig, Gunawan, Prinz, & Gredebäck, 2012; Gredebäck & Melinder, 2010). Reactive gaze behaviour indicates that infants are able to encode the goal of an observed action by 6 months of age (e.g., Daum, Prinz, & Aschersleben, 2011; Woodward, 1998). The on-line anticipation of action goals has been reported primarily by the end of the first year of life (Falck-Ytter et al., 2006; Gredebäck & Melinder, 2010; Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; but see Kanakogi & Itakura, 2011). Many studies, typically using manual actions, indicate a close relationship between infants' ability to anticipate an action goal and active experience, that is, their own ability to perform that action (Kanakogi & Itakura, 2011; Gredebäck & Kochukhova, 2010; Gredebäck & Melinder, 2010; Cannon et al., 2012). Similarly, post-hoc measures have also been found to correspond to infants' own experience (Daum et al., 2011; Sommerville, Woodward, & Needham, 2005) though not consistently so (Daum, Prinz, & Aschersleben, 2009; Gredebäck & Melinder, 2010). Anticipatory measures thus seem to rely more unambiguously on infants' own experience than post-hoc measures. The occurrence of anticipatory gaze is often interpreted as evidence for simulation processes in the observer (e.g. Falck-Ytter et al., 2006; Gredebäck & Kochukhova, 2010; Kanakogi & Itakura, 2011), which are thought to 'aid in effortless, efficient prediction of ongoing movements' (Cross, Stadler, Parkinson, Schütz-Bosbach, & Prinz, 2013), whereas the mechanisms that allow infants to infer action goals in a post-hoc paradigm are generally less clear (Aslin, 2007).

A further theoretical consideration is that measuring anticipatory gaze takes into account anticipatory processes in the observer. Because anticipatory processes are important for the perception, as well as the production of actions, anticipatory measures of action perception warrant better comparability to action production than post-hoc measures. For the above described advantages, we measured anticipatory gaze in the current work. In the following section, I will summarise empirical findings of previous research, most of which used post-hoc measures. The empirical and theoretical observations about anticipatory and post-hoc measures have to be kept in mind when previous findings about children's action perception are interpreted.

1.4. Action Perception in Children

As previously described, the production and the perception of an action are closely linked. Thus, the actions that can be used to explore an influence of experience on action perception are dependent on children's own cognitive and motor skills at a certain age. For example, in order to investigate infants' understanding of a reaching action, one has to consider the age at which infants typically start reaching themselves. In the following, previous findings concerning infants' action perception and experience will be summarised, organised according to the different areas: nonverbal actions performed by one individual, nonverbal actions performed jointly, and conversations, as a case of 'verbal joint action'.

1.4.1. Nonverbal individual action

In a majority of studies, reaching-to-grasp actions have been utilised to investigate the interplay between the perception and the production of actions, a skill that emerges at the age of 3 to 4 months (e.g. White, Castle, & Held, 1964). For example, if 3-month-old infants' reaching-to-grasp skills were trained through the use of 'sticky mittens' (i.e. Velcro mittens that objects stick to), they were subsequently able to encode the goal of an actor's reaching action, when tested via a visual habituation paradigm (Sommerville et al., 2005). Without this training, infants only mastered this task at 6 months (Woodward, 1998). The use of sticky mittens to promote grasping behaviour in 3-month-olds also induced more general cognitive developments, such as increased visual exploration behaviour of agents and objects (Libertus & Needham, 2010), or increased face preference (Libertus & Needham, 2011). Furthermore, there was a correspondence between 6-month-old infants' grasping skills (palmar vs. thump opposition) and their looking times to unexpected grasping actions (i.e. large hand aperture for small objects and vice versa; Daum et al., 2011; see also Loucks & Sommerville, 2012). And only infants with advanced fine motor skills (in this case, the extent to which infants grasped or manipulated objects) were able to discriminate between biologically possible and impossible grasping movements at 8 months of age (Reid, Belsky, & Johnson, 2005). Examining anticipatory gaze behaviour, it was found that grasping skills (measured by the reaching angle between the two hands) correlated with gaze latency towards the goal of human grasping actions in a group of 4- to 10-month-old infants (Kanakogi & Itakura, 2011).

Other manual actions have also been utilised. For example, 9-month-olds were only able to attribute a goal to a pointing action if they could use the pointing gesture themselves (Woodward & Guajardo, 2002). Using a means-end task (pulling a cloth to retrieve a toy), a correlation between infants' own 'planful behaviour' (e.g. fixating the toy and grasping it) and their ability to encode the goal of a similar observed task was found in 10-month-olds (Sommerville & Woodward, 2005). Also in 10-month-olds, active training with a means-end task (using a cane to retrieve a toy) facilitated the ability to identify the goal of a similar tool-use event (Sommerville, Hildebrand, & Crane, 2008). Furthermore, the extent to which 12-month-old infants spontaneously produced containment actions (i.e., placing objects in a bucket) corresponded to their gaze latency when subsequently watching them (Cannon et al., 2012).

Crawling has also been used to demonstrate a relationship between infants' motor skills and their cognitive abilities. It was found that already crawling infants at 7 months looked longer at self-propelled objects than not yet crawling 7-month-olds (Cicchino & Rakison, 2008), which indicates a more sophisticated perception of self-propelled objects once crawling is achieved. And finally, children's goal anticipation during the observation of a puzzle being solved was dependent on their own ability to solve the puzzle at 25, but not yet at 18 months of age (Gredebäck & Kochukhova, 2010).

The above described studies have generally demonstrated that there is either a link between infants' own production skills and perceived individual action, or an influence of action experience on perception. For many manual actions, this link seems to develop at some point during the first year of life. However, it remains unclear whether a link between production and perception is established as soon as the ability to produce an action emerges, or whether more experience is necessary.

1.4.2. Nonverbal joint action

Infants engage in face-to-face interactions with their caregivers from birth. These interactions are guided and shaped by adults during the first year of life. The ability to engage in *coordinated* joint action emerges only during the second year of life (for an overview, see Brownell, 2011). However, infants observe others' coordinated joint action from very early on. How they perceive these interactions has recently become a question of interest (e.g., Henderson & Woodward, 2011; Schmitow & Kochukhova, 2013). Using a habituation paradigm, it has been shown that 14-month-old infants are able to infer the joint goal of two actors collaboratively retrieving a toy from a closed box (Henderson & Woodward, 2011). Using the same paradigm, 10-month-olds were not yet able to infer the joint goal, not even after visual experience (Henderson, Wang, Matz, & Woodward, 2013). However, if 10-month-olds actively experienced the joint action themselves prior to the habituation task, they were able to infer the joint goal of the two observed agents (Henderson et al., 2013). Using gaze latency, it has been shown that 10- and 18-month-olds' gaze shifts towards joint goals are modulated by their own experience with the respective manual actions, such as placing things in a bucket, building a tower, or give and take actions (Schmitow & Kochukhova, 2013). When presented with two agents feeding each other, infants at 12, but not 6 months, anticipated the goal of this action (i.e., the mouth), and this was modulated by their own experience with being fed (Gredebäck & Melinder, 2010). Notably, already 6-month-olds were able to anticipate that food will be brought to the mouth if one agent fed herself (Kochukhova & Gredebäck, 2010). And last, 18-month-old infants were able to anticipate a joint goal of two agents who sequentially placed blocks, and did so more often if the actors were socially engaged (Fawcett & Gredebäck, 2013).

The above reported studies (using both anticipatory and post-hoc measures) suggest that infants' ability to infer observed joint goals typically emerges around their first birthday, or shortly afterwards. This ability seems to depend on infants' own experience with the manual action (Schmitow & Kochukhova, 2013), and their own experience with joint action (Henderson et al., 2013), respectively. Apart from this apparent analogy to the perception of individual actions, the feeding studies point towards a difference between the perception of individual and joint action in infants, because 6-month-olds were able to anticipate an individually but not a jointly performed feeding action. It has not yet been investigated systematically whether infants understand actions performed by one agent differentially from actions performed by two agents.

1.4.2. 'Verbal joint action' in conversations

Infants start understanding words at around 8 months of age (Fenson et al., 1994; Harris, Yeeles, Chasin, & Oakley, 1995), and they utter their first word usually around their first birthday (e.g., Waxman & Markow, 1995). Their vocabulary increases rapidly and the ability to form sentences develops, so that they are able to engage in simple and more complex conversation by 3 years of age (Clark, 2003). The development of language production and perception has been a huge research area of linguists and psychologists for a long time (e.g., McCarthy, 1933). Likewise, how conversations work has been a question of interest (e.g., Sacks, Scheglof, & Jefferson, 1974; for an overview, see Hutchby & Wooffitt, 2008). But only recently, the idea originated to visually present prelinguistic and linguistic children with dyadic conversations and measure their gaze behaviour (von Hofsten, Uhlig, Adell, & Kochukhova, 2009). The way children shift their gaze between speakers can shed light on how they perceive conversations, even if they are not yet able to speak themselves. For example, already 6-month-olds could follow a conversation more easily if the agents interacted in a face-to-face manner, as opposed to a back-to-back interaction, where the speakers looked into opposite directions (Augusti, Melinder, & Gredebäck, 2010). In another study, children's ability to anticipate turn-taking (i.e., a change of speaker) in natural conversations was explored. It was found that 3-year-olds anticipated nearly twice as many turns as 1-year-olds (von Hofsten et al., 2009), which could indicate that the developing language experience influenced this ability.

The two above mentioned studies give only a brief glimpse of how visually presented conversations are perceived by prelinguistic and linguistic children. It remains to be investigated how language experience (e.g., semantic skills) and other linguistic factors influence the perception of conversations.

1.5. Research Questions

The preceding overview highlighted the role of experience in children's action perception in three distinct areas, namely, nonverbal individual action, nonverbal joint action, and verbal joint action. The main focus of the present dissertation was to further investigate how experience shapes children's anticipation of action in these areas. Accordingly, the present work addresses three main research questions that will be detailed in the following.

1.5.1. The developing link between production and perception of individual action

It has been elaborated previously that the perception of individual action is largely dependent on children's own action skills. A missing detail of this research is, however, *when* action anticipation becomes linked to production during typical development. The first study of this dissertation ('common representation study') aimed to specify whether a common representation of perception and action develops as soon as an action emerges, or whether more active experience is necessary for its development. To this end, 6- and 12-month-old infants were presented with videos of contralateral manual actions (e.g., reaching across the body midline) and their gaze latency towards action goals was measured. Additionally, infants' own contralateral reaching skills were tested using a task adopted from Bruner (1969). Contralateral reaching emerges in the middle of the first year of life and slowly improves over the next months. It is thus particularly useful to determine when a common representation between perception and action is established. If a correlation between the two tasks is already present in 6-month-olds, this suggests that action perception links to action production as soon as infants start producing an action. If a correlation is only present in 12-month-olds, this suggests that more active experience is necessary. The study outlined above will be described in detail in Chapter 2 of this dissertation (Melzer, Prinz, & Daum, 2012).

1.5.2. The difference between individual and joint action

As highlighted previously, the research on how children perceive others' joint action by measuring gaze behaviour is relatively new. The few conducted studies have mainly focused on the questions when children are able to infer joint goals, and how this is influenced by experience. In order to interpret those and future findings, it would be helpful to know whether infants' perception of joint action is essentially different from individual action, or whether they follow the same developmental trajectory. In the second study ('joint action study'), we addressed this question by presenting infants and adults with videos of a block-stacking action that was either performed by one agent or two agents, and compared their gaze behaviour towards action goals. The overarching goal was identical in both conditions ('to build a tower'), only the sub-goals ('to grasp a block', 'to stack it') differed in that they were performed by one or two agents. Infants were 9 and 12 months old, and had little or no experience with coordinated joint action themselves, whereas adults are usually very experienced in coordinating their actions with others. This experience could influence the level of representation that is used to guide an observer's gaze shifts. If participants' perception of the actions is based on their representation of the overarching goal, their gaze behaviour should not be affected by the number of agents. If, by contrast, participants represent each agent's individual sub-goals in isolation, switching between the sub-goals of the two agents in the joint-action condition could cause delayed gaze shifts. Chapter 3 will present the study summarised here (Keitel, Prinz, & Daum, submitted).

1.5.3. The role of semantics and intonation in the perception of conversations

Previous research on the perception of conversations using gaze behaviour has shown that 3-year-old children anticipate more turns in observed conversations than 1-year-olds (von Hofsten et al., 2009), which suggests that the older children were better able to anticipate the course of the conversation. However, what influences children's anticipation of turns in observed conversations is largely unknown. The third study ('conversation study') aimed to further explore the perception of conversations by investigating the role of two linguistic factors for the anticipation of turns in children with more or less language experience. The first factor, semantics (i.e., language understanding), was addresses by testing prelinguistic (6 and 12 months) and linguistic (24 and 36 months) children, and a control group of adults. The age at which children are able to reliably anticipate a speaker's next turn indicates how much language experience is required in order to understand the course of a conversation. The second factor, the prosodic cue intonation (i.e., the rise and fall of the voice in speaking), was addressed by presenting two conversations, one with normal intonation and one with flattened intonation. Intonational differences are already processed by newborns (Nazzi, Floccia, & Bertoncini, 1998; Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008) and intonation plays a role in early word learning (Thiessen, Hill, & Saffran, 2005). It could therefore also support children's perception of conversations. This study will be reported in detail in Chapter 4 (Keitel, Prinz, Friederici, von Hofsten, & Daum, 2013).

1.6. Summary of the Main Results

Before presenting the studies which address above research questions, I will first provide a brief overview of the main results. The common representation study (Melzer et al., 2012) addressed the question when during infancy a link between the production and the perception of an action is established. To this end, the perception and the production of contralateral reaching were tested in 6- and 12-month-old infants. The results showed that, as expected, the 12-month-olds performed contralateral reaching actions more often than the 6-month-olds. Furthermore, the 12-month-olds showed anticipatory gaze shifts towards action goals, whereas the 6-month-olds showed reactive gaze shifts (i.e., their gaze arrived at the goal after the action was completed). And, most importantly, a correlation between the two tasks was only present in 12-month-olds, but not yet in 6-month-olds. These findings suggest that a common representation between action perception and production during development is not instantly present. Instead, the formation of such a link seems to depend on a certain amount of active experience, or 'training', in performing an action.

The joint action study (Keitel et al., submitted) aimed to determine whether infants and adults perceive joint action per se different from individual action. We presented 9- and 12-month-old infants and adults with a block-stacking action that was either performed by one agent or two agents. It was found that adults anticipated goals in both conditions significantly faster than infants, and their gaze latencies did not differ between conditions. By contrast, infants showed faster anticipation of goals in the individual condition than in the joint condition. This difference was more pronounced in the younger age group of 9-month-olds. Thus, infants with virtually no coordinated joint action skills themselves were unable to use a representation of the overarching joint goal of two agents. Infants possibly represented the sub-goals of the block-stacking action in isolation, which led to delayed gaze shifts in the joint condition, when they had to switch between the representations of the two agents' sub-goals. Adults, however, were able to infer the overarching joint goal of two agents, which led to comparable gaze behaviour in both conditions. These findings suggest a modulating influence of experience on the perception of joint action.

The conversation study (Keitel et al., 2013) addressed the influence of semantics and intonation on the perception of conversations in prelinguistic and linguistic children. For this purpose, children of four age groups (6, 12, 24 and 36 months), and adults were presented with videos of two dyadic conversations, one with normal and one with flattened intonation. The first main finding was that only the 3-year-olds and the adults were able to reliably anticipate a speaker's next turn. Younger children shifted their gaze between the speakers regardless of the turn-taking. This indicates that extensive language experience is necessary, before semantics are developed sufficiently to anticipate the course of a conversation. The second main finding was that only 3-year-olds benefited from intonation; in younger age groups, as well as adults, the anticipation of turns was not affected by intonation. This suggests that intonation only has a supporting role on conversation perception, when language understanding is well developed but still not as sophisticated as that of adults. Thus, language experience alters the proficiency to use this prosodic cue.

1.7. Considerations, Limitations, and Perspectives

The present dissertation aimed to investigate how the increasing experience in performing nonverbal and verbal actions and interactions influences children's and adults' perception of others' actions in the respective areas. The findings generally suggest a significant role of experience in the perception of others' actions. First, in the area of individual action, where a vast amount of studies has focused already on the link between experience and action perception, we could add to the existing literature that such a link is not instantly present but needs active experience to develop. Second, the perception of visually presented joint action is a relatively new research area. We could show that the perception of joint action by infants with virtually no experience in coordinated joint action themselves, is essentially different from that of individual action. Third, the perception of visually presented conversations is also a rather new research area. Here, we could show that it takes up to 3 years of language experience, for children to be able to reliably anticipate the course of a conversation, and that the age of 3 years is special concerning the use of intonation. Thus, each of the studies provided an exciting new piece of the puzzle on how children perceive nonverbal and verbal actions and interactions

The investigated actions differ in the cognitive requirements that are necessary to perform and to understand them. The perception of individual action necessitates representing the goals of ‘simple motor actions’ by one agent (or elementary motor acts, Csibra, 2007). The perception of joint action involves representing the joint goal of multiple agents, in addition to the representations of the agents’ simple motor actions. The perception of conversations involves representing the semantic and syntactic information of a conversation, as well as using prosodic cues, such as intonation. The findings of the present dissertation thus demonstrate a significant role of experience on action perception on different cognitive levels, from simple motor actions to complex conversations. Experience therefore plays a special role during development, when new actions and skills are learned constantly. However, it seems natural to conclude that experience continues to play a significant role for action perception throughout one’s life, which is supported by studies with older and younger adults (Cross et al., 2006; Diersch, Cross, Stadler, Schütz-Bosbach, & Rieger, 2012).

A critical aspect that concerns the present work and others’ that address the influence of experience on action perception, is that it is difficult to pinpoint exactly what active and passive experience contribute to children’s action perception. During typical development, both are always entangled. Children learn to produce new actions through active experience, and this is known to modulate their action perception, but they likewise observe other people’s actions constantly. Empirically, when the production and the perception of an action were measured and the order was counterbalanced, production has been found to have a measurable effect on subsequent action perception in infants, whereas this was not the case vice versa (Cannon et al., 2012; Sommerville et al., 2008; Sommerville, Blumenthal, Venema, & Braun, 2011; Hauf & Prinz, 2004). A notable exception forms observational learning of some means-end, or action-effect relations, which has previously been found in infants and young children (Provasi, Dubon, & Bloch, 2001; Abravanel & Gingold, 1985; Hauf & Aschersleben, 2008). For the common representation study (chapter two), the production of contralateral reaching was tested and thus active experience targeted. The study showed a modulating influence of infants' production skills on their perception, which corroborates the predominant role of active experience. In the joint action study (chapter three), no such measure was obtained but it was assumed that older children and adults have more experience than younger children. Here, the role of active experience on the perception of joint action is thought to be comparable to that of individual action (Henderson et al., 2013; Schmitow & Kochukhova, 2013). Yet, the influence of extensive visual experience on the production, and consequentially on the perception of actions during normal development should not be underestimated. For example, the notion of a positive impact of action observation on the motor system has been supported by an interesting study concerning rehabilitation after stroke in adults (Ertelt et al., 2007). Here, physical training was combined with a concomitant action observation therapy, in which participants observed everyday actions over a period of 4 weeks. Participants' motor functions improved significantly compared with a control group, which received only physical training. The reason for this was thought to be a reactivation of motor areas through visual training. In a study with elite basketball players and 'expert watchers' (basketball coaches and journalists), it has been found that, although the players could predict basketball shots earlier and more accurately than watchers and novices, both players and watchers showed an increase in corticospinal excitability during observation of basketball actions, whereas novices did not (Aglioti, Cesari, Romani, & Urgesi, 2008). This suggests that the motor system of individuals with extensive visual experience was activated comparably to that of individuals with active experience when observing domain-specific actions. The findings of the above studies show a persistent influence of visual experience on the motor system and thus on action production. Infants also show an activation of motor areas when observing actions that are implemented in their own motor repertoire (e.g., Southgate, Johnson, El Karoui, & Csibra, 2010), which could support an influence of visual experience on action production during typical development. Although this could not yet be determined experimentally – perhaps because in experiments, actions are usually only presented for a few seconds – a reciprocal relationship between perception and production of actions in infants, and therefore between active and visual experience, is likely. This is also predicted by the common coding approach (Prinz, 1990, 1997), and accepted in adults.

Similar to manual actions, the production and the perception of language are inextricably linked (Pickering & Garrod, 2013). However, the ability to understand language (i.e., semantics) precedes production during development (Benedict, 1979). Further, semantics is assumed to predominantly modulate the ability to predict the end of turns (De Ruiter, Mitterer, & Enfield, 2006; Magyari & de Ruiter, 2012), which modulated gaze behaviour in the conversation study (chapter four). Thus, there seems to be a predominant role of passive experience (language perception, or semantics) over active experience (language production) in the perception of conversations. Although the role of active and passive experience on the perception of nonverbal and verbal actions seems to be reversed, the main message remains the same: production and perception are linked, and thus active *and* passive experience can influence action perception.

As the results of the present work were obtained experimentally, they are necessarily reductionist, and the aspect of ecological validity has to be considered critically. For a topic as inherently social as action perception, this reductionism might appear somewhat unfortunate. However, the advantages of this approach outweigh concerns about its ecological validity. One issue is the missing context in the present studies, especially in the studies concerning nonverbal action (chapter two and three). The context in which an action is embedded usually provides additional information on others' intentions (Stapel, Hunnius, & Bekkering, 2012; Kilner et al., 2007) and supports action understanding (Iacoboni et al., 2005; Fogassi et al., 2005). Thus, the present results relate to the essential parts of action perception (that of the movement and the goal), neglecting contextual information. Disregarding the context provided unambiguous and valid information on how children anticipate others' action goals. Another issue is that participants watched videos of actions instead of a 'live performance'. The social aspect of action perception is reduced in videos, which could be due to the two-dimensional nature of stimuli and/or because observers have no motivation to react towards the actors. Further, adults and infants have been found to show decreased activation of motor areas when observing video stimuli, compared with live stimuli (Shimada & Hiraki, 2006). For these reasons, it might have been more difficult for participants to build representations of the actions observed in video sequences. Although other infant studies have implemented live performance of actions (e.g., Woodward, 1998), those have analysed looking time. In order to measure anticipatory gaze, the exact analysis of eye movements is essential, which precludes the use of live performances if remote eye trackers are to be used. Besides this technical issue, the use of videos in the present work provided further benefits, including improved objectivity and economic advantages.

We chose to use a single-method approach to investigate the development of action perception, because the use of anticipatory gaze provides equal insight into infants', children's and adults' anticipation of action goals, and can be realised readily with any age group. In future research, our findings could be extended by using additional methods, for example, electrophysiological measures, such as electroencephalography (EEG), or spectroscopic measures, such as near infrared spectroscopy (NIRS). Both can also be adopted with infants, and could provide additional information on action perception. With EEG and NIRS, it is possible to measure neural activity related to motor activation and thus participants' motor simulation during action perception.

Furthermore, the present findings give rise to a more fine-grained investigation of the developmental trajectories of action perception in the future. In the common representation study (chapter two), we found a link between perception and action of contralateral reaching in 12-, but not in 6-month-olds. It is possible that such a link is already established earlier, with less than 6 months active experience in contralateral reaching. In the joint action study (chapter three), the older infant group of 12-month-olds still showed a differential perception of individual and joint action, although less so than 9-month-olds. We have reason to believe that children in the first half of their second year of life would show a gaze behaviour comparable to that of adults (Henderson & Woodward, 2011; Fawcett & Gredebäck, 2013), but within the present paradigm, this remains to be determined experimentally. Similarly, in the conversation study (chapter four), 3-year-old children's gaze behaviour was approaching that of adults regarding their general ability to anticipate turns, but in contrast to adults, they showed differential perception of conversations with and without intonation. It is an interesting question at which age children are able to compensate the missing intonation by relying on language understanding. Thus, in the same way as the studies of this dissertation provide new and intriguing findings, they open a variety of new questions.

Coming back to the quote by Julius Caesar 'Experience is the teacher of all things', our findings suggest that experience 1) teaches our brain to link action perception to action production, 2) teaches us to perceive joint action similarly to individual action, and 3) teaches us to perceive conversations based on language understanding as opposed to formal prosodic cues such as intonation. Furthermore, the present dissertation added to the increasing literature, showing that infants have remarkable skills in identifying others' goals, and become 'familiar' with others' actions more quickly than Jean Piaget's solipsistic view implied. In the following, the studies summarised above will be presented in detail.

Chapter 2

Production and perception of contralateral reaching: A close link by 12 months of age

Shortened title: Production and perception of reaching

Anne Melzer[1], Wolfgang Prinz[1] and Moritz M. Daum[1]

[1]Max-Planck-Institute for Human Cognitive and Brain Sciences

Infant Behavior and Development, 2012, 35, pp. 570 – 579.

Abstract

The goal of the present study was to measure infants' action production and perception skills with tasks that both include goal anticipation, in a within-subject design. In the production task, the frequency of 6- and 12-month-old infants' contralateral reaching movements was examined. In the perception task, videos of contralateral movements being performed were presented to the same infants and anticipatory eye movements were analysed. The main findings were: (1) 12-month-olds used their contralateral hand more frequently than 6- month-olds; (2) 12-month-olds mainly anticipated the goals of observed actions, whereas 6-month-olds mainly followed the action; finally, and most importantly, (3) at 12 months, production and perception were linked, but at 6 months, this was not yet the case. Our results show that anticipatory eye movements do not instantly reflect infants' reaching production. A certain amount of experience is required to establish a common representation of the production and the perception of reaching movements.

