# **Intentional Communication in Infants and Toddlers**

Edmund Anthony Donnellan

A thesis submitted for the degree of Doctor of Philosophy

Department of Psychology, University of Sheffield

June 2017

**Abstract**

Theories of intentional communication suggest that when we communicate we have two types of intention. Our *communicative* intention is that someone recognises *that* we are communicating, while our *informative* intention is that someone infers *what* we are communicating about. In this thesis we investigate the development of these two kinds of intention.

In the first two empirical chapters, we investigated whether toddlers understand the intentional structure of communication. In Chapter 2, we explored whether toddlers pretend not to hear their caregivers' requests (deliberately not recognising a communicative intention). Though caregivers reported that toddlers from 18 months pretended not to hear, we were unable to demonstrate this experimentally. In Chapter 3 we investigated whether toddlers detect a difference between when someone doesn't notice another's request, and when someone ignores it. We do not observe differences in toddlers' responses that suggest that they are sensitive to the intentional structure of communication.

In the latter three empirical chapters, we investigated infants' informative intentions. In Chapter 5, we tested whether 11- and 12-month-olds' vocalisations, gestures or gesture-vocal combinations are coordinated with gaze to their caregiver's face, indicating an understanding that informative intentions are met by directing the attention of others. Gestures, vocalisations and combinations co-occurred with gaze above chance levels. In Chapter 6, we considered whether these infant behaviours predict later expressive vocabulary. Fine-grained analyses of specific types of gesture and vocalisation revealed that the frequency of some behaviours (particularly gaze-coordinated CV vocalisations) were strong predictors of 15-month expressive vocabulary. Predictive relations with 24-month vocabulary were less strong, but gestures were better predictors in this case. Finally, in Chapter 7 we considered caregiver responses to infant behaviours and found that in many cases these mediated the relationship between these behaviours and later vocabulary. Collectively these findings support the hypothesis that infants have informative intentions around their first birthday, and caregivers' responses to intentional communication drive vocabulary development.

## Acknowledgements

Firstly, I would like to thank Danielle Matthews for her supervision. It is a great honour to work with you. Thank you for your tireless work and support in producing this thesis, and also for all your guidance in helping me learn the academic ropes.

I would also like to thank Katie Slocombe who also helped supervise the PhD. Thank you for your support throughout, and especially for organising for me take a break to study the chimpanzees.

As well as providing supervision on the PhD, I would like to thank both Danielle and Katie for their help from 2011, in trusting someone with no background in psychology to work on their projects. Without this, I may never have been able to transition to psychology, and the world this opens up, so my deepest thanks to the both of you for this too.

Thanks also to Elena Hoicka for her part on my supervisory team, and to Richard Moore for the fruitful discussions over the studies in the first half of this thesis. Thanks also to Colin Bannard for his assistance with statistics throughout the thesis. I am truly grateful to Danielle and Michelle McGillion for allowing and trusting me to use the rich dataset that they worked so hard to collect for the studies in the second half of this thesis. Also, thanks are due to Andy Ham for his technical assistance in constructing materials vital for testing.

I owe a massive debt gratitude to all the families that generously devoted their time to participate in the studies contained within this thesis.

I would like to thank a number of people who have helped with the collection and coding of data within this thesis. Thanks to Dominique Gardner, Ashley Taafe, Arthur Nye, Michelle and Miklos Kurthy for their patient and skilful assistance in the experiments in chapters 2-3. Thanks also to Isobel Dunnett-Orridge for her meticulous coding on the data included in chapters 5-7.

I have been very fortunate to have been part of the Sheffield University babylab since 2011. I would like to thank all members of the babylab, but in particular Michelle, Lily Fitzgibbon and Emma Blakey who were there early on to provide reassurance, sage advice and great chat over a cup of tea. I would also like to thank other members of the babylab (past and present), especially Andrea Díaz-Barriga Yáñez, Birsu Kandemirci, Lowri Bridgett-Thomas, Anna Ryder, Kiera Solaiman, Simone Bijvoet-van den Berg, Sophie Turnbull, Ciara Kelly, Steph Powell and Gemma Stephens. I'm also fortunate to have worked with students at the University of York both before and during my PhD. Thanks especially to Pawel Fedurek for having me as a collaborator, and for fascinating discussions on all things chimpanzee, and to Cloud Wilke for having me as a research assistant for five fantastic months in the rainforest.

Thank you to George Botterill and Stephen Laurence, who helped me complete my MA in Cognitive Studies, and encouraged me to move on towards the psychology department. I am indebted to you both for your advice and support throughout my undergraduate degree, masters and in the intervening years before I began this PhD. Thanks also to Tom Stafford and Pia Nystrom who also helped aid my transition out of the philosophical armchair.

Outside of academia, I would like to thank everyone at Beanies for providing me with employment before, and during my studies. I'd especially like to thank Matt West for providing me with work on his farm on days when the PhD was tough and I wanted to be outside in the fresh air, covered in soil and hounded by dogs. And to all my friends, flatmates and bandmates who I have often shamefully neglected during the PhD: thank you for the pub trips, music, juggling, cycling and laughter. To the McGillions and the Corcorans (Brian, Kate, Tracy, Barry, Sean, Ben and Ollie) thank you for the warmest of welcomes I've been given every time I've come across the Irish Sea.

To my family, especially my Mum and Dad, I could not have done this without you. Thank you for quite literally, everything. To JP, Ben and Francesca, thank you for the unending silliness that seems to characterise our every interaction, I couldn't have asked for better siblings. Thank you also to Zoe and Ikue for putting up with the rest of us at loud family gatherings, and to my wonderful and fun niece and nephew Nene and Sosuke, who have provided fun distractions from work.

Finally, to Michelle, thank you. Thank you again not only for all the practical and material help you have given to me to enable me to complete this thesis. You've helped in so many more ways, with your love and support. I'm so lucky to have had you around and I'm forever grateful for all that you have done for me. There is no way I could have done this without you.

## Declaration

I hereby declare that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Edmund Anthony Donnellan

**Presentations**

Studies in this thesis have been presented as oral presentations at the following conferences:

Donnellan, E., Moore, R., Hoicka, E., Slocombe, K.E., & Matthews, D.E. (2014) 'Do toddlers understand the intentional structure of communication?' *University of Sheffield Psychology Postgraduate Conference*, May 12-15th, Sheffield, England.

Donnellan, E., McGillion, M.L., Slocombe, K.E., & Matthews, D.E. (2016) 'Intentional Communication in Human Infants' *Scottish Primate Research Group Annual Meeting*, March 4th-6th, The Burn, Glenesk, Scotland; and *University of Sheffield Psychology Postgraduate Conference*, May 16th-19th Sheffield, England.

Donnellan, E., McGillion, M.L., Slocombe, K.E., & Matthews, D.E. (2017) 'Intentional Vocal and Gestural Communication and Early Vocabulary in Human Infants' *Scottish Primate Research Group Annual Meeting*, March 10th-12th, The Burn, Glenesk, Scotland.

Studies in this thesis have been presented as posters at the following conferences:

Donnellan, E., Moore, R., Hoicka, E., Slocombe, K.E., & Matthews, D.E. (2015) 'Pretending not to hear: potential demonstrations of toddlers' understanding of the communicative intention. *Reciprocity and Social Cognition Symposium*, March 23rd-25th, Berlin School of Mind and Brain, Germany; *University of Sheffield Psychology Postgraduate Conference*, May 18th-21st Sheffield, England; *Child Language Symposium*, July 20th-21st, University of Warwick, England; and *Perspectives on the Ontogeny of Mutual Understanding workshop*, Max Planck Institute for Psycholinguistics, October 1st-2nd, Nijmegen, Netherlands.

Donnellan, E., McGillion, M.L., Slocombe, K.E., & Matthews, D.E. (2017) 'Do Prelinguistic Vocalisations and Gestures Predict Later Language Because They Are Early Instances of Intentional Communication?' *Society for Research in Child Development Biennial Meeting*, April 6th-8th, Austin, Texas, USA.

**Contents**

# 1. Intentional Communication in Infants and Toddlers

This thesis focuses on intentional communication in the first years of life, and how infants' understanding of intentional communication, and their ability to produce intentionally communicative behaviour aids their transition to language. Before articulating exactly what questions are to be explored within the thesis, this opening review chapter has four aims. Firstly, we give an account of what intentional communication is. Secondly, we review the literature that has considered when in development children begin to produce intentionally communicative behaviour, and how this can be measured (a question we return to for a more detailed review in Chapter 4). Thirdly, we review the literature on when children understand when others are intentionally communicating. Fourthly, we review the literature on specifically how children understand the complex intentional structure of communication, and how this could be tested. Finally, when we have covered all these key areas, we consider the two main questions that will be asked in this thesis.

## What is Intentional Communication?

The idea that there are multiple layers of intention underlying human communication has been well established and refined, building on the work of Paul Grice. Grice (1957, p. 383) defined conditions for an intentionally communicative act so that, for a speaker A producing an utterance *x*:

> '*A* must intend to induce by *x* a belief in an audience, and he must also intend his utterance to be recognized as so intended.'

Following Grice, a number of authors have distinguished between two separable types of intention in intentionally communicative behaviours (Bruner, Roy, & Ratner, 1982; Gómez, 1994; R. Moore, 2016; Sperber & Wilson, 1986). These are termed the *informative intention*, and the *communicative intention*. Informative intentions, serve 'to inform the audience of something', while communicative intentions serve 'to inform the audience of one's informative intention' (Sperber & Wilson, 1986, p. 29). These intentions are separable to an audience, as demonstrated when attempts at communication are misunderstood. For example, Sperber & Wilson (1986) argued that an audience can recognises that you are communicating (correctly detecting your communicative intention), but might not correctly infer what you want (the informative intention).

It is worth noting that the informative intention, on some accounts, might involve two further separable levels of intention. As the informative intention is not necessarily 'informative' in function (i.e., we might not intend to inform another person about something), this has led some to reclassify this as a *social* intention involving two separable elements (Tomasello, 2008). Tomasello (2008) splits the informative intention into two types of intention: the social intention, and the referential intention. The referential intention is what we intend our communication to be about, while the social intention is what we intend our communication to achieve *concerning* that referent (so in most respects it is effectively an informative intention). Similar distinctions have also been made by others at this level (Greenfield & Dent, 1979, p. 568; Prizant & Wetherby, 1987). While recognising these nuances, for the purpose of addressing the questions of this thesis, we will not distinguish between referential and social intentions, instead opting to consider these as part of the informative intention (R. Moore, 2016).

Often inferring just what another's informative intention is can be complicated. If I point to a kettle, I might be implying any number of things about the kettle, e.g., that you turn it on, that it's an especially nice kettle, that I like the colour of the kettle, or, that the kettle is located *there*. Tomasello (2008) distinguishes three types of informative intention, which might be thought of as types of *functions* of an intentionally communicative act – exactly what it is that you intend to happen as a result of your actions. These are requestive (e.g., turn on the kettle), informative (e.g., there is the kettle you were looking for) and expressive (e.g., that's a nice kettle). This framework is illustrative, however non-exhaustive, as there can of course be finer-grained categories of each type of intention. For example, requests can be for objects, for someone to interact with an object, or for support in doing some activity (Bruner et al., 1982). The key point is that communicators' informative intentions are not necessarily inferable just from the communicator's act alone (in this case, extension of the index-finger towards an object). An audience is required to infer a communicator's informative intention using prior knowledge, or context.

There is a debate in the literature over how to characterize the intentional structure of communication in terms of how the two types of intention interact. This has implications for how it is possible to learn about intentional communication, and so we review this debate here. Consider the following examples:

1. A child is searching for their favourite ball. The child's mother remembers seeing the ball in a corner that the child has neglected to search. The mother points to the corner.
2. A child is searching for their favourite ball. The child's father remembers seeing the ball in a corner that the child has neglected to search. The father says, "I saw the ball in the corner".

Even though the mother points, and the father vocalises, both have the same intentional structure behind their actions. The two types of intention in both the mother's pointing, and the father's utterance can be characterized as follows:

*Informative Intention*: I intend for you to know the location of the ball

*Communicative intention*: I intend for you to recognise that {I intend for you to know the location of the ball}

The communicative intention is often characterized as an intention *about* the informative intention. Thus, this has led some to argue for an embedded structure of intentions in communication. Sperber (2000) puts forward the scenario where a woman (Mary) is eating berries, but doing so with the intention to communicate to a man (Peter) that they are edible. Peter must then hold the following belief: 'Mary intends that he [Peter] should believe that she intends that he should believe that these berries are edible' (Sperber, 2000). Sperber (2000) describes this as a 'fourth-order metarepresentational belief'(see also Dennett, 1983), something of the form:

| | |
|---|---|
| 4th order | Mary intends |
| 3rd order | that Peter believes |
| 2nd order | that Mary intends |
| 1st order | that Peter believes |
| | that the berries are edible |

A serious concern about this complicated structure is that it is perhaps beyond the scope of younger children to make inferences about the higher-order intentions of others. However, communication does not have to have such a complex metarepresentational form (R. Moore, 2016; Robert Thompson, 2014). The trouble arises from embedding intentions into a single structure with higher-order levels, so that recognising that someone is communicating requires higher-order inferences rather than just inferring simple intent from someone's actions (e.g., they intend to eat berries). Instead, it is possible that inferring *that* someone else's actions are communicative, is a separate process from inferring *what* it is that they are communicating about. This then only requires making two 'second-order' inferences and not higher level ones, which is less cognitively demanding (R. Moore, 2016). To reformulate the example above, a communicator acts with the following two separable sets of intentions, each more easily inferable by the observer:

| | |
|---|---|
| 2nd order | Mary intends |
| 1st order | that Peter recognizes |
| | that Mary is addressing an action towards him |

*and:*

| | |
|---|---|
| 2nd order | Mary intends |
| 1st order | that Peter believes |
| | that the berries are edible |

This dual intentional structure makes understanding of communication much more tractable, and increases the likelihood that younger children and even non-human animals might understand how communication works. We only have to detect *that* someone might be communicating (this normally involves them soliciting our attention in some way), before attempting to infer *what* they are communicating about. Similarly, in order to make requests, we have to indicate *that* we are communicating, and *what* we are communicating about (Bruner et al., 1982). In fact, R. Moore (2016) argues that in many cases, explicitly having to acknowledge that someone is communicating, or explicitly marking your behaviour as communicative is unnecessary if physical orientation already implies that actions will be communicatively addressed, such as in face-to-face interactions. As such it may be that the communicative intention requires more behavioural manifestation when we initiate an interaction, compared to when we sustain one. This does not demonstrate that having or detecting communicative intentions is unnecessary for communication, just that a communicator may not engage in any extra behaviour to mark their communicative intention in an ongoing interaction. Equally an audience may assume that an interlocutor has a communicative intention from the context of the interaction.

In sum, throughout this thesis, we assume that intentional communication consists of actions that are produced with two separable types of intention: the communicative intention (intending *that* someone recognises your action as communicative), and the informative intention (intending *what* it is that you are communicating about). We assume that such intentional communication does not depend on higher-order mindreading abilities. In the first half of this thesis, we will refer to the *intentional structure of communication*. This refers specifically to the relationship between communicative and informative intentions, in that communication functions because when you detect an interlocutor's communicative intention, you attempt to infer their informative intentions (i.e., on recognition that someone is trying to communicate, you attempt to infer what they are communicating about).

In what follows, we review evidence from studies of communication in infancy and early childhood. Sometimes these studies have sought to distinguish the two levels of intention, other times this distinction is not made but rather intentional communication is, for example, contrasted to vocal and gestural acts that produced without the expectation of having an effect on the attention of an interlocutor. Given the hidden nature of intentions it is hard to be certain of what lies behind any given act and we bear this in mind during the review. In the next part of this review, we focus on young children's production of acts that are thought to be intentionally communicative, and we discuss whether it is possible to determine if they have a communicative intention, or an informative intention. Later in this chapter, we focus on young children's comprehension of others' intentional communication in relation to their informative intentions. Finally we focus on whether there is evidence that young children understand the intentional structure of communication, that is to say, whether they understand the relation between communicative and informative intentions in communication.

## When Do Children Begin to Produce Intentionally Communicative Behaviours?

Prior to investigating what evidence there is that infants produce intentional communication (either at the level of informative or communicative intentions), we briefly provide a review of the emergence of vocalisations and gestures produced by infants over development. We then review literature on how it is possible to demonstrate empirically whether infants' actions are intentionally communicative and argue that this is only possible at the level of informative intentions. Finally, we review the literature that demonstrates when young children produce intentionally communicative behaviours in this sense.

### Early Communicative Behaviours

Infants vocalise from birth. Many of these vocalisations are vegetative, such as burps, coughs and sneezes, and crying behaviours which persist throughout life. However, the emergence of non-vegetative and non-cry infant vocalisations can be classified into broad developmental stages (Oller, 2000; Stark, 1980; Vihman, 1996). During the 'phonation' (Oller, 2000) or 'reflexive vocalisation' (Vihman, 1996) stage, 0- to 2-month-olds start to make short vowel-like vocalisations. Following this, from around 1 to 4 months, infants begin cooing and laughing. Subsequently, during the 'expansion' stage (Oller, 2000) or 'vocal play' stage (Stark, 1980; Vihman, 1996), 3- to 8-month-olds begin to make fully formed vowel sounds, start to blow raspberries and experiment with closing their vocal tract while vocalising, resulting in sounds known as marginal babbling. They begin to explore the pitch (squealing and growling) and dynamics of vocalisations (yelling and whispering). These aspects of prosody are required for later speech, and develop in both early vowel sounds and later babble. There is also a suggestion that infants crying behaviour shows developments in prosody that might help infants learn the sorts of vocal control required for speech (Wermke, Mende, Manfredi, & Bruscaglioni, 2002).

From around 5 months, infants begin to engage in canonical babbling, where they produce repetitive sequences of consonant-vowel (CV) syllables, e.g., baba, dada (Oller, 2000). Further, after 10 months, infants start to engage in variegated babble, where two different CV syllables are used in a sequence. This sets up the infant to produce their first words (Stark, 1980; Vihman, 1996). It is thought that much of what drives word learning is caregivers shaping their child's babble into words. For example, a sequence of babbled 'ba' is initially responded to by a mother asking if the child is saying 'boot', before later revising this to 'bubble' (Ramsdell, Oller, Buder, Ethington, & Chorna, 2012, p. 1268). These prelinguistic vocalisations and how they relate to later language are of particular importance in the second half of this thesis, and will be discussed further in Chapter 4.

Later in development, as well as producing vocalisations, infants begin to produce a number of gestures. Around the end of the first year of life, they begin engaging in giving and showing gestures, extending objects of interest towards adults, or holding these up to adults' faces (Cameron-Faulkner, Theakston, Lieven, & Tomasello, 2015; Harding & Golinkoff, 1979; Leung & Rheingold, 1981; Masur, 1983). One gesture that has been the focus of much research is index-finger pointing. Around the end of the first year of life, infants begin to extend their arm fully, form an index-finger pointing shape and direct it towards an area of space which usually contains an object (E. Bates, Camaioni, & Volterra, 1975; Bruner, 1975; Masur, 1983). Other types of pointing have been observed, from open-hand points (Cochet & Vauclair, 2010), to rarer lip-pointing in Laos (Enfield, 2001) and Panama (Sherzer, 1973) and nose-pointing in Papua New Guinea (Cooperrider & Núñez, 2012), but these are not universal. Indeed, while most studies of index-finger pointing have taken place in the UK (e.g., Matthews, Behne, Lieven, & Tomasello, 2012), the USA (e.g., Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998), Italy (e.g., Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004) and Germany (e.g., Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004), this specific gesture has been tentatively suggested to be universal, with recent cross-cultural study identifying similar age of onset of index-finger pointing in infants in Papua New Guinea, Indonesia, Japan, Peru, Mexico, Canada, Holland and Nepal (Lieven & Stoll, 2013; Liszkowski, Brown, Callaghan, Takada, & de Vos, 2012; Salomo & Liszkowski, 2013), although there is evidence that in at least one culture pointing is incredibly rare or non-existent (Wilkins, 2003). The exact age of onset of gestures, especially index-finger pointing, is discussed in more detail in Chapter 4, but it is fair to say infants begin to produce these around their first birthday.

Around this time too vocalisations emerge that seem to have a stable function, despite not being part of the lexicon in which the child is being raised, for example communicative grunts while reaching towards objects (Karousou & López-Ornat, 2013; Vihman, 1996). It is typically during the second year of life that infants begin to use words (the conventional forms that make up the lexicon of the community in which they are being raised) and more importantly start to use words referentially (Fenson et al., 1994). It is useful to note that although infants may mimic words that they have heard, on some definitions a 'word' isn't used as such unless it actually has a symbolic function. For example, infants may use a term whilst participating in act, e.g., shouting '*bam*' (Italian for 'boom') as they knock down a tower. However, symbolic use of the word comes when an infant says '*bam*' to announce their intention to knock down a tower, prior to doing so (Vihman, 1996).

**How Do We Identify Intentional Communication?**

The presence of specific behavioural markers can indicate that a child is intentionally communicating. We will first present these behavioural markers as applied to intentional action in general, before assessing what level of intention they indicate when applied to intentional communication. These criteria have their origins in the work of Jerome Bruner, who set out what might demonstrate that behaviour is intentional. These criteria have since been developed to apply specifically to intentional communication, but they were not necessarily applied in this way. Originally, Bruner (1973, 1975) set out criteria for intentional action as follows:

> 'Intention, viewed behaviourally, has several measurable features: anticipation of the outcome of an act, selection among appropriate means for achievement of an end state, sustained direction of behaviour during deployment of means, a stop order defined by an end state, and finally some form of substitution rule whereby alternative means can be deployed for correction of deviation or to fit idiosyncratic conditions' (Bruner, 1973, p.2).

These criteria require unpacking. The first two criteria (anticipating the outcome, and selection of a method of achieving that outcome) make behaviour goal-directed (de Wit & Dickinson, 2009; Heyes & Dickinson, 1990; Klossek, Russell, & Dickinson, 2008). They both mirror criteria that de Wit & Dickinson (2009) argue makes action that is goal-directed, *intentionally* goal-directed. The distinction between goal-directed action and intentionally goal-directed action is the mental representation of these two criteria – one has to believe that there is a specific outcome of the act, and has to desire that outcome (de Wit & Dickinson, 2009, p. 464; Heyes & Dickinson, 1990, pp. 92–94). Action is not intentional if it is not directed towards a goal with *both* a belief that the action will achieve the goal *and* a desire to achieve that goal. These might seem hard to demonstrate behaviourally, but anticipation of the outcome on performing the act might provide evidence that someone has a belief that the act will bring about that outcome. Similarly, choosing an appropriate course of action among a number of alternatives in order to achieve a certain outcome might provide evidence that someone has a desire for a specific outcome. In either case, one doesn't need to posit more complex mental states, behaviour is intentional if it is goal-directed (regardless of whether the goal can be demonstrated to be consciously represented).

The third and fourth criteria (sustained direction of behaviour and a stop order) concern persistence and stopping rules (Bruner, 1975, pp. 7–8). If behaviour is intended towards achieving a specific goal, then we would expect that it should be repeated if the goal is not met. In addition, if the goal *is* met, the behaviour should be expected to cease. The fifth criteria (alternative means) concerns elaboration or flexibility of behaviour. When initial attempts at achieving a goal are not met, behaviour should be modified flexibly to enhance the possibility of achieving the goal, if it is intentional.

As intentional communication is a form of intentional action, it is reasonable to assume that this criteria for intentional action applies to intentional communication. Some researchers have combined elements of these criteria for intentional action to provide evidence that infants are intentionally communicating by observing children's behaviour. For example, communicative repair, where an infant is misunderstood by an adult (and does not get what they want), they persist in their behaviour until they get what they want, and often elaborate on the behaviour as they persist (Golinkoff, 1986, 1993; Liszkowski, Carpenter, & Tomasello, 2007). This provides an impressive demonstration of three of Bruner's criteria in one stroke (Golinkoff, 1986, 1993; Shatz, 1981). In Chapter 4, we discuss the problems in applying

Brunerian criteria to real-world observations of children's behaviour in order to determine whether they are intentionally communicative.

These criteria distinguish actions that are performed with an intention from those that are involuntary. As studying intentional communication relies on inferring others' intentions from their actions, these are useful criteria. However, they don't necessarily imply that such behaviours have the more complex intentional structure of communication defined at the outset of this chapter. There is nothing within the criteria that specifically indicates that an action has been performed with a communicative intention (i.e., that someone recognises that your action is communicative).

Many researchers argue that what really indicates that a communicator has a communicative intention (that another recognises that their action is communicative) is ostension. Ostensive behaviour is behaviour performed in a way that is deliberately designed to make others draw an inference about the communicative intention that lay behind it. So our communicative intention (that others recognize that our behaviour is an attempt to get them to infer our informative intention) is indicated often, though not exclusively, through our use of ostension (R. Moore, 2016). Raising eyebrows, speaking in 'motherese', direct eye contact and engaging in contingent turn-taking behaviour have all been highlighted as acts of ostension in the infant literature (Behne, Carpenter, & Tomasello, 2005, p. 494; Csibra, 2010, p. 144). Some argue that infants have a preference for behaviours that are ostensive. On 'natural pedagogy' accounts, infants should attend more to ostensive behaviours (Csibra, 2010; Csibra & Gergely, 2009, 2011; Sperber, 2000).

Of all the ostensive behaviours, making eye contact has received most attention as a potential marker of communicative intent. Gaze-checking involves briefly looking at another's eyes whilst engaged in, prior to, or just after, a communicative act. Many researchers claim that when infants make eye contact while performing other actions (e.g., vocalising, reaching or pointing) this is a behavioural marker of their communicative intentions (E. Bates et al., 1975; Bruner et al., 1982; Camaioni et al., 2004; Franco & Butterworth, 1996; Harding & Golinkoff, 1979; Liszkowski et al., 2007; Maljaars, Noens, Jansen, Scholte, & van Berckelaer-Onnes, 2011; Masur, 1983; Tomasello, Carpenter, & Liszkowski, 2007; Zinober & Martlew, 1985).

However, it is unclear whether this constitutes evidence of intentional communication at the level of the communicative intention (that another recognises *that* I am communicating). For some, this behaviour indicates something like the presence of the communicative intention, in that eye contact demonstrates that communicative behaviour is ‘for’ someone, i.e., it is ‘intentionally addressed’ (Franco & Butterworth, 1996, p. 332). However there are two reasons to be cautious in settling on such a rich interpretation of eye contact. On the one hand, absence of eye contact doesn’t indicate the absence of a communication intention, and on the other hand, the presence of eye contact does not guarantee that someone has a communicative intention. Regarding the former, eye contact (or other ostensive behaviours) is not necessary for intentional communication, if we are sure that others are attending to us, we may not mark our action in any special way. This might be especially true in joint activities, such as book reading, whereby infants do not need to indicate that their action is communicative, as they can assume that any action will be interpreted as such because they are in a communicative context (Cochet & Vauclair, 2010; Murphy, 1978). Regarding the latter, infants might not be looking to faces to indicate that their action is communicative, as they look at faces for other reasons. For example, infants preferentially look at human faces over other objects in their environment from a very early age (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991) and display a preference for faces that are making eye contact with them (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Menon, & Johnson, 2006). Admittedly, some of these preferences are found at birth, and as Johnson et al. (1991) note the spontaneous reflex of tracking faces observed within an hour of birth declines in a matter of weeks. However, while this spontaneous ability may decline, Frank, Vul, & Johnson (2009) found that infants slowly become more and more interested in looking at faces in between 3 and 9 months.

Given this, we argue here, and in the second half of the thesis (see Chapter 4), that gaze coordination may not be evidence of intentional communication, in terms of the communicative intention, i.e., intending that others recognise that actions as communicative (Shatz, 1981). Instead, we argue that gaze checking indicates that an action might be intentionally communicative in a more primitive way, in that infants’ are aware that others need to attend to their action so that they can get what they want. It effectively demonstrates an awareness that communication functions when others are attending (D'Odorico, Cassibba, & Salerni, 1997; Desrochers, Morissette, & Ricard, 1995; Franco & Butterworth, 1996; Gómez, 1994; Harding & Golinkoff, 1979; Masur, 1983; Snyder, 1978). Eye contact therefore indicates that infants are trying to direct the attention of others in order to achieve their informative intention. In this sense it is a better indicator of an infants' informative intention (that they want something to happen) than the communicative intention.

In this respect, Bates et al. (1975, p. 217) documented a watershed moment in one of their participants at around 12 to 13 months, when she oriented towards an object, pointed whilst vocalising, then turned to point at the adult, before pointing back at the object. It was Bates et al.'s (1975) view that this was one of the first instances of intentional communication, and that the second point in the sequence (to the adult) was later replaced by eye contact whilst pointing fixedly at the object. The sequence of points thus read something like the following: '*THIS!*' (first point) , '*YOU!*', (second point) '*LOOK AT THIS!*' (third point). Thus, Bates et al. (1975) claim that eye contact takes the place of the second point, meaning that pointing whilst making eye contact means something like, '*YOU! LOOK AT THIS!*'. Given this, gaze alternation (i.e., between the object and the face of the audience) is seen as a particularly clear marker that infants are intending to redirect the attention of others to achieve their informative intention (Liszkowski et al., 2007, p. 713; Matthews et al., 2012, p. 820). All these arguments articulate the idea that eye contact may indicate that the infant understands that communication functions when it is attended to by others, providing evidence that infants have an informative intention.

The end of the first year of life (specifically at the onset of pointing) could herald the point at which infants begin to intentionally communicate in this way. However, there is not necessarily a watershed moment where infants begin to intentionally communicate. Firstly, it is possible that infants may intentionally communicate in some regards whilst not in others (Harding, 1984; Prizant & Wetherby, 1987; Scoville, 1984), for example, some vocalisations may be intentionally communicative at a specific age, whilst others are not. Therefore, finding evidence that *some* types of vocalisations or gestures are intentionally communicative, does not mean that all infants' vocalisations and gestures are intentionally communicative. Secondly, it is likely that early communicative behaviours might demonstrate some of the behavioural criteria for intentional communication, but not all, and that there is likely to be a gradual development in infant's communicative abilities. In fact, this may be the process by which intentional communication develops. Infants produce behaviours with some of the behavioural markers of communicative intent, that are then *interpreted as intentionally communicative* by caregivers, who respond to them as if they are intentionally communicative. This caregiver responsiveness then scaffolds the emergence of intentional communication (Bruner, 1975, 1976; Bruner et al., 1982; Casby & Cumpata, 1986; Crais, Douglas, & Campbell, 2004; Gros-Louis, West, Goldstein, & King, 2006; Harding, 1982, 1984; C. Moore & Corkum, 1994; O'Connell & Farran, 1982; Prizant & Wetherby, 1987; Scoville, 1984; Shatz, 1981; Snyder, 1978; Sugarman, 1984; Trevarthen & Hubley, 1978; Zinober & Martlew, 1985).