1. Introduction

The production of an action and the perception of the same action are closely linked in adults. The interplay of these two skills is described in detail by the Common Coding Principle (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1990, 1997). It specifies that production and perception of actions share a common abstract representation, providing a basis for a bidirectional influence between both action production and perception. In adults, it is well established that action perception can facilitate or interfere with action production, which is usually examined in shortened or prolonged reaction times in the execution of actions (Brass, Bekkering, & Prinz, 2001; Brass, Bekkering, Wohlschläger, & Prinz, 2000; Craighero, Bello, Fadiga, & Rizzolatti, 2002; Kilner, Paulignan, & Blakemore, 2003). In the same manner, action production can facilitate or interfere with the processing of observed actions (Hamilton, Wolpert, & Frith, 2004; Jacobs & Shiffrar, 2005; Miall et al., 2006; Wühr & Müsseler, 2001). The idea of a common representation of action production and perception gained support with the discovery of the mirror-neuron system (MNS) in monkeys and humans, which consists of neurons that fire both when an action is produced and when it is observed (for a review, see Rizzolatti & Craighero, 2004). In contrast to the well-documented common representation of action production and perception in adults, less is known about its development. Some authors consider this link innate (Bertenthal & Longo, 2007; Lepage & Théoret, 2007) whereas others think it develops at some point early in life (Del Giudice, Manera, & Keysers, 2009; Keysers & Perrett, 2004).

In infant research, the link between action production and action perception is often only theoretically addressed. Commonly, the results obtained in an observation paradigm (e.g. looking times) are related to a behavioural skill that is typically present in older but not in younger infants, without measuring individual behavioural skills (Daum & Gredebäck, 2011; Daum, Vuori, et al., 2009; Falck-Ytter et al., 2006; Longo & Bertenthal, 2006). However, to investigate the development of this link in infants, it would be preferable to measure both action production and perception, (a) within the same infants, and (b) using highly comparable tasks.

An increasing number of studies have investigated the link between action production and perception directly, using within-subject designs. For example, Daum et al. (2011) found a link between 6-month-old infants' grasping level (palmar vs. thump opposition) and their looking times to unexpected grasping actions (i.e. large hand aperture for small objects and vice versa; see also Loucks & Sommerville, 2012). Sommerville & Woodward (2005) found a link between infants' performance in a means-end task (pulling a cloth to retrieve a toy) and their perception of such an action in a visual habituation task. Here, infants' own level of planful behaviour was found to be correlated with their dishabituation time. A modulating effect of action production on perception was found in several studies, for example, using different means-end paradigms (Needham, Barrett, & Peterman, 2002; Cannon et al., 2012), which also supports the idea of a link between both competencies. Sommerville et al. (2005) used grasping actions and demonstrated an influence of action production on action perception in infants as young as 3 months of age. An effect of action perception on infants' own action production (usually in observational learning paradigms) has also been shown using actions such as button presses (Hauf & Aschersleben, 2008), tool-use paradigms (Abravanel & Gingold, 1985) or means-end paradigms (Provasi et al., 2001).

A fundamental problem that concerns most of these studies is the comparability of tasks for action production and perception. Action production per se includes the anticipation of a goal (von Hofsten, 2004). Performing a goal-directed action automatically entails the anticipation of a future event or a goal state. In contrast, many of the tasks used to measure action perception include a post hoc evaluation of an action. That means, the outcome of an action is presented and infants' reactions (often looking times) to this outcome are measured after the action is completed. Examples are the preferential looking paradigm (Cicchino & Rakison, 2008), habituation paradigms (Sommerville et al., 2008; Sommerville & Woodward, 2005) and observational learning (Abravanel & Gingold, 1985; Hauf & Aschersleben, 2008). Recent studies have found a dissociation between the post hoc evaluation, assessed by looking times or pupil dilation, and anticipatory measures assessed by anticipatory gaze shifts (Daum, Attig, Gunawan, Prinz, & Gredebäck, in press;

Gredebäck & Melinder, 2010). Measures that assess the *post hoc evaluation* of an action outcome were found to indicate much earlier comprehension of action goals in infants than measures that assess the *anticipation* of an action outcome.

The goal of the present study was to bridge this gap. Infants' processing of goal-directed reaching actions in the perception task was studied via anticipatory eye movements, a measure that involves similar anticipatory components as the production of a goal-directed reaching action. Both of our tasks included the same action: contralateral reaching. In the action production task, infants' own ability to reach contralaterally was tested; in the action perception task, participants watched a model perform contralateral reaching and transport movements.

Reaching within one side of the body midline (i.e. ipsilateral reaching) develops at about 3–4 months of age (Morange & Bloch, 1996; White et al., 1964). Reaching across the body midline (i.e. contralateral reaching) develops somewhat later. Bruner (1969) called the observation that infants do not reach across the body midline from the beginning the 'mysterious midline barrier'. He wrote that 'if a toy is held before the hand of an infant after he has already grasped something in that hand, the contralateral hand will not reach across the midline to get it.' (p. 276). He suggested that the midline barrier disappears at 7 months (i.e. that infants suddenly reach contralaterally at this age). Supporting this, the onset of spontaneous midline crossing with one hand was found around 7 months of age (Morange & Bloch, 1996). However, if the ipsilateral arm was restrained, contralateral reaching was already found at the age of 4 months (Provine & Westerman, 1979). In the context of bimanual reaching (i.e. using both hands to grasp a large object on one side of the body midline), spontaneous midline crossing could be observed in infants as young as 3 months; hence, bimanual reaching can be considered a precursor of unimanual contralateral reaching (van Hof, van der Kamp, & Savelsbergh, 2002). The different results concerning the age of onset can be accounted for by differences in the methods that were used (i.e. spontaneous vs. restrained, and unimanual vs. bimanual reaching). Interestingly, even though the age of onset varied, all studies found a rather slow increase in the use of contralateral reaching, as opposed to a strict threshold proposed by Bruner. It is only at the age of 8–9 years that an adult-like level of spontaneous midline crossing is achieved (Cermak, Quintero, & Cohen, 1980; Schofield, 1976).

In the present study, infants' contralateral reaching was tested adopting the paradigm described by Bruner (i.e. occupying one hand and presenting a second toy). This method presents a compromise between restricted and spontaneous midline crossing. An extensive pilot study was conducted to get an exact developmental outline of contralateral reaching with this method. Fifty-eight participants in four age groups were tested (6, 7, 8 and 14 months). Results yielded a linear increase in contralateral reaching,[1] indicating that this method makes it possible to differentiate between levels of contralateral reaching production in infants. At 6 months, infants are well able to reach ipsilaterally. We measured how often infants 'broke the habit' and used the contralateral hand. Thus, the task quantified their skill, or propensity, to reach contralaterally. It was, however, not appropriate to examine differences in infants' ipsilateral reaching skills as a decrease in ipsilateral reaching with this paradigm does not mean that infants' ipsilateral reaching skills deteriorate. The slow increase of contralateral reaching with this paradigm is particularly useful to study the development of a link between action production and perception.

In the perception task, participants' gaze behaviour was evaluated. Measurement of anticipatory eye movements has been used as an indicator for action perception in adults (Flanagan & Johansson, 2003) as well as in infants (Falck-Ytter et al., 2006; Gredebäck et al., 2009; von Hofsten et al., 2009). In their seminal study, Flanagan and Johansson (2003) recorded eye movements of participants during the execution of a block-stacking task and during their perception of the same task. They found highly similar gaze patterns engaged in the production of actions and the perception of others' actions, which strongly supports a common underlying mechanism for both competences. This method is easily adaptable to young infants (Falck-Ytter et al., 2006) and provides a powerful instrument for assessing infants' action perception.

[1] Mean percent values and standard deviations of contralateral reactions in four age groups: 6-month-olds (N = 14): M = 19.58%, SD = 14.00%; 7-month-olds (N = 15): M = 23.76%, SD = 12.34%; 8-month-olds (N = 15): M = 28.50%, SD = 13.72%; 14-month-olds (N = 14): M = 42.37%, SD = 10.67. Linear trend: $R^2 = .31$, $p < .001$.

The idea for the complementary perception task in this study was based on the study by Falck-Ytter et al. (2006), who presented videos of a model reaching for a ball and placing it in a bucket, and recorded each participant's gaze. They found that 6-month-olds showed reactive eye movements towards the goal of the action, whereas 12-month-olds, as well as adults, showed anticipatory eye movements. We modified the paradigm by adding a second condition with contralateral arm movements performed by the model. Since we exclusively measured the progress of contralateral reaching in the production task, only observed contralateral movements were analysed to match the production task. Additionally, and in contrast to Falck-Ytter et al. (2006), not only transport movements but also reaching movements were analysed.

The action production and perception tasks included similar anticipatory components of the same action. This allowed us to compare the production and the perception of this action in a within-subject design and gain more insight into the development of a common representation of both competences. Based on the study by Falck-Ytter et al. (2006), two age groups of infants were tested: 6- and 12-month-olds. An additional group of adults participated in the perception task to provide information about the developmental ‘end point’ of action anticipation. These data were not included in the statistical analysis, however, as not to distort the results of children's development.

We expected to find a correlation between the two tasks if action production and perception of contralateral reaching already share a common representation in the first year of life. The crucial question was, however, at what point in development such a common representation develops. It is possible that a common representation is present as soon as infants start performing an action, or that a certain amount of experience is necessary. In the former case, we would find a correlation at 6 and 12 months of age, in the latter case, such a correlation would only be present at 12 months of age.

2. Method

2.1. Participants

In each of the three age groups, 24 participants successfully took part in the study. Groups consisted of 6-month-old infants (15 girls, 9 boys; mean age = 6 months; 0 days; age range = 5;15 to 6;15), 12-month-old infants (10 girls, 14 boys; mean age = 12;3; age range = 11;19 to 12;13), and adults (12 female, 12 male; mean age = 25.4 years; age range = 20–33).

Additionally, twenty-eight 6-month-old infants were tested but were not included in the final data analysis, with two infants not providing enough trials in either task. In the production task, a total of three infants did not complete the task due to fussiness. In the perception task, data from twenty-seven 6-month-olds (including two also excluded from the production task) were excluded from the final analysis because they were inattentive to the stimuli ($n = 19$) or because of insufficient signal quality ($n = 8$). The high exclusion rate in the present study is partially due to a block design (see Procedure Section 2.3.). However, it is still comparable with the one of other eye tracking studies with young infants, where it is often found to be approximately 50% (for an overview, see Haith, 2004). Ten 12-month-old infants were not included in the final data analysis because they did not provide enough trials in the production task ($n = 6$) or in the perception task ($n = 4$; 2 were inattentive and 2 had insufficient signal quality). The data from one additional adult who participated in the action perception task had to be excluded due to insufficient signal quality. Contact information of infants was obtained from public birth records; the group of adult participants consisted of students. Infants, as well as adults, received a small gift as compensation for their participation. The study was approved by the local ethics committee of the University of Leipzig, and conducted in accordance with the Declaration of Helsinki.

2.2. Apparatus and Stimuli

Both tasks were conducted in the same testing room, which was divided into two separate parts by an opaque curtain. The room was unfurnished apart from the testing equipment.

For the production task, infants and parents sat on a chair in front of a table with a cut out area (see Figure 2.1), and the experimenter sat opposite them. To occupy one of the infant's hands, a small red cube (width = 2.5 cm) was used. Three different toys (roughly 6 cm × 6 cm), attached to 25 cm long sticks, were used as second object to be held in front of the infant. The toys were designed to be very appealing for infants; they were all multi-coloured and made different sounds (rattling and different chiming sounds) when they were moved.

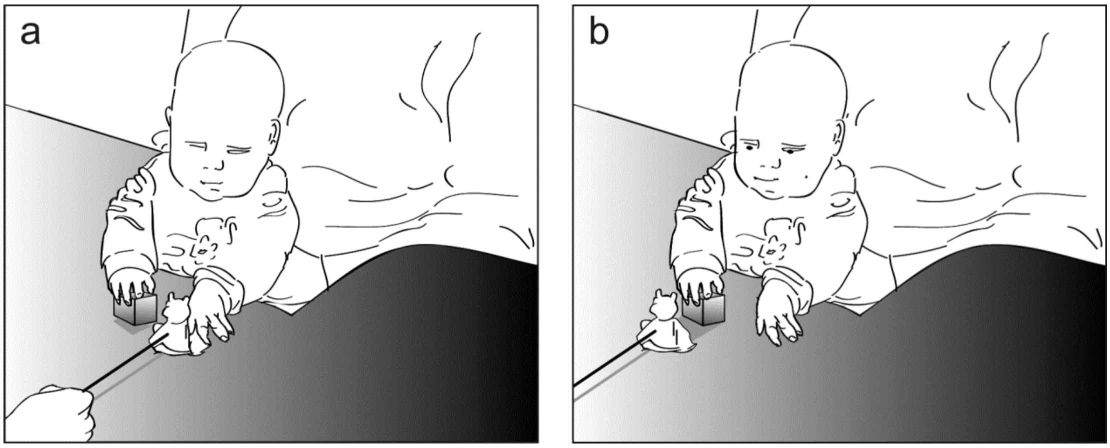

Figure 2.1. Schematic drawing of the production task. The infant was given a cube (occupation object) in either his/her left or right hand. Subsequently, a second toy (target object) was held either (a) in front of the empty hand to elicit an ipsilateral reaction (ipsilateral presentation) or (b) in front of the occupied hand to elicit a contralateral reaction (contralateral presentation).

Videos in the perception task were presented on a corneal reflection eye tracker (Tobii 1750, Stockholm, Sweden) with an infant add-on (precision: 1°, accuracy: 0.5°). Videos subtended a visual angle of approximately 27.5° × 15.2°. Stimulus presentation and data acquisition were accomplished using the software ClearView 2.7.1. Eye movements were recorded at a rate of 50 Hz. A 9-point infant calibration was used. Infants sat in a safety car seat (Maxi Cosi Cabrio) at a distance of approximately 60 cm from the monitor.

Each video showed a model reaching for a ball and dropping it into a bucket; the action therefore consisted of two separable movements: reaching and transporting. The bucket had a chicken head on it which moved and made a sound when the ball was dropped into it to make the goal of the entire action more salient (Falck-Ytter et al., 2006). Two different versions of the video were presented. One showed the model starting with an ipsilateral reaching movement, followed by a contralateral transport movement from the location of the ball to the bucket in the transport phase (referred to as ipsi-first trial; Figure 2.2 a). The other video showed the model starting with a contralateral reaching movement, with the corresponding transport phase consisting of an ipsilateral transport phase (referred to as contra-first trial; Figure 2.2 b). Only the contralateral reaching or transport movement of each trial was analysed. The colour of the ball differed between trials (red, blue and yellow); in all other respects, the videos were identical. Mirrored versions of all videos were produced to counterbalance potential hand and side effects. Each video had a total duration of 5280 ms. The reaching movement started 400 ms after video onset (during the first 400 ms, a still picture was shown, see left panel of Figure 2.2). Reaching movements took 760 ms and transport movements took 1120 ms. Between trials, salient attention grabbers were presented, consisting of a moving toy accompanied by a brief attention-grabbing sound (visual angles approximately 7° × 8°).

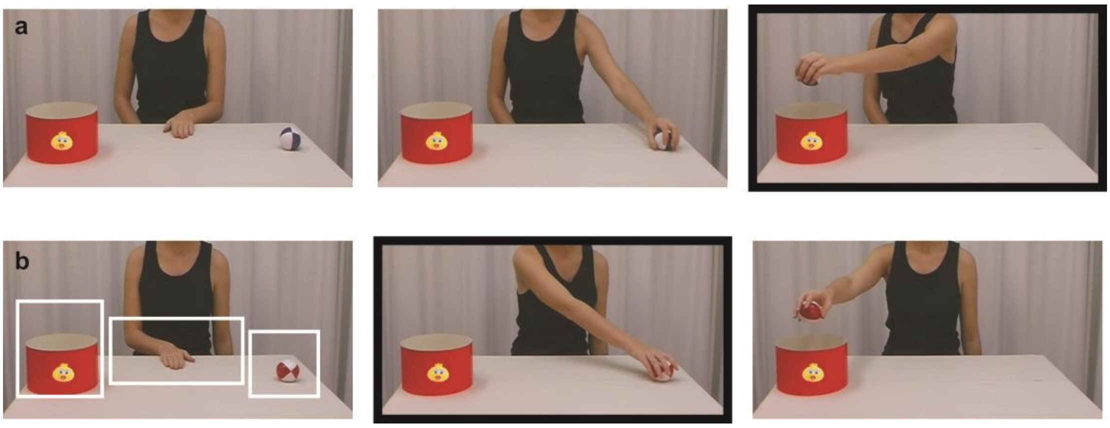

Figure 2.2. Still pictures of stimulus videos in the perception task. Pictures show the starting position of the reaching movement (left column), end position of the reaching movement/starting position of the transport movement (middle column) and end position of the transport movement (right column). Upper row (a) displays an ipsi-first trial, bottom row (b) a contra-first trial including AOIs indicated by white boxes. The pictures in black frames indicate end positions of the contralateral movements that were analysed.

2.3. Procedure

Upon arrival, parents were informed about the procedure (without disclosing the exact purpose of the study) and gave informed consent. After the infant was familiarised with the experimenter, the parent and infant were led to the testing room. All infants were tested individually with one parent present. The perception task was always conducted before behavioural testing of the infant's reaching production. This procedure has also been used in previous studies (Daum, Vuori, et al., 2009; Daum & Gredebäck, 2011). More importantly, we were not interested in immediate effects of action production on action perception or vice versa, but in an underlying, existing common representation. For this matter, it was advisable to conduct the perception task first in order to weaken potential interference of any immediate learning effects. Observational experience has been found to be less likely to have an influence on infants' action production than vice versa (Cannon et al., 2012; Sommerville et al., 2011). It is only in this section that the perception task will be described before the production task.

Perception task. The parent placed the infant in the safety car seat and sat down on a chair behind the infant. Once infant and parent were comfortable, the stimulus presentation was started. Parents were allowed to comfort their child if necessary but were otherwise asked to remain silent. After the calibration procedure, the experimenter started the video presentation. A maximum of 30 trials were presented (each consisting of a reaching and a transport movement) in order to obtain a reliable measure of infants' gaze behaviour. Ipsi-first and contra-first trials were presented in two separate blocks of 15 trials each. The order of the presentation of the blocks was counterbalanced across participants. After three consecutive trials, the attention grabber was shown until the infant looked at the screen before the next three trials were presented. The task was terminated if infants became fussy during presentation. The calibration procedure took 1–3 min, and testing of the perception task took approximately 3.5 min.

Production task. Subsequently, the production task assessing the infant's own contralateral reaching behaviour was conducted by the same experimenter in the other part of the testing room. The infant sat on his or her parent's lap and was given a small cube (occupation object) in either their left or their right hand (order counterbalanced). Subsequently, a second toy (target object) was held alternately either in front of the infant's empty hand (to elicit an ipsilateral reaction; Figure 2.1 a) or in front of their occupied hand (to elicit a contralateral reaction with the empty hand; Figure 2.1 b). After the infant reacted, or after a waiting period of at least 5 s, this procedure was repeated. After six trials, the occupation object was placed in the other hand and the procedure repeated. This resulted in a total of 12 trials (6 ipsilateral and 6 contralateral presentations). The reactions to both forms of presentations were analysed. Ipsilateral presentations were not absolutely necessary for assessing contralateral reaching skills but enhanced the number of valid trials per infant and were also encouraging infants to continue with the task. A trial was repeated immediately if the infant dropped the occupation object or gave it away before the target object was presented. The behavioural testing procedure took approximately 6 min. Overall, each parent-infant pair spent approximately 1 h in the laboratory.

The adults only participated in the perception task, since adults are capable of reaching contralaterally (Helbig & Gabbard, 2004) and no variance in their reaching production was expected within the present paradigm. After they were informed about the procedure, they gave informed consent and were led to the testing room. Adults watched the same videos as the infants.

2.4. Coding and Analysis

Production task. The frequency of ipsilateral and contralateral reactions in the production task was analysed from video recordings. Infants who showed fewer than six valid trials were excluded from further analysis. A trial was categorised to be valid if the infant focused on the toy when reaching for it to avoid counting non-goal-directed behaviour. Furthermore, we only included trials where the infant unmistakably held the cube in one hand until a reaction was shown. As reactions, we considered taking, grasping or touching the toy with one hand or with both hands. Bimanual reactions were also included in the category of contralateral reactions, since they always include the contralateral hand. We furthermore considered trials as valid where the infant looked at the toy for at least 5 s but did not move; these trials were labelled as ‘no reaction’.

The mean number of valid trials was 8.9 (74.2%) for 6-month-olds and 9.0 (75.0%) for 12-month-olds. The fact that we obtained a nearly identical number of valid trials supports the assumption that the task could be managed equally well by both infant age groups. Furthermore, there was no difference between the number of valid ipsilateral and contralateral presentations that infants received ($p > .33$ in both age groups). Thus, the reactions to both forms of presentations could be merged. The number of different reactions was calculated as relative frequency, resulting in percent values of contralateral, ipsilateral and no reactions for each infant. For example, if an infant were to always grasp the toy with the empty hand, it would get an ipsilateral reaching score of 50%, a contralateral reaching score of 50%, and 0% no reactions. This means, a contralateral reaching score of 50% reflects the most sophisticated reaching production.[2] If, however, an infant were to touch the toy with the occupied hand (or drop the occupation object in order to grasp the toy) in 3 trials, it would get an ipsilateral score of 75% and a contralateral score of 25%. All trials were coded by two independent raters, who assigned the same category in 92.2% of the trials.

Perception task. Eye movements were analysed using Matlab 7.1 (The MathWorks Inc.). Three areas of interest (AOIs) were defined (see Figure 2.2 b) surrounding the starting point of the reaching movement (visual angle 10.5° × 5.6°), the ball (5.5° × 5.6°) and the bucket (6.8° × 8.3°). The time it took for the gaze to arrive (gaze arrival time) at goal areas (i.e. the ball AOI and the bucket AOI) relative to the time the model's hand arrived was calculated. Eye movements were categorised as anticipatory if the gaze preceded the arrival of the hand (positive gaze arrival times). If the gaze followed the hand, eye movements were classed as reactive (negative gaze

[2]A score of 50% contralateral and 50% ipsilateral reactions could, in theory, also be achieved if an infant reacted in a different, and perhaps non-sophisticated manner, namely, if he or she always reacted with the occupied hand. However, the unusual case that infants reacted with their occupied hand during ipsilateral presentation (Figure 2.1 a) only occurred in 1.9% of presentations.

arrival times). In addition to these exact gaze arrival times, the relative frequency of anticipatory eye movements is reported (i.e. number of anticipated movements divided by the total number of observed movements). Only observed contralateral arm movements were taken into account. These were subsumed regardless of their occurrence in the reaching or transport phase. Trials were only included when participants looked at the starting position of the movement before they shifted gaze to the ball AOI or bucket AOI. The minimum time participants had to look at the starting position of the movement was set to 100 ms. This restriction ensured that only trials where the arm movement was watched from the beginning were taken into account, and reduced the likelihood of including random gaze shifts in the data analysis. Only participants who attentively watched at least 10 of the 30 contralateral movements were included in further analyses to ensure the reliability of data. The mean number of valid contralateral movements observed was 15.7 (52.2%) for 6-month-olds, 18.5 (61.6%) for 12-month-olds and 24.4 (81.4%) for adults.

Statistical analyses of the differences between age groups in the production task and the perception task were calculated using *t*-tests and Chi-Square tests. For analyses of the relationship between the production and the perception task, Pearson correlations were calculated. As correction for multiple testing, a Fisher's omnibus test was used. We report *p*values two-tailed throughout (except for the Fisher's omnibus test, where it is generally one-tailed); a significance level of .05 was adopted for all statistical tests in this study.

3. Results

3.1. Production Task

The relative frequency of contralateral, ipsilateral and no reactions for 6- and 12-month-old infants are shown in Figure 2.3. The frequency of contralateral reactions increased with age (from $M_{6\ months} = 18.88\%$, SD = 15.93% to $M_{12\ months} = 30.72\%$, SD = 15.35%, $t(46) = 2.26$, $p = .01$. Likewise, the frequency of ipsilateral reactions decreased (from $M_{6\ months} = 80.42\%$, SD = 15.43% to $M_{12\ months} = 67.59\%$, SD = 15.82%), $t(46) = 2.85$, $p < .01$. No difference was found in the frequency of no reactions (from $M_{6\ months}$ = 0.69%, SD = 3.40% to $M_{12\ months}$ = 1.69%, SD = 5.08%), $t(46) = 0.80$, $p = .43$.