In sum, detecting genuine instances of intentional communication in infancy is complicated. One has to determine the point at which children have moved from a stage in development whereby they don't realise that their behaviours are effective because others interpret them in some way, to when they do (E. Bates et al., 1975; Harding, 1982; Snyder, 1978; Sugarman, 1984). There are a number of behavioural criteria that might indicate that an action is intentionally communicative at the level of informative intention. As all the individual behavioural criteria for intentional communication (indicating a desire for a specific outcome, selecting an appropriate means, persisting, elaborating, making eye contact) can each be subject to leaner interpretation, some researchers favour approaches whereby the presence of more behavioural markers make the richer interpretation more likely. Studies therefore often combine different behavioural markers in order to determine whether infants are communicating intentionally. Behaviour is intentionally communicative for example, if infants make eye contact with their caregiver, specifically refer to something, persist in their signal and terminate it when a (seemingly satisfactory) response is made by their caregiver (Bruner et al., 1982; Harding & Golinkoff, 1979, pp. 35–36). The implication is that behaviours that demonstrate more behavioural markers that might indicate intentional communication, are more likely to be intentional (Schel, Townsend, Machanda, Zuberbühler, & Slocombe, 2013).

## The Development of Intentional Communication

Here we review the literature to provide an account of when infants begin to intentionally communicate, in terms of directing the attention of others (i.e., they have an informative intention, but not necessarily coupled with a communicative intention in the fuller sense), and what the evidence is for this. We focus initially on vocalisations, before moving on to gestures.

**Vocalisations.** Prior to 6 months of age, infants' vocalisations are unlikely to be intentionally communicative in the sense that infants are trying to direct the attention of others to achieve their informative intention. However, that is not to say that they do not use them in a way that is uninteresting with regards to how later intentional communication may be learned. As early as 12 weeks, infants maintain eye contact whilst vocalising (Kaye & Fogel, 1980), and are  more likely to make speech-like syllabic sounds than vocalic sounds when maintaining eye contact with their caregiver from 4 weeks to 6 months (Hsu, Fogel, & Messinger, 2001). Furthermore, they vocalise whilst gazing at their caregiver while expressing positive or negative facial emotions (Colonnesi, Zijlstra, van der Zande, & Bögels, 2012). These experiments do not tell us anything about infants' communicative intentions as indicated by eye gaze, as infants did not have to solicit their caregivers' attention, because the experimental set up involved infants being face-to-face with their caregivers.  However, it does demonstrate that from early on, infants are forming contingencies between vocalisations and looking at the face of others, which plays a key role in later intentional communication.

As young as 3 months, infants seem to use their vocalisations in social contexts. In the very simplest case, infants vocalise when an adult is present more than when the adult is not. Interestingly infants at this age vocalise more when they can see an adult's eyes (when the adult is making eye contact with them), than when the adult's eyes are not visible (Bloom, 1975; Kaye & Fogel, 1980). In addition, infants use more syllabic speech-like sounds than vocalic non-speech like sounds when an adult engages in conversation-like turn-taking behaviour, responding contingently to the infants vocalisations, than when the adult responds non-contingently (Bloom, Russell, & Wassenberg, 1987). Further study has found that infants from 1 to 6 months are more likely to produce syllabic (speech-like) sounds than vocalic sounds when they are smiling and when their mother is smiling (Hsu et al., 2001). There is evidence that suggests that infants have learnt that their vocalisations are effective at eliciting reactions from others by 5 months. This is demonstrated by an increase in vocalisations towards a still face after it stops interacting contingently, which is taken as evidence of an extinction burst demonstrating infants have learned an association between their vocalisations and others' reactions (Goldstein, Schwade, & Bornstein, 2009). This suggests that from quite early on in development, infants are using their vocalisations in a way that suggests that they understand that these vocalisations elicit responses from others. This also represents an important developmental step towards using vocalisations intentionally communicatively.

In the months before an infant's first birthday, some argue that infants begin to use their vocalisations for different functions. This might be considered as evidence for intentional communication, in that they are expressing a desire for a specific outcome (see Bruner's criteria above). For example, caregivers can discriminate two vocalisations produced by 7- to 12-month-olds on the basis of function, either emotive ('she is pleased') or communicative ('he is trying to get my attention'). These two types of vocalisations seem to have distinct acoustic features (Papaeliou, Minadakis, & Cavouras, 2002), which might seem to suggest that infants have different vocalisations to deliberately get caregiver attention. Further, 9- to 10-month-olds seem to have acoustically distinct vocalisations for functions that their caregivers label either communicative or investigative, for example where an infant is exploring their environment (Papaeliou & Trevarthen, 2006). In a further analysis of this distinction, Esteve-Gibert and Prieto (2013) found that 7- to 11-months-olds gradually increased the amount of communicative vocalisations compared to investigative. Additionally, they investigated pragmatic functions within only communicative vocalisations, such as infants expressing discontent or satisfaction, requesting something or responding to a stimulus. At 7 months, infants expressed only discontent or satisfaction, and it was only at 11 months that these extra categories began to be expressed. This suggests that there is an upwards developmental trajectory for vocalisations that caregivers interpret as communicative towards the end of the first year of life. Arguably, this is not convincing evidence that infants produce intentionally communicative vocalisations, however it does suggest instead that caregivers are interpreting some behaviours as communicative, responding appropriately, which in turn promotes infants' production of these, and might scaffold the transition to intentional communication.

Only a handful of studies have investigated whether vocalisations around the end of the first year of life are accompanied by eye contact. The findings and limitations of these studies are discussed in detail in Chapter 4, but suffice to say that there is some evidence that infants around the first year of life produce vocalisations whilst making eye contact with their caregivers (Harding & Golinkoff, 1979; Miller & Lossia, 2013). D'Odorico et al. (1997) argue that initially vocalisations during the first year are produced as expressions of internal need (hunger, affect etc.), and children look at faces because they are dynamically interesting. Because caregivers treat both of these acts as intentional efforts to communicate (see also Gros-Louis et al., 2006), infants gradually learn to communicate intentionally – they intentionally solicit attention through gazing at their mothers face, and then vocalise. They found that both 12- and 20-month-olds vocalise while making eye contact with an adult, but this is more frequent at 20 months (D'Odorico et al., 1997).

**Gestures.** The majority of investigations into infants' gestures, and whether they are intentionally communicative, has focused on index-finger pointing, which emerges early in the second year of life. There is a large experimental literature suggesting that infants' pointing may be intentionally communicative from its onset. At 12 months, infants point more (thus demonstrating persistence) when adults attend to a referent that they were not pointing at (Liszkowski et al., 2004, 2007), and seem to produce points with the goal of informing adults of the location of missing items (Liszkowski, Carpenter, Striano, & Tomasello, 2006). However, these findings should be treated with caution, as only infants who were already reported to be pointing by their caregivers were included in the study. As such, it is possible that these infants were precocious in their pointing ability and further down a developmental trajectory than the general population, so it is not clear that pointing is necessarily intentionally communicative in this way at its onset.

This is important because, as with vocalisations, some argue that early production of index-finger pointing is not intentionally communicative, and that such early pointing provides learning opportunities for infants to transition to intentionally communicative pointing later on

(Aureli, Perucchini, & Genco, 2009; E. Bates et al., 1975; Carpendale & Carpendale, 2010; Carpendale & Lewis, 2004; Harding & Golinkoff, 1979; Morissette, Ricard, & Décarie, 1995; Povinelli & O'Neill, 2000; Sodian & Thoermer, 2004; Triesch, Teuscher, Deák, & Carlson, 2006). Infants are capable of forming the hand-shape required for a pointing gesture as young as 15 weeks old (Fogel & Hannan, 1985), and even within the womb (Marschik, Prechtl, Prayer, Peyton, & Einspieler, 2013), but these do not appear to be attempts to intentionally communicate. Moreover, when they begin to extend their arm with the index-finger shape around the end of the first year of life, infants sometimes point when there is no one around (E. Bates et al., 1975, p. 217; Carpendale & Carpendale, 2010; Delgado, Gómez, & Sarriá, 2011), suggesting that it may not be intentionally communicative from onset.

Infants may initially engage in pointing in order to elicit certain behavioural responses in others, but not begin to use it in a sophisticated intentionally communicative way (with the intention of directing the attention of others) until towards the end of the second year of life (Povinelli & O'Neill, 2000). Specifically, they engage in pointing because it attracts positive affect towards themselves (C. Moore & Corkum, 1994). When caregivers see their infant point, they are likely to look at the infant and try and determine what the infant is referring to, which is likely to involve positive affect. This provides a reliable association between pointing production and positive affect for infants. Testing this claim, Moore & D'Entremont (2001) found that between 12 and 16 months, infants point more to an object of interest when adults are already looking at them, than when adults are looking at the object or in another direction. This was compared with the pointing behaviour of 2-year-olds, who pointed at the object of interest more sensitively to whether the adult has already seen the object. However, it could be argued that the goal of infants' pointing might be to provoke a comment from caregivers, in which case, this would suggest that infants are intentionally communicating, by demonstrating that they understand an adult has to be attending to their pointing in order for it to function, hence pointing more (Liszkowski et al., 2004).

While evidence concerning early infant pointing does not unambiguously demonstrate that it is intentionally communicative, later evidence is less equivocal. Gaze checking behaviour, though rare early on (see Chapter 4), starts to become more abundant during the second year, with infants beginning to clearly check that adults are attending to them prior to producing points around 16 to 18 months (E. Bates et al., 1975, p. 217; Franco & Butterworth, 1996; Masur, 1983).

From about 18 months, there is good evidence that toddlers have begun to communicate intentionally using pointing, in the sense that they attempt to get others to infer their informative intention. Grosse, Behne, Carpenter, & Tomasello (2010) demonstrated that 18-, 24-, and 30-month-olds all repaired a misunderstanding in communication. Though this has been previously demonstrated in younger infants (Golinkoff, 1986), the important distinction here was that children repaired communication even after they had achieved their material goal. In this scenario, by 'happy accident' a researcher passed the child an object that they had requested, thus the child had achieved their material goal. However, the researcher was still attempting to reach an object that the infant had not requested, and thus had 'misunderstood' the child's request (Grosse et al., 2010, pp. 1714–1715). Around 80% of 30-month-olds, over half of the 24-month-olds and 45% of 18-month-olds attempted to engage in some sort of repair behaviour. The interpretation is that these toddlers were concerned with ensuring that their communication had been understood, not just that they had got what they wanted. This at least goes some way to demonstrating that toddlers produce communication intentionally, in that they understand that communication works by others *correctly* inferring their informative intention.

In summary, while initially, infant pointing might not be intentionally communicative (around 12 months of age), it seems that by 18 months there is good evidence to suggest that infants point in a way that marks an attempt to get others to recognize their intentions (Desrochers et al., 1995).

## Summary: When Do Children Begin to Produce Intentionally Communicative Behaviours?

We have outlined the criteria which might indicate when infants begin to produce intentional communication, and demonstrated that such criteria demonstrate the presence of informative intentions but not necessarily communicative intentions. There have been attempts to apply these criteria to infants' production of behaviour during the first years of life. On balance, vocalisations are unlikely to be intentionally communicative in the sense that they are intended to direct the attention of others prior to the end of the first year of life. This comes from evidence that infants begin to more clearly coordinate gaze to their caregiver's face with their vocalisations around this time (suggesting that they understand that others have to attend to their communicative actions in order for them to be effective), and from evidence that infants are trying to express specific informative intentions with their prelinguistic vocalisations that caregivers' recognise. With regards to gestures, by the middle of the second year, infants' index-finger pointing seems to be intentionally communicative at the level of the informative intention. This comes from evidence that infants more clearly coordinate gaze to their caregiver's face with their pointing around this time, and that they seem to demonstrate that they want specific communicative outcomes from their pointing. Such evidence is lacking or ambiguous when index-finger pointing first begins. Other gestures have not been studied with respects to intentionality in such detail (see Chapter 4).

While it is clear that midway through the second half of the second year of life, children have started to produce behaviours that appear to be intentionally communicative at the level of informative intentions, the evidence at the end of the first year of life is unclear that this is the case. We return to this point at the end of this chapter.

## When Do Children Interpret the Behaviours of Others as Intentionally Communicative?

The third part of this chapter focuses on whether infants' understanding of the communicative acts of others demonstrates that they understand something about the communicative intentions of others. The majority of evidence concerns how young children understand the informative intentions of others, as demonstrating that they can detect the communicative intention of others, or understand the role of the communicative intention in communication is much more complex. The first part of this review focuses on how children understand the informative intentions of others, and is split into three sections, concerning adult speech, gestures and gaze behaviour. After this, we move on to the handful of studies that demonstrate whether children are sensitive to others' communicative intentions, and how the two levels of intention in communication are structured. These studies are of crucial importance to Chapters 2 and 3 of this thesis, where we investigate whether infants' production of a non-responsive behaviour, or comprehension of a communication breakdown between others demonstrates that they understand something about the structure of intentional communication.

**Understanding Adult Speech**

Around 6 to 9 months, infants seem to understand that specific words used by adults refer to specific objects. In a "looking-while-listening" paradigm, 6- to 9-month-olds demonstrated that they had mapped some nouns to referents, by reliably fixating on a picture of an object when they heard its label, and not on a picture of a different object presented simultaneously, e.g., fixating on a picture of an apple not a mouth, when hearing the word "apple" (Bergelson & Swingley, 2012). However, it is not clear whether such mapping of words to referents indicates that infants understand that such words are used intentionally communicatively (whether they are socially learned), or whether they are associatively learned sound-object mappings. If these are sound-object mappings, then infants do not understand such utterances by adults as intentionally communicative. It is known that vocabulary is associatively learned by 2-year-olds, but that they only recognise those that are learned socially as words for objects (Bannard & Tomasello, 2012). So it is unclear whether such evidence of early mapping of words to objects indicates that infants understand that others are intentionally communicating.

More recently, evidence has been put forward that at 6 months, infants understand that speech communicates information (Vouloumanos, Martin, & Onishi, 2014). In this study, in a familiarisation phase a researcher (the communicator) expressed preference for one of two objects (by repeatedly grasping only one of them). In the test phase, the communicator (now constrained behind a wall) either uttered a novel word, or coughed, and a second researcher (the recipient) either picked up the object that the communicator had expressed a preference for, or the object that they had not. In cough trials, infants did not look differently at either outcome. However, they looked longer when the communicator produced a word, and the recipient picked up an object that the communicator hadn't expressed a preference for, compared to when the recipient picked up the object that the communicator had expressed a preference for. This study demonstrates that 6-month-olds might expect that speech can be a label for objects, but again it does not necessarily indicate that infants understand anything about communicative intentions.

There is evidence to suggest that infants understand that adults' speech conveys some sort of information to other adults. In one looking time study, there was a familiarisation phase, where 12-month-olds observed a researcher (the communicator) failing to stack a ring onto a funnel, because the funnel was out of reach (Vouloumanos, Onishi, & Pogue, 2012). Then, in a test phase, a second researcher (the recipient) was also present, with both ring and funnel available, whilst the communicator was now completely out of reach of both. The communicator uttered either a novel word, or coughed. Following this, the recipient either stacked the ring on the funnel, failed to stack the ring, or performed an unrelated action (that would not conceivably be the goal of the communicator during the familiarisation). In cough trials, infants looked equally long at the scenario for all outcomes. However, when the communicator had used a word, infants looked significantly longer (indicating a violation of expectation) when the recipient failed to stack the ring or performed an unrelated action, than when they stacked the ring on the funnel. The authors took this as evidence that 12-month-olds understood that speech (but not non-speech) can communicate intentions of the communicator.