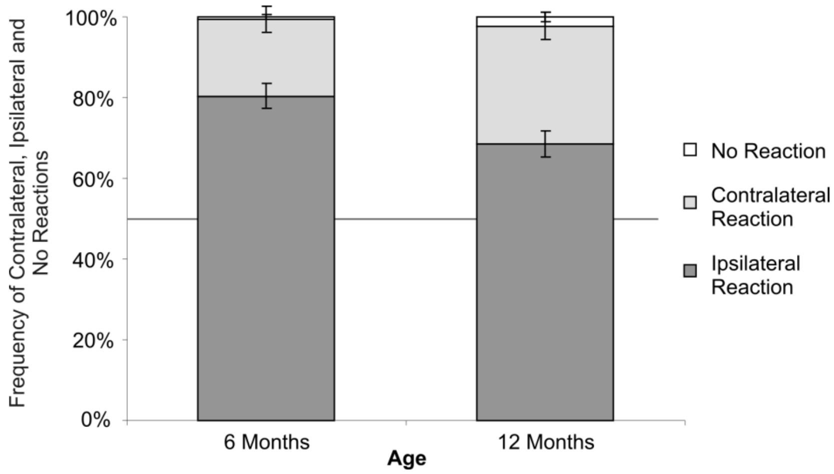

Figure 2.3. Infants' reaching production. Mean frequency of infants' reactions in the production task with standard errors. A frequency of 50% each for ipsilateral and contralateral reactions would represent the most sophisticated reaching production possible. From 6 to 12 months, the frequency of contralateral reactions increased (p = .03).

3.2. Perception Task

Gaze arrival times and relative frequencies of anticipated contralateral movements in the perception task of 6- and 12-month-old infants, and also for adults, are shown in Figure 2.4. A *t*-test between 6- and 12-month-old infants for mean gaze arrival times reached significance $t(2,46) = 8.40$, $p < .001$; 12-month-old infants showed earlier gaze arrival times than 6-month-olds. Likewise, a *t*-test for relative frequencies of anticipated contralateral movements showed that 12-month-olds anticipated more contralateral movements than 6-month-olds ($M_{6\ months}$ = 19.1%, SD = 3.2%, $M_{12\ months}$ = 61.8%, SD = 3.8%), $t(2,46) = 8.46$, $p < .001$.

Six-month-old infants showed mainly reactive eye movements (i.e. gaze arrival times smaller than zero; M = –207.12 ms, SD = 179.11 ms), $t(23) = 5.67$, $p < .001$. By contrast, 12-month-old infants showed mainly anticipatory eye movements (i.e. gaze arrival times larger than zero; M = 197.85 ms, SD = 154.05 ms), $t(23) = 6.29$, $p < .001$, as did the adults (M = 413.66 ms, SD = 113.66 ms), $t(23) = 17.83$, $p < .001$. The number of participants who anticipated the goals of observed movements confirmed these data. In the group of 6-month-old infants, 21 of 24 showed overall gaze arrival times smaller than zero, $\chi^2 = 13.5$, $p < .001$. In the group of 12-month-old infants, 22 of 24 showed gaze arrival times larger than zero, $\chi^2 = 16.67$, $p < .001$; as did all adults, $\chi^2 = 24$, $p < .001$.

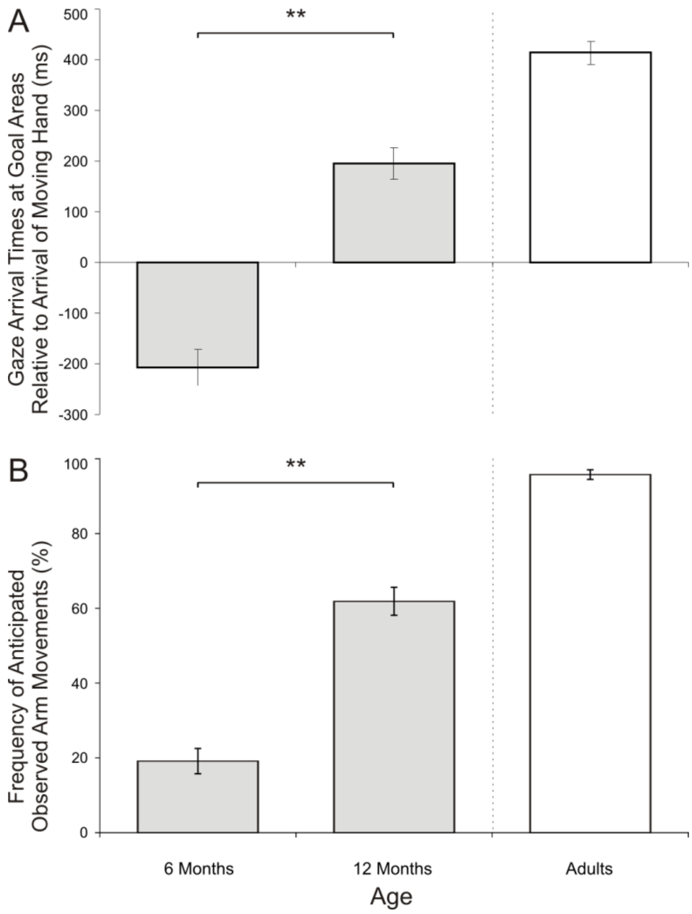

Figure 2.4. Gaze behaviour with respect to contralateral movements observed. (a) Bars show mean gaze arrival times relative to the arm movement observed, with standard errors. Time point zero on the y-axis refers to the moment in the video when the hand enters the goal area (i.e. either ball AOI or bucket AOI). Positive values indicate that gaze arrived at the goal areas before the hand did; negative values indicate that gaze arrived after the hand. (b) Bars show mean relative frequency of anticipated arm movements, with standard errors. The difference between 6- and 12-month-olds reached significance (**p < .001). Gaze behaviour of adults is informative and was not included in statistical analyses.

3.3. Interrelation of Production and Perception Task

Correlations between the performance in the production task and gaze behaviour in the perception task were separately calculated for both infant age groups. As measurement of the performance in the production task, we calculated the ratio of contralateral to ipsilateral reactions: $N_{contralateral}/N_{ipsilateral}$. This resulted in a value between 0 and 1, with 1 being the most sophisticated reaching production.

As measurement of gaze behaviour in the perception task, gaze arrival times at goal areas for contralateral movements were used. For the 6-month-old infants, no significant correlation was found, $r = .22$, $p = .31$. By contrast, for the 12-month-old infants, a significant correlation between their reaching production and observed contralateral movements was found, $r = .45$, $p = .03$. This means, the more sophisticated infants' reaching production was at 12 months, the earlier they anticipated other's contralateral movements.

To rule out the possibility that the correlation between the action production and perception task in 6-month-olds did not become significant because their data was noisier, a Levené's test for variance homogeneity was conducted. Participants in the two infant age groups did not show differences in variance in either the production task or the perception task ($F(1,46) = 0.00$, $p = .96$; $F(1,46) = 1.58$, $p = .22$, respectively). Furthermore, to assess the accuracy of correlations in the two infant age groups, a permutation technique was used (i.e. the jackknife; Efron, 1979). Here, one pair of the sample *n* is left out and a correlation coefficient is calculated with the remaining sample ($n - 1$). This procedure is repeated with each pair left out once. The resulting *n* correlation coefficients can be used to estimate a bias. This bias describes the 'tendency of the sample correlation to overestimate or underestimate the true, unknown correlation' (Efron & Tibshirani, 1998). In particular with infant participants, where the variability is usually very high, this is also a way to assess the impact of outliers. For 6-month-olds, our sample correlation ($r = .22$) probably overestimated the true correlation by $bias_{jack} = 0.004$ (resulting in a correlation of $r = .21$). For 12-month-olds, our sample correlation ($r = .45$) probably underestimated the true correlation by $bias_{jack} = -0.023$ (resulting in a correlation of $r = .47$). This can be interpreted as follows: the correlations we found in our samples are in good accordance with the true correlations in the population. In 12-month-old infants, the true correlation is probably slightly higher than in our sample.

Additionally, we calculated the correlation between the performance in the production task and the frequency of anticipated contralateral movements for both infant age groups. Again, there was no significant correlation for 6-month-olds ($r = .16$, $p = .47$). For 12-month-olds, the correlation approached significance ($r = .35$, $p = .09$). Using the jackknife technique again, a bias was estimated that indicated how accurate these correlations were. In 6-month-olds, the true correlation was probably underestimated by $\text{bias}_{\text{jack}} = -0.018$ (resulting in a correlation of $r = .17$). In 12-month-olds, the true correlation was probably underestimated as well, by the value $\text{bias}_{\text{jack}} = -0.029$ (resulting in a correlation of $r = .38$). This means, our sample correlations were close to, or slightly lower than, the true correlations in the population.

To account for multiple testing, a Fisher's omnibus test was used, combining the *p* values of the two correlations in 12-month-olds (Haccou & Meelis, 1994). This procedure revealed significance ($\chi^2 = 11.95$, df = 4, $p < .02$), which indicates that these results were not a by-product of multiple testing, but that indeed there was a correlation between action production and perception in 12-month-olds.

4. Discussion

In the present study, the aim was to shed more light on the development of a common representation of action production and perception with two tasks that focused on anticipatory components and were more comparable than those reported in previous studies. For this purpose, we combined a paradigm for testing the production of contralateral reaching with a paradigm for testing the perception of contralateral reaching and transport movements in a within-subject design. The tasks were conducted with 6- and 12-month-old infants. In the production task, contralateral reaching was tested using a paradigm proposed by Bruner (1969); this yielded an increase in contralateral reaching with age. In the perception task, infants watched videos of a model performing reaching and transport movements while their eye movements were recorded. At 12 months, most infants were able to anticipate the goal of contralateral movements, whereas at 6 months, infants showed mainly reactive eye movements. Production and perception of contralateral reaching movements were correlated at 12 months of age. The more sophisticated 12-month-olds' reaching production was, the better they anticipated other people's contralateral movements. Importantly, perception and production were not yet correlated at 6 months. The lack of a significant correlation was neither due to a larger variance in the younger infant group nor to the influence of a bias in our sample. Accordingly, our findings suggest that a link between production and perception of contralateral arm movements, and possibly therefore a common representation, develops in the second half of the first year of life.

Our findings extend those of previous studies which showed a link between action production and perception in infants (Abravanel & Gingold, 1985; Hauf & Aschersleben, 2008; Provasi et al., 2001; Sommerville et al., 2008; Sommerville & Woodward, 2005). However, our results did not yield a link in the group of 6-month-old infants. Daum et al. (2011) found a link in 6-month-old infants, and a modulating influence of action production on action perception has been reported in infants as young as 3 months (Needham et al., 2002; Sommerville et al., 2005). These studies found an effect on the post hoc evaluation of actions like dishabituation time (Sommerville et al., 2005) or total looking time (Daum et al., 2011; Needham et al., 2002). These measures show that very young infants are able to evaluate the goal-directedness and rationality of an action after the action is completed. This ability does not necessarily depend on infants' own experience (Csibra, 2008; Gredebäck & Melinder, 2010). It is unclear which cognitive processes post hoc measures of action perception measure; very different factors like surprise, familiarity, etc. could have contributed to previous results. Anticipatory eye movements are a more suitable indicator to assess the link between action production and perception, because (1) the same underlying mechanism controls eye movements during the production and perception of actions (Cannon & Woodward, 2008; Flanagan & Johansson, 2003), and (2) anticipatory eye movements are more dependent on behavioural experience than post hoc measures (Gredebäck & Melinder, 2010). The present results indicate that a common representation of action production and perception that includes similar anticipatory demands requires some experience with the task. This is in line with previous findings which suggest a dissociation between post hoc and anticipatory measures (Gredebäck & Melinder, 2010).

The developmental outline of the interplay between production and perception of contralateral reaching can be illustrated as follows: At 6 months, infants' own reaching production is limited; contralateral reaching is still at a very early stage of development. Although infants are able to anticipate at 6 months, for example a re-occurring object behind an occluder (Kochukhova & Gredebäck, 2007), or feeding actions to the mouth (Kochukhova & Gredebäck, 2010), they rarely anticipate the goal of contralateral reaching and transport movements (see also Falck-Ytter et al., 2006). Furthermore, 6-month-olds' production and perception behaviour was not yet correlated. This means, eye movements do not instantly reflect infants' behavioural skills. We can conclude further that the emerging ability to produce an action does not instantly feed a common representation of action production and perception.

At 12 months, infants' reaching production has become more sophisticated but is still not perfect. At this age, most infants are able to anticipate the goal of observed contralateral movements (although, compared to adults, their gaze arrival at goals is slower and they show fewer anticipatory eye movements). At 12 months, a link between the production and the perception of contralateral movements is established.

Between 6 and 12 months of age, infants become more experienced in the *production* of reaching per se and of contralateral reaching in particular. It is likely that this increase in behavioural experience feeds the common representation of the competences. However, they likewise become more experienced in the *perception* of reaching movements. It is possible that experience in observing someone else reaching has a similar supportive function as the production of the same movements. Further research is needed to clarify this important issue.

In adults, the ability to anticipate goals of contralateral reaching and transport movements has further developed and gaze arrival is more than twice as fast as in 12-month-olds. We did not test the adults' ability to reach, however, because our paradigm was not suitable for showing differences in reaching production in adults;

any link between production and perception of contralateral reaching in adults would have been masked by ceiling effects. In adults, this link can be found reliably using other paradigms (for an overview, see Prinz et al., 2009).

Related to the common coding approach (Prinz, 1990, 1997), our findings suggest that a common representation of the production and the perception of a developing action is not instantaneously present but needs time and experience to develop. Whether or not the corresponding mirror-neuron system is present from birth cannot be answered with this study. We did not find a link between action production and perception at 6 months. This could mean that such a system is not present from birth (Del Giudice et al., 2009; Keysers & Perrett, 2004), or that the 'hardware' is innate (Bertenthal & Longo, 2007; Lepage & Théoret, 2007), but needs to be fed with specific actions.

To conclude, the present findings show that the perception of an action does not instantly reflect infants' own behavioural skills. Given similar anticipatory components included in both the action production and perception task, it appears that a considerable amount of experience is necessary to establish a common representation of action production and perception.

Acknowledgements

We thank Stephanie Trapp for the additional coding of behavioural data and Wiebke Rahmlow and Franziska Jung for performing the action sequences. Furthermore, we would like to thank Stephan Liebig for drawing the methods figure and Christian Keitel for valuable comments on the manuscript. This research was supported by the Deutsche Forschungsgemeinschaft (GRK 1182).

References

Abravanel, E., & Gingold, H. (1985). Learning via observation during the 2nd year of life. *Developmental Psychology, 21*(4), 614-623.

Bertenthal, B. I., & Longo, M. R. (2007). Is there evidence of a mirror system from birth? *Developmental Science, 10*(5), 526-529.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*(1-2), 3-22.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*(2), 124-143. doi: 10.1006/brcg.2000.1225

Bruner, J. S. (1969). Eye, Hand, and Mind In D. Elkind & J. H. Flavell (Eds.), *Studies in Cognitive Development: Essays in Honor of Jean Piaget*. Oxford: University Press, Inc.

Cannon, E. N., & Woodward, A. L. (2008). *Action anticipation and interference: A test of prospective gaze.* Paper presented at the Proceedings of the Meeting of the Cognitive Science Society, Austin, TX.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42. doi: 10.1111/j.1467-7687.2011.01095.x

Cermak, S. A., Quintero, E. J., & Cohen, P. M. (1980). Developmental age trends in crossing the body midline in normal children. *American Journal of Occupational Therapy, 34*(5), 313-319.

Cicchino, J. B., & Rakison, D. H. (2008). Producing and processing self-propelled motion in infancy. *Developmental Psychology, 44*(5), 1232-1241. doi: 2008-12114-004 [pii] 10.1037/a0012619

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*(5), 492-502.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705-717. doi: S0010-0277(07)00215-6 [pii] 10.1016/j.cognition.2007.08.001

Daum, M. M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (in press). Actions seen through babies' eyes: A dissociation between looking time and predictive gaze. *Froniers in Psychology*.

Daum, M. M., & Gredebäck, G. (2011). The development of grasping comprehension in infancy: Covert shifts of attention caused by referential actions. *Experimental Brain Research, 208*.

Daum, M. M., Prinz, W., & Aschersleben, G. (2011). Perception and production of object-related grasping in 6-month-old infants. *Journal of Experimental Child Psychology, 108*, 810 - 814.

Daum, M. M., Vuori, M. T., Prinz, W., & Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6-and 9-month-old infants. *Developmental Science, 12*(6), 854-862. doi: DOI 10.1111/j.1467-7687.2009.00831.x

Del Giudice, M., Manera, V., & Keysers, C. (2009). Programmed to learn? The ontogeny of mirror neurons. *Developmental Science, 12*(2), 350-363. doi: DOI 10.1111/j.1467-7687.2008.00783.x

Efron, B. (1979). Computers and the theory of statistics - Thinking the unthinkable. *Siam Review, 21*(4), 460-480.

Efron, B., & Tibshirani, R. J. (1998). *An introduction to the bootstrap*. Boca Raton: Chapman & Hall.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424*(6950), 769-771.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*(2), 197-206. doi: DOI 10.1016/j.cognition.2009.09.004

Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., Rosander, K., & von Hofsten, C. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology, 45*(4), 1190-1194. doi: Doi 10.1037/A0015667

Haccou, P., & Meelis, E. (1994). *Statistical analysis of behavioural data*. New York: Oxford University Press.

Haith, M. M. (2004). Progress and standardization in eye movement work with human infants. *Infancy, 6*(2), 257-265.

Hamilton, A. F., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14*(6), 493-498. doi: 10.1016/j.cub.2004.03.007

Hauf, P., & Aschersleben, G. (2008). Action-effect anticipation in infant action control. *Psychological Research-Psychologische Forschung, 72*(2), 203-210. doi: DOI 10.1007/s00426-006-0101-3

Helbig, C. R., & Gabbard, C. (2004). What determines limb selection for reaching? *Research Quarterly for Exercise and Sport, 75*(1), 47-59.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*(5), 849-937.

Jacobs, A., & Shiffrar, M. (2005). Walking perception by walking observers. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 157-169. doi: 10.1037/0096-1523.31.1.157

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences, 8*(11), 501-507. doi: DOI 10.1016/j.tics.2004.09.005

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*(6), 522-525. doi: 10.1016/s0960-9822(03)00165-9

Kochukhova, O., & Gredebäck, G. (2007). Learning about occlusion: Initial assumptions and rapid adjustments. *Cognition, 105*(1), 26-46. doi: DOI 10.1016/j.cognition.2006.08.005

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development, 81*, 1729-1738.

Lepage, J.-F., & Théoret, H. (2007). The mirror neuron system: Grasping others' actions from birth? *Developmental Science, 10*(5), 513-523.

Longo, M. R., & Bertenthal, B. I. (2006). Common coding of observation and execution of action in 9-month-old infants. *Infancy, 10*(1), 43-59.

Loucks, J., & Sommerville, J. A. (2012). The role of motor experience in understanding action function: the case of the precision grasp. *Child Development, 83*(3), 801-809. doi: 10.1111/j.1467-8624.2012.01735.x

Miall, R. C., Stanley, J., Todhunter, S., Levick, C., Lindo, S., & Miall, J. D. (2006). Performing hand actions assists the visual discrimination of similar hand postures. *Neuropsychologia, 44*(6), 966-976. doi: 10.1016/j.neuropsychologia.2005.09.006

Morange, F., & Bloch, H. (1996). Lateralization of the approach movement and the prehension movement in infants from 4 to 7 months. *Early Development & Parenting, 5*(2), 81-92.

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky mittens' enhances young infants' object exploration skills. *Infant Behavior & Development, 25*(3), 279-295. doi: Pii S0163-6383(02)00097-8

Prinz, W. (1990). A common coding approach to perception and action In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167 - 201). Berlin: Springer Verlag.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129-154.

Prinz, W., Aschersleben, G., & Koch, I. (2009). Cognition and action In E. Morsella, J. A. Bargh & P. M. Gollwitzer (Eds.), *Oxford handbook of human action.* (pp. 35-71). Oxford: Oxford University Press.

Provasi, J., Dubon, C. D., & Bloch, H. (2001). Do 9-and 12-month-olds learn means-ends relation by observing? *Infant Behavior & Development, 24*(2), 195-213.

Provine, R. R., & Westerman, J. A. (1979). Crossing the midline: limits of early eye-hand behavior. *Child Development*, *50*(2), 437-441.

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience. Vol 27 2004, 169-192.*

Schofield, W. N. (1976). Hand movements which cross body midline - Findings relating age-differences to handedness. *Perceptual and Motor Skills, 42*(2), 643-646.

Sommerville, J. A., Blumenthal, E., Venema, K., & Braun, K. (2011). The body in action: The impact of self-produced action on infants' action perception and understanding. In V. Slaughter & C. Brownell (Eds.), *Early development of body representations*. Cambridge: Cambridge University Press.

Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: the impact of doing versus watching on infants' subsequent perception of tool-use events. *Dev Psychol, 44*(5), 1249-1256. doi: 2008-12114-006 [pii] 10.1037/a0012296

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition, 95*(1), 1-30.

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

van Hof, R., van der Kamp, J., & Savelsbergh, G. J. (2002). The relation of unimanual and bimanual reaching to crossing the midline. *Child Development, 73*(5), 1353-1362.

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences, 8*(6), 266-272.

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders, 3*(2), 556-569. doi: 10.1016/j.rasd.2008.12.003

White, B. L., Castle, P., & Held, R. (1964). Observations on the development of visually-directed reaching. *Child Development, 35*(2), 349-364.

Wühr, P., & Müsseler, J. (2001). Time course of the blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 27*(5), 1260-1270.

Chapter 3

Infants' and adults' perception of individual and joint action

Shortened title: Perception of individual and joint action

Anne Keitel[1], Wolfgang Prinz[1] and Moritz M. Daum[1,2]

[1]Max-Planck-Institute for Human Cognitive and Brain Sciences

[2]University of Zurich

Submitted 04/10/2013 to *Cognition*.

Abstract

Infants and adults frequently observe actions performed jointly by more than one person. Research in action perception, however, has focused largely on actions performed by an individual person. Here, we explore how infants and adults perceive the same action performed by one agent (individual condition) or two agents (joint condition). We used eye tracking to measure the latency of participants' gaze shifts towards action goals. Adults anticipated goals in both conditions significantly faster than infants, and their gaze latencies did not differ between conditions. By contrast, infants showed faster anticipation of goals in the individual condition than in the joint condition. This difference was more pronounced in 9-month-olds. These results suggest that adults are able to infer the joint goal of two agents, whereas infants might represent the respective sub-goals in isolation.

1. Introduction

Practically from birth, infants observe the behaviour of the people around them, and they learn to anticipate the goals of others' actions during their first year of life (e.g., Falck-Ytter, Gredebäck, & von Hofsten, 2006). Recently, interest in how infants passively perceive others' *interactions* emerged, that is, actions performed jointly by more than one person (Schmitow & Kochukhova, 2013). It is as yet an unsolved question whether the perception of joint action is essentially consistent to individual action, or whether they follow different developmental trajectories. The present research aimed to investigate this question by presenting infants and adults with a block-stacking action that was either performed by one or two agents.

An important aspect during one's own performance and the perception of actions is the anticipation of the future end state of the action (von Hofsten, 2004). The occurrence of anticipatory gaze shifts indicates that an observer has built a representation of the observed action goal that allows one to predict the outcome of the action before it is completed, and it is typically modulated by infants' production skills with the respective action (e.g., Melzer, Prinz, & Daum, 2012). The anticipation of actions has been investigated extensively both in adults (Ambrosini, Sinigaglia, & Costantini, 2012; Costantini, Ambrosini, & Sinigaglia, 2012; Elsner, Falck-Ytter, & Gredebäck, 2012; Flanagan & Johansson, 2003; Rotman, Troje, Johansson, & Flanagan, 2006) and infants (Ambrosini et al., 2013; Daum, Attig, Gunawan, Prinz, & Gredebäck, 2012; Falck-Ytter et al., 2006; Henrichs, Elsner, Elsner, & Gredebäck, 2012; Melzer et al., 2012). In these studies, the perception of individually performed manual actions was assessed such as reaching-to-grasp an object (Henrichs et al., 2012; Melzer et al., 2012), containment of an object (Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; Falck-Ytter et al., 2006), or eating (Kochukhova & Gredebäck, 2010). Depending on the task, infants start to anticipate action goals at around 6 months (Hunnius & Bekkering, 2010; Kanakogi & Itakura, 2011), and by the end of their first year of life, infants are able to anticipate the goal of many manual actions (e.g., Falck-Ytter et al., 2006; Melzer et al., 2012). However, in our social world, actions are often performed jointly by more than one person.

These joint actions vary from involving two interaction partners (e.g., in a face-to-face conversation) to a multitude of cooperating or competing interaction partners (e.g., in musical or sport performances). Although frequently observed in everyday life, little research has addressed the question of how infants and adults perceive these interactions.

1.1. Joint Action in Adults and Infants

Adults generally coordinate their actions easily to achieve a joint goal such as preparing a dinner together (for an overview see, Sebanz, Bekkering, & Knoblich, 2006). To do so, adults represent and predict not only their own actions, but also their interaction partner's actions (Kourtis, Sebanz, & Knoblich, 2013; Sebanz et al., 2006). Performance of simple tasks is often improved if another person is present, a phenomenon called social facilitation (e.g. Triplett, 1898; Zajonc, 1965), whereas more complex tasks can lead to performance impairment (Allport, 1920). Studies on task sharing have also demonstrated more specific interferences in situations where two adults acted according to complementary task rules (e.g., Atmaca, Sebanz, & Knoblich, 2011; Sebanz, Knoblich, & Prinz, 2003). More precisely, when an agent represented a task rule of a co-agent that interfered with their own task rule (i.e., when they required conflicting actions), response times were larger compared to conditions without conflicting task representations. In general, adults are exceptionally capable of actively engaging in coordinated joint action.

Infants participate in parent-child exchanges practically from birth (for an extensive overview of the first two years, see Brownell, 2011). During the first months of life, these face-to-face interactions become increasingly coordinated with respect to their timing and structure (Bigelow & Walden, 2009). Importantly, in early interactions, infants are not required to represent the interaction partner's intentions or goals. In the second half of the first year of life, the adult-infant dyads include external objects and events, which is referred to as joint attention (Liszkowski, Carpenter, Striano, & Tomasello, 2006). Around their first birthday, infants begin to initiate joint action (Liszkowski et al., 2006), and between 14 and 18 months children begin to autonomously engage in coordinated joint action with adults (Bakeman & Adamson, 1984; Warneken & Tomasello, 2007). Thus, during the first year of life, infants participate in joint action, but it is only by the second year of life that they actively coordinate their actions with others.

1.2. Perception of Nonverbal and Verbal Interactions

Infants do not only engage in joint action with their parents or their siblings. Given their limited motor repertoire in the first year of life, they also observe interactions between other people without being directly involved, for example their parents having a conversation or playing cards. It remains a largely unexplored question how infants in their first year of life perceive jointly performed actions, at an age when they are not yet able to engage in coordinated joint action themselves.

In one of the few studies that investigated the perception of a nonverbal interaction, 6- and 12-month-olds were presented with videos of one agent feeding another (Gredebäck & Melinder, 2010). The 12-month-olds anticipated the goal of the feeding action (i.e., that the food would be brought to the mouth of the second agent), whereas the 6-month-olds did not. By contrast, 6-month-old infants anticipated that food would be brought to the mouth if one agent fed herself (Kochukhova & Gredebäck, 2010). These studies suggest that 6-month-olds are able to anticipate an individually performed feeding action, but not yet an interactively performed one. It is important to note, however, that these results have to be compared carefully due to different visual and timing aspects. A further aspect that has been investigated is the role of infants' experience when observing manual interaction. Comparable to infants' anticipation of individual actions, their perception of interactions seemed to depend on their own active experience with the manual action (Schmitow & Kochukhova, 2013). Regarding experience with joint action, it was demonstrated that 10-month-olds were able to infer the joint goal of two collaborative partners if they actively experienced the joint action prior to observing it in a habituation paradigm (Henderson, Wang, Matz, & Woodward, 2013). Without this active experience, the joint goal could only be inferred by 14-month-olds (Henderson & Woodward, 2011).

Furthermore, 18-month-olds inferred a joint goal that two agents performed sequentially (Fawcett & Gredebäck, 2013). It is also noteworthy that, in the related field of verbal interactions (i.e., conversations between two agents), it was demonstrated that even infants anticipated the course of a conversation at least to some extent (Keitel, Prinz, Friederici, von Hofsten, & Daum, 2013; von Hofsten, Uhlig, Adell, & Kochukhova, 2009). Although the described studies investigated the perception of interaction, they do not answer the question of whether the perception of joint action is essentially different from individual action in infants and adults. In order to do just this, we conducted a study in which we systematically manipulated the number of agents involved.

1.3. The present study

In the present study, we presented infants and adults with an action that can easily be performed by one or two agents and that is familiar to infants: building a tower of wooden blocks, or 'block-stacking'. We tested 9- and 12-month-old infants, when practically no coordinated joint action capabilities are present (see, Brownell, 2011), and adults who are typically very skilled at coordinating their actions with others (e.g., Sebanz et al., 2006). The participants observed videos of a toy tower being built by either one agent (individual condition) or alternatingly by two agents (joint condition). Both conditions involved identical action goals, and the arrival of participants' gaze shifts at goals was analysed. The presented action involved one overarching goal (to build a tower) and a number of sub-goals (to grasp a block; to stack it). The overarching goal was identical in both conditions; sub-goals were identical concerning their position and sequence, but varied in that they were performed by one or two agents. If participants' perception of the actions is based on their representation of the overarching goal, their gaze behaviour should not be affected by the number of agents. If, by contrast, participants represent each agent's individual sub-goals in isolation, switching between the goals of the two agents in the joint condition could cause differential gaze behaviour of the observer compared to the perception of the individual action.

2. Method

2.1. Participants

The final sample consisted of 23 9-month-old infants (*M* = 9 months 6 days; range: 9;2 to 9;12; 12 female), 23 12-month-old infants (*M* = 12 months 2 days; range: 11;15 to 12;15; 11 female), and 14 adults (*M* = 23.4 years; range 21 to 28; 6 female). Seven more 9-month-olds and six more 12-month-olds were tested but did not complete enough trials to be included in the analyses due to fussiness in one or both conditions. One additional adult participant had to be excluded from analyses due to a technical error. All infants were born at full term. Infants received a toy for their participation, and adults received monetary compensation. The study was approved by a local ethics committee and conducted in accordance with the Declaration of Helsinki.

2.2. Apparatus and stimuli

Two videos were recorded, showing how a tower of coloured wooden blocks was built by either one agent (individual condition) or two agents (joint condition, see Figure 1). In both conditions, the complete tower consisted of six blocks, which were placed to the left and right of the base. The agent(s) alternatingly grasped one block at a time from the left and from the right and placed it on the base (‘stacking’). Once the tower was complete, the blocks were replaced in their initial position in reverse order (‘unstacking’). Goals were identical in both conditions. This resulted in a total of 24 reaching and transport movement sequences (trials) per video during which the participants’ gaze behaviour was analysed. To increase the participants’ attention towards the stimulus presentation, a ‘swooshing sound’ was presented during the transport sequences. During the recording session, a metronome ticked at the rate of 1 Hz to pace the actors’ movements, and to make the timing in the two conditions as similar as possible, Accordingly, the tower was built rhythmically, and each movement (reaching for a block; transporting a block) lasted approximately 1 s (see Figure 1 for details). The difference in the mean durations of movements between the two conditions was minimal (10 ms, i.e., 0.5%). The length of each action sequence video was approximately 40 s. Conditions only differed in the number of agents; all other aspects (number and position of blocks, timing of movements, background, lighting, etc.) were analogous.

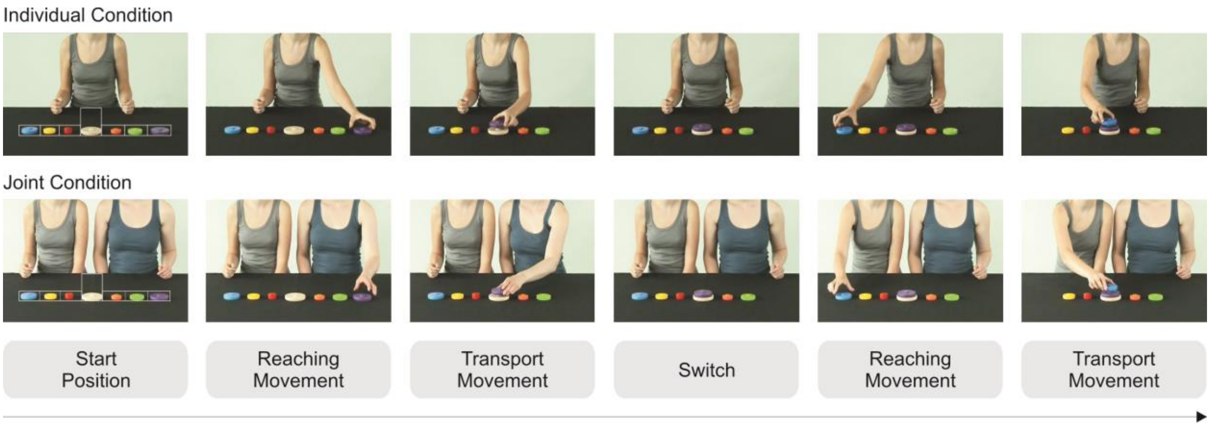

Figure 3.1. Snapshots of individual and joint conditions. The white boxes in the left panel indicate AOIs. The average duration (and standard deviation) in the individual condition were $M = 970$ ms ($SD = 66$ ms) for reaching and $M = 987$ ms ($SD = 62$ ms) for transport movements. In the joint condition these were $M = 990$ ms ($SD = 39$ ms) for reaching and $M = 987$ ms ($SD = 142$ ms) for transport movements.

Videos were presented on a 17-inch monitor and subtended a visual angle of approximately 28.3° × 19.8°. Gaze was measured using a remote corneal reflection eye tracker (Tobii 1750, Stockholm, Sweden; sampling rate: 50 Hz; software: Clear View 2.7.1) with an infant add-on (precision: 1°, accuracy: 0.5°). We used a 9-point-infant calibration.

2.3. Procedure

Written informed consent was obtained from the adult participants and from infants' parents prior to testing. After the calibration sequence, which took approximately 30 s, videos of the two conditions were presented. Order of conditions was counterbalanced across participants. Before the start of each video, a salient attention grabber was shown (videos of colourful toys that moved and made sounds). The presentation of each video was repeated in order to collect more valid trials. The stimulus presentation took approximately 3 min.

2.4. Data analysis

Gaze data was analysed using Matlab 7.1 (The MathWorks Inc.). Areas of Interest (AOIs) surrounded the positions of the blocks as well as the tower (see white boxes in Figure 1). AOIs for the block positions ranged from 4.8° to 5.1° horizontal visual angle and covered a vertical visual angle of 2.2°. The tower AOI covered a visual angle of 4.7° × 4.9°.

We computed the arrival of gaze shifts at goal AOIs relative to the arrival of the moving hand for each trial. Positive values represented anticipatory gaze shifts whereas negative values represented reactive gaze shifts. The time interval for anticipatory gaze shifts began with the movement of the hand and ended with the arrival of the hand at the goal area. At this point, the time interval for reactive gaze shifts began; it ended 1 s after the movement was finished. An individual trial was considered to be valid if a gaze shift was preceded by a fixation at the previous AOI for at least 100 ms. This ensured that actions were observed attentively. Only participants with at least 12 valid trials (6 per condition) where included in final analyses. On average, 9-month-olds provided 40.6 ($SD = 13.4$), 12-month-olds 50.3 ($SD = 21.2$), and adult participants 70.6 ($SD = 22.2$) valid trials.

3. Results

Initial analyses did not suggest any evidence for a main effect or interaction effects of presentation order (all ps > .32); those data were thus collapsed. Infants' and adults' gaze behaviour was anticipatory on average in both conditions (see Fig. 2 and Table 1). Performed t-tests against zero confirmed that participants shifted their gaze to the action goals significantly ahead of the agent's hand (9-month-olds: $t_{indiv}(22) = 5.13$, $p < .001$, $d = 1.07$; $t_{joint}(22) = 2.28$, $p = .03$, $d = 0.48$; 12-month-olds: $t_{indiv}(22) = 9.45$, $p < .001$, $d = 1.97$; $t_{joint}(22) = 4.73$, $p < .001$, $d = 0.99$; adults: $t_{indiv}(13) = 28.54$, $p < .001$, $d = 7.63$; $t_{joint}(13) = 27.14$, $p < .001$, $d = 7.25$).

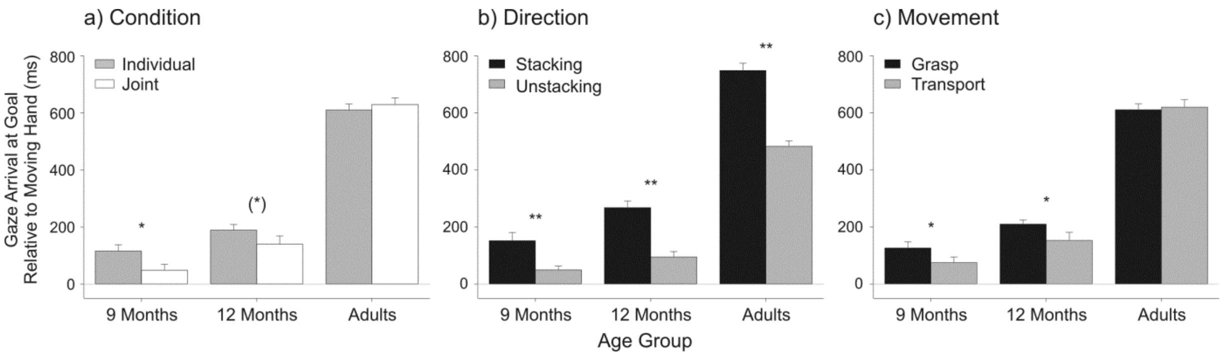

Figure 3.2. Mean gaze arrival at goals for all age groups a) in both conditions, b) for stacking direction, and c) movement type (with standard errors). Grey line at zero displays arrival of the hand at goal areas. Positive values indicate that gaze was anticipatory. Asterisks denote difference between a) individual and joint conditions, b) the two different directions and c) both movement types (**: p < .01; *: $p < .05$; (*): $p < .10$).

A 3 × 2 (Age [9 months, 12 months, adults]) × Condition [individual, joint]) Analysis of Variance with gaze latency yielded significant main effects of age, $F(2,57) = 167.89$, $p < .001$, $\eta^2_G = .80$, and condition, $F(1,57) = 4.50$, $p = .04$, $\eta^2_G = .004$, as well as a marginally significant interaction between both, $F(2,57) = 2.59$, $p = .08$, $\eta^2_G = .005$ (generalised eta squared values are presented to ensure comparability with other studies, see Bakeman, 2005; Olejnik & Algina, 2000). The main effect of age was caused by significant differences between all age groups (all ps < .009, Bonferroni-corrected). Participants anticipated action goals faster the older they were. Paired t-tests showed a significant difference between the individual and the joint action condition in 9-month-olds, $t(22) = 2.40$, $p = .03$, $d = 0.50$, a marginally significant difference in 12-month-olds, $t(22) = 2.07$, $p = .05$, $d = 0.43$, and no difference in adults, $p > .34$. Thus, infants showed faster gaze latencies in the condition with one agent, whereas adults anticipated both conditions equally fast. This pattern was confirmed non-parametrically: Eighteen 9-month-olds showed faster anticipations in the individual condition, compared with only 5 who did so in the joint condition, $\chi^2(1) = 7.35$, $p < .01$. Similarly, 15 out of 23 12-month-olds anticipated actions faster in the individual condition, $\chi^2(1) = 2.13$, $p = .14$, as did 6 out of 14 adults, $p = .59$.

	Individual		Joint	
	M	*SD*	*M*	*SD*
9 Months	115.47	107.85	48.12	101.25
12 Months	188.88	95.84	139.40	141.45
Adults	609.99	79.96	629.44	86.78

Table 3.1. Mean values and standard deviations of gaze latency (in ms) for both conditions for infants and adults.

We further explored how the different types of stacking direction (stacking vs. unstacking) and movement (reach vs. transport) affected gaze arrival times. Stacking the blocks was anticipated faster than unstacking by all age groups (all *p*s < .003, Figure 2b); and infants, but not adults, anticipated grasping faster than transport actions (infants: *p*s < .05; adults: *p* = .67, Figure 2c). Further analyses, for example, of condition and stacking direction or movement type, were not recommended because not all participants delivered data in the corresponding trials, and often only a single trial was acquired. These limitations would lead to highly unreliable results.

4. Discussion

The aim of the current study was to explore how the perception of individual and joint actions develops. Accordingly, we presented infants and adults with the same block-stacking action that was performed by either one or two agents. Our main findings were that 1) adults anticipated both conditions equally fast, and they generally initiated gaze shifts towards action goals very quickly, and 2) infants anticipated action goals in the individual condition faster than the joint condition, and their gaze shifts towards goals were initiated later than those of adults. One approach that possibly explains the present findings is that adults and infants represented the observed actions on different hierarchical levels, namely the level of overarching or sub-goals (Csibra, 2007). On a higher level, the overarching goal of our agent(s) was to alternatingly build a tower from the left and right, and this was identical in both conditions. However, if the actions were represented on the lower level of sub-goals, a number of differences would arise between conditions. The sub-goals were performed by either one agent or two different agents. This resulted in less certainty about which agent would act in the joint condition. Furthermore, there was an inevitable increase in visual stimulus complexity in the joint condition because two agents filled the scene instead of one. Thus, depending on the level of action representation, the conditions were either comparable or quite different.

4.1. Adults represent joint goals

The adults in our study did not show differential gaze behaviour towards the action goals in the individual and joint condition. This suggests that they inferred the overarching goal of the agent(s) to alternatingly build a tower from the left and right, and did not represent goals related to the performing agent in the joint condition. This higher-level representation could then be used to quickly initiate gaze shifts towards sub-goals in a top-down manner in both conditions. Adults usually make use of higher-level information, such as goals and intentions, that guide their anticipatory gaze shifts (Eshuis, Coventry, & Vulchanova, 2009). Such a higher-level representation leads to fast initiation of gaze shifts because the location of the next sub-goal can be inferred before the agent has started a movement. It is thus partly independent of low-level visual information such as visual stimulus complexity or movement kinematics. Such a predominantly top-down processing can explain why adults' anticipatory gaze was initiated very quickly and not affected by the number of agents that performed the block-stacking action.

There is, however, an alternative explanation as to why adults did not show differential gaze behaviour in the individual and joint condition. Because the observed action was undoubtedly quite simple, adults could have performed at ceiling, and this could have covered up underlying differences between conditions. We cannot rule out that adults would show delayed initiation of gaze shifts if observing a more demanding joint action. This remains subject to further research. However, adults are generally able to represent overarching, joint goals (Sebanz et al., 2006), so that a comparable gaze behaviour towards individual and joint action seems possible even in a more demanding task.

4.2. Infants represent sub-goals

The infants in our study anticipated individual action faster than joint action. This suggests that they could not make use of a representation of the overarching, joint goal in the condition with two agents that could guide their gaze towards sub-goals top-down. Instead, infants probably had to infer the sub-goal of each reaching or transport movement, while the actions were in progress, based on observable information in a bottom-up manner. Infants in their first year of life have been found to represent the sub-goals of an action, instead of the overarching goal (Woodward, Sommerville, & Guajardo, 2001). Furthermore, if children aged 9 and 12 months learned the goal of an animated agent, they subsequently predicted the agent to choose a goal based on its previous movement path, whereas children aged 3 years, and adults, made predictions based on the agent's previous goal (Daum et al., 2012). Thus, infants seem to rely primarily on low-level visual cues that need to be analysed instantaneously, such as a path, or a trajectory (Johnson, Amso, & Slemmer, 2003; Meichler & Gratch, 1980; Nelson, 1971, 1974), or the hand aperture in reaching actions (Ambrosini et al., 2013; Daum, Vuori, Prinz, & Aschersleben, 2009). This would lead to later initiation of gaze shifts in the joint condition for a number of reasons. First, switching between the representations of the two agents leads to a processing delay that would affect gaze latency (e.g., Altmann, 2011). Second, if no overarching goal representation was present, infants could not know which agent would act, and this uncertainty would further delay the initiation of gaze shifts. Third, the increased stimulus complexity in the joint condition would affect gaze behaviour if the action was analysed bottom-up, based on low-level visual information. Taken together, the present data suggests that infants' gaze shifts were guided predominantly bottom-up by low-level visual information that allowed them to infer the agent(s) sub-goals. This led to a generally later initiation of gaze shifts and a differential perception of individual and joint action.

4.3. From low-level to higher-level processing

In the present study, the infant groups anticipated goals in the individual condition better than in the joint condition, and this difference was more distinct in the younger infant group. As described previously, this suggests that infants probably could not make use of a representation of the overarching joint goal of two agents, whereas adults could. These findings suggest that the younger the infants, the more they depended on observable visual information (e.g., movement kinematics) to infer an action goal. This low-level visual information is less important in a top-down processing where the goal is inferred before a movement has started. One of the key reasons for the development from predominantly low-level to higher-level processing is very likely experience with manual actions on the one hand, and joint action on the other hand. Such a link between anticipatory gaze shifts and experience has been shown in infants (Kanakogi & Itakura, 2011; Melzer et al., 2012; Schmitow & Kochukhova, 2013) and adults (e.g., Sailer, Flanagan, & Johansson, 2005). It is likely that during their second year of life, children learn to anticipate joint action as well as individual action because they become more experienced in coordinating their actions with others (Brownell, 2011). This notion is corroborated by findings showing that 14- and 18-month-olds could infer a joint goal (Fawcett & Gredebäck, 2013; Henderson & Woodward, 2011). Due to their extensive active experience, adults are able to infer overarching joint goals and are less dependent on low-level visual information. It is has been shown, however, that adults still make use of low-level information, when a priori predictions are not possible, for example when they observe unusual or unpredictable actions (Rotman et al., 2006).

An interesting detail of our results is that even the 9-month-olds anticipated action goals on average. Usually, this gaze behaviour is rarely found in infants below 12 months of age (but see, Hunnius & Bekkering, 2010; Kanakogi & Itakura, 2011). In our study, the rhythmic turn-taking nature of movements could have supported infants' anticipatory gaze shifts (Haith, Hazan, & Goodman, 1988). Notably, this supported infants' anticipation of goals even in the joint condition, although 9-month-olds were not yet capable of engaging in joint action themselves (Warneken & Tomasello, 2007).

It is further important to note the bystander nature of the paradigm used in the present research. Participants observed the actions passively without being involved. The obvious benefit of this approach is that we were able to investigate infants that were not yet capable of engaging in joint action themselves. At the same time, infants might have been more attentive and motivated to make sense of our block-stacking if they had been involved. It is probably also due to this bystander paradigm (and to the fact that no conflicting representations were built) that adults did not show an interference effect as reported in previous research using paradigms where two agents performed a task jointly (Atmaca et al., 2011; Sebanz et al., 2003).

In another line of results, we found differences between the two directions of stacking (stacking vs. unstacking), and the two movement types (reach vs. transport). Stacking was anticipated faster by all age groups than unstacking. During stacking, all goals where defined by salient goals (i.e., the coloured blocks during reaching, and the tower during transport actions). During unstacking, the blocks were replaced in their initial location but there was no visible goal for these transport actions, which led to later initiation of gaze shifts (Becker & Fuchs, 1969). Furthermore, infants but not adults anticipated reaching faster than transport actions. This was probably due to the lack of active experience in infants, and the impact of experience on anticipatory gaze (e.g., Melzer et al., 2012). The ability to reach emerges at 3 or 4 months of age (White, Castle, & Held, 1964), which means that the 9- and 12-month-old infants in our study had had some experience with reaching actions. The ability to stack blocks, however, develops at around 12 months (e.g., Hayashi & Takeshita, 2009), which means that our infants had had little to no experience. This difference in active experience between the movement types most likely led to a differential perception of reaching and transport actions. Adults had already gained extensive experience in reaching and all sorts of manipulative behaviour so they perceived these actions similarly.

In conclusion, infants in their first year of life perceive individual and joint action differently. Infants are probably not yet able to infer the overarching joint goal of two agents and have to make use of low-level visual information. Adults, by contrast, anticipate individual and joint goals equally fast, because they are able to infer the joint goal of two agents. This development from low-level to higher-level processing is most likely due to first-hand experience in coordinated joint action.

Acknowledgements

We would like to thank Anja Gampe and Anna Maria Pleser for performing the actions, and Silke Atmaca for comments on an earlier version of this manuscript. AK gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (Graduate School ‘Function of Attention in Cognition’, GRK 1182) and a *Wiedereinstiegsstipendium* fellowship from the Freistaat Sachsen through the University of Leipzig.

References

Allport, F. H. (1920). The influence of the group upon association and thought. *Journal of Experimental Psychology, 3*, 159-182.

Altmann, E. M. (2011). Testing probability matching and episodic retrieval accounts of response repetition effects in task switching. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(4), 935-951. doi: 10.1037/a0022931

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS One, 8*(7), e67916. doi: 10.1371/journal.pone.0067916

Ambrosini, E., Sinigaglia, C., & Costantini, M. (2012). Tie my hands, tie my eyes. *Journal of Experimental Psychology-Human Perception and Performance, 38*(2), 263-266. doi: Doi 10.1037/A0026570

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Exp Brain Res, 211*(3-4), 371-385. doi: 10.1007/s00221-011-2709-9

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379-384.

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.

Becker, W., & Fuchs, A. F. (1969). Further properties of the human saccadic system: eye movements and correction saccades with and without visual fixation points. *Vision Research, 9*(10), 1247-1258.

Bigelow, A. E., & Walden, L. M. (2009). Infants' response to maternal mirroring in the still face and replay tasks. *Infancy, 14*(5), 526-549. doi: Doi 10.1080/15250000903144181

Brownell, C. A. (2011). Early developments in joint action. *Review of Philosophy and Psychology, 2*(2), 193-211. doi: 10.1007/s13164-011-0056-1

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42. doi: 10.1111/j.1467-7687.2011.01095.x

Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Out of your hand's reach, out of my eyes' reach. *Quarterly Journal of Experimental Psychology, 65*(5), 848-855. doi: 10.1080/17470218.2012.679945

Csibra, G. (2007). Action mirroring and action understanding: an alternative account In P. Haggard, Y. Rosetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII* (pp. 435 - 459). Oxford: Oxford University Press.