Caregivers report that infants begin to understand words and phrases from around 8- to 10-months of age (Fenson et al., 1994). It is not clear if such understanding is prefaced on their understanding of the communicative intentions of others, or whether appropriate responses to commonly used phrases are learned by infants. For example, another proposed mechanism for some early word learning is through sound symbolism. Sound symbolism is the non-arbitrary link between the form of a word and its meaning. Some languages contain classes of sound symbolic words, e.g., ideophones in Japanese. In others, certain sounds are shared across words with similar meanings, e.g., phonoesthemes in English (see Imai & Kita, 2014 for a review).14-month-olds use sound symbolism to map referents to words, learning referents when they were a sound symbolic map for a label (e.g., 'kipi' for a spiky object, 'moma' for a rounded object) more easily then when these were mismatched (Imai et al., 2015). Thus, evidence that infants understand words may not necessarily demonstrate that they understand the communicative intentions of others, as in some cases, words could be learnt through sound symbolism (which does not necessarily require understanding that words are intentionally communicative).

Certainly, by around two years of age, toddlers seem to demonstrate an understanding of the communicative intentions behind adult speech. One clear demonstration of this comes from Tomasello & Akhtar (1995). In this study, a researcher put an object in an apparatus and performed an action using the apparatus while saying, "Widget [child's name], widget". They either gaze alternated between the child and the object *or* the apparatus itself, thus the novel word ("widget") could refer to either an object or an action. Children (aged between 24 and 30months) were asked to "show me widget", and depending on whether the action had been highlighted, or the object responded by either performing the action or holding up the object. This demonstrates that children can infer adults' informative intentions behind their utterance when such utterance is underspecified. However, this is fairly indirect evidence for quite an important principle, and more direct evidence is needed to determine when children begin to demonstrate understanding of adults informative intentions (Ambridge & Lieven, 2011, p. 76).

**Understanding Adult Gestures**

As with adult speech, there is purportedly some evidence that infants prior to their first birthday understand that adult gestures convey information. In a similar setup to that described above (Vouloumanos et al., 2014), one researcher (the communicator) expressed a preference for one object over another (by grasping it repeatedly), and then pointed to it when a recipient was present (Krehm, Onishi, & Vouloumanos, 2014). Nine- to 11-month-olds looked significantly longer at scenarios where the recipient picked up the object that the communicator had not expressed a preference for. However, it should be noted that it is not clear how sophisticated this understanding might be, as it only demonstrates that infants are surprised when someone does not pick up an object that has been pointed to (as the communicator only pointed to the object that they had expressed a preference for), and this is therefore perceptually more salient.

Some researchers claim that infants' responses to adults' pointing indicate that they understand something about adults' communicative intentions in terms of the informative intention. Much of the evidence for this claim comes from similar experimental setups in which an object that an infant wants is hidden in one of a number of locations, and then a location is pointed to by an experimenter. Early studies suggested that 3-year-olds search in the location that is pointed to well above chance levels (Tomasello, Call, & Gluckman, 1997). Since then, using similar experimental set ups, researchers have demonstrated that infants search in the location that adults point to as early as 12 to 14 months (Behne et al., 2005; Behne, Liszkowski, Carpenter, & Tomasello, 2012; Gräfenhain, Behne, Carpenter, & Tomasello, 2009; Kristin Liebal, Behne, Carpenter, & Tomasello, 2009). Because infants search for objects in the location that adults have pointed at, this is often taken as evidence that infants understand that the adult's pointing gesture was intended to inform them of the location of the hidden object (Behne et al., 2012). Thus, it demonstrates evidence that infants understand others' actions in terms of intentional communication as young as 12 months. However, there are reasons to be sceptical of such an interpretation. It is possible in these studies that infants are likely to follow the direction of points because they are cued in the direction that the pointing hand indicates, without any understanding that this person had an informative intention. 14-month-olds search in locations that have been pointed to, even when pointing is provided for another person when infants are third party to the interaction (Gräfenhain et al., 2009). This might be taken evidence of a failure of infants to understand the communicative intent behind the pointing, because the point was not intended *for them*, and yet they went and looked in the location that was pointed to anyway, perhaps because solely because they were cued to search in that location by the pointing hand.

More convincing evidence that infants understand communicative intentions behind pointing comes from when they are able to follow pointing to a location outside the field of vision that the pointing hand is in. A longitudinal study has demonstrated that infants could only follow index-finger pointing prior to 15 months when both the object and the pointing hand was in the same field of vision, meaning that it is possible that infants are cued to an object in their field of vision (Morissette et al., 1995). After 15 months, infants could follow pointing to outside of their field of vision, which arguably is more unambiguous evidence that infants understand that others are trying to inform, and they are not just being cued to a location within the visual field (Butterworth & Grover, 1988; Morissette et al., 1995).

Furthermore, following pointing at 12 months could be a precocious ability, and this could be an artefact of the pointing studies. Elsewhere, a longitudinal study investigating infants' ability to follow adult pointing found that at 12 and 14 months, infants did not search in the location that had been pointed at above the level predicted by chance levels, and it was not until 16 to 18 months that the majority of infants searcher in the location that was pointed to (Pfandler, Lakatos, & Miklósi, 2013). Interestingly, this study was specifically designed to remove standard markers of ostension, so points were not repeated more than once and there was no gaze alternation or verbal communication (Pfandler et al., 2013, p. 712). In contrast, in the studies mentioned above (Behne et al., 2005, 2012) ostension *was* utilized during pointing.

By 18 months, there is less ambiguous evidence that infants infer adult intentions from their index-finger pointing. Liebal et al. (2009) provided evidence that 18-month-olds infer an adult's intentions from their index-finger pointing. In their study, an infant played a jigsaw game with an experimenter, followed by a clean-up game with a second experimenter. One of the experimenters then pointed to a jigsaw puzzle piece in the middle of the floor. Depending on which of the two experimenters had pointed, 18-month-olds reacted differently. When the pointing experimenter was the one that they had previously been playing the jigsaw game with, they took the piece to the jigsaw. However if the pointing experimenter was the one who they had played the clean-up game with, they put the piece in a basket with the other toys. 14-month-olds did not perform so well on this task, but on an easier version (with fewer memory demands) their responses to the same pointing gesture varied as to whether they had played a clean-up game with the experimenter who pointed, or whether it was an unfamiliar adult. This demonstrates that 14-and 18-month-olds understand that pointing gestures can indicate an informative intention that can be inferred from context (in this case, what shared game the participant had engaged in). However, it should be noted that this more complex ability to infer an informative intention from a point is certainly not robust at 14 months, the data from Liebal et al. (2009) demonstrated that children still did not actually clean up the target object when pointed to by the experimenter (with whom they had been playing a cleaning-up game) over

50% of the time. So, we should be cautious in saying that 14 months is a robust onset age for this sort of understanding.

### Understanding Adult's Communicative Gaze

Some evidence for understanding of early communicative intentions comes from the literature on gaze following. Gaze following is the ability to follow the line of someone else's gaze, effectively looking where another person is looking. This can be done communicatively, it is possible to look in an intentionally communicative way, normally by engaging in ostensive behaviours whilst gazing in the direction of a referent. Some rich interpretations of infants' gaze following behaviour hold that when infants follow the gaze of others, they understand that others have communicative intentions (Behne et al., 2005; Tomasello, 2008).

One of the studies mentioned above that looked at pointing, also investigated infants' ability to follow the direction of a researchers' gaze to a location where an object was hidden (Behne et al., 2005). Specifically, it focused on cases when gaze was given ostensively, and absent-mindedly. For ostensive gaze, the researcher raised their eyebrows, and alternated gaze between the location and the infant a number of times. For absent-minded gaze, the researcher maintained a neutral face, rested their chin on their hands in a bored manner, and did not engage in eye contact with the infant. Infants as young as 14 months chose the box that had been gazed at ostensively above levels predicted by chance, but only at chance when gaze was absent-minded. This suggests that infants followed gaze more reliably when it was given in an intentionally communicative way, suggesting that they may understand that this was intentionally communicative.

Further evidence to support this view is that aspects of what Behne et al., (2005) referred to as ostension seem to be critical to elicit gaze following in infants. For example, unless a researcher engaged in a period of eye contact with 9-month-olds prior to gazing at a location, these infants did not follow the resulting gaze (Senju, Csibra, & Johnson, 2008). Additionally, looking multiple times at an object means that 9-month-olds are more likely to look at it, than if it is only looked at once (S. C. Johnson, Ok, & Luo, 2007). Finally, one study found that 16- to 24-month-olds failed to follow gaze of adults, and did not involve the use of these ostensive behaviours (Povinelli, Reaux, Bierschwale, Allain, & Simon, 1997). This suggests that behaviours that indicate that an adults' gaze is intentionally communicative seem to be critical for infants to successfully follow adults' gaze.

However, one concern with this interpretation is that it might not be that an adult's ostensive behaviours indicate to an infant that the gaze is intentionally communicative, more that such ostensive behaviours make infants attend more closely to the adult, as these are salient behaviours. Advocates for a leaner interpretation of infants' gaze following behaviour claim that infants learn associations between adults' gaze shifts and the introduction of novel and interesting objects during play. Thus, when their caregiver performs a salient behaviour (making eye contact, or raising eyebrows) and then looks in some direction, the infant also looks in that direction in expectation of something of interest (C. Moore & Corkum, 1994). This does not rely on infants interpreting their caregivers' actions as intentionally communicative. Evidence to support this comes from training studies, whereby 8- to 9-month-olds (who did not initially demonstrate that they followed gaze) were taught to follow the gaze of adults, by associating those adult head turns with interesting sights (Corkum & Moore, 1998; C. Moore, Angelopoulos, & Bennett, 1997). Further evidence comes from computational modelling, in which gaze following abilities can arise from a basic set of structures and mechanisms (e.g., infant preferences, reward-driven learning and an environment where caregivers look at objects that an infant finds interesting), none of which rely on understanding the intentions of others (Triesch et al., 2006). Thus, one alternative view is that gaze following in general at a younger age is attributable to expectations of objects of interest and infants' following of gaze that is given in an intentionally communicative way is more successful because it is given in a more salient way.

Following the gaze of others may demonstrate an understanding of intentional communication in other ways, if it can be demonstrated that the infant is engaging in perspective taking (Butterworth & Grover, 1988; Butterworth & Jarrett, 1991). Evidence from longitudinal studies of gaze following are especially illuminating in this regard. Early gaze following (around 6 months) requires objects that are being gazed at to be in the same field of vision as the adult who is gazing (Butterworth & Jarrett, 1991). It is only around 18 months that infants are able to gaze at objects that are quite some distance from their caregivers (Morissette et al., 1995), and crucially, infants can turn to look at object that is being gazed at that is behind them (Butterworth & Jarrett, 1991). As discussed previously, in regard to pointing, this is better evidence that infants understand that others are trying to inform them of something (which they can search for, considering another's perspective), and not just trying to cue them to a location within their visual field. Thus, findings that infants can follow gaze from as early as 3 months to proximal targets (D'Entremont, Hains, & Muir, 1997) are less important when considering how infants might understand intentional communication.

**Summary: When Do Children Interpret the Behaviours of Others as Intentionally Communicative?**

Infants' seem to understand that speech can convey information about adults' intentions from around 12 months, and this is established around two years of age. Around the same time, adult pointing gestures and ostensive gaze can cue infants' attention to locations; however whether this requires that they understand that the adult intended to communicate is not clear. By the middle of the second year of life, infants can follow adults' pointing and ostensive gaze to locations out of their visual field that the adult is also in, suggesting that they may be engaging in some perspective taking. By 18 months, infants respond to the pointing gestures of adults in a way that suggests that they understand informative intentions, responding flexibly depending on their previous interactions with that adult. It is likely that during the end of the second year of life, infants understand that others have informative intentions, i.e., that they when they perform certain actions, they intend that you do something, and there is tentative evidence that this could be understood earlier. What has not been demonstrated in these studies however, is what infants understand about the role of the communicative intention (i.e., the intention that an action is communicative), which we deal with in detail below.

## The Communicative Intention and the Intentional Structure of Communication

The fourth section of this review focuses on two related questions about the communicative intention, i.e., the intention that people recognise that your action is communicative. First, we ask what evidence there is that infants understand the intentional structure of communication: specifically whether they understand that communication functions because others detect that you are communicating, and infer what you are communicating about. Second, we ask if there is any evidence that infants understand that the communication of others functions because they detect the communicative intentions, and infer what they are communicating about.

**Do Infants Understand the Intentional Structure of Communication?**

One way of demonstrating understanding that someone understands the intentional structure of communication is through cases of hidden authorship. Hidden authorship is where you attempt to make someone infer your informative intention without detecting your communicative intention (Gómez, 1994; Grosse, Scott-Phillips, & Tomasello, 2013; Sperber & Wilson, 1986; Tomasello, 2008). An often cited example of this is the case of a dinner party guest who wants the host to fill up their empty glass, and so deliberately (but seemingly idly) pushes it forward into the eye-line of the host (Sperber & Wilson, 1986; Tomasello, 2008). In this scenario, the guest has the informative intention, "my empty wine glass needs filling", which they hope that the host infers. However, they make every attempt to avoid the host recognizing that this is their intention, by avoiding eye contact with them, and making the idle push of the glass as natural as possible so that it doesn't look ostensive. By this manipulation of the intentional structure, we can demonstrate that as adults we do have an implicit understanding of the distinction between the two levels of intention.

There is some evidence to suggest that 3- to 5-year-olds engage in hidden authorship (Grosse et al., 2013). In this study, a researcher repeatedly stated that they wanted to finish a puzzle without any help, while the toddler was informed of the location of a piece of the puzzle. Toddlers informed the researcher of the location of the piece without letting the adult know that this is what they were doing (i.e., they successfully made the adult find the object, but not infer that they intended the adult to infer the location of the object through recognition of their intention to communicate). In terms of the intentional structure of communication, they hid their communicative intention while simultaneously performing an action that was seemingly designed for the recipient to obtain information. While both age groups suppressed their ostensive behaviour when trying to inform the researcher of the location of the piece, 5-year-olds were more likely to produce no ostension, whilst 3-year-olds were more likely to reduce ostension, suggesting that this was a more convincing behaviour at 5 years.

There is no evidence that we know of to suggest that understanding of the intentional structure occurs prior to 3 years. Indeed there is scepticism that such understanding can emerge before this time (Tomasello, 2008, pp. 132–133; Tomasello et al., 2007, p. 715). One key problem with hidden authorship studies is that they cannot be used earlier in development, as they require a high level of language comprehension from participants (who have to understand that the researcher wants to finish the puzzle by themselves, which is difficult to express non-linguistically).

**Do Infants Understand The Specific Role of the Communicative Intention?**

One way to investigate what young children understand about the communicative intention is to provide them with a scenario in which someone attempts to communicate with them in a novel way that they are unfamiliar with. Recently, two studies have attempted this (R. Moore, Liebal, & Tomasello, 2013; R. Moore, Mueller, Kaminski, & Tomasello, 2015). In the first of two studies, Moore et al. (2013) set up a variation on the hidden object paradigms described above (e.g., Behne et al., 2005). In their variation, the researcher referred to the location of an object without using a referential word, gesture or gaze, but instead by using a button on the desk in front of them which turned a light and sound mechanism at one of two locations. They compared toddlers success at finding the object between when the experimenter intentionally pressed the button, saying, "and now…", compared to accidental conditions when they accidentally pressed button and expressed surprise saying "whoops!" (R. Moore et al., 2013, p. 68). They also tested ostensive and non-ostensive communication. In the ostensive communication condition, the researcher engaged in eye contact with the toddler before saying, "and now…", while in the non-ostensive condition they did not make eye contact. 3-year-olds were unable to locate the object above chance in the accidental conditions, but in the ostensive intentional condition, children located the object significantly above chance. In the non-ostensive intentional condition, children located the object marginally significantly above chance. The impressive element of this study was that the pattern of results suggested children understood something about the communicative intention. When an adult points to a location or refers to it using a name, a child's search in that location doesn't necessarily imply that they understand such acts are communicative, it could mean that they have been cued to the location (as discussed above). However, in a novel set up such as this one, children have to detect the adult's communicative intention in order to succeed, i.e., they have to detect *that* the adult is communicating. If the children were purely being cued to a location where the lights flashed, they would pick this location even when the experimenter accidentally pressed the button, however instead in these situations they picked at chance.