Daum, M. M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (2012). Actions seen through babies' eyes: A dissociation between looking time and predictive gaze. *Frontiers in Psychology, 3*, 370. doi: 10.3389/fpsyg.2012.00370

Daum, M. M., Vuori, M. T., Prinz, W., & Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6-and 9-month-old infants. *Developmental Science, 12*(6), 854-862. doi: DOI 10.1111/j.1467-7687.2009.00831.x

Elsner, C., Falck-Ytter, T., & Gredebäck, G. (2012). Humans anticipate the goal of other people's point-light actions. *Frontiers in Psychology, 3*, 120. doi: 10.3389/fpsyg.2012.00120

Eshuis, R., Coventry, K. R., & Vulchanova, M. (2009). Predictive eye movements are driven by goals, not by the mirror neuron system. *Psychological Science, 20*(4), 438-440.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879.

Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science, ePub 02.July 2013*, 1-9. doi: 10.1111/desc.12074

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424*(6950), 769-771.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*(2), 197-206. doi: 10.1016/j.cognition.2009.09.004

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development, 59*(2), 467-479.

Hayashi, M., & Takeshita, H. (2009). Stacking of irregularly shaped blocks in chimpanzees (Pan troglodytes) and young humans (Homo sapiens). *Animal Cognition, 12 Suppl 1*, S49-558. doi: 10.1007/s10071-009-0273-5

Henderson, A. M., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy, 18*(1), 10-39. doi: 10.1111/j.1532-7078.2012.00126.x

Henderson, A. M., & Woodward, A. L. (2011). ‘Let's work together’: what do infants understand about collaborative goals? *Cognition, 121*(1), 12-21. doi: 10.1016/j.cognition.2011.05.008

Henrichs, I., Elsner, C., Elsner, B., & Gredebäck, G. (2012). Goal salience affects infants' goal directed gaze shifts. *Froniers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00391

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: a study of infants' visual anticipations during action observation. *Developmental Psychology, 46*(2), 446-454. doi: 10.1037/a0016543

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America, 100*(18), 10568-10573. doi: 10.1073/pnas.1630655100

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*. doi: 10.1038/Ncomms1342

Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C., & Daum, M. M. (2013). Perception of conversations: The importance of semantics and intonation in children's development. *Journal of Experimental Child Psychology, 116*, 264-277. doi: 10.1016/j.jecp.2013.06.005

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development*, *81*, 1729-1738.

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: an EEG study. *Social Neuroscience, 8*(1), 31-42. doi: 10.1080/17470919.2012.694823

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*(2), 173-187. doi: DOI 10.1207/s15327647jcd0702_2

Meichler, M., & Gratch, G. (1980). Do 5-month-olds show object conception in Piaget's sence? *Infant Behavior & Development*, *3*, 256-281.

Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and observation of contralateral reaching: A close link by 12 months of age. *Infant Behavior & Development, 35*(3), 570-579. doi: 10.1016/j.infbeh.2012.05.003

Nelson, K. E. (1971). Accommodation of visual tracking patterns in human infants to object movement patterns. *Journal of Experimental Child Psychology, 12*(2), 182-196.

Nelson, K. E. (1974). Infants' short-term progress towards one component of object permanence. *Merril-Palmer Quarterly, 20*, 3-8.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25*(3), 241-286. doi: 10.1006/ceps.2000.1040

Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology, 96*(3), 1358-1369. doi: 10.1152/jn.00227.2006

Sailer, U., Flanagan, J. R., & Johansson, R. S. (2005). Eye-hand coordination during learning of a novel visuomotor task. *Journal of Neuroscience, 25*(39), 8833-8842. doi: Doi 10.1523/Jneurosci.2658-05.2005

Schmitow, C., & Kochukhova, O. (2013). How infants look at others' manual interactions: The role of experience. *Infant Behavior & Development*, *36*(2), 223-227. doi: DOI 10.1016/j.infbeh.2013.01.002

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70-76. doi: 10.1016/j.tics.2005.12.009

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition, 88*(3), B11-21.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *American Journal of Psychology, 9*, 507-533.

von Hofsten, C. (2004). An action perspective on motor development. *Trends in Cognitive Sciences*, *8*(6), 266-272.

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders, 3*(2), 556-569. doi: 10.1016/j.rasd.2008.12.003

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

White, B. L., Castle, P., & Held, R. (1964). Observations on the development of visually-directed reaching. *Child Development, 35*(2), 349-364.

Woodward, A. L., Sommerville, J. A., & Guajardo, J. J. (2001). How infants make sense of intentional action In B. F. Malle, L. J. Moses & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 149-169). Cambridge, MA: The MIT Press.

Zajonc, R. B. (1965). Social Facilitation. *Science, 149*(3681), 269-274.

Chapter 4

Perception of conversations: The importance of semantics and intonation in children's development

Shortened title: Perception of conversations

Anne Keitel[1], Wolfgang Prinz[1], Angela D. Friederici, Claes von Hofsten[2] and Moritz M. Daum[1,3]

[1]Max-Planck-Institute for Human Cognitive and Brain Sciences

[2] Uppsala University

[3]University of Zurich

Journal of Experimental Child Psychology, 2013, 116, pp. 264 – 277.

Abstract

In conversations, adults readily detect and anticipate the end of a speaker's turn. However, little is known about the development of this ability. We addressed two important aspects involved in the perception of conversational turn taking: semantic content and intonational form. The influence of semantics was investigated by testing prelinguistic and linguistic children. The influence of intonation was tested by presenting participants with videos of two dyadic conversations: one with normal intonation and one with flattened (removed) intonation. Children of four different age groups—two prelinguistic groups (6- and 12-month-olds) and two linguistic groups (24- and 36-month-olds)—and an adult group participated. Their eye movements were recorded, and the frequency of anticipated turns was analyzed. Our results show that (a) the anticipation of turns was reliable only in 3-year-olds and adults, with younger children shifting their gaze between speakers regardless of the turn taking, and (b) only 3-year-olds anticipated turns better if intonation was normal. These results indicate that children anticipate turns in conversations in a manner comparable (but not identical) to adults only after they have developed a sophisticated understanding of language. In contrast to adults, 3-year-olds rely more strongly on prosodic information during the perception of conversational turn taking.

1. Introduction

During social interactions, we are confronted with a large amount of verbal and nonverbal information. To act and react quickly and appropriately, the incoming flow of information needs to be analyzed on-line and upcoming events need to be anticipated. This holds especially true for conversations. Here, the principle of taking turns is helpful and most fundamental (Sacks, Scheglof, & Jefferson, 1974). In the current study, we investigated the development of the ability to anticipate turns during the perception of a conversation between two people.

When engaged in a conversation, it is easy to identify the end of the turn of a conversation partner and the beginning of one's own turn. The end of a speaker's turn is accompanied by a variety of different cues. On the one hand, language comprehension (i.e., the semantic content or utterance content) seems to be the most important factor for detecting the end of a turn (de Ruiter, Mitterer, & Enfield, 2006; Magyari & de Ruiter, 2012). de Ruiter et al. (2006) presented adult participants with audio recordings of isolated turns from natural Dutch telephone conversations and asked them to press a button when they suspected the turn end. The participants were instructed not to wait until the turn was finished but rather to anticipate its ending. The results showed that responses were very reliable and that the average response time was 200 ms before a turn was finished. This indicated that the participants not only were very accurate in detecting the end of turns they were even able to anticipate a turn end. Importantly, they were equally able to do so when listening to recordings where the intonation had been removed but leaving semantics and syntax intact.

On the other hand, in natural conversation, a turn end is usually accompanied by a number of acoustically marked prosodic boundary cues (Gerken & McGregor, 1998) such as intonation, syllable length, and pauses. In general, prosodic boundary cues help to segment linguistic units (Gerken & McGregor, 1998), making them an important feature in the acquisition of language (Gerken, 1996).

At the end of conversational turns, the pitch (i.e., intonation) rises or falls, the last vowel is lengthened, and pauses are longer compared with the end of clauses or phrases. It has been shown that adults were, in principle, able to use these cues to identify a speaker's turn. When utterances are made unintelligible, with only prosodic cues (notably intonation) still intact, participants could identify the end and beginning of turns at above chance level (de Ruiter et al., 2006; Schaffer, 1983). But performance was better when participants could rely on both prosodic cues and utterance content to detect a turn end, compared with a condition where only prosodic cues were available (de Ruiter et al., 2006). Although it is unusual in natural conversation that only prosodic cues are available (except, e.g., when listening to a conversation through a wall or from far away), these experimental studies suggest that adults can use prosody to better anticipate the end of a sentence but mainly do so once neither semantic nor syntactic information is available (Grosjean & Hirt, 1996).

Taken together, the results of these studies indicate that adults are able to detect the end of a conversational turn even before the previous speaker has finished. If only linguistic cues are available, then they do so by focusing predominantly on the utterance content. Prosodic cues, such as intonation, primarily have a supportive function.

As children develop, conversations become increasingly important, and the principle of taking turns appears to be already relevant at a young age. Infants as young as 3 months reacted with adapted timing and more speech-like vocalizations if their mother (Masataka, 1993) or an experimenter (Bloom, Russell, & Wassenberg, 1987) interacted with them in a turn-taking pattern (as opposed to a random, temporally noncontingent pattern). Even 2-month-old infants were found to discriminate between contingent (turn-taking) and noncontingent interaction with their mothers (Murray & Trevarthen, 1985), and they were found to be more interactive and content in the turn-taking condition. Note, however, that this study has been criticized (Rochat, Neisser, & Marian, 1998) and that other studies failed to replicate its results (Muir & Hains, 1993; Rochat et al., 1998). Up to now, only a few studies have addressed children's perception of turn taking between other people. When watching two people having a conversation, infants at 6 months of age and above could follow the conversation more easily if speakers were engaged in a face-to-face interaction as opposed to an interaction where the speakers looked in opposite directions, not facing each other (Augusti, Melinder, & Gredebäck, 2010). This result suggests that already prelinguistic infants are able to perceive (and make use of) relevant aspects of turn taking in conversations. von Hofsten, Uhlig, Adell, and Kochukhova (2009) presented 1- and 3-year-old children with a performed everyday conversation and analyzed how often gaze was shifted to the next speaker before they started speaking. Their results showed that the anticipation of turns improved significantly with age, from 33.8% in 1-year-olds to 62.2% in 3-year-olds. This study indicates that, at least by 3 years of age, children have a general ability to anticipate the next turn in an observed conversation. However, so far it has not yet been investigated which linguistic factors influence the development of anticipation of turns or, more specifically, what role semantic and prosodic cues play in infants' and toddlers' ability to anticipate turns. As described above, in adults semantics plays the predominant role in the perception of turn taking. However, it is unknown which factors drive the perception of turn taking earl y in life when little or no semantic and syntactic understanding is present. Accordingly, in the following, we briefly review the early development of semantic and syntactic understanding.

Infants start understanding words at around 8 months of age (Fenson et al., 1994; Harris, Yeeles, Chasin, & Oakley, 1995). By 16 months, infants comprehend approximately 169 words (Fenson et al., 1994). After 16 months, the assessment of infants' rapidly evolving receptive vocabulary is difficult (Harris & Butterworth, 2002). However, by 30 months, infants' expressive vocabulary has reached nearly 600 words (Fenson et al., 1994). Apart from increases in the lexical inventory, children between 2 and 3 years of age learn to produce simple and more complex sentences (Clark, 2003), which also involves at least some basic knowledge about sentence construction (i.e., syntax). A more sophisticated understanding of syntactic schemes is not achieved until 3.5 or 4 years of age (Tomasello, 2000). The influence of the developing semantic and syntactic understanding abilities on the perception of turn taking has yet to be investigated.

Concerning the sensitivity to prosody, it has been shown that 6-month-old infants already detect syntactic units, such as clauses (Nazzi, Nelson, Jusczyk, & Jusczyk, 2000; Seidl, 2007) and phrases (Soderstrom, Seidl, Nelson, & Jusczyk, 2003), using prosodic cues. In most of these studies, the role of prosodic cues was explored in general without focusing on the role of individual cues. However, one prosodic cue, namely intonation, is of specific importance for infants; newborns are already able to extract intonation in speech (Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008; Nazzi, Floccia, & Bertoncini, 1998), and infants generally prefer infant-directed speech over adult-directed speech (Fernald, 1985). Pronounced intonation has been identified as the key reason for this preference in 4-month-olds (Fernald & Kuhl, 1987). Furthermore, pronounced intonation facilitates the segmentation, and therefore the learning of new words, for infants (Thiessen, Hill, & Saffran, 2005). Thus, intonation is one of the prosodic cues that plays a special role in infants’ perception of spoken language and, consequently, might play an essential role in their anticipation of turns. Accordingly, the primary aim of this study was to investigate the influence of semantic content and intonational form on the perception of turn taking during development.

The secondary issue approached in the current study was a methodological one. The above-mentioned study by von Hofsten et al. (2009) raised one issue. Their findings showed that the anticipation of turns improved significantly between 1 and 3 years of age. However, the interpretation of these results is difficult because there is no norm, baseline, or chance level to which the anticipation frequency can be compared. Thus, it is not possible to statistically evaluate the quality of performance. In other words, a reliability measure is not available. Accordingly, the second aim of the current study was to analyze the gaze data in more detail and to develop a statistical method to assess whether the anticipation of turns was reliable or merely a consequence of random eye movements.

Taken together, previous research shows that in perceived conversations, the ability to anticipate the onset of turns improves with age (von Hofsten et al., 2009). The evolving semantic and syntactic development most likely contributes to the development of the ability to anticipate turns (de Ruiter et al., 2006), which is closely linked to the ability to follow the course of a conversation. Furthermore, because infants are already able to detect turn ends using prosodic cues (Nazzi et al., 2000; Soderstrom et al., 2003), these cues may also support children's ability to identify and anticipate turns in natural dyadic conversations. Here, we investigated whether the prosodic cue intonation facilitates children's anticipation. Additional analyses of gaze behavior will provide the possibility of analyzing the reliability of anticipated turns.

In the current study, we adopted the paradigm used by von Hofsten et al. (2009) to evaluate the role of semantics and intonation in children's ability to anticipate turns in observed conversations. Participants were presented with recordings of actors performing casual everyday conversations while their gaze was measured. The influence of semantics was investigated by testing different age groups: two prelinguistic groups (6- and 12-month-olds), two linguistic groups (24- and 36-month-olds), and a control group of adults. The role of intonation was tested by presenting the participants in each group with two conversations: one with normal intonation and one with flattened intonation. We expected to find developmental differences between younger and older children with respect to intonation. On the one hand, it seems plausible that prelinguistic infants might be more sensitive to intonation cues than older children because this facilitates early speech processing and word learning. On the other hand, if children's use of intonational cues depends on the development of the respective communicative functions, then intonation might be more helpful for older linguistic children because conversations play a more important part in their lives. Furthermore, we refined the analysis of participants' gaze behavior during the conversations. We analyzed not only how often the onset of a turn was anticipated but also how often gaze shifts were unrelated to turns or random. With this procedure, we intended to obtain a more detailed picture of the development of the ability to anticipate conversational turns. We expected a reliable anticipation of turns only in older children and adults because a sophisticated understanding of the utterance content seems to be the most important ability for the anticipation of turns (de Ruiter et al., 2006; Magyari & de Ruiter, 2012).

In a first step, a pilot study was conducted with 82 children from 6 to 36 months of age. This study revealed that only the 3-year-olds anticipated more turns in the condition with normal, as opposed to flattened, intonation.[1] However, the actors in the conversation were not trained and spoke rather artificially. Furthermore, actors moved a lot while talking, which could serve as an additional cue to turn taking. Hence, to analyze this interesting effect in more detail, we recorded new stimulus material with trained actors (see Method) who could provide ecologically valid conversations (e.g., reasonably natural speech) while not overly moving.

2. Method

2.1. Participants

A total of 120 participants, 24 in each of the five age groups, completed the study and were included in the final analyses: 6-month-olds (10 female and 14 male, mean age = 6 months 2 days, range = 5 months 25 days to 6 months 13 days), 12-month-olds (12 female and 12 male, mean age = 12 months 8 days, range = 11 months 21 days to 12 months 19 days), 24-month-olds (7 female and 17 male, mean age = 24 months 17 days, range = 24 months 9 days to 24 months 25 days), 36- month-olds (14 female and 10 male, mean age = 36 months 16 days, range = 36 months 0 days to 37 months 0 days), and adults (13 female and 11 male, mean age = 25 years, range = 20–32 years). An additional 12 6-month-olds, 13 12-month-olds, 13 24-month-olds, and 6 36-month-olds were tested but excluded from data analysis because they did not watch one or both of the conversations attentively. Contact information of children was obtained from public birth records. Families received a small gift for their participation. The study was approved by a local ethics committee and conducted in accordance with the Declaration of Helsinki.

[1] In total, 22 6-month-olds, 21 12-month-olds, 23 24-months-olds, and 16 36-month-olds completed the pilot study. An additional 22 children were tested but not included in the final analyses due to fussiness. A 4 (Age: 6, 12, 24, or 36 months) × 2 (Intonation: normal or flattened) analysis of variance (ANOVA) revealed that children anticipated more turns the older they were ($F = 12.14$, $p < .001$, $\eta^2_p = .32$) and no other effects. Paired *t* tests showed that only 36-month-olds anticipated more turns in the normal condition, $t(15) = 2.77$, $p = .014$, $d = 0.69$. Children in younger age groups did not show a difference between normal and flattened conversations (all *p*s > .24).

2.2. Apparatus and stimuli

Two German conversations were video-recorded with four female actors recruited from an acting school. They were instructed to move their heads and bodies as little as possible, to keep their voices natural, and to speak in an adult-directed manner. The actors sat on chairs facing each other. At the beginning of each video, one actor was present and said ‘‘Hello’’ into the camera to greet the participants. The second actor then entered, sat down, and also greeted the participants before starting the conversation with the first actor. Contents of the conversations were everyday topics such as leisure activities (Conversation A) and holiday preparations (Conversation B). Each conversation consisted of 29 turns (i.e., 28 turn taking) between actors that were analyzed (greeting of the children at the beginning was excluded). The actors’ rate of speaking was moderate (see Table 4.1 for details). Conversations were designed to have similar properties. For example, the number of questions was equal for each speaker and, accordingly, for each conversation. The conversations did not contain ‘‘continuers’’ (Schegloff, 1982) or ‘‘back-channels’’ (Yngve, 1970) between turns, such as ‘‘m-hm.’’ Furthermore, both conversations were of similar length. The duration of speech (and pauses) between the two conversations differed by only approximately 11% (8%), and this difference in duration means was smaller than 20% (16%) of 1 standard deviation. Conversations were recorded with a directed microphone (SennheiserK6/ME66).

	Number of Turn-Taking	Length (s) Total	Ø Speech	Ø Pauses	Ø Words / Turn
Conversation A	28	100.76	2.26	0.93	6.6
Conversation B	28	105.36	2.52	0.86	6.9

	Normal Intonation M (Hz)	SD (Hz)	Flattened Intonation M (Hz)	SD (Hz)
Conversation A	198.87	47.28	198.71	0.63
Conversation B	190.41	50.67	190.54	1.02

Table 4.1. Details of the two conversations. Upper panel: number of turn-taking; total length of conversations, mean length of speech (i.e. turns) and pauses between turns; mean number of words per turn. Lower panel: mean and standard deviation of pitch for normal and flattened conditions.

The conversations were presented with either normal or flattened intonation. For the flattened intonation conversations, the variations of the fundamental frequency (F0) were removed and averaged to the mean frequency of the conversations using the software Praat (Boersma &Weenink, 2010). Specifically, the pitch contour of the conversations was extracted and segmented into pitch points at a rate of 100 Hz. The pitch points were removed, and a new pitch contour was created with the average frequency of the respective conversation using PSOLA (pitch synchronous overlap and add) resynthesis. This resulted in clearly less intonated monotone speech (see lower panel of Table 4.1 and Figure 4.1).

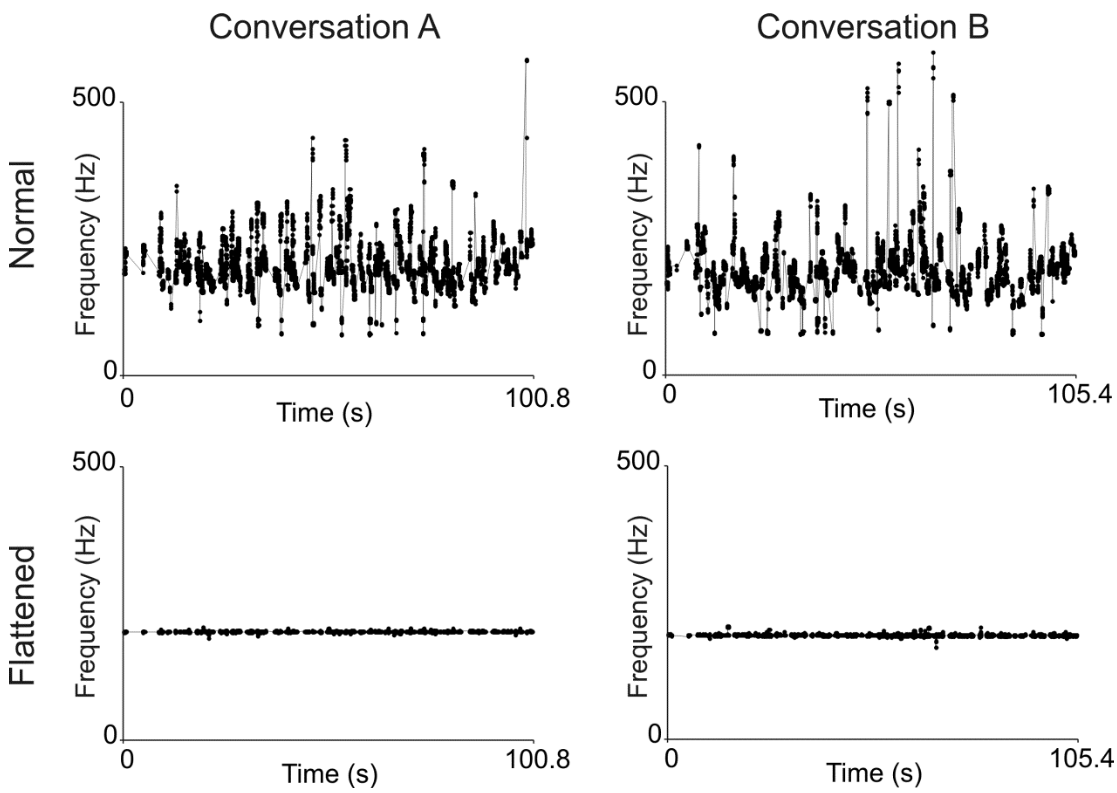

Figure 4.1. Pitch contour of both conversations in the normal and flattened conditions.

The videos were presented on a 17″- monitor (resolution: 800 × 600 pixels), and gaze was measured using a remote corneal reflection eye tracker (Tobii 1750, Stockholm, Sweden; ClearView 2.7.1 software; sampling rate: 50 Hz) with an infant add-on (precision: 1°; accuracy: 0.5°). The participants sat at a distance of approximately

60 cm from the monitor. Videos subtended a visual angle of 27.5° × 15.2°. A 9-point infant calibration was used.

2.3. Procedure

The experimenter explained the procedure and method to infants' parents and obtained their informed consent. The exact purpose of the study was disclosed after testing so as not to influence participants' and parents' behavior. All infants were tested individually with one parent present. After the calibration sequence, participants watched two conversations: A and B. One of the conversations was shown with normal intonation and one with flattened intonation. The order of conversations and the intonation of each conversation were counterbalanced across participants. Before the start of each conversation, a salient attention grabber (videos of toys that moved and made noises, e.g., a spinning and laughing starfish) was presented to focus the participant's attention on the monitor. The calibration procedure took approximately 1 min, and presentation of the videos took approximately 4 min.

2.4. Data analysis

Eye movements were analyzed using Matlab 7.1 (MathWorks). For all analyses, only gaze shifts toward the areas of interest (AOIs) were included. AOIs surrounded the faces from the top of the head to the chin and from the tip of the nose to the rear of the ear (see boxes in Figure 4.2). AOIs ranged from 3.3° to 4.2° horizontal visual angle and from 5.5° to 6.2° vertical visual angle.