In a second study, with 2-year-olds, the paradigm was modified slightly, in that instead of a button lighting up a buzzer, a researcher pulled and shook one of two ropes that raised and shook one of two buckets (R. Moore et al., 2015). This change in the paradigm was because the study was also being run for dogs. The performance of the 2-year-olds was similar to that of the 3-year-olds in the previous study, in that they picked above chance in conditions where the action was intentional (both ostensively and non-ostensively). This result should be treated with caution however, as the act of pulling the rope which pulled the bucket (all of which was in view of the infant) could be interpreted as a gesture (much like shaking an object with an extended arm), and is therefore not necessarily non-referential on the part of the adult.

In both studies, it should be noted that ostension (gazing to the infant prior to pressing the button or shaking the rope) did not affect whether infants interpreted this action as communicative. However, as acknowledged by the authors, it is possible that such an effect was not found because the researcher was communicating ostensively prior to the communicative trials, and so infants may have interpreted their actions as intentionally communicative (so long as they weren't accidental) on the assumption that the communicative situation was still ongoing.

A second way to investigate children's understanding about the communicative intention, is to determine how they understand indirectly communicative acts, in the sense that the informative intention is not so tightly bound to the referential act (Schulze & Tomasello, 2015). In their study, Schulze & Tomasello (2015) played a game in which a required object for the game was locked in a box behind the experimenter (which the infant had been familiarized with). At a certain point, the game would have to halt until the object was retrieved. At this point, the researcher either held up the key to the box whilst making eye contact with the child before placing it in front of them (ostensive condition), dropped the key on the floor, then held it up whilst not making eye contact with the child before placing it in front of them (intentional condition) or accidentally pushed the key towards the child whilst pointing to the toy (accidental condition). Both 18- and 26-month-olds retrieved the object (by taking the key and opening the box) significantly more often in the ostensive condition. Crucially, though the children's attention was brought to the key in both the ostensive and intentional condition, it was only in the ostensive condition that infants retrieved the object from the box. This was in contrast to the mixed findings above (R. Moore et al., 2013, 2015). The authors claim that this is evidence that infants detected the communicative intention (through ostension) which led to a non-obvious inference of the informative intention that the researcher intended for them to open the box.

### Summary: The Communicative Intention and the Intentional Structure of Communication

There is evidence that infants are sensitive to the communicative intention, in interpreting actions as communicative that are either completely novel, or only indirectly related to the referent of the informative intention by 18 months. This demonstrates that infants may be able to detect the communicative intention of others, and subsequently infer what their informative intention is. Studies of hidden authorship go one step further in demonstrating that through their ability to manipulate the intentional structure of communication; children understand how it is structured from around 3 years of age.

## Questions for this Thesis

This literature review has highlighted two outstanding questions concerning the development of intentional communication that we address in this thesis.

Firstly, it is not clear what point children begin to understand the intentional structure of communication with respect to the two levels of intentions (communicative and informative intentions). 3-year-olds seem to understand the intentional structure of communication well enough to manipulate it, as in cases of hidden authorship (Grosse et al., 2013). Hidden authorship studies cannot be used earlier in development because they require more complex language skills than toddlers would typically possess. However, 18-month-olds seem to understand something about communicative intentions, recognising when others have a communicative intention, and inferring their informative intention (R. Moore et al., 2013, 2015; Schulze & Tomasello, 2015). Thus, a yet unanswered question is whether toddlers around this age fully understand the intentional structure of communication in terms of the relation between communicative and informative intentions to the level demonstrated by hidden authorship. This is the subject of Chapters 2 and 3.

Secondly, it is not clear when infants begin to intentionally communicate, in terms of communicating with the intention of directing others' attention. This is a crucial step towards adult-like intentional communication where children have both communicative and informative intentions. What is required is an account that considers all infants vocalisations and gestures around their first birthday and investigates whether these are intentionally communicative and how they relate to later language use. This is the subject of Chapters 4-7.

# 2. Do Toddlers Pretend Not to Hear?


## Abstract

In this chapter we sought to test whether children younger than 3 years could demonstrate an understanding of the relation between levels of intention in communication through their engagement in a deceptive manipulation of ostensive listening behaviour, namely, pretending not to hear requests from their caregivers. We conducted two experiments where caregivers asked their children to do things that they did not want to do whilst their children were playing, and recorded the child's responses as well as obtaining caregiver report measures. Caregivers reported that 18- to 30-month-olds pretended not to hear in the home, however we were unable to verify this under laboratory conditions. Our results did not unambiguously suggest that children understood the intentional structure of communication, and avenues of further research are proposed.


## Introduction

Human communication is thought to have a dual intentional structure consisting of informative and communicative intentions. When we communicate, we both intend to communicate what we want, but equally we intend that the other person recognizes that we are communicating (Gómez, 1994; R. Moore, 2014; Sperber & Wilson, 1986; Tomasello, 2008). As active recipients of communication, we detect that someone is communicating, and then infer what they are communicating about. It is not simply the case that when my friend points at an object, I give it to her just because I infer that she intended me to do so. I also recognize that she intended for me to recognize that her action was communicative in the first place - it is this *extra* level of intention that makes human communication work so effectively.

Humans mark actions as intentionally communicative by engaging in a suite of pragmatic behaviours (often called ostension, or ostensive behaviours) to indicate to another individual that we are intentionally communicating (Sperber, 2000). In general, these behaviours are designed to direct another's attention towards our actions, and consist of things such as eye contact, or raising our eyebrows. Broadly, our downstream inference of what a person is trying to communicate to us is triggered by our recognition of that person's intention that we attend to their actions, expressed by their attempts to draw our attention to their actions. It is thought that this intentional structure makes humans' ability for language possible (Tomasello, 2008).

While studies demonstrate that, in the second year of life, children can produce acts that seem to have dual intentional structure; gaze alternating while vocalising (D'Odorico et al., 1997; Harding & Golinkoff, 1979) and gesturing (E. Bates et al., 1975; Camaioni et al., 2004; Franco & Butterworth, 1996; Masur, 1983; Zinober & Martlew, 1985), and can detect a communicative intention in a novel communicative act (R. Moore et al., 2013), it is difficult to establish if they understand the *structure* of these intentions in communication. That is to say that there is no evidence to suggest that toddlers understand the relation between communicative and informative intentions; that communication functions because detecting someone's communicative intention usually leads us to infer and engage with their informative intention. One potential avenue for exploring understanding of the intentional structure of communication is to investigate instances of deception, where acts are produced that manipulate this structure, as in cases of hidden authorship. A recent study demonstrated that 3-years-old may understand the intentional structure of communication through their ability to engage in hidden authorship (Grosse et al., 2013), as reviewed in the previous chapter. However, if the intentional structure of communication underpins the capacity for language, we should be able to demonstrate that children understand this structure before language learning begins in earnest. Unfortunately, studies of hidden authorship cannot reasonably demonstrate such understanding with children younger than 3 years, because fairly advanced language skills are required to participate in the tasks.

A second potential avenue exists for exploring an understanding of intentional structure of communication. Whilst hidden authorship focuses on deception when signaling intent, it is possible that a person's behaviour when actively detecting another's intentions can reveal what they understand about the structure of intentional communication. When someone intentionally communicates with us, we detect *that* they are intending to communicate, and infer *what* it is that they intend to communicate. Normally, acknowledgement of someone's intent to communicate involves making eye contact, or orienting our head or body in such a way as to indicate that we are receptive. This ostensive listening indicates our recognition of another's communicative intention. Ostensive listening indicates to the signaller that we have recognized that they are communicating with us, and that we will attempt to infer their informative intentions. Listener behaviour therefore provides a mirror to the intentional structure of a signaller's communication. Manipulations of ostensive listening behaviour could provide evidence for toddlers' understanding of the intentional structure of communication, as much as manipulations of ostensive signalling behaviour do. However, this intriguing possibility has not yet been studied.

The current study focuses on a manipulation of ostensive listening behaviour. We focus specifically on toddlers who pretend not to hear something that their caregiver has said. Pretending not to hear another person involves not giving any indication that we have heard them, in other words, refraining from ostensive listening behaviour. By not acknowledging that another person is communicating (their communication intention), one does not have to comply with what they are communicating about (their informative intention). If we indicate to a communicator that we have detected their communicative intent, we normally have to comply with their request, or else we appear uncooperative, which can lead to punitive consequences (Tomasello, 2008, p. 92). To avoid appearing uncooperative, one strategy available to young children faced with a request to do something they'd rather not do, is to pretend they simply didn't hear it. A key aspect of pretending not to hear behaviour is that it necessarily involves detecting another's communicative intention. It also demonstrates an understanding that *showing* in your actions that you have detected it (engaging in ostensive listening behaviour) indicates to them that you will infer their informative intention (and conversely, hiding detection allows one to avoid compliance). Thus it indicates an awareness of the relation between two levels of intention in communication, through a deceptive manipulation of this structure.

Children are able to engage in pretence at around 18 to 24 months (Leslie, 1987; Rakoczy, 2006, 2008). On Rakoczy's (2006) account, at this age toddlers can intentionally and playfully pretend that a counterfactual proposition is true and act appropriately *as if* it was the case. Pretending not to hear involves pretending that a counterfactual proposition (that they have not heard their caregiver) is true. It is reasonable to expect 18-month-olds to engage in an active pretence. It is theoretically possible that social pretense of this type could emerge earlier (Kovacs, pers comm.) however this has never been tested.

We argue that pretending not to hear behaviour is most likely to occur when caregivers request that the toddler does something that they dislike doing. Demonstrating that a toddler is pretending not to hear involves showing that they do not respond to a request that they are very likely to have heard. A particularly strong indication of this would be if the child conspicuously appeared busy, i.e., pretended to be busy. We further argue that such demonstration would provide evidence that children understand the structure of intentional communication much earlier than previously thought. Though this would still not provide evidence that children understand the intentional structure of communication *prior* to the onset of conventional language (as this typically occurs much earlier), it has the potential to help provide a framework to explore the understanding of intentional structure of young children, and non-human animals without the need for sophisticated linguistic understanding.

This chapter has two main aims. We hypothesise that, in the months preceding their second birthday, toddlers begin to pretend not to hear when their caregiver communicates with them. Our primary aim is to demonstrate this both experimentally and through caregiver report. As we are specifically investigating toddler's non-responses to things that their caregivers say, a second aim of this chapter is to provide convincing evidence to distinguish pretending not to hear from competing explanations of the same phenomenon. These competing explanations are that toddlers genuinely might not hear things that their caregivers say, or equally, they might hear their caregiver, but ignore what they say as irrelevant (and so do not engage in ostensive listening behaviour because they don't believe the communication is relevant to them).

## Experiment 1

In this experiment, we sought to demonstrate that toddlers pretend not to hear requests from their caregiver for something that they don't want to do. Firstly, to provide evidence that toddlers pretend not to hear, we have to initially provide evidence that they are capable of hearing and responding to their caregivers when they are asked about something that they would like to do under experimental conditions. Second, we have to provide evidence that children do not indicate to their caregiver that they have heard them (by making eye contact, or responding vocally) when asked to do something that they don't want to do. These are necessary conditions for pretending not to hear. Finally, we sought to distinguish toddlers' pretending not to hear behaviour from other types of behaviour, e.g., by showing engagement in an active pretence that they had not heard what their caregiver said (by engaging in a new activity, or interacting with someone who is not their caregiver).

We predicted that 24-month-olds would be more likely to pretend not to hear than the 18-month-olds.

## Method

### Participants

Twelve 18-month-olds (4 girls, 8 boys, *M* = 18 months and 3 days, *SD* = 12 days), and thirteen 24-month-olds (6 girls, 7 boys, *M* = 24 months and 7 days, *SD* = 10 days) were included in the final sample.

A further 11 children were tested, but were excluded for bilingualism (*n* = 2), and during coding for not responding to either trial (*n* = 7), and for caregivers using the child's name in the request (*n* = 2).

### Apparatus and Materials

Two Sony CX-280 video cameras were used to record testing sessions. One camera was mounted in the top corner of the room to provide a scene view, while another was hidden behind a curtain on a tripod facing the participant. A low 50cm high wall was constructed in the middle of the room out of cardboard boxes, and covered in curtains, with a small gap to pass through at one end. Participants played with 7 x 80mm, 220g rubbery balls in red and blue, and four rolls of masking tape, which were used as ball holders (used in Experiment 1 in Chapter 3 of this thesis).

All video data was coded in ELAN (Slotjes & Wittenburg, 2008).

**Procedure**

Prior to taking part in this experiment, participants played in a warm up room with the researchers for 10-20 minutes, and taken part in a comprehension experiment (Experiment 1 in Chapter 3 of this thesis), and so were already in the testing room, and familiar with the apparatus and materials used in this experiment.

Participants were encouraged to play with Researcher 2 and a number of brightly coloured balls on a table facing away from their caregivers. They were separated from their caregivers by the dividing wall (described above) which had an opening at one end that allowed passage past it. Caregivers were sat with Researcher 1 on the other side of the room, facing the back of the participant (Figure 1).