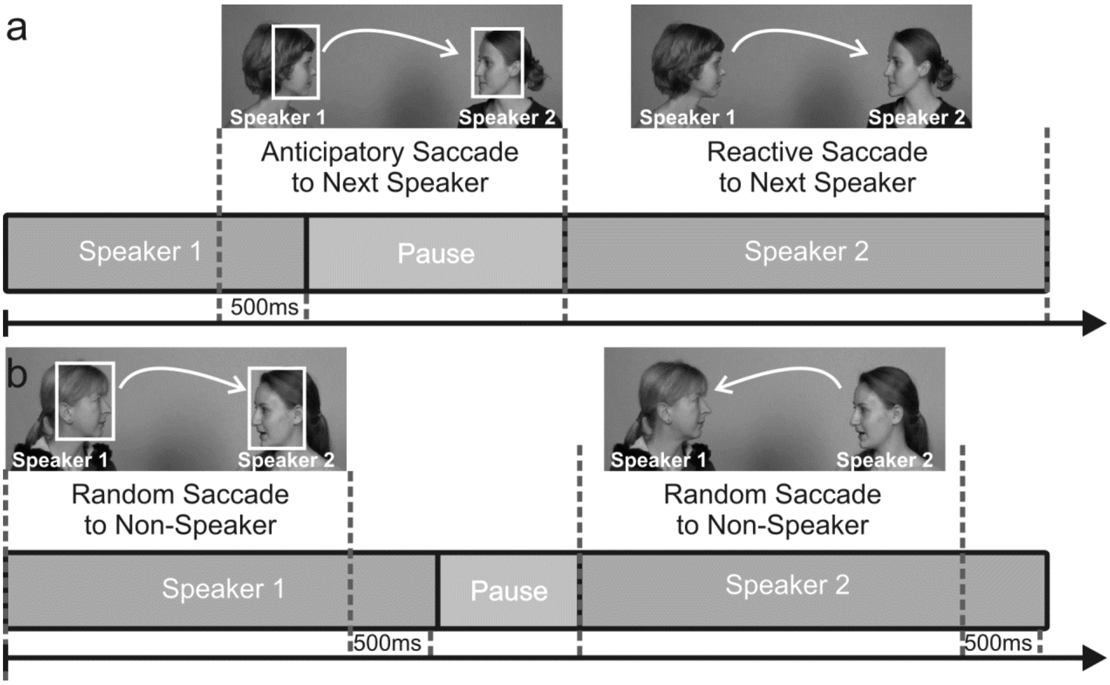

Figure 4.2. Illustration of the time intervals and direction of anticipatory and reactive gaze shifts (A) and of random gaze shifts (B). Areas of interest (AOIs) are indicated by white boxes around faces in the left pictures.

First, gaze shifts related to turns were identified as anticipatory or reactive in order to calculate the relative frequency of anticipated turns. This measure was used by von Hofsten et al. (2009), as well as in a number of other action perception studies (Daum, Attig, Gunawan, Prinz, & Gredebäck, 2012; Johnson, Amso, & Slemmer, 2003; Melzer, Prinz, & Daum, 2012), and represents an intuitive and simple indication of participants' performance. The number of anticipated turns was divided by the total number of attentively watched turns:

$$Relative\ Frequency_{Anticipated\ Turns} = \frac{N_{Anticipated\ Turns}}{N_{Total\ Turns}}$$

To be classified as an anticipatory gaze shift, a gaze needed to be shifted from the current speaker to the next speaker before she had begun to speak (see Figure 4.2 A). The respective time interval began 500 ms prior to the end of the current speaker's turn, included the pause between turns, and ended with the start of the next speaker's turn. The 500-ms interval prior to the pause ensured that a gaze shift was classified as anticipatory even if participants shifted their gaze to the next speaker while the current speaker had not yet finished. This takes into account that adults are able to anticipate the end of the current speaker's turn as opposed to responding to a perceived turn end (de Ruiter et al., 2006). Gaze shifts were classified as reactive if the gaze was shifted to the next speaker after she had begun to speak. Gaze shifts were included in the analysis only where participants had looked at the current speaker for at least 100 ms before shifting gaze to the next speaker. This ensured that turns were watched attentively and that gaze shifts were related to a turn in the conversation. In some cases, children did not notice a change of speaker and did not shift the gaze to the next speaker during her turn but kept fixating on the nonspeaker until it was her turn again. Such gaze behavior was not dismissed but rather was included in the number of turns in which infants watched the conversation attentively in order to calculate the relative frequency of anticipated turns.

Second, gaze behavior was analyzed in more detail. The previous analysis of anticipatory gaze shifts focused on gaze shifts that were most likely related to a change of speaker. However, during conversations, in addition to turn-related gaze shifts, turn-unrelated or random gaze shifts occur. These random gaze shifts between speakers can distort the frequency of anticipated turns. To account for this, we calculated the occurrence rate of anticipatory and random gaze shifts. For these indexes, the total number of anticipatory or random gaze shifts was divided by the respective time interval during which such a gaze shift could theoretically occur:

$$Occurrence\ Rate_{Gaze\ Shifts} = \frac{N_{Gaze\ Shifts}}{t}$$

Anticipatory gaze shifts were identified using the same criteria as in the previous analysis of their relative frequency. Gaze shifts were classified as random if the gaze was shifted from the current speaker to the nonspeaker (excluding a 500-ms interval prior to a pause; see Figure 4.2 B for an overview and specifications of the exact time intervals and direction of anticipatory and random gaze shifts). The sum of the anticipatory and random gaze shifts was then divided by the total duration that a participant had fixated on both of the speakers' faces in the respective time intervals (in general, these were pauses for anticipatory gaze shifts, and speech for random gaze shifts, apart from the 500-ms intervals prior to a pause during which an anticipatory gaze shift could occur as well). These durations represent the time during which the conversation was watched attentively.

Occurrence rates can be regarded as a probability to make a gaze shift during the respective time intervals. Their analysis enables a direct comparison of the occurrence of anticipatory and random gaze shifts. If the occurrence rate of anticipatory gaze shifts were statistically larger than the occurrence rate of random gaze shifts, then we could infer reliable anticipation of turns in conversations. In other words, this analysis examines whether or not turn-related gaze shifts are significantly different from chance level. This allows for conclusions about how well an observer can anticipate the course of a conversation.

In addition, the occurrence rates allow for a more detailed assessment of the influence of intonation on the perception of conversational turn taking. If the occurrence rate of anticipatory gaze shifts were decreased in the condition with flattened intonation, then intonation would have a supportive influence on the perception of turn taking in conversations. If, however, the occurrence rate of random gaze shifts were increased in the condition with flattened intonation, then participants were likely to be distracted by the missing intonation. Therefore, this analysis provides an important control of undesirable effects.

Participants were included in further analyses when they attentively followed at least four turns (14%) in each conversation (see calculation of anticipation frequency above for criteria of ‘‘attentively followed turns’’). On average, participants in the different age groups attended to 38% (6-month-olds), 40% (12-month-olds), 41% (24-month-olds), 57% (36-month-olds), and 94% (adults) of turns. Bonferroni-corrected *t* tests (all tests reported are two-tailed) showed that 36-month-olds attended to significantly more turns than the younger age groups (all ps $< .01$) and adults differed from all child age groups (all ps $< .001$). There was no difference among 6- to 24-month-olds. At 3 years of age, children are more interested in conversations than younger children and show some basic communication skills themselves (Nadel, Guerini, Peze, & Rivet, 1999). Greater interest and attention span, as well as advanced semantics, can explain the higher number of attended trials in older children and adults.

3. Results

3.1. Relative frequency of anticipated turns

Initial analyses did not reveal any effect of the order of presentation. Consequently, those data were collapsed for further analyses. Figure 4.3 and Table 4.2 show the relative frequency of anticipated turns, as indicated by gaze shifts in intonationally normal and flattened conditions for each age group. The distribution of anticipatory gaze shifts during the 500-ms interval before the pause and during the pause is illustrated in Figure 4.4. The histograms show that a considerable amount of anticipatory gaze shifts was performed while a speaker was still speaking, in line with previous findings (de Ruiter et al., 2006).

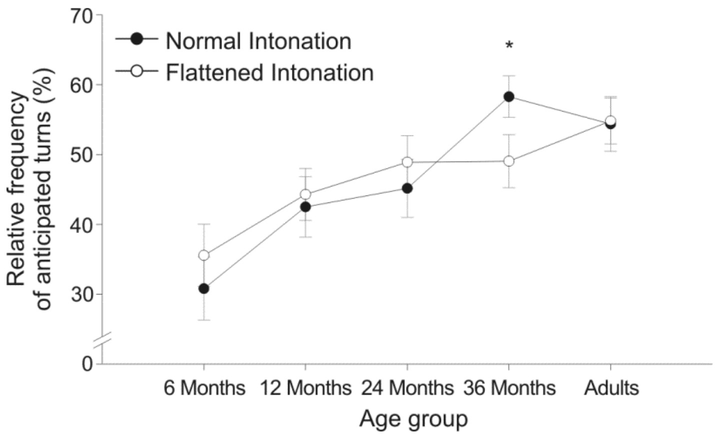

Figure 4.3. Relative frequency of anticipated turns, as indicated by gaze shift in both conversations for all age groups (with standard errors). An asterisk indicates a significant difference between normal and flattened conditions (*$p < .05$).

A 4 (Age: 6, 12, 24, or 36 months) × 2 (Intonation: normal or flattened) analysis of variance (ANOVA) was conducted with relative frequency of anticipations as the dependent variable. The adult sample served as a control group, reflecting the ‘‘developmental end state’’ of the perception of conversations; therefore, the adult data were only informative and not included in the ANOVAs. The main effect of age in the four groups of children reached significance, $F(3,92) = 6.48$, $p = .001$, $\eta^2_p = .17$. In general, participants anticipated more turns the older they were.

Age Group	6 Months		12 Months		24 Months		36 Months		Adults	
	M	SD	M	SD	M	SD	M	SD	M	SD
Normal Intonation (%)	30.8	22.1	42.5	21.2	45.2	20.3	58.3	14.6	54.4	19.2
Flattened Intonation (%)	35.6	22.0	44.3	18.3	48.9	18.5	49.0	18.6	54.8	16.2
Anticipatory Gaze Shifts	0.23	0.13	0.32	0.11	0.34	0.15	0.42	0.18	0.42	0.15
Random Gaze Shifts	0.24	0.12	0.36	0.14	0.32	0.09	0.31	0.10	0.26	0.06

Table 4.2. Mean frequency of anticipated turns (in %) in the normal and flattened conditions, and mean occurrence rates for anticipatory and random gaze shifts (both with standard deviations and for all age groups).

Bonferroni-corrected post hoc *t* tests yielded significant differences between 6-month-olds and 24-month-olds and between 6-month-olds and 36-month-olds (both *p*s < .03). Differences between the other groups of children were not significant (all *p*s > .20). There was no main effect of intonation, $F(1,92) < 1$, but the interaction between age and intonation approached significance, $F(3,92) = 2.24$, $p = .09$, $\eta^2_p = .07$.

Paired *t* tests indicated a difference between normal and flattened intonation conditions only in 36-month-olds, $t(23) = 2.66$, $p = .014$, $d = 0.54$. Of the 24 children in this age group, 17 anticipated more turns when intonation was normal, $\chi^2(1) = 5.26$, $p = .02$ (1 child showed no difference between normal and flattened conditions). In all other age groups, including the adults, there was no difference between conditions (all *p*s > .30).

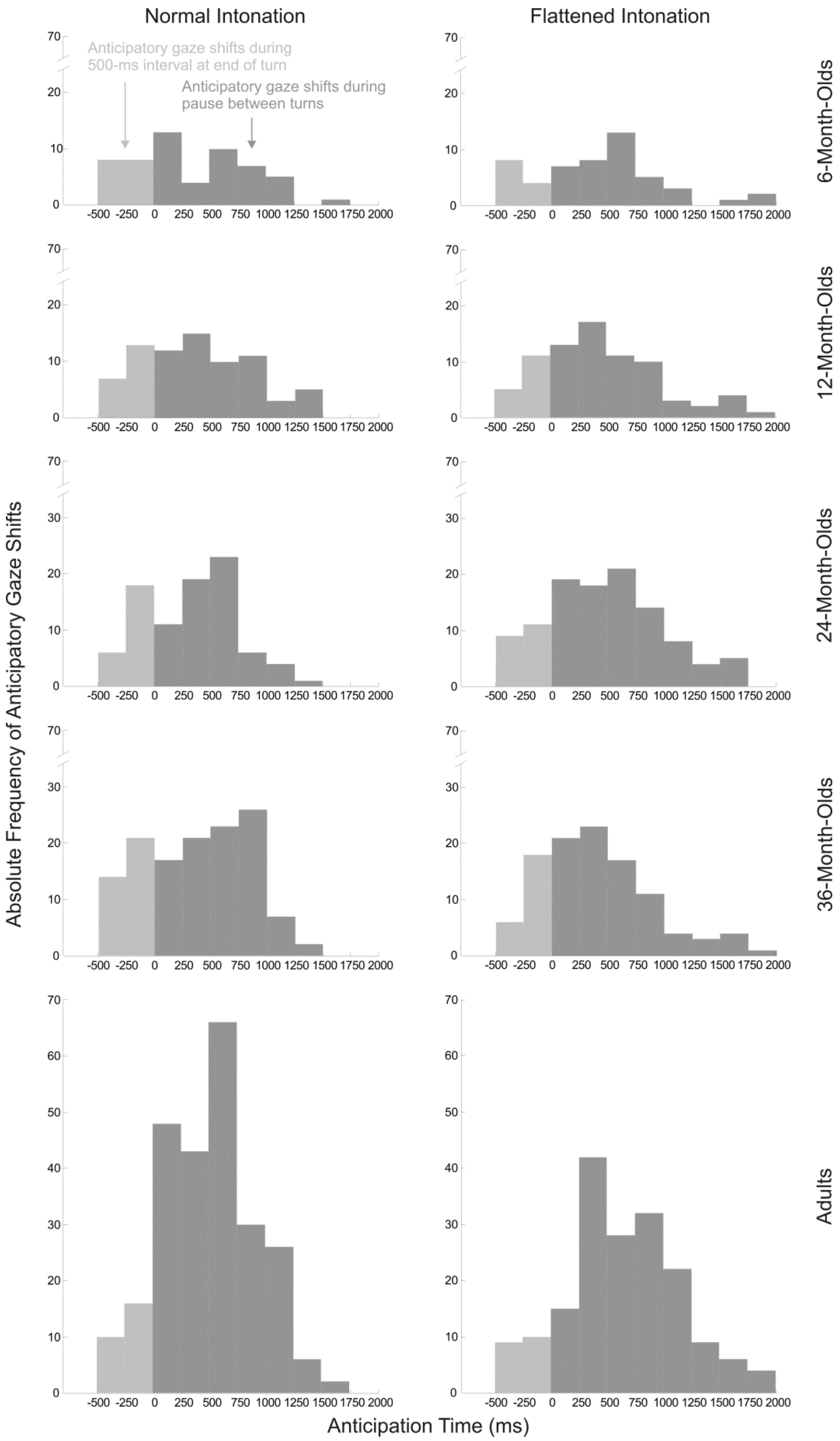

Figure 4.4. Histograms of anticipatory gaze shifts (absolute values) during the 500-ms interval prior to a pause (light gray bars) and during a pause (dark gray bars) for both conditions and all age groups. Bin size is 250ms. Note that reactive gaze shifts are not included.

3.2. Occurrence rates of anticipatory and random gaze shifts

The values of the occurrence rates of anticipatory and random gaze shifts were averaged over both conversations (see lower part of Table 4.2). A 4 (Age: 6, 12, 24, or 36 months) × 2 (Occurrence Rate: anticipatory or random) ANOVA yielded a significant main effect of age, $F(3,92) = 7.77$, $p < .001$, $\eta^2_p = .20$, and a significant Age × Occurrence Rate interaction, $F(3,92) = 3.35$, $p = .02$, $\eta^2_p = .099$. There was no main effect of occurrence rate ($F = 1.39$, $p = .24$). The significant interaction can be explained by paired t tests for each age group. As illustrated in Figure 4.5, children from 6 to 24 months of age showed anticipatory gaze shifts as often as random gaze shifts during the time they watched the conversations attentively (all *p*s > .14). The 36-month-olds showed more anticipatory gaze shifts than random gaze shifts, $t(23) = 2.13$, $p < .05$, $d = 0.43$. For the adult group, this difference was even larger, $t(23) = 4.28$, $p < .001$, $d = 0.88$.

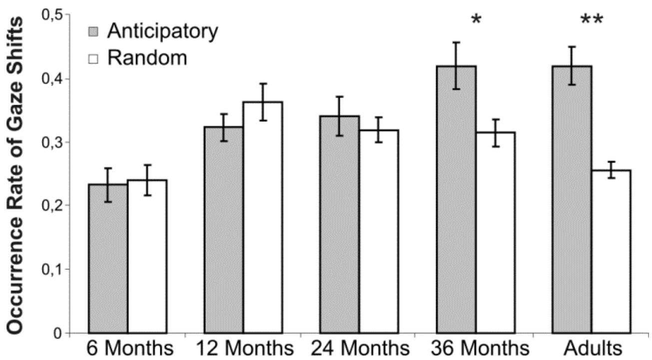

Figure 4.5. Occurrence rates of anticipatory and random gaze shifts for all age groups averaged over both conditions (with standard errors). Asterisks indicate a significant difference between the occurrence rates of anticipatory and random gaze shifts (*p < 05; **p < .01).

In addition, we evaluated the effect of intonation on the occurrence rates of anticipatory and random gaze shifts. In line with results of the relative frequencies, anticipatory gaze shifts occurred more often in the conversations with normal intonation only in 36-month-olds, $t(23) = 2.63$, $p = .02$, $d = 0.54$. All other age groups showed no difference between normal and flattened conditions (all *p*s > .12). Importantly, the occurrence rate of random gaze shifts did not differ between normal and flattened conditions (in 36-month-olds: $p = .46$; in all other age groups: *p*s > .10), indicating that children did not generally make more or fewer gaze shifts in one of the conditions.

4. Discussion

The aim of the current study was to investigate the development of the ability to anticipate turns in conversations. The influence of semantics was evaluated by studying prelinguistic and linguistic children and adults. The importance of the prosodic feature intonation and its role in anticipating turns was investigated by varying the presence and absence of intonation in the conversations. We presented videos of dyadic conversations with normal and flattened intonation and recorded participants' eye movements. The combined analyses of the relative frequency of anticipated turns and of the occurrence rate of random and anticipatory gaze shifts allowed for a statistical analysis of infants' gaze behavior, providing information on the quality of the anticipation of turns. The main findings can be summarized as follows. First, the main study and the pilot study both provided converging evidence that intonation influenced the perception of conversations only in 3-year-olds. Second, the anticipation of turns was reliable only from 3 years of age onward.

4.1. Intonation and perception of turn taking

Intonation did not have a measurable effect on the anticipation of conversational turn taking by infants from 6 to 24 months of age. This is interesting given that intonational differences are already processed by newborns (Nazzi et al., 1998; Sambeth et al., 2008) and intonation plays an important role in early word learning (Thiessen et al., 2005). There are two possible explanations for the lack of an effect that are not necessarily mutually exclusive. The first explanation concerns infants' general ability to anticipate turns in observed conversations, which might be limited primarily due to semantic limitations. Previous studies have reported a predominant influence of semantic skills on the anticipation of turns in adults (de Ruiter et al., 2006; Magyari & de Ruiter, 2012). The semantic and syntactic understanding of the young children in our study might have been too rudimentary to reveal intonation effects. In addition, limited attentional capacities could have impaired the ability to use intonational cues and/or anticipate turns. For example, it might have been more difficult for infants to integrate the high amount of verbal and visual information. Hence, a potential effect of intonation could have been covered up by infants' general inability to follow the conversation. This line of explanation gains some support from our result that young children under 3 years did not yet show a reliable anticipation of turns. We discuss the reliability aspect below.

The second explanation as to why young children did not show a difference between intonationally normal and flattened conditions is related to the specific function of the intonational cues. Electrophysiological evidence suggests that the processing of prosodic boundary cues, including intonation, differs fundamentally between older and younger children (Männel & Friederici, 2009, 2010). Similar to adults, older children from 3 years onward demonstrate a brain response indexing the recognition of the linguistic function of prosodic boundary cues in sentences that is independent of the acoustically salient pause. Younger children are not able to process the linguistic function of the intonational cues. Instead, they need the pause for boundary detection and show obligatory electrophysiological responses that indicate low-level acoustic processes (Männel & Friederici, 2009, 2010). Thus, although younger children process intonation on a general level (Sambeth et al., 2008), they do not have the ability to use it in its function as an indicator of the end of a phrase in a sentence (Männel & Friederici, 2009, 2010) or the end of turns in conversations (current data).

Only 36-month-olds benefited from the additional information provided by the presence of intonation as one important prosodic feature. They showed more anticipatory eye movements, and thus better anticipation of the next turn, when intonation was present. This was not due to a generally higher occurrence of gaze shifts during the conversation (which could have implied that participants were distracted by the missing intonation) but rather was caused by a higher occurrence of anticipatory gaze shifts when intonation was present compared with when it was not. This is in line with the finding that children at 3 years of age have learned to use prosodic boundary cues to indicate higher level linguistic aspects, as implicated by recent electrophysiological data on sentence processing (Männel & Friederici, 2010).

However, their language abilities are not yet as sophisticated as those of adults, which may explain why at this point in development the additional information provided by intonation effectively supports the perception of conversations and the anticipation of a speaker's next turn.

In adults, intonational cues do not affect the anticipation of turns in conversations. Adults rely mainly on other cues to predict speakers' turns, most likely related to a sophisticated, lexico-syntactic driven understanding of language (de Ruiter et al., 2006; Magyari & de Ruiter, 2012). This result is in line with the notion that adults use prosody more efficiently to predict the end of a sentence (or a turn) if semantic and syntactic information is lacking (Grosjean & Hirt, 1996).

4.2. Reliability of perception of turn taking

Regarding the question of how well turns in conversations can be anticipated, the relative frequency of anticipated turns only allows the conclusion that there is an increase with age, replicating previous findings (von Hofsten et al., 2009). This increase is very likely the result of a better comprehension of language with age because semantics and syntax are major factors in predicting the end of a turn (de Ruiter et al., 2006; Grosjean & Hirt, 1996; Magyari & de Ruiter, 2012). The conducted analyses of the occurrence rates of anticipatory and random gaze shifts refine and extend these results with respect to a crucial aspect, namely, the reliability of turn-taking-related gaze shifts. Children from 6 to 24 months of age made anticipatory (turn-related) or random (turn-unrelated) gaze shifts equally often. This means that the young children shifted their gaze between the two speakers regardless of their turn taking and, therefore, did not show the ability to anticipate the course of the conversations reliably.

Only from 36 months of age onward did anticipatory gaze shifts occur more often than random gaze shifts. Accordingly, reliable anticipation of turns (and therefore the ability to anticipate the course of conversations) is present only by 3 years of age. These analyses reveal that there is no fundamental difference in the quality of gaze shifts between prelinguistic (6- and 12-month-old) and linguistic (24- and 36-month-old) children, but there is a change in gaze behavior from 2 to 3 years of age. By 3 years, gaze shifts are qualitatively related to turn taking. It is at this age that children can produce complex sentences (Clark, 2003; Fenson et al., 1994) and master basic verbal communication skills (Nadel & Fontaine, 1989). They become more engaged in adult-like dyadic conversations and start using communicative strategies (Haslett & Samter, 1997). In addition to advanced semantic and syntactic development, it is possible that only from 3 years of age onward can children comprehend the principle of taking turns and apply it to perceived conversations. However, it has been shown that 3-month-olds can interact nonverbally in turns (Bloom et al., 1987), which is sometimes regarded as a precursor to more complex forms of verbal communication exchanges (Billard, 2002; Nadel et al., 1999). More research is required to fully understand all of the factors relevant in the development of the ability to anticipate the course of an observed verbal conversation.

A further interesting and not necessarily expected result of the current study is that the adult participants anticipated only little more than half of the observed turns. Moreover, their anticipation frequency did not differ from that of the 3-year-olds. In this case, the occurrence rates of random and anticipatory gaze shifts are particularly valuable as a reliability measure and can help to assess the adults' performance. Adults showed a comparable occurrence rate of anticipatory gaze shifts to the 3-year-olds. At the same time, they made much fewer random gaze shifts, resulting in a bigger difference between the two occurrence rates. Hence, adults showed considerably more anticipatory gaze shifts than random gaze shifts ($p < .001$) and a very reliable skilled anticipation of the course of the conversations.

The simple and straightforward design of the current study provides a first step to study children's perception of everyday conversations and their use of linguistic cues such as intonation when anticipating conversational turns. In the same way as it provides initial insight into the processing of observed conversations, it opens a variety of new questions. The most obvious question is when an adult-like perception of conversations is achieved. Other open questions are whether infant-directed speech facilitates young children's anticipation of turns and to what extent natural variations in pitch at the end of turns can be related to gaze shifts. Furthermore, the relationship between anticipatory gaze shifts and children's semantic skills has yet to be proven experimentally and could be achieved by correlating children's language skills to anticipatory gaze shifts. In addition, the relationship between gaze shifts and general cognitive skills, such as children's attention capacity, could provide further indications about underlying mechanisms.

For the current study, it is important to point out that participants were passive bystanders and not actively involved in the conversations. It is possible that sensitivity to turn-taking cues was reduced compared with interactions where one is required to react. Another factor that might have limited especially the children's performance is the adult nature of conversational topics and demeanor. However, it is important to note that the conversations represented normal speech and pause duration between adults as they are regularly observed by adults as well as children. For example, the pauses in our study were approximately 900 ms on average; a mean pause duration shorter than 500 ms is not unusual in normal adult conversation (see, e.g., Heldner & Edlund, 2010). The average speaking rate (in English) is 4 to 7 syllables per second (Huggins, 1967). In our conversations, the average speaking rate was at the lower end of this range (4.3 syllables/s). Furthermore, the duration of turns in our study ranged from 320 ms (for ‘‘Bye’’ at the end of the conversation) to 5760 ms. Even the shortest duration of speech does not pose any challenges to the human auditory system to extract the pitch and compare it with previous values. Assuming a female voice with an average fundamental frequency of 200 Hz (as in our conversations), adults can theoretically detect and compare the pitch of a vowel as short as 20 to 25 ms (Lee, 1994; Lee & Bacon, 1997). Even the auditory system of newborn infants is already able to detect the pitch contour of words (Nazzi et al., 1998), and the fact that newborns and infants are able to detect a rapid change of pitch similarly to adults is vastly documented by psychophysiological studies (e.g., Alho, Sainio, Sajaniemi, Reinikainen, & Näätänen, 1990; Carral et al., 2005; Haden et al., 2009; Stefanics et al., 2009). Furthermore, although our results suggest that the ability to anticipate conversational turns is related to language comprehension and production (because only 3- year-olds and adults anticipated turns reliably), this also poses a limitation on the interpretation of the young children's results; because 6- to 24-month-olds shifted their gaze between speakers independently of their turn taking, their data need to be interpreted carefully.