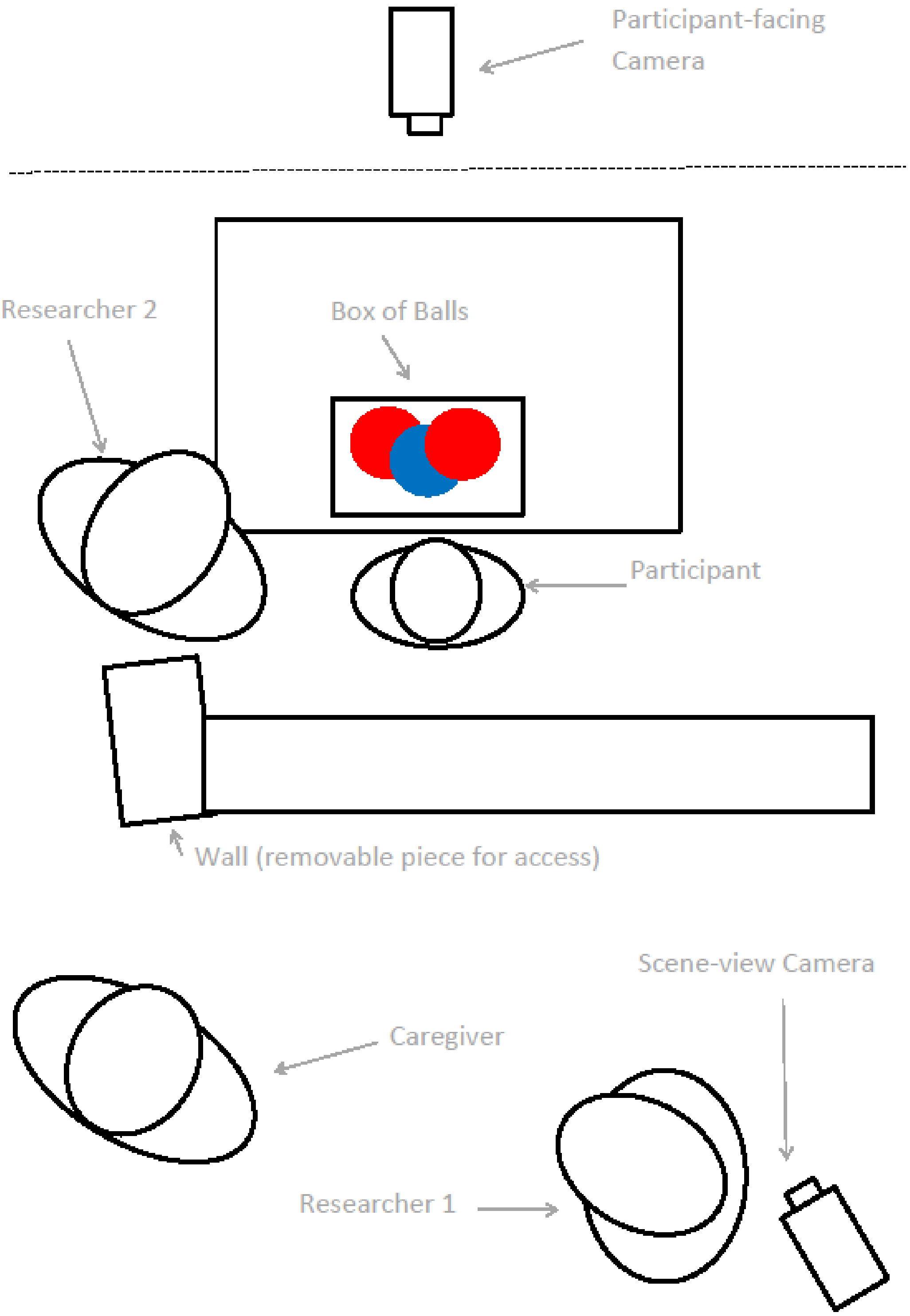


*Figure 1*. Room setup for Experiment 1

When the participant was engaged in the game and not facing their caregiver, caregivers were instructed by Researcher 1 to give their child either a positive request (e.g., ‘Do you want a snack?’) or a negative request (e.g., ‘Do you want to change your nappy?’). These requests were discussed prior to the beginning of the experiment. Caregivers were asked to bring a snack for their child, and think of something that they believe that their child would not usually want to comply with. Caregivers were asked to produce the request in a natural way as they would at home, using language that they thought their child would understand. However, they were also asked not to say the participant’s name at the beginning of the request, so as to allow the participant to hear the content of the request before responding, instead of orienting as soon as their name was called. As caregivers were not told prior to the experiment that we were interested in whether or not the participant responded, we gave no instructions to the caregivers on whether to repeat the request or not.

With the exception of one participant, negative requests all related to nappy changing ($n$ = 24). The remaining request involved asking the participant to put on sunglasses ($n$ = 1). Positive requests were all food and drink related ($N$ = 25), with requests to come and have snacks or juice.

During the 7s response period (starting from the beginning of the request) Researcher 2 was instructed not to initiate any new play, and only resume play if the participant initiated it. There was normally a period of time in between requests where the participant was given time to settle back into the game with the Researcher 2. The order of the requests was counterbalanced, with some participants having the positive request first (18-month-olds $n$ = 5, 24-month-olds $n$ = 8), and others having the negative request first (18-month-olds $n$ = 7, 24-month-olds $n$ = 5).

**Coding**

We coded measures of responsiveness (i.e., indicating that they had heard), and behaviours that co-occurred with non-responsiveness that could indicate that toddlers were pretending not to hear. Participants’ responses were coded from the beginning of the caregiver’s utterance to either 7s after this point or until the participant had responded. If there was no response, the response window lasted 7s, or to the point at which either the caregiver or the researcher next spoke.

**Responsiveness**. Participants were judged to have responded when they made eye contact with their caregiver, said ‘no’, ‘yes’, or made another negative or positive vocalisation identifiably related to the request.

**Pretending not to hear.** Participants who did not engage in any of the responding behaviours, and continued to play with the toy without engaging with the caregiver for the full 7s were categorised as *non-responders*. For all non-responders, we coded two switching behaviours that we argue provide evidence that a participant was pretending not to hear. Participants’ responses were coded for whether they began a new activity with the toy (or something else) or began a new interaction with Researcher 2 (the researcher next to them). Non-responders who engaged in this behaviour were categorized as *pretenders*.

Only participants who responded to at least one request (either positive or negative) were included in the analysis. If participants did not respond to either request, we could not guarantee that they could hear in the experimental set up. Participants responding to neither request were excluded from analysis (see Participants).

Caregivers were instructed not to use the participant’s name when making the requests and participants were excluded when their name was used in the request (see Participants).

## Results

### Responsiveness

Participants often responded to their caregiver’s requests. Table 1 shows the percentage of participants who responded as a function of condition and age.

Table 1

*Percentage of Participants Responding as a Function of Condition and Age*

| *Age* | *Positive* | *Negative* |
|---|---|---|
| 18-month-olds (*n* = 12) | 100% | 42% |
| 24-month-olds (*n* = 13) | 92% | 54% |

We fitted binary logistic regression models to analyze whether responsiveness (1 = responded, 0 = no response) differed as a function of condition and age group. To account for between-participant differences in responsiveness, participant (*N* = 25) was included as a random effect for all models. A model with condition as a predictor was a significantly better fit that the null model (containing no predictors), $x^2$ (2) = 16.28, $p < .001$, whereas a model with age group as a predictor was not. A model with both predictors did not improve on the model with only condition as a predictor. Therefore, responsiveness did not differ as a function of age. The model with condition as a predictor (LLRI = .27, C = .80, Dxy = .60), revealed that participants were significantly less likely to respond in the negative condition than in the positive condition (*b* = -3.26, SE = 1.10, *z* = -2.97, p = .003).

**Pretending Not to Hear**

Table 2 shows the percentage of participants who did not respond, and engaged in switching behaviours (actively engaging in a new activity or interaction)  as a function of age and condition.

Table 2

*Percentage of Participants That Engaged in Switching Behaviours Without Responding as a Function of Condition and Age*

| *Age* | *Positive* | *Negative* |
|---|---|---|
| 18-month-olds (*n* = 12) | 0% | 25% |
| 24-month-olds (*n* = 13) | 0% | 31% |

As participants of both ages were not observed to engage in switching behaviours in the positive condition, assumptions were not met for conducting binary logistic regression. Fisher's exact test revealed that 24-month-olds were not significantly more likely to engage in the switching behaviours than the 18-month-olds in response to the negative request.

**Requests**

As caregivers were not instructed on whether to repeat the request, requests were often repeated, especially when participants did not respond. This was especially the case in for negative requests. Significantly more utterances (defined as separate sentences) were given during the negative request trials ($M$ = 1.48) than during the positive request ($M$ =1.12), $t(24)$ = 2.57, $p$ = .017.

## Discussion

We demonstrated that toddlers engage in behaviour that suggests that they are pretending not to hear requests from their caregivers, when their caregiver has requested that they do something that the toddler does not want to do. We found no evidence that 24-month-olds were more likely to engage in this behaviour than 18-month-olds. We demonstrated that, at both ages, toddlers selectively respond to requests from their caregivers, and they are more likely to respond to requests for things that they might want to do, than things that they might not want to do.

Some 18- and 24-month-olds did not respond to negative requests, and seemed to engage in the switching behaviours that we associate with pretending not to hear (engaging in a new non-request related activity, or engaging in a new interaction with the researcher), and so our interpretation is that they were pretending not to hear. However, we did not sufficiently rule out alternative possibilities for their non-responsiveness. It is arguably the case that some toddlers did not hear requests in the negative condition, whereas they did in the positive condition. For this to be likely there would need to be something that made positive requests more salient than negative requests. During this experiment, often caregivers had snacks in their hand, or proceeded to remove them from their bag as they spoke, which may have provided extra reason for toddlers to turn around and attend to their caregiver. Therefore these results could be accounted for by an imbalance between the positive and negative requests, because positive requests were backed up with the physical incentive (which may or may not have cued participants to respond). Also, the exclusion rate for the task was high (around 20% of children who took part in the testing session) for participants who did not respond to either request. This raises the concern that the set up may have been difficult at least for some toddlers to hear that their caregiver was communicating with them.

The experiment also gives no reason to discount the competing explanation that children might just ignore things that they believe are irrelevant to them. Toddlers may not have responded to negative requests because they didn't believe they were directed at them, as they weren't pertinent to what they were engaged in (e.g., playing with the toy). Equally, it is possible that when toddlers heard something irrelevant, they momentarily lost concentration in their play, and when they continued did something different (i.e., engaging in a switch behaviour), leading us to code them as engaging in pretence.

There were a number of other potentially confounding variables across conditions. Participants were often not in the same spatial location when caregivers made the request. Although broadly being behind the wall, facing away from their caregivers, the balls could roll around, and did not keep participants in a specific place in the room. Caregiver utterances were not constrained enough, with caregivers using different means of asking their toddlers about things across conditions, (e.g., saying, "Do you want something to eat?" versus "Shall we change your nappy?"). Caregivers also often repeated requests in the negative condition, when they did not do so in the positive condition. While it is not entirely clear how these confounds might account for the results, it does highlight the need to ensure better experimental control of caregiver's utterances and participant's spatial location.

We did not find a difference in pretending not to hear behaviour as a function of age as we had predicted, despite a higher percentage of 24-month-olds engaging in this behaviour than 18-month-olds. This could be due to low sample size, or alternatively, pretending not to hear may occur earlier than we anticipated, or follow an idiosyncratic developmental schedule.

Further work should aim to replicate this result on a larger sample, across a broader age range. In addition, we should seek to address the alternative explanations more fully. Firstly, we need to ensure that caregivers do not accidentally cue their children into turning around by reaching for snacks in their bag or similar activities. Secondly, we need to include a control phrase of child-directed, irrelevant speech and a baseline phase where there is no child-directed speech. These would give us an indication of how toddlers behave when caregivers say something irrelevant to them, or don't say anything to them, with which to compare their behaviour in the negative and positive conditions. In addition, more work should be done to remove confounding variables, surrounding variability of caregiver's utterances, to avoid caregivers repeating utterances in one condition and not the other, and constraining participants to a specific spatial location. This would give us greater confidence that they are equally likely to hear their caregivers across conditions, and so make instances of pretending not to hear behaviour more convincing.

## Experiment 2

The aim of Experiment 2 was to address methodological questions raised in Experiment 1, to provide a more reliable measure of toddlers' pretending not to hear behaviour. We added a control condition, to determine whether toddlers' behaviour differed between occasions when their caregivers asked them to do things that they didn't want to do (negative condition), and caregivers saying things that were irrelevant to toddlers' play but were still child-directed (control condition). In addition, we coded a baseline phase to determine how much the switching behaviours that we associate with pretending not to hear behaviour occur when there is *no* child-directed speech. Caregivers were given a script in order to control variability in requests across conditions, and to avoid repetition, and the toy was changed to one that would keep children in a more constrained spatial location across trials. We extended the period for children to give a response (from 7 to 10 seconds), in order to maximize the possibility that they would engage in pretence. Finally, caregivers were told not to reveal the snack until after children responded to them in the positive condition.

# Method

## Participants

**Main sample.** Twenty-five 18- to 20-month-olds (10 girls, 15 boys, *M* = 19 months and 0 days, *SD* = 29 days), twenty-seven 24- to 26-month-olds (13 girls, 14 boys, *M* = 25 months and 7 days, *SD* = 25 days), and twenty-three 30- to 32-month-olds (11 girls, 12 boys, *M* = 30 months and 25 days, *SD* = 29 days) were included in the main sample.

A further 9 children were tested, but excluded, as only participants with a valid control, positive and negative trial were included for the analysis. Participants were excluded for failure to achieve the correct spatial set up (*n*= 2), when caregivers used their child's name during the request (*n* = 2), sibling interference (*n* = 2) and experimenter error (*n* = 1), during one or more trials. Participants were also excluded for bilingualism (*n* = 1) and not responding to any of the three experimental trials (*n* = 1).

**Baseline phase.** A subset of participants from the main sample was included for analyses of the baseline phase. Participants were only included when there was no child-directed speech towards them during the 10s from either the caregiver or the researchers. This subset consisted of twenty 18- to 20-month-olds (8 girls, 12 boys, *M* = 18 months and 28 days, *SD* = 29 days), twenty-three 24- to 26-month-olds (10 girls, 13 boys, *M* = 25 months and 12 days, *SD* = 24 days), and twenty 30- to 32-month-olds (9 girls, 11 boys, *M* = 30 months and 26 days, *SD* = 29 days).

**Caregiver report**. All caregivers (plus an additional 7 caregivers whose children took part in the pilot studies) completed a questionnaire. After exclusions we had questionnaires from caregivers of twenty-nine 18- to 20-month-olds (10 girls, 19 boys, *M* = 19 months and 3 days, *SD* = 31 days), caregivers of thirty-four 24- to 26-month-olds (18 girls, 16 boys, *M* = 25 months and 2 days, *SD* = 27 days) and caregivers of twenty-five 30- to 32-month-olds (11 girls, 14 boys, *M* = 30 months and 21 days, *SD* = 25 days). Exclusions were for bilingualism (*n* = 1), for leaving substantial blanks (*n* = 1), for not answering the first question (*n* = 1) and for failure to follow questionnaire instructions (*n* = 1).

## Apparatus and Materials

Two Sony CX-280 video cameras were used to record testing sessions. One camera was mounted in the top corner of the room to provide a scene view, while another was hidden in a fabric covered box on a tripod facing the participant. Participants played with a wooden activity cube toy (ELC: http://www.elc.co.uk/Mini-Wooden-Activity-Cube/136105,default,pd.html) on a custom-made wooden bench (50cm high).

We administered our questionnaire on responsiveness and pretending not to hear (Appendix A). All video data was coded in ELAN (Slotjes & Wittenburg, 2008).

## Procedure

Prior to taking part in this experiment, participants had played in a warm up room with the researchers for 10-20 minutes, they had also taken part in a comprehension experiment (Experiment 2 in Chapter 3 of this thesis), so were already familiar with the testing room and the researchers.

Participants were encouraged to play with Researcher 2 and a toy on a bench facing away from their caregivers. Caregivers were with Researcher 1 on the other side of the room, facing the back of the participant (Figure 2).

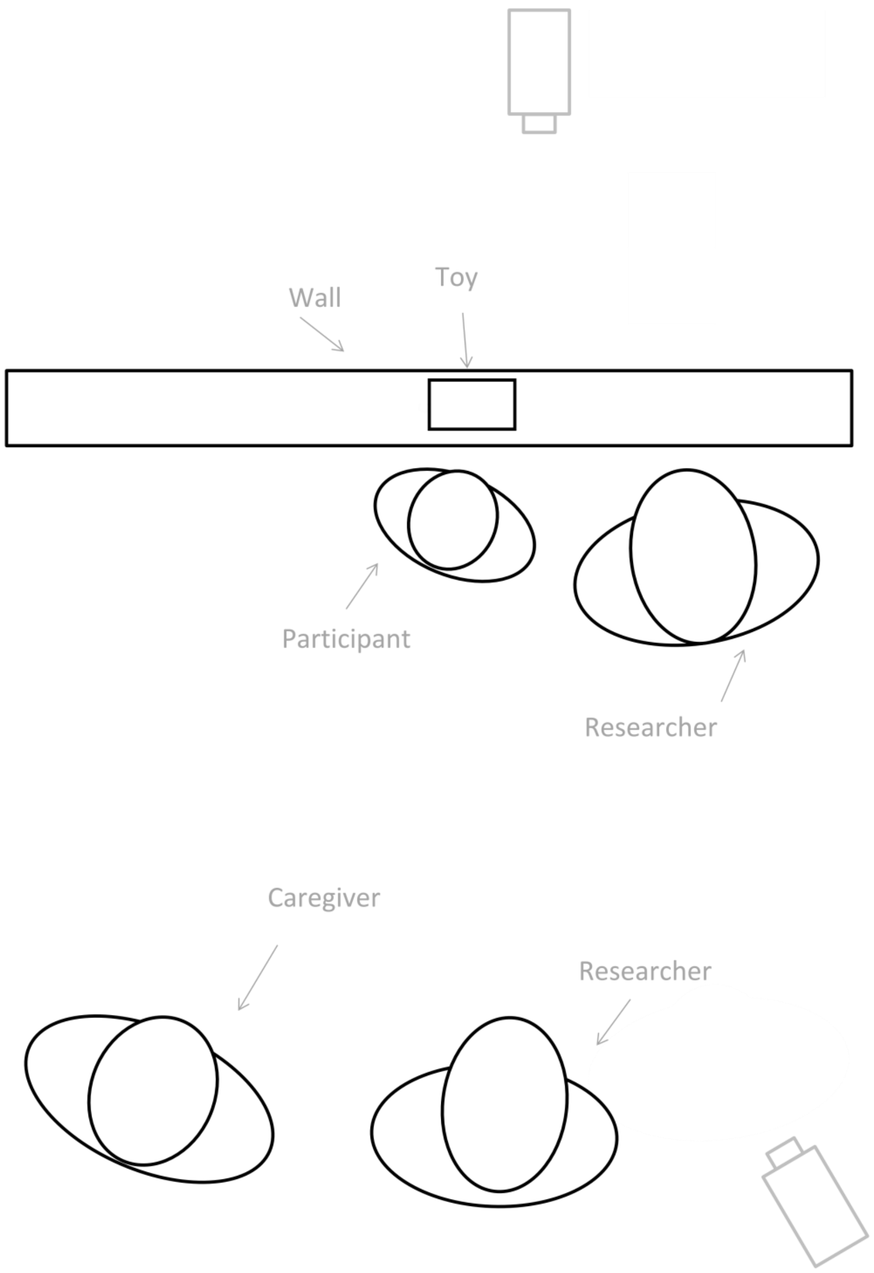


*Figure 2.* Room setup for Experiment 2

When the participant was engaged in playing with the toy and was facing away from the caregiver, caregivers were instructed to make either a positive request (‘Snack time! Do you want a snack?’), a negative request (‘Nappy time! Do you want to change your nappy?’), or a control phrase (‘Oh! That’s a nice cup!’) to the participant. Caregivers were asked to produce all utterances in a natural way as they would at home, as if they were directed to the participant. However, they were asked not to say the participant’s name at all during the request, and not to repeat the request. Some requests were adapted for participant’s comprehension, available resources and through consultation with the caregiver. Caregivers were asked whether they thought their child would understand ‘snack’, and if not, they provided an alternative. Some caregivers did not bring food to the testing session, and others reported that food or drink was not a strong enough incentive, or nappy changing a strong enough disincentive. The order of the utterances was counterbalanced.

During the 10s response period (starting from the beginning of the utterance) the researcher playing with the child was instructed not to initiate any new play, and only play if the participant initiated it.

***Positive condition*.** For two of the 18- to 20-month-olds, and one 24- to 26-month-old, caregivers used the word to describe the snack instead of ‘snack’ because they were unsure of their child’s comprehension of ‘snack’, e.g., “Blueberry time! Do you want some blueberries?”. For two 24- to 26-month-olds, and two 30- to 32-month-olds, the request referred to something other than a snack, referring to juice, stickers, the caregiver’s phone or the participant’s doll.

***Negative condition*.** For one of the 24- to 26-month-olds, and two 30- to 32-month-olds, caregivers asked about going to the toilet (as the participant was not wearing a nappy)**.** For one 30- to 32-month-old, the caregiver asked the participant to put on their coat.

**Caregiver report.** At the end of the experiment, caregivers filled out our questionnaire on responsiveness and pretending not to hear (see Appendix A), prior to being debriefed. After finishing the questionnaire, caregivers were informed that we were interested in children’s non-responsiveness and pretending not to hear behaviour, but were not informed of this prior to filling out the questionnaire.

The questionnaire consisted of 10 questions. Questions 1 to 8 required participants to give a response on an ordinal scale (1-5) indicating frequency of a given behaviour (1 = Never, 2 = Less Than Half the Time, 3 = About Half the Time, 4 = More Than Half the Time, 5= Always). Questions 1 and 5 only contained one item, whereas questions 4,6,7 and 8 contained 2 items, and questions 2 and 3 contained 5 items each. Therefore, there was a maximum of 20 items per participant using this scale. Questions 9 and 10 asked about older siblings and other languages spoken in the home.

***Reliability check questions.*** During piloting of the caregiver report measure (a structured interview of caregivers conducted in 2014), caregivers often gave idiosyncratic responses – e.g., reporting that their children said 'no' when they were pretending not to hear – which categorically would not be pretending not to hear behaviour on our account.

We included questions where participants had to rate how frequently they observed a number of behaviours (their child making eye contact, saying "no" or shaking their head, getting angry, laughing), when their child did, or did not respond to them. We planned to exclude caregivers when they reported responding behaviours (making eye contact, saying no/shaking their head, getting angry, laughing or smiling) more frequently in contexts where we had explicitly asked about their child's non-responding behaviour than in contexts where we asked about their responding behaviour. While 11 participants would have been excluded using this criteria, in hindsight the question asking about non-response was somewhat confusing, so may have lead these caregivers to idiosyncratic responses. These participants were therefore not excluded from the analyses, and their inclusion did not significantly affect any results.

**Coding**

We coded measures of responsiveness, and behaviours that co-occurred with non-responsiveness that could indicate that toddlers were pretending not to hear. Participants responses were coded from the beginning of the caregiver's utterance to either 10s after this point or until the participant had responded. If there was no response, the response window lasted 10s, or to the point at which either the caregiver or the researcher next spoke.

**Responsiveness**. Participants were judged to have responded when they made eye contact with their caregiver (however, see Experiment 2b), said ‘no’, ‘yes’, made another negative or positive vocalisation identifiably related to the request, or nodded or shook their head.

**Pretending not to hear.** Participants who did not engage in any of the responding behaviours, and continued to play with the toy and not engage with the caregiver for the full 10s were categorised as *non-responders*. For all non-responders, we coded two switching behaviours that we argue provide evidence that a participant is pretending not to hear. Participants’ responses were coded for whether they began a new activity with the toy (or something else) or began a new interaction with researcher 2 (the researcher next to them). Non-responders who engaged in this behaviour were categorized as *pretenders*.

To check how likely the switching behaviours would be to occur in the absence of any requests from the caregiver, we coded a baseline non-request phase where possible. That is, we coded 10 seconds where there was no child directed speech from the caregiver or the researchers.

As in Experiment 1, only participants who responded to at least one caregiver utterance were included in the analysis, and excluded if they did not (see Participants). Caregivers were instructed not to use their child’s name during the utterances, and participants were excluded if their name was used (see Participants).

## Results

### Responsiveness

Table 3 shows the percentage of participants that responded to the request across the positive and negative conditions, by age.

Table 3

*Percentage of Participants Responding as a Function of Condition and Age*

| *Age* | *Positive* | *Negative* | *Control* |
|---|---|---|---|
| 18- to 20-month-olds | 100% | 84% | 24% |
| 24- to 26-month-olds | 93% | 96% | 44% |
| 30- to 32-month-olds | 96% | 100% | 57% |

We fitted binary logistic regression models to analyze whether responsiveness (1 = responded, 0 = no response) differed as a function of condition and age group. To account for between-participant differences in responsiveness, participant ($N$ = 75) was included as a random effect for all models. A model with condition as a predictor was a significantly better fit than the null model (containing no predictors), $x^2$ (2) = 80.10, $p$ < .001, whereas a model with only age group as a predictor was not. A model with both predictors was a significantly better fit than the model with only condition, $x^2$ (2) = 6.71, $p$ = .035, while an interaction model did not result in a significant improvement in fit.

Figure 3 shows the mean log odds and the Bayesian credible interval (proxy for a 95% confidence interval) for condition and age group pairs calculated using Markov chain Monte Carlo method from the best fitting model with both condition and age as predictors ($x^2$ (2) = 86.61, $p$ <.001. LLRI = .36. C = 91, Dxy = .81). Significant differences occur when the credible intervals for one pair do not overlap with the mean of another pair (at alpha .05, two-tailed).

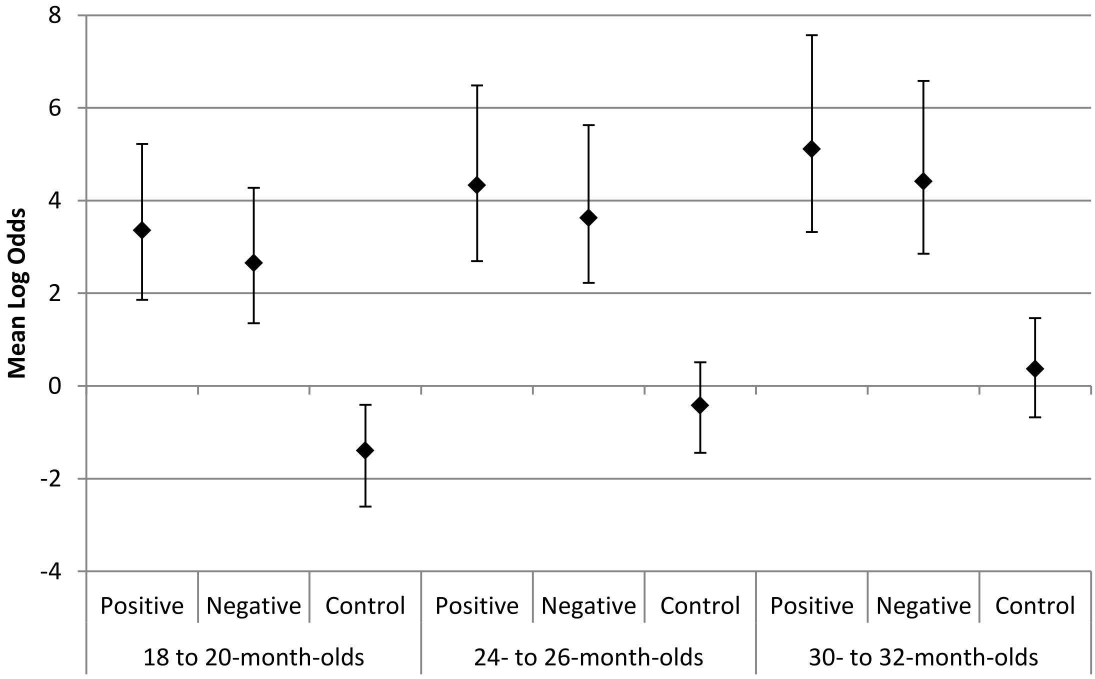


*Figure 3.* Mean log odds of responsiveness for all combinations of condition and age. Error bars represent 95% Bayesian credible intervals about the mean.

Figure 3 reveals that participants of all ages are significantly less likely to respond in the control condition compared to the negative and positive condition. In addition, 18- to 20-month-olds were least likely to respond in the control condition (mean log odds = -1.39), and were significantly less likely to respond than the 30- to 32-month-olds (mean log odds = 0.37, $p$ = .010). 18- to 20-month-olds were also the least likely to respond in the negative condition (mean log odds = 2.66) and were significantly less likely to respond than the 30- to 32-month-olds (mean log odds = 4.42, $p$ = .010). No other differences as a function of age or condition were significant.

**Pretending Not to Hear**

Table 4 shows the percentage of participants who did not respond, and engaged in switching behaviours, as a function of age and condition.

Table 4

*Percentage of Participants Who Engaged in Switching Behaviours Without Responding as a Function of Age and Condition*

| *Age* | *Positive* | *Negative* | *Control* |
|---|---|---|---|
| 18- to 20-month-olds | 0% | 4% | 24% |
| 24- to 26-month-olds | 4% | 4% | 37% |
| 30- to 32-month-olds | 0% | 0% | 30% |

We fitted binary logistic regression models to analyze whether the participants engaging in switching behaviour (engaged in switching behaviour = 1, or not = 0) differed as a function of condition and age group, with participant ($N$ = 75) as a random effect for all models. A model with condition as a predictor was a significantly better fit than the null model (with no predictors), $x^2$ (2) = 39.57, $p$ < .001, whereas a model with only age group as a predictor was not. A model with both predictors was not a significant improvement in fit. Therefore, the likelihood that a participant would engage in switching behaviour did not differ as a function of age. The best fitting model with condition as the only predictor (LLRI = .25. C = .95, Dxy = .89) reveals that participants were significantly more likely to engage in pretending not to hear in the control condition than in the negative condition ($b$ = 2.79, SE = 0.80, $z$ = 3.48, $p$ < .001) but were not more likely to engage in pretending not to hear behaviours in the negative than in the positive condition.

**Baseline Measures**

We compared the experimental conditions with child-directed speech (positive, negative and control), with the baseline phase where there was *no child-directed speech*; to determine how frequently switching behaviours that we associate with pretending not to hear (changing activity, or spontaneously engaging with researcher 2) occur. For the subset of the data where we have a valid baseline phase, Table 5 reports the percentage of participants who engaged in these switching behaviours across the three experimental conditions and the baseline phase.

Table 5

*Percentage of Participants Engaging in Switching Behaviour Without Responding as a Function of Condition/Baseline Phase*

| *Age Group* | *Positive* | *Negative* | *Control* | *Baseline* |
|---|---|---|---|---|
| 18- to 20-month-olds (*n* = 20) | 0% | 0% | 30% | 35% |
| 24- to 26-month-olds (*n* = 23) | 4% | 4% | 35% | 48% |
| 30- to 32-month-olds (*n* = 20) | 0% | 0% | 40% | 35% |

We fitted binary logistic regression models to analyze whether participants engaging in switching behaviours (engaged in switching behaviour = 1, or not = 0) differed as a function of condition and age group, with participant (*n* = 63) as a random effect for all models. A model with condition as a predictor was found to be a significantly better fit than the null model (with no predictors), $x^2$ (2) = 62.71, $p < .001$, whereas a model with only age group as a predictor was not an improvement from the null. A model with both predictors was not an improvement in fit on the model with just condition. The best fitting model with condition as the only predictor (LLRI = .25. C = .95, Dxy = .90) reveals that participants were significantly more likely to engage in switching behaviours in the baseline phase ($b$ = 3.89, SE = 1.06, $z$ = 3.66, $p < .001$) and in the control condition ($b$ = 3.66, SE = 1.06, $z$ = 3.46, $p < .001$) than they were in the negative condition.

**Caregiver Report**

Caregivers were asked to think about occasions when they asked their child to do something that their child did not want to do. For each age group, median and inter-quartile range of responses and number of respondents to all items on the questionnaire are reported in Appendix B, Table 38.

Caregiver report was analyzed using Kruskal-Wallis tests to determine if caregivers' responses differed as a function of the age of their child. The only item on the questionnaire that yielded a difference as a function of age asked caregivers whether their child responded by saying something other than 'no' when asked to do something that they didn't want to do, $H$(2)

= 9.89, $p = .007$. Caregivers with younger children were more likely to report that they did not to do this. For every other item on the questionnaire however, there were no significant age differences. Results from caregiver report are therefore collapsed by age for the following descriptives.