In conclusion, the current study shows that, even though language evolves rapidly during the first years of life, it is only by 3 years that children develop the ability to reliably anticipate turns in an observed conversation between adults. Furthermore, the results indicate that intonation is of particular importance for 3-year-olds in order to anticipate the course of such observed conversations. Thus, children seem to use the prosodic information to a greater extent to anticipate conversational turns when their language comprehension is well developed but still not as sophisticated as that of adults.

Acknowledgments

We thank Claudia Männel and Jonas Obleser, as well as three anonymous reviewers and the editor Catherine Mondloch, for helpful comments on an earlier version of the manuscript. This research was supported by the Deutsche Forschungsgemeinschaft (GRK 1182).

References

Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., & Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology, 77*, 151-155.

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

Billard, A. (2002). Imitation: A means to enhance learning of a synthetic protolanguage in autonomous robots. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 281-310). London, England: MIT Press.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language, 14*, 211-227.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (Version 5.1.32).

Carral, V., Huotilainen, M., Ruusuvirta, T., Fellman, V., Näätänen, R., & Escera, C. (2005). A kind of auditory 'primitive intelligence' already present at birth. *The European Journal of Neuroscience, 21*, 3201-3204.

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Daum, M. M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (2012). Actions seen through babies' eyes: A dissociation between looking time and predictive gaze. *Frontiers in Psychology, 3*, 370.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82*, 515-535.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*.

Fernald, A. (1985). 4-month-old infants prefer to listen to motherese. *Infant Behavior & Development, 8*, 181-195.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development, 10*, 279-293.

Gerken, L. (1996). Prosody's role in language acquisition and adult parsing. *Journal of Psycholinguistic Research, 25*, 345-356.

Gerken, L., & McGregor, K. (1998). An overview of prosody and its role in normal and disordered child language. *American Journal of Speech-Language Pathology, 7*, 38-48.

Grosjean, F., & Hirt, C. (1996). Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. *Language and Cognitive Processes, 11*, 107-134.

Haden, G. P., Stefanics, G., Vestergaard, M. D., Denham, S. L., Sziller, I., & Winkler, I. (2009). Timbre-independent extraction of pitch in newborn infants. *Psychophysiology, 46*, 69-74.

Harris, M., & Butterworth, G. (2002). *Developmental psychology*. East Sussex, UK: Psychology Press Ltd.

Harris, M., Yeeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22*, 1-18.

Haslett, B. B., & Samter, W. (1997). *Children communicating: The first 5 years.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics, 38*, 555-568.

Huggins, A. W. (1967). Distortion of the temporal pattern of speech by syllable-tied alternation. *Language and Speech, 10*, 133-140.

Johnson, S. P., Amso, D., & Slemmer, J. A. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 10568-10573.

Lee, J. (1994). Amplitude modulation rate discrimination with sinusoidal carriers. *The Journal of the Acoustical Society of America, 96*, 2140-2147.

Lee, J., & Bacon, S. P. (1997). Amplitude modulation depth discrimination of a sinusoidal carrier: effect of stimulus duration. *The Journal of the Acoustical Society of America, 101*, 3688-3693.

Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology, 3*, 376cc.

Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience, 21*, 1988-2006.

Männel, C., & Friederici, A. D. (2010). Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, 1-13.

Masataka, N. (1993). Effects of contingent and noncontingent maternal stimulation on the vocal behaviour of three- to four-month-old Japanese infants. *Journal of Child Language, 20*, 303-312.

Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and observation of contralateral reaching: A close link by 12 months of age. *Infant Behavior & Development, 35*, 570-579.

Muir, D. W., & Hains, S. M. (1993). Infant sensitivity to perturbations in adult facial, vocal, tactile, and contingent stimulation during face-to-face interactions. In B. de Boysson-Bardies (Ed.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 171-183). Amsterdam, North Holland: Elsevier.

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month olds and their mothers. In T. M. Field & N. A. Fox (Eds.), *Social Perception in Infants* (pp. 177-197). Norwood, NJ: Ablex Publishers.

Nadel, J., & Fontaine, A.-M. (1989). Communication by imitation: A developmental and comparative approach to transitory social competence. In B. H. Schneider, G. Attili, J. Nadel & R. P. Weissberg (Eds.), *Social competence in developmental perspective* (pp. 131-144). Dordrecht, Boston, London: Kluwer Academic Publishers.

Nadel, J., Guerini, C., Peze, A., & Rivet, C. (1999). The evolving nature of imitation as a format for communication. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 209-234). Cambridge, UK: Cambridge University Press.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior & Development, 21*, 779-784.

Nazzi, T., Nelson, D. G. K., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy, 1*, 123-147.

Rochat, P., Neisser, U., & Marian, V. (1998). Are young infants sensitive to interpersonal contingency? *Infant Behavior & Development, 21*, 355-366.

Sacks, H., Scheglof, E., & Jefferson, G. (1974). Simplest systematics for organization of turn-taking for conversation. *Language, 50*, 696-735.

Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology, 119*, 332-341.

Schaffer, D. (1983). The role of intonation as a cue to turn taking in conversation. *Journal of Phonetics, 11*, 243-257.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen (Ed.), *Analyzing Discourse: Text and Talk* (pp. 71-93). Washington, DC: Georgetown University Press.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language, 57*, 24-48.

Soderstrom, M., Seidl, A., Nelson, D. G. K., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49*, 249-267.

Stefanics, G., Haden, G. P., Sziller, I., Balazs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology, 120*, 304-308.

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

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders, 3*, 556-569.

Yngve, V. H. (1970). On getting a word in edgewise. *Chicago Linguistic Society, 6*, 567-578.

References (Chapter 1)

"perception, n. (2005). *The Oxford English Dictionary* (Third edition). Oxford University Press. Retrieved 23 August 2012, from http://www.oed.com/view/Entry/140560

"teleological, a. (2011). *The Oxford English Dictionary* (Second edition). Oxford University Press. Retrieved 14 October 2013, from http://www.oed.com/view/Entry/198708

Abravanel, E., & Gingold, H. (1985). Learning via observation during the 2nd year of life. *Developmental Psychology, 21*(4), 614-623.

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109-1116.

Aslin, R. N. (2007). What's in a look? *Developmental Science, 10*(1), 48-53. doi: DOI 10.1111/j.1467-7687.2007.00563.x

Augusti, E. M., Melinder, A., & Gredebäck, G. (2010). Look who's talking: Pre-verbal infants' perception of face-to-face and back-to-back social interactions. *Frontiers in Psychology, 1*, 161. doi: 10.3389/fpsyg.2010.00161

Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language, 6*(2), 183-200.

Bíró, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Developmental Science, 10*(3), 379-398. doi: 10.1111/j.1467-7687.2006.00544.x

Bourgeois, K. S., Khawar, A. W., Neal, S. A., & Lockman, J. J. (2005). Infant manual exploration of objects, surfaces, and their interrelations. *Infancy, 8*(3), 233-252. doi: DOI 10.1207/s15327078in0803_3

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*(1-2), 3-22.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*(2), 124-143. doi: 10.1006/brcg.2000.1225

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: inferential processes versus action simulation. *Current Biology, 17*(24), 2117-2121. doi: 10.1016/j.cub.2007.11.057

Brownell, C. A. (2011). Early developments in joint action. *Review of Philosophy and Psychology, 2*(2), 193-211. doi: 10.1007/s13164-011-0056-1

Bruner, J. S. (1969). Eye, Hand, and Mind In D. Elkind & J. H. Flavell (Eds.), *Studies in Cognitive Development: Essays in Honor of Jean Piaget*. Oxford: University Press, Inc.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42. doi: 10.1111/j.1467-7687.2011.01095.x

Cicchino, J. B., & Rakison, D. H. (2008). Producing and processing self-propelled motion in infancy. *Developmental Psychology, 44*(5), 1232-1241. doi: 2008-12114-004 [pii] 10.1037/a0012619

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*(5), 492-502.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *Neuroimage, 31*(3), 1257-1267. doi: DOI 10.1016/j.neuroimage.2006.01.033

Cross, E. S., Liepelt, R., F., H. A., Parkinson, J., Ramsey, R., Stadler, W., & Prinz, W. (2012). Robotic movement preferentially engages the action observation network. *Human Brain Mapping, 33*(9), 2238-2254. doi: 10.1002/hbm.21361

Cross, E. S., Stadler, W., Parkinson, J., Schütz-Bosbach, S., & Prinz, W. (2013). The influence of visual training on predicting complex action sequences. *Human Brain Mapping, 34*(2), 467-486. doi: 10.1002/hbm.21450

Csibra, G. (2007). Action mirroring and action understanding: an alternative account In P. Haggard, Y. Rosetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII* (pp. 435 - 459). Oxford: Oxford University Press.

Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*(2), 255-259. doi: 10.1111/1467-7687.00039

Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, *124*(1), 60-78. doi: 10.1016/j.actpsy.2006.09.007

Daum, M. M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (2012). Actions seen through babies' eyes: A dissociation between looking time and predictive gaze. *Frontiers in Psychology, 3,* 370. doi: 10.3389/fpsyg.2012.00370

Daum, M. M., Prinz, W., & Aschersleben, G. (2009). Means-end behavior in young infants: The interplay of action perception and action production. *Infancy, 14*(6), 613-640. doi: Doi 10.1080/15250000903263965 Pii 917026622

Daum, M. M., Prinz, W., & Aschersleben, G. (2011). Perception and production of object-related grasping in 6-month-old infants. *Journal of Experimental Child Psychology, 108,* 810 - 814.

De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language, 82*(3), 515-535.

Diersch, N., Cross, E. S., Stadler, W., Schütz-Bosbach, S., & Rieger, M. (2012). Representing others' actions: the role of expertise in the aging mind. *Psychological Research*, *76*(4), 525-541. doi: 10.1007/s00426-011-0404-x

Elsner, B. (2007). Infants' imitation of goal-directed actions: the role of movements and action effects. *Acta Psychologica, 124*(1), 44-59. doi: 10.1016/j.actpsy.2006.09.006

Ertelt, D., Small, S., Solodkin, A., Dettmers, C., McNamara, A., Binkofski, F., & Buccino, G. (2007). Action observation has a positive impact on rehabilitation of motor deficits after stroke. *NeuroImage, 36 Suppl 2*, T164-173. doi: 10.1016/j.neuroimage.2007.03.043

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879.

Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science, ePub 02.July 2013*, 1-9. doi: 10.1111/desc.12074

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5).

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424*(6950), 769-771.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science, 308*(5722), 662-667. doi: 10.1126/science.1106138

Gallese, V. (2006). Intentional attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research, 1079*, 15-24. doi: 10.1016/j.brainres.2006.01.054

Gallese, V. (2009). Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood. *Psychological Research, 73*(4), 486-498. doi: 10.1007/s00426-009-0232-4

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*(9), 396-403. doi: 10.1016/j.tics.2004.07.002S1364-6613(04)00184-6 [pii]

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

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

Glenberg, A. M. (2007). Language and action: creating sensible combination of ideas In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics*. Oxford: University Press.

Gredebäck, G., & Kochukhova, O. (2010). Goal anticipation during action observation is influenced by synonymous action capabilities, a puzzling developmental study. *Experimental Brain Research, 202*(2), 493-497. doi: 10.1007/s00221-009-2138-1

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*(2), 197-206. doi: 10.1016/j.cognition.2009.09.004

Hamilton, A. F., & Grafton, S. T. (2008). The motor hierarchy: from kinematics to goals and intentions In P. Haggard, Y. Rosetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII* (pp. 381 - 407). Oxford: Oxford University Press.

Hamilton, A. F., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14*(6), 493-498. doi: 10.1016/j.cub.2004.03.007

Harris, M., Yeeles, C., Chasin, J., & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22*(1), 1-18.

Hauf, P., & Aschersleben, G. (2008). Action-effect anticipation in infant action control. *Psychological Research-Psychologische Forschung, 72*(2), 203-210. doi: DOI 10.1007/s00426-006-0101-3

Hauf, P., & Prinz, W. (2004). The understanding of own and others' actions during infancy: "You-like-Me" or "Me-like-You?". *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 6*(3), 429-445.

Henderson, A. M., Wang, Y., Matz, L. E., & Woodward, A. L. (2013). Active experience shapes 10-month-old infants' understanding of collaborative goals. *Infancy, 18*(1), 10-39. doi: 10.1111/j.1532-7078.2012.00126.x

Henderson, A. M., & Woodward, A. L. (2011). "Let's work together": what do infants understand about collaborative goals? *Cognition, 121*(1), 12-21. doi: 10.1016/j.cognition.2011.05.008

Hickok, G. (2013). Do mirror neurons subserve action understanding? *Neuroscience Letters*, *540*, 56-58. doi: 10.1016/j.neulet.2012.11.001

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849-937.

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis*. Cambridge, UK: Polity Press.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol, 3*(3), e79. doi: 10.1371/journal.pbio.0030079

Jacob, P. (2013). How from action-mirroring to intention-ascription? *Consciousness and Cognition*, *22*(3), 1132-1141. doi: 10.1016/j.concog.2013.02.005

Jeannerod, M. (1994). The representing brain - Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences, 17*(2), 187-202.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *NeuroImage, 14*(1 Pt 2), S103-109. doi: 10.1006/nimg.2001.0832

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*. doi: 10.1038/Ncomms1342

Keitel, A., Prinz, W., & Daum, M. M. (submitted). Infants' and adults' perception of individual and joint action.

Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C., & Daum, M. M. (2013). Perception of conversations: The importance of semantics and intonation in children's development. *Journal of Experimental Child Psychology, 116*, 264-277. doi: 10.1016/j.jecp.2013.06.005

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences*, *15*(8), 352-357. doi: 10.1016/j.tics.2011.06.005

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing, 8*(3), 159-166. doi: 10.1007/s10339-007-0170-2

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*(6), 522-525. doi: 10.1016/s0960-9822(03)00165-9

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child Development*, *81*, 1729-1738.

Kosonogov, V. (2012). Why the mirror neurons cannot support action understanding. *Neurophysiology, 44*(6), 499-502. doi: DOI 10.1007/s11062-012-9327-4

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 352*(1358), 1231-1239. doi: 10.1098/rstb.1997.0105

Libertus, K., & Needham, A. (2010). Teach to reach: the effects of active vs. passive reaching experiences on action and perception. *Vision Research, 50*(24), 2750-2757. doi: 10.1016/j.visres.2010.09.001

Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science, 14*(6), 1355-1364. doi: 10.1111/j.1467-7687.2011.01084.x

Loucks, J., & Sommerville, J. A. (2012). The role of motor experience in understanding action function: the case of the precision grasp. *Child Development, 83*(3), 801-809. doi: 10.1111/j.1467-8624.2012.01735.x

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16*(8), 601-608. doi: 10.1111/j.1467-9280.2005.01582.x

Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology, 3,* 376cc. doi: 10.3389/fpsyg.2012.00376

McCarthy, D. (1933). Language development In C. Murchison (Ed.), *A handbook of child psychology* (Vol. 1). New York, US: Russell & Russell/Atheneum Publishers.

Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and observation of contralateral reaching: A close link by 12 months of age. *Infant Behavior & Development, 35*(3), 570-579. doi: 10.1016/j.infbeh.2012.05.003

Miall, R. C., Stanley, J., Todhunter, S., Levick, C., Lindo, S., & Miall, J. D. (2006). Performing hand actions assists the visual discrimination of similar hand postures. *Neuropsychologia, 44*(6), 966-976. doi: 10.1016/j.neuropsychologia.2005.09.006

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks, 9*(8), 1265-1279.

Nazzi, T., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior & Development, 21*(4), 779-784.

Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). New York: Basic Books.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-347. doi: 10.1017/S0140525X12001495

Prinz, W. (1990). A common coding approach to perception and action In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167 - 201). Berlin: Springer Verlag.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129-154.

Prinz, W. (2012). *Open minds*. Cambridge, Massachusetts: The MIT Press.

Provasi, J., Dubon, C. D., & Bloch, H. (2001). Do 9-and 12-month-olds learn means-ends relation by observing? *Infant Behavior & Development, 24*(2), 195-213.

Reid, V. M., Belsky, J., & Johnson, M. H. (2005). Infant perception of human action: Toward a developmental cognitive neuroscience of individual differences. *Cognition, Brain, Behaviour, 9*(2), 35 - 52.

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience. Vol 27 2004, 169-192.*

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience, 11*(4), 264-274. doi: nrn2805 [pii] 10.1038/nrn2805

Rosenbaum, D. A., & Krist, H. (1996). Antecedents of action In H. Heuer & S. W. Keele (Eds.), *Handbook of Perception and Action* (Vol. 2: Motor Skills). London: Academic Press.

Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of Neurophysiology, 96*(3), 1358-1369. doi: 10.1152/jn.00227.2006

Sacks, H., Scheglof, E., & Jefferson, G. (1974). Simplest systematics for organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology, 119*(2), 332-341. doi: DOI 10.1016/j.clinph.2007.09.144

Schmitow, C., & Kochukhova, O. (2013). How infants look at others' manual interactions: The role of experience. *Infant Behavior & Development, 36*(2), 223-227. doi: DOI 10.1016/j.infbeh.2013.01.002

Shimada, S., & Hiraki, K. (2006). Infant's brain responses to live and televised action. *Neuroimage, 32*(2), 930-939.

Sommerville, J. A., Blumenthal, E., Venema, K., & Braun, K. (2011). The body in action: The impact of self-produced action on infants' action perception and understanding. In V. Slaughter & C. Brownell (Eds.), *Early development of body representations*. Cambridge: Cambridge University Press.

Sommerville, J. A., Hildebrand, E. A., & Crane, C. C. (2008). Experience matters: the impact of doing versus watching on infants' subsequent perception of tool-use events. *Dev Psychol, 44*(5), 1249-1256. doi: 2008-12114-006 [pii] 10.1037/a0012296

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition, 95*(1), 1-30.

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

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition, 107*(3), 1059-1069. doi: 10.1016/j.cognition.2007.10.002

Southgate, V., Johnson, M. H., El Karoui, I., & Csibra, G. (2010). Motor system activation reveals infants' on-line prediction of others' goals. *Psychol Sci, 21*(3), 355-359. doi: 10.1177/0956797610362058

Stapel, J. C., Hunnius, S., & Bekkering, H. (2012). Online prediction of others' actions: the contribution of the target object, action context and movement kinematics. *Psychological Research, 76*(4), 434-445. doi: 10.1007/s00426-012-0423-2

Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: revisited. *Human Movement Science, 28*(5), 566-577. doi: 10.1016/j.humov.2009.01.002

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

von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders, 3*(2), 556-569. doi: 10.1016/j.rasd.2008.12.003

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257-302. doi: 10.1006/cogp.1995.1016

White, B. L., Castle, P., & Held, R. (1964). Observations on the development of visually-directed reaching. *Child Development, 35*(2), 349-364.

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

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development, 17*(1), 1061-1084. doi: 10.1016/S0885-2014(02)00074-6

Wühr, P., & Müsseler, J. (2001). Time course of the blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1260-1270.

List of Figures

Figure 2.1. Schematic drawing of the production task. The infant was given a cube (occupation object) in either his/her left or right hand. Subsequently, a second toy (target object) was held either (a) in front of the empty hand to elicit an ipsilateral reaction (ipsilateral presentation) or (b) in front of the occupied hand to elicit a contralateral reaction (contralateral presentation). 35

Figure 2.2. Still pictures of stimulus videos in the perception task. Pictures show the starting position of the reaching movement (left column), end position of the reaching movement/starting position of the transport movement (middle column) and end position of the transport movement (right column). Upper row (a) displays an ipsi-first trial, bottom row (b) a contra-first trial including AOIs indicated by white boxes. The pictures in black frames indicate end positions of the contralateral movements that were analysed. 36

Figure 2.3. Infants' reaching production. Mean frequency of infants' reactions in the production task with standard errors. A frequency of 50% each for ipsilateral and contralateral reactions would represent the most sophisticated reaching production possible. From 6 to 12 months, the frequency of contralateral reactions increased ($p = .03$). 41

Figure 2.4. Gaze behaviour with respect to contralateral movements observed. (a) Bars show mean gaze arrival times relative to the arm movement observed, with standard errors. Time point zero on the y-axis refers to the moment in the video when the hand enters the goal area (i.e. either ball AOI or bucket AOI). Positive values indicate that gaze arrived at the goal areas before the hand did; negative values indicate that gaze arrived after the hand. (b) Bars show mean relative frequency of anticipated arm movements, with standard errors. The difference between 6- and 12-month-olds reached significance (**$p < .001$). Gaze behaviour of adults is informative and was not included in statistical analyses. 42

Figure 3.1. Snapshots of individual and joint conditions. The white boxes in the left panel indicate AOIs. The average duration (and standard deviation) in the individual condition were M = 970 ms (SD = 66 ms) for reaching and M = 987 ms (SD = 62 ms) for transport movements. In the joint condition these were M = 990 ms (SD = 39 ms) for reaching and M = 987 ms (SD = 142 ms) for transport movements. 60

Figure 3.2. Mean gaze arrival at goals for all age groups a) in both conditions, b) for stacking direction, and c) movement type (with standard errors). Grey line at zero displays arrival of the hand at goal areas. Positive values indicate that gaze was anticipatory. Asterisks denote difference between a) individual and joint conditions, b) the two different directions and c) both movement types (**: p < .01; *: *p* < .05; (*): *p* < .10). 62

Figure 4.1. Pitch contour of both conversations in the normal and flattened conditions. 84

Figure 4.2. Illustration of the time intervals and direction of anticipatory and reactive gaze shifts (A) and of random gaze shifts (B). Areas of interest (AOIs) are indicated by white boxes around faces in the left pictures. 86

Figure 4.3. Relative frequency of anticipated turns, as indicated by gaze shift in both conversations for all age groups (with standard errors). An asterisk indicates a significant difference between normal and flattened conditions (**p* < .05). 89

Figure 4.4. Histograms of anticipatory gaze shifts (absolute values) during the 500-ms interval prior to a pause (light gray bars) and during a pause (dark gray bars) for both conditions and all age groups. Bin size is 250ms. Note that reactive gaze shifts are not included. 91

Figure 4.5. Occurrence rates of anticipatory and random gaze shifts for all age groups averaged over both conditions (with standard errors). Asterisks indicate a significant difference between the occurrence rates of anticipatory and random gaze shifts (*p < 05; **p < .01). 92

List of Tables

Table 3.1. Mean values and standard deviations of gaze latency (in ms) for both conditions for infants and adults. 63

Table 4.1. Details of the two conversations. Upper panel: number of turn-taking; total length of conversations, mean length of speech (i.e. turns) and pauses between turns; mean number of words per turn. Lower panel: mean and standard deviation of pitch for normal and flattened conditions. 83

Table 4.2. Mean frequency of anticipated turns (in %) in the normal and flattened conditions, and mean occurrence rates for anticipatory and random gaze shifts (both with standard deviations and for all age groups). 90

List of Abbreviations

ANOVA	Analysis of Variance
AOI	Area of Interest
AON	Action-Observation Network
EEG	Electroencephalography
MNS	Mirror-Neuron System
NIRS	Near-Infrared Spectroscopy

Curriculum Vitae

Name	Anne Keitel (geb. Melzer)
Familienstand	verheiratet, 2 Töchter
Geburtsdatum	12.07.1980
Geburtsort	Berlin

Seit 04/2013	Stipendiatin des Freistaates Sachsen (‘Wiedereinstiegsstipendium’)
01/2013 – 01/2014	*Elternzeit*
07/2012 – 12/2012	Doktorandin am Max-Planck-Institut für Kognitions- und Neurowissenschaften
06/2011 – 06/2012	*Elternzeit*
10/2008 – 05/2011	Doktorandin am Max-Planck-Institut für Kognitions- und Neurowissenschaften
10/2008 – 09/2012	Mitglied im Graduiertenkolleg ‘Funktion von Aufmerksamkeit bei kognitiven Prozessen’
10/2000 – 04/2008	Diplom in Psychologie, Technische Universität Berlin (Abschlussnote: 1,0)
09/1992 – 06/2000	Allgemeine Hochschulreife, 1. Gymnasium Prenzlauer Berg, Berlin

List of Publications

Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and observation of contralateral reaching: A close link by 12 months of age. *Infant Behavior & Development, 35*(3), 570-579.

Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C., & Daum, M. M. (2013). Perception of conversations: The importance of semantics and intonation in children's development. *Journal of Experimental Child Psychology, 116*, 264-277.