When caregivers reported frequency of a behaviour for their child on our five point scale by reporting 3 – 5 (3 “About Half the Time”, 4 “More Than Half the Time” or 5 “Always”), we report this as *frequently* engaging in that behaviour. Where participants report 2 (“Less Than Half the Time”), we report this as *infrequently* engaging in this behaviour, and when they report 1 (“Never”), we report this as *never* engaging in this behaviour.

**Avoiding responding.** 64% of caregivers reported that they felt that their child frequently avoided eye contact, and 62% reported that they felt their child frequently avoided saying anything when asked to do something that they did not want to do.

**Carrying on playing.** 98% of caregivers reported that their children would frequently carry on playing when they hadn’t responded to being asked to do something that they didn’t want to do. 87% reported that their children frequently continued playing in the same way. We hypothesized that pretending not to hear could be indicated by children playing more intently. 50% reported that their child did this infrequently or never.

**Deliberate non-response.** When asked about occasions where their child did not respond, 52% reported that they believed that their child was frequently too engaged in what they were doing. However, on a separate item 75% of caregivers reported that they felt that their child frequently deliberately did not respond to them.

**Ignoring and pretending not to hear.** Caregivers were asked about cases where their children deliberately did not respond, and carried on playing when they were asked to do something that they didn’t want to do. In these cases, 81% of caregivers reported that they felt their children frequently deliberately *ignored* them. On a separate item, asking specifically about whether children engaged in an active pretence, 70% reported that they felt that their children frequently actively pretended not to hear them (by pretending to be very busy), while 20% reported that this was infrequent, and 10% reported that this was never the case.

## Discussion

Children in Experiment 2 almost always responded to both positive and negative requests, whereas they responded to the control condition speech about half the time. Likewise, when considering whether children engaged in switching behaviours upon hearing an utterance, it was only in the control condition that such behaviours occurred. In a baseline phase, we also found that these switching and engaging behaviours were very likely to occur naturally (far more so than they did in response to a negative request). Finally, caregiver report indicated that caregivers think their children pretend not to hear them with some frequency.

The primary aim of this study was to address competing explanations for children's responses to negative requests in Experiment 1. Specifically, we aimed to address whether toddlers had not heard the negative request or judged it to be irrelevant to them. In this study, overwhelmingly, toddlers responded to negative requests, rendering these questions moot. It is not obvious why children responded to negative requests so much more in this study than Experiment 1. One possibility is that by controlling the request that caregivers had to produce ("Nappy time! Time to change your nappy!') we presented children with relatively unfamiliar language (compared to the way caregivers naturally asked in Experiment 1) and that by the time the children had processed the request, they had already responded in some way to the caregiver. It is also possible that children in the first experiment genuinely had trouble hearing the requests, whereas they did not in the second. Exclusions for when a toddler did not respond to *any* requests, were substantially reduced from around 20% in Experiment 1, to around 1% (1 participant in 84 tested) in Experiment 2.

The additional controls in this study nonetheless provided some interesting observations. First, children generally ignored the control utterances that were presumably irrelevant to them. Second, the switching and engaging behaviours we took to be indicators of pretending not to hear are in fact much more common in the absence of a negative request than they are in response to the negative request. In the baseline phase, where there was no child directed speech, these switching behaviours were 10 times more likely than they were in response to the negative condition. They are thus not sufficiently clear markers of actively pretending not to hear. This suggests that identifying pretending not to hear behaviour is only possible when you can determine that the child has definitely heard and understood what is being asked of them.

There is a divergence between what we were able to observe of children's pretending not to hear behaviour, and what caregivers reported. Caregivers reported that children often continued playing without responding to them when they asked them to do something that they didn't want to do, and they often felt that this was deliberate avoidance behaviour. Around 75% of caregivers reported that their child frequently deliberately did not respond to them when asked to do something that they didn't want to do. Around 70% of caregivers further reported that they felt that their child was actively pretending not to hear them frequently on occasions where they deliberately did not respond. The discrepancy in caregiver reports and lab findings is almost certainly due to the novelty of the testing setup to children, including the way children were asked the requests. Children may not have been able to resist turning around to respond to these relatively unfamiliar utterances in an unfamiliar setting. It is possible that in the natural environment, there would be more avoidance behaviours that children could engage in, or more opportunities to exploit environmental factors to avoid responding. For example, under our set up, toddlers could not run away from their caregivers to avoid responding (as the only entrance to the room was behind their caregivers). Additionally, the presence of other toys, siblings, or distractions could allow toddlers to more convincingly engage in pretence. Future studies should therefore look at more naturalistic settings for this research.

One final issue that bears discussion concerns the nature of the responses children made. During testing, it was observed that a number of children made eye contact, but then went back to playing without making any further explicit response to their caregivers' requests. It is possible that while undoubtedly these children are not accomplished at pretending not to hear (as you would not make eye contact if you were pretending not to hear), that the salience of the requests was too strong to resist turning around, and we may be discounting theoretically interesting responses by not distinguishing fully between explicit responses and more ambiguous responses. This possibility is explored in the post-hoc analysis of Experiment 2 (see below).

A final intriguing possibility is that in reality, toddlers don't pretend not to hear utterances that their caregivers make. It is possible, as with other early infant communicative behaviours, that our tendency to treat conspecifics as Gricean communicators leads us to over interpret their actions, which while helping scaffold these interactions, is *not* evidence that our assumptions are warranted (Stephens & Matthews 2015, p.15).

**Experiment 2b (Post-hoc Analysis)**

In Experiment 2, a number of children made eye contact in response to their caregiver but then appeared to give no *explicit* response such as making affirmative or negative vocalisations, or nodding or shaking their head, often conspicuously returning back to play. This led to observation of an unintended, but nonetheless interesting behaviour that may also shed light on toddler's understanding of communicative intentions.

If toddlers make eye contact with their caregiver, but then do not acknowledge what has been said and go back to what they are doing, this could equally be interpreted both as a response or a non-response. In terms of our ostensive listening model, though the toddler fails to hide their acknowledgement of the communicative intention (and thus engages in ostensive listening behaviour), they make no effort to engage with the informative intention. On a lean interpretation, this failure to hide acknowledgement of the communicative intention, but lack of explicit engagement with the informative intention might demonstrate that toddlers *lack* the understanding of intentional structure, i.e., they don't understand that acknowledgment of someone's communicative intention (through ostensive listening behaviour) implies a tacit agreement to engage with their informative intention – in fact they may not understand the communicative intention at all. However, a rich interpretation might be that children are trying to pretend that they haven't heard their caregiver, but are not accomplished at inhibiting orienting responses to their caregiver's vocal utterances, so fail to complete the deception. This would be consistent with findings that three-year-olds fail in tasks where they are required to deceptively point, as a result of poor inhibitory control (Carlson, Moses, & Hix, 1998).

The evidence that would distinguish between the lean and rich interpretation has to be the behaviour that the toddlers engage in after eye contact has been made. Do toddlers make an attempt to pretend that they haven't heard requests to do things that they don't want to do, even after indicating that they have heard (by making eye contact)? The aim of this post-hoc analysis is to analyze trials from Experiment 2 to determine whether toddlers who made eye contact only in response to the negative request went back to their activity, or attempted to pretend that they hadn't heard, by interacting with the researcher or by engaging in new activity. While this behaviour does not constitute the accomplished pretending not to hear behaviour investigated in the previous two studies, it gives valuable insight into whether toddlers might understand the intentional structure of communication.

## Method

### Participants

Participants from Experiment 2 were included for secondary analysis. Participants were excluded when participants made eye contact initially, but were then prompted by their caregiver (*n* = 17; for explicit criteria, see *Coding* below). This left twenty 18- to 20-month-olds (8 girls, 12 boys, *M* = 19 months and 3 days, *SD* = 31 days), twenty 24- to 26-month-olds (12 girls, 8 boys, *M* = 25 months and 10 days, *SD* = 22 days) and eighteen 30- to 32-month-olds (8 girls, 10 boys, *M* = 30 months and 27 days, *SD* = 30 days).

### Coding

Experiment 2 was not designed to discriminate between cases where participants made eye contact only, and those whose response involved other more explicit responding behaviours. The aim of Experiment 2 was to determine if children pretended not to hear. Our logic was that children who make eye contact give away that they have heard, so we had no reason to treat these cases separately from those where participants responded more explicitly. Therefore, caregivers were not instructed on whether to respond to eye contact differently as to the more explicit responses. In our original instructions to caregivers, as soon as participants had responded, we allowed caregivers to interact freely. As a result, caregivers often prompted their child when they had only made eye contact, but before they had given a more explicit response (saying yes/no, nodding shaking their head). In twenty-six trials, participants' caregivers prompted them after they only made eye contact. These trials were invalid for this analysis, and participants without all three valid trials were excluded ($n$ = 17).

**Explicit response.** Participants were coded as responders if they *explicitly* responded to requests from their caregiver, saying ‘no’, ‘yes’, making another negative or positive vocalisation identifiably related to the request, or nodding or shaking their head. Participants who made eye contact with their caregiver were only coded as responders if (1) eye contact co-occurred with one of the other two responding behaviours (i.e., vocalising or nodding/shaking their head), or (2) if they approached their caregiver without continuing to play, or (3) if they engaged in another communicative act (e.g., gesturing or smiling at their caregiver).

**No explicit response**. Participants who only made eye contact initially and went back to playing *without* approaching their caregiver, or engaging in another communicative act with their caregiver were coded as *partial-responders*. Participants who did not engage in the explicit responding behaviours, and also did not make eye contact, were coded as *non-responders*. Together, participants in these categories gave *no explicit response*.

**Pretending not to hear.** For all those participants who gave no explicit response we coded two sub-types of behaviour which we originally took as evidence that a participant is pretending not to hear. Participants’ responses were coded for whether they began a new activity with the toy (or something else) or began a new interaction with researcher 2 (the researcher next to them).

Participants responses were coded from the beginning of the caregiver’s utterance to either 10s after this point or until the participant had unambiguously responded positively or negatively (by either vocalising related to the question or shaking/nodding their head). In cases where participants made eye contact initially, the response window continued either to when the participant approached their caregiver, or when they gave some other communicative response, or, if there was no such response, the response window lasted 10s, or to the point at which either the caregiver or the researcher next spoke.

## Results

### Responsiveness

Participants often responded to their caregiver by responding verbally to the request, nodding or shaking their head. When participants initially only made eye contact (which happened in 34 trials), on eighteen occasions, these participants stopped playing and approached their caregiver. A further two engaged in another communicative response, with one participant smiling ostensively at their caregiver before turning back, while another showed their caregiver a part of the toy using a showing gesture (in response to the control utterance) – both of whom are included as responders. Table 6 shows the percentage of responders (and conversely the percentage of those who gave no explicit response) as a function of age and condition.

Table 6

*Percentage of Responders as a Function of Condition and Age*

| *Age Group* | *Positive* | *Negative* | *Control* |
|---|---|---|---|
| 18- to 20-month-olds (*n* = 20) | 85% | 60% | 20% |
| 24- to 26-month-olds (*n* = 20) | 85% | 95% | 30% |
| 30- to 32-month-olds (*n* = 18) | 94% | 94% | 28% |

We fitted binary logistic regression models to analyze whether responsiveness (1 = responded, 0 = no explicit response) differed as a function of condition and age group. To account for between-participant differences in responsiveness, participant (*n* = 58) was included as a random effect for all models. A model with condition as a predictor was a significantly better fit than the null model (containing no predictors), $x^2$ (2) = 63.26, $p < .001$, whereas a model with only age group as a predictor was not. A model with both age group and condition as predictors was a significantly better fit than the model with only condition, $x^2$ (2) = 6.00, $p$ = .050. An interaction model did not result in a significant improvement in fit.

Figure 4 shows the mean log odds and the Bayesian credible interval (proxy for a 95% confidence interval) for condition and age group pairs calculated using Markov chain Monte Carlo method from the best fitting model with both condition and age as predictors ($x^2$ (4) =

69.26, $p$ <.001. LLRI = .31. C = .90, Dxy = .79). Significant differences occur when the credible intervals for one pair do not overlap with mean of another pair (at alpha .05, two-tailed).

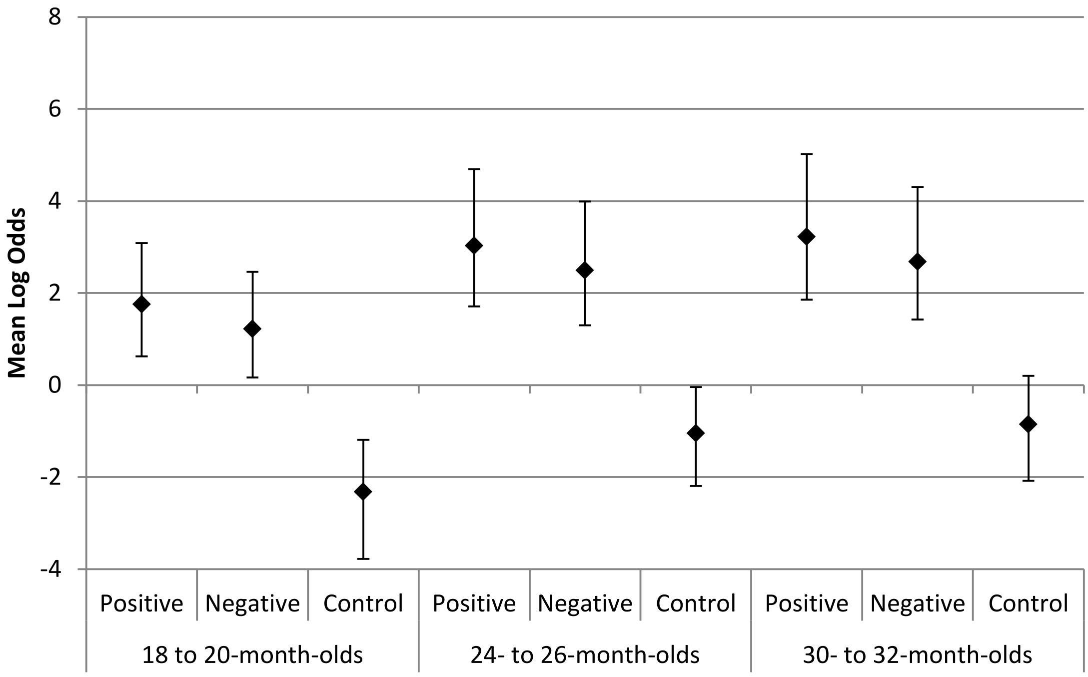


*Figure 4.* Mean log odds of responsiveness for all combinations of condition and age. Error bars represent 95% Bayesian credible intervals about the mean

Figure 4 reveals that all age groups are significantly more likely to respond in the positive and negative conditions than in the control condition.

In the negative condition, 18- to 20-month-olds were the least likely to respond (mean log odds = 1.22), and were significantly less likely to respond than the 24- to 26- (mean log odds = 2.49, $p$ = .028), and 30- to 32-month-olds (mean log odds = 2.68, $p$ = .019). In the positive condition, 18- to 20-month-olds were again, least likely to respond (mean log odds = 1.75), and significantly less likely to respond than the 30- to 32- (mean log odds = 3.22, $p$ = .019), but not the 24- to 26-month-olds. In the control condition, 18- to 20-month-olds were again least likely to respond (mean log odds = -2.32) and were significantly less likely to respond than the 24- to 26- (mean log odds = -1.05, $p$ = .028) and 30- to 32-month-olds (mean log odds = -0.86, $p$ = .019).

**Partial Response**

A number of participants were coded as partial-responders, that is to say that they made eye contact, but then returned to playing without engaging in any more communicative acts with their caregiver, and without approaching their caregiver. Table 7 shows the percentage of partial-responders from Experiment 2 as a function of age and condition.

Table 7

*Percentage of Partial-responders as a Function of Condition and Age*

| *Age