Atmaca, S., Stadler, W., **Keitel, A.**, Ott, D., Lepsien, J., & Prinz, W. (in press). Prediction processes during multiple object tracking (MOT): Involvement of dorsal and ventral premotor cortices. *Brain and Behavior*.

Keitel, A., Prinz, W., & Daum, M. M. (submitted). *Infants' and adults' perception of individual and joint action.*

SELBSTSTÄNDIGKEITSERKLÄRUNG

Hiermit erkläre ich, dass die vorliegende Arbeit ohne unzulässige Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt wurde und dass die aus fremden Quellen direkt oder indirekt übernommenen Gedanken in der Arbeit als solche erkenntlich gemacht worden sind.

Anne Keitel (geb. Melzer)

Leipzig, den 21. Oktober 2013

Bibliographic Details & Kurzzusammenfassung

BIBLIOGRAPHIC DETAILS

Anne Keitel (geb. Melzer)

Action perception in development: The role of experience

Fakultät für Biowissenschaften, Pharmazie und Psychologie

Universität Leipzig

Dissertation

141 pages, 107 references (Chapter 1), 11 figures, 3 tables

The perception of an action and its production are inextricably linked. This entails that, during development, the skills that children are able to perform influence their perception of other's actions. The present dissertation aimed to investigate the role of children's experience on the perception of actions in three distinctive areas: manual actions performed by one person (individual action), manual actions performed by two people (joint action), and a conversation between two people. In order to succeed in each of the three areas, children have to acquire new skills and do so successively during their first three years of life. The methodological approach of this work was to measure the gaze behaviour of children, aged 6 months to 3 years, and adults during the observation of visually presented actions, which provided information on whether they were able to anticipate action goals.

The findings obtained generally show an influence of experience on the anticipation of action goals in each of the three areas. First, a link between action and perception is not established as soon as an action emerges. There is at least some experience necessary for its development. Second, infants with no coordinated joint-action skills themselves anticipate the goals of joint action less well than those of individual action. Adults with considerable joint-action skills anticipate both equally well. And third, the course of a conversation can only be reliably anticipated by children aged 3 years and adults, whereas younger children shift their gaze between speakers randomly. Furthermore, only at the age of 3 years, did intonation support children's anticipation of conversations.

Summary / Zusammenfassung

Summary

Introduction

The production of an action and its perception are inextricably linked (for an overview, see Prinz, Aschersleben, & Koch, 2009). More precisely, there is a common representational domain between planned and observed actions (Prinz, 1990, 1997; Hommel, Müsseler, Aschersleben, & Prinz, 2001). This entails that, during development, the skills that children learn to perform influence their perception of other's actions (e.g., Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; Kanakogi & Itakura, 2011). Generally, actions can be performed by one person, such as simply grasping a cup of coffee ('nonverbal individual action'), but they can also be performed jointly by more people, such as two people preparing a dinner together ('nonverbal joint action'), or having a conversation ('verbal joint action'). In order to succeed in the performance of those nonverbal and verbal actions and interactions, children have to obtain different skills: Infants learn to perform various manual actions during their first year of life (e.g., Bourgeois, Khawar, Neal, & Lockman, 2005). During their second year of life, they learn to coordinate those individual actions with others in joint action (Brownell, 2011). During their third year of life, children master to verbally interact with others (Clark, 2003). Thus, each of the actions in the different areas requires new skills that are learned successively during the first 3 years of life.

Summary of the Dissertation

The present dissertation aimed to investigate how the increasing experience in performing nonverbal and verbal actions and interactions influences children's and adults' perception of others' actions in the respective areas. To this end, we visually presented participants of different age groups with actions and measured their gaze. The way that participants shift their gaze towards action goals (nonverbal actions) or between speakers of a conversation (verbal interaction) reveals whether they are able to anticipate goals, or the course of a conversation, respectively. The occurrence of anticipatory gaze shifts indicates that an observer has built a representation of the observed action goal that allows one to predict the outcome of the action before it is completed. For example, infants have been shown to anticipate the goals of many manual actions by 12 months of age (e.g., Cannon, et al., 2012; Falck-Ytter, Gredebäck, & von Hofsten, 2006). This ability is typically modulated by their own experience with the respective action (e.g., Gredebäck & Melinder, 2010).

The dissertation comprises three studies, addressing a nonverbal manual action performed by one individual (study one), a nonverbal manual action performed by two people (study two), and a visually presented conversation between two people (study three). In each study, different age groups were tested to assess the role of increasing experience on the anticipation of action goals.

Study 1: Common representation of individual action

The first study addressed the question whether a link (i.e., a common representation) between the perception and the production of individual action is established as soon as an action emerges during development, or whether more active experience is necessary for its formation. To this end, 6- and 12-month-old infants were presented with videos of contralateral manual actions (e.g., reaching across the body midline) and their gaze shifts towards action goals were measured. Additionally, infants' own contralateral reaching skills were tested using a task adopted from Bruner (1969). Contralateral reaching emerges in the middle of the first year of life and slowly improves over the next months. It is thus particularly useful to determine when a common representation between perception and action is established. The results showed that, as expected, the 12-month-olds performed contralateral reaching actions more often than the 6-month-olds. Furthermore, the 12-month-olds showed anticipatory gaze shifts towards action goals, whereas the 6-month-olds showed reactive gaze shifts (i.e., their gaze arrived at the goal after the action was completed). And, most importantly, a correlation between the two tasks was only present in 12-month-olds, but not yet in 6-month-olds. These results suggest that a common representation between action perception and production during development is not instantly present. Instead, the formation of such a link seems to depend on a certain amount of active experience, or ‘training’ in performing an action.

Study 2: The difference between individual and joint action in development

The second study concerned the perception of joint action. Children learn to coordinate their actions with others during the second year of life (Brownell, 2011). However, they passively observe others’ joint action from very early on. It is yet an unsolved question whether infants’ perception of joint action is essentially different from individual action, or whether both follow the same developmental trajectory. We addressed this question by presenting infants and adults with videos of a block-stacking action that was either performed by one agent or two agents, and compared their gaze behaviour towards action goals. The overarching goal was identical in both conditions (‘to build a tower’); only the sub-goals (‘to grasp a block’, ‘to stack it’) differed, in that they were performed by one or two agents. The tested infants were 9 and 12 months old, and had little or no experience with coordinated joint action themselves, whereas adults usually are very experienced in coordinating their actions with others. It was found that infants differed in their perception of individual and joint action, in that they anticipated individual action faster, whereas adults could anticipate both actions equally well. Infants possibly represented the sub-goals of the block-stacking action in isolation, which led to delayed gaze shifts in the joint condition, when they had to switch between the representations of the two agents’ sub-goals. Adults, however, were able to infer the overarching joint goal of two agents, which led to comparable gaze behaviour in both conditions. These findings suggest a modulating influence of experience on the perception of joint action.

Study 3: Conversation perception by prelinguistic and linguistic children

The third study investigated how increasing language experience (e.g., semantic skills) and the prosodic factor intonation (i.e., the rise and fall of the voice in speaking) influence the perception of conversations in prelinguistic and linguistic children. For this purpose, children of four age groups (6, 12, 24 and 36 months), and adults were presented with videos of two dyadic conversations, one with normal and one with flattened, monotone intonation. It was analysed how often participants were able to anticipate a turn (i.e., how often gaze was shifted to the next speaker before she started speaking). The first main finding was that only the 3-year-olds and the adults were able to reliably anticipate a speaker's next turn. This indicates that extensive language experience is necessary, before semantics are developed sufficiently to anticipate the course of a conversation. The second main finding was that only 3-year-olds benefited from intonation. Neither in the younger age groups nor in adults was the anticipation of turns affected by intonation. This suggests that intonation only has a supporting role on conversation perception, when language comprehension is well developed but still not as sophisticated as that of adults. Thus, language experience alters the proficiency to use this prosodic cue for the perception of conversations.

Conclusions

This dissertation investigated the influence of increasing experience on action perception in three distinctive areas, namely, individual action, joint action and conversations. Different levels of cognitive development are necessary to succeed in the performance of actions in the respective areas. Likewise, the perception of those distinct actions depends on the skills (or experience) that children have gained during development. First, experience alters the perception of individual manual action. Their understanding involves representing the goals of 'simple motor actions' (or 'elementary motor acts', Csibra, 2007). Specifically, experience not only supports anticipation of such individual action goals, but also promotes a common representation between individual action and perception. Second, experience supports the perception of joint action. This involves representing others' joint goals, in addition to the representations of simple motor actions performed by the agents. And third, experience supports the perception of conversations. This involves representing the semantic and syntactic information of a conversation, as well as using prosodic information, such as intonation.

The findings of the present dissertation thus demonstrate a significant role of experience on action perception on different cognitive levels, from simple motor actions to complex conversations. Experience plays a special role during development, when new actions and skills are learned constantly. However, it seems natural to conclude that experience continues to play a significant role for action perception throughout one's life, which is supported by studies with older and younger adults (Cross, Hamilton, & Grafton, 2006; Diersch, Cross, Stadler, Schütz-Bosbach, & Rieger, 2012).

Zusammenfassung

Allgemeine Einführung

Die Ausführung und die Wahrnehmung einer Handlung sind untrennbar miteinander verbunden (für einen Überblick, Prinz, Aschersleben, & Koch, 2009). Genauer gesagt, teilen beide Bereiche eine gemeinsame Repräsentation ('common representation', Prinz, 1990, 1997). Daraus resultiert auch, dass die Fähigkeiten, die Kinder während ihrer Entwicklung erlernen, ihre Wahrnehmung von Handlungen anderer Personen beeinflussen (z.B. Cannon, Woodward, Gredebäck, von Hofsten, & Turek, 2012; Kanakogi & Itakura, 2011). Im Allgemeinen können Handlungen von einer einzelnen Person ausgeführt werden, wie zum Beispiel das Greifen nach einer Tasse Kaffee ('nonverbale individuelle Handlung'). Sie können aber auch von mehreren Personen gemeinsam ausgeführt werden, so wie das gemeinsame Vorbereiten des Abendessens ('nonverbale gemeinsame Handlung', im Englischen bekannt als 'Joint Action') oder ein dyadisches Gespräch ('verbale gemeinsame Handlung'). Um solche nonverbalen und verbalen, individuellen und gemeinsamen Handlungen erfolgreich auszuführen, müssen Kinder verschiedene Fähigkeiten erlernen: Säuglinge lernen im ersten Lebensjahr verschiedene manuelle Handlungen auszuführen (z.B. Bourgeois, Khawar, Neal, & Lockman, 2005). Im zweiten Lebensjahr lernen Kinder, solche individuellen Handlungen mit anderen Personen zu gemeinsamen Handlungen zu koordinieren (Brownell, 2011). Im dritten Lebensjahr meistern Kinder die verbale Interaktion mit Anderen (Clark, 2003). Die Handlungen in den verschiedenen Bereichen erfordern daher neue Fähigkeiten, die Kinder allmählich während der ersten drei Lebensjahre lernen.

Die vorliegende Dissertation hatte zum Ziel, den Einfluss der zunehmenden Erfahrung in der Ausführung von nonverbalen und verbalen, individuellen und gemeinsamen Handlungen auf die Handlungswahrnehmung (bzw. das Handlungsverständnis) zu untersuchen. Dazu wurden Teilnehmern verschiedenen Alters Handlungen visuell präsentiert und ihr Blickverhalten gemessen. Die Art, wie Versuchsteilnehmer ihren Blick auf ein Handlungsziel (nonverbale Handlungen) oder zwischen zwei Gesprächspartnern (verbale Handlung) verschieben, kann darüber Aufschluss geben, ob sie die Handlungsziele, bzw. den Gesprächsverlauf, antizipieren können. Das Auftreten von antizipatorischem Blickverhalten indiziert, dass ein Beobachter eine Repräsentation des Handlungsziels aufgebaut hat, die es ihm ermöglicht, den Ausgang einer Handlung vorherzusagen, bevor sie vollständig ausgeführt wurde. Säuglinge können zum Beispiel die Ziele von vielen manuellen Handlungen im Alter von 12 Monaten antizipieren (z.B. Cannon, et al., 2012; Falck-Ytter, Gredebäck, & von Hofsten, 2006). Diese Fähigkeit wird typischerweise von der eigenen Erfahrung der Kinder moduliert (z.B. Gredebäck & Melinder, 2010).

Zusammenfassung der Dissertation

Diese Dissertation beinhaltet drei Studien, in welchen die Wahrnehmung einer individuellen manuellen Handlung (Studie 1), als auch die einer gemeinsamen manuellen Handlung (Studie 2), und die eines visuell präsentierten Gesprächs (Studie 3) untersucht werden. In jeder Studie wurden verschiedene Altersgruppen getestet, um den Einfluss der zunehmenden Erfahrung auf die Antizipation von Handlungszielen zu untersuchen.

Studie 1: Die gemeinsame Repräsentation von individuellen Handlungen

In der ersten Studie wurde untersucht, ob ein Zusammenhang zwischen der Wahrnehmung und der Ausführung einer individuellen Handlung (d.h. eine gemeinsame Repräsentation) gebildet wird, sobald eine Handlung während der Entwicklung erlernt wird, oder ob mehr aktive Handlungserfahrung dafür nötig ist. Dazu wurden sechs- und zwölfmonatigen Säuglingen Videos von kontralateralen Handlungen präsentiert (z.B. Greifen quer über die Körpermittellinie) und ihre Blickbewegungen zu Handlungszielen gemessen. Zusätzlich wurde die Fähigkeit von

Säuglingen selbst kontralateral zu greifen mit einem Paradigma gemessen, das von Bruner (1969) entwickelt wurde. Kontralaterales Greifen entsteht in der Mitte des ersten Lebensjahres und entwickelt sich über die nächsten Monate langsam weiter. Daher ist dies ein besonders geeignetes Paradigma, um die Entstehung eines Zusammenhangs zwischen Handlungswahrnehmung und –ausführung zu untersuchen. Die Ergebnisse zeigen, dass Zwölfmonatige wie erwartet öfter kontralaterale Handlungen ausführten als Sechsmonatige. Ausserdem konnten die Zwölfmonatigen die Handlungsziele antizipieren, während die Sechsmonatigen reaktive Blickbewegungen zeigten (d.h. ihr Blick erreichte das Ziel erst nachdem die Handlung bereits beendet war). Das wichtigste Ergebnis war jedoch, dass eine Korrelation zwischen beiden Aufgaben erst bei Zwölfmonatigen nachzuweisen war, jedoch noch nicht bei Sechsmonatigen. Diese Ergebnisse deuten darauf hin, dass eine gemeinsame Repräsentation zwischen Handlungswahrnehmung und –ausführung nicht unmittelbar gebildet wird. Stattdessen ist für die Herausbildung eines solchen Zusammenhangs ein gewisses Ausmaß an aktiver Erfahrung oder ‘Training’ mit der Handlung notwendig.

Studie 2: Der Unterschied zwischen individueller und gemeinsamer Handlung

Die zweite Studie betraf die Wahrnehmung von gemeinsamen Handlungen. Kinder erlernen während des zweiten Lebensjahres, ihre Handlungen gemeinsam mit anderen zu koordinieren (Brownell, 2011). Sie können jedoch von Anfang an gemeinsame Handlungen beobachten. Es ist noch unerforscht, ob die Wahrnehmung von gemeinsamen Handlungen sich essenziell von individuellen Handlungen unterscheidet, oder ob sie den gleichen Entwicklungsverlauf nimmt. Wir haben diese Frage adressiert, indem wir Säuglingen und Erwachsenen eine Turmbauhandlung präsentiert haben, die entweder von einer oder zwei Personen ausgeführt wurden. Das Blickverhalten der Versuchsteilnehmer zu den beobachteten Handlungszielen konnte dann zwischen beiden Bedingungen verglichen werden. Das übergeordnete Ziel dieser Handlungen war in beiden Bedingungen gleich (‘einen Turm bauen’), die Unterziele (‘einen Stein greifen’, ‘ihn auf den Turm legen’) unterschieden sich jedoch darin, dass sie von ein oder zwei Personen ausgeführt wurden. Die getesteten Säuglinge waren neun und zwölf Monate alt und verfügten über wenig bzw. gar keine eigene Erfahrung mit koordinierten gemeinsamen Handlungen. Erwachsenen sind dagegen normalerweise sehr erfahren darin, ihre Handlungen mit anderen zu koordinieren. Es konnte gezeigt werden, dass Säuglinge individuelle und gemeinsame Handlungen unterschiedlich wahrnahmen. Sie antizipierten Handlungsziele der individuellen Handlung schneller als die der gemeinsamen. Die Erwachsenen konnten beide Handlungsziele gleich gut antizipieren. Säuglinge haben möglicherweise die Unterziele der Turmbauhandlung isoliert repräsentiert, was zu verzögerten Blickbewegungen geführt hat, wenn zwischen den Repräsentationen der Handlungsziele zweier Personen gewechselt werden musste. Die Erwachsenen waren dagegen dazu in der Lage, das übergeordnete Ziel beider Handlungen zu inferieren, was zu einem ähnlichen Blickverhalten in beiden Bedingungen geführt hat. Diese Ergebnisse deuten auf einen modulierenden Einfluss von Erfahrung auf die Wahrnehmung von gemeinsamen Handlungen.

Studie 3: Gesprächswahrnehmung durch vorsprachliche und sprechende Kinder

In der dritten Studie wurde untersucht wie sich die zunehmende Spracherfahrung (z.B. semantische Fähigkeiten) und der prosodische Aspekt der Intonation (d.h. die Sprachmelodie) auf die Wahrnehmung von Gesprächen durch vorsprachliche und sprechende Kinder auswirken. Zu diesem Zweck wurden Kindern aus vier Altersgruppen (6, 12, 24 und 36 Monate) und Erwachsenen Videos von zwei dyadischen Gesprächen präsentiert, eines mit normaler Intonation und eines mit abgeflachter, monotoner Intonation. Es wurde ausgewertet, wie oft die Versuchsteilnehmer einen Sprecherwechsel antizipieren konnten (d.h. wie oft der Blick zur nächsten Sprecherin wechselte, bevor diese angefangen hatte zu sprechen). Der erste Hauptbefund war, dass nur die Dreijährigen und die Erwachsenen einen Sprecherwechsel zuverlässig vorhersagen konnten. Dies deutet darauf hin, dass umfassende Spracherfahrung notwendig ist, damit das Sprachverständnis weit genug entwickelt ist, um den Verlauf eines Gespräches zu antizipieren. Der zweite Hauptbefund war, dass nur die Dreijährigen vom Vorhandensein normaler Intonation profitierten. Weder die jüngeren Altersgruppen noch die Erwachsenen waren bei ihrer Antizipation der Sprecherwechsel von Intonation beeinflusst. Dies deutet auf eine unterstützende Rolle der Intonation für die Gesprächswahrnehmung, wenn das

Sprachverständnis schon weit entwickelt, jedoch noch nicht so fortgeschritten ist, wie das von Erwachsenen. Spracherfahrung wirkt sich somit auf die Fähigkeit aus, den prosodischen Aspekt der Intonation für die Antizipation eines Gesprächsverlaufs zu nutzen.

Schlussfolgerungen

In dieser Dissertation wurde untersucht, welchen Einfluss die zunehmende Handlungserfahrung auf die Wahrnehmung von Handlungen in drei unterschiedlichen Bereichen hat, nämlich individuelle Handlungen, gemeinsame Handlungen und Gespräche. Es sind verschiedene kognitive Entwicklungsstufen notwendig, um Handlungen in den Bereichen erfolgreich auszuführen. Aber auch die Wahrnehmung dieser unterschiedlichen Handlungen hängt von den Erfahrungen und Fähigkeiten ab, die Kinder während ihrer Entwicklung gewinnen. Erstens, Erfahrung ändert das Verständnis von individuellen, manuellen Handlungen. Dies beinhaltet die Repräsentation von 'einfachen motorischen Handlungen' (oder 'elementary motor acts', Csibra, 2007). Erfahrung unterstützt nicht nur die Antizipation von solchen einfachen Handlungszielen, sondern fördert auch eine gemeinsame Repräsentation von Handlungswahrnehmung und –ausführung. Zweitens, Erfahrung fördert das Verständnis von gemeinsamen Handlungen. Zusätzlich zur Repräsentation von einfachen Handlungszielen erfordert dies auch die Repräsentation von gemeinsamen Handlungszielen. Drittens, Erfahrung unterstützt das Verständnis von Gesprächen. Dies beinhaltet die Repräsentation von semantischen und syntaktischen Informationen des Gesprächs, als auch die Nutzung von prosodischer Information wie Intonation.

Die Befunde der vorliegenden Dissertationen zeigen demzufolge einen bedeutenden Einfluss von Erfahrung auf die Wahrnehmung und das Verständnis von Handlungen auf verschiedenen kognitiven Ebenen, von einfachen motorischen Handlungen bis hin zu komplexen Gesprächen. Erfahrung spielt demnach eine besondere Rolle während der Entwicklung, wenn fortwährend neue Handlungen und Fähigkeiten erlernt werden. Es scheint jedoch naheliegend, dass Erfahrung auch im weiteren Lebensverlauf eine bedeutende Rolle für die Handlungswahrnehmung spielt. Dies wird von Studien bestätigt, die junge und ältere Erwachsene untersuchten (Cross, Hamilton, & Grafton, 2006; Diersch, Cross, Stadler, Schütz-Bosbach, & Rieger, 2012).

References

Bourgeois, K. S., Khawar, A. W., Neal, S. A., & Lockman, J. J. (2005). Infant manual exploration of objects, surfaces, and their interrelations. *Infancy, 8*(3), 233-252. doi: DOI 10.1207/s15327078in0803_3

Brownell, C. A. (2011). Early developments in joint action. *Review of Philosophy and Psychology, 2*(2), 193-211. doi: 10.1007/s13164-011-0056-1

Bruner, J. S. (1969). Eye, Hand, and Mind In D. Elkind & J. H. Flavell (Eds.), *Studies in Cognitive Development: Essays in Honor of Jean Piaget*. Oxford: University Press, Inc.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science, 15*(1), 35-42. doi: 10.1111/j.1467-7687.2011.01095.x

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Cross, E. S., Hamilton, A. F., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *Neuroimage, 31*(3), 1257-1267. doi: DOI 10.1016/j.neuroimage.2006.01.033

Csibra, G. (2007). Action mirroring and action understanding: an alternative account In P. Haggard, Y. Rosetti & M. Kawato (Eds.), *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII* (pp. 435 - 459). Oxford: Oxford University Press.

Diersch, N., Cross, E. S., Stadler, W., Schütz-Bosbach, S., & Rieger, M. (2012). Representing others' actions: the role of expertise in the aging mind. *Psychological Research*, *76*(4), 525-541. doi: 10.1007/s00426-011-0404-x

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience, 9*(7), 878-879.

Gredebäck, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition, 114*(2), 197-206. doi: 10.1016/j.cognition.2009.09.004

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*(5), 849-937.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature Communications, 2*. doi: 10.1038/Ncomms1342

Prinz, W. (1990). A common coding approach to perception and action In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167 - 201). Berlin: Springer Verlag.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129-154.

Prinz, W., Aschersleben, G., & Koch, I. (2009). Cognition and action. In E. Morsella, J. A. Bargh & P. M. Gollwitzer (Eds.), *Oxford handbook of human action.* (pp. 35-71). Oxford: University Press.

Confirmation of contribution of co-authors:

Title	Production and perception of contralateral reaching: A close link by 12 months of age
Journal	*Infant Behavior and Development*
Authors	Anne Melzer, Wolfgang Prinz, Moritz M. Daum

Contribution of Anne Keitel (nee Melzer)

- project conception
- experimental design
- preparation of experimental stimuli
- data acquisition
- data analysis
- writing of the manuscript

Contribution of Wolfgang Prinz

- project conception
- writing of the manuscript

Contribution of Moritz M. Daum

- project conception
- experimental design
- writing of the manuscript

Anne Keitel (nee Melzer)

Wolfgang Prinz

Moritz M. Daum

Confirmation of contribution of co-authors:

Title	Infants' and adults' perception of individual and joint action
Journal	submitted to *Cognition*
Authors	Anne Keitel, Wolfgang Prinz, Moritz M. Daum

Contribution of Anne Keitel

- project conception
- experimental design
- preparation of experimental stimuli
- data acquisition
- data analysis
- writing of the manuscript

Contribution of Wolfgang Prinz

- writing of the manuscript

Contribution of Moritz M. Daum

- project conception
- writing of the manuscript

Anne Keitel

Wolfgang Prinz

Moritz M. Daum

Confirmation of contribution of co-authors:

Title	Perception of conversations: The importance of semantics and intonation in children's development
Journal	*Journal of Experimental Child Psychology*
Authors	Anne Keitel, Wolfgang Prinz, Angela D. Friederici, Claes von Hofsten, Moritz M. Daum

---

Contribution of Anne Keitel

- project conception
- experimental design
- preparation of experimental stimuli
- data acquisition
- data analysis
- writing of the manuscript

Contribution of Wolfgang Prinz

- project conception
- writing of the manuscript

Contribution of Angela D. Friederici

- project conception
- writing of the manuscript

Contribution of Claes von Hofsten

- project conception
- writing of the manuscript

Contribution of Moritz M. Daum

- project conception
- writing of the manuscript

Anne Keitel

Wolfgang Prinz

Angela D. Friederici

Claes von Hofsten

Moritz M. Da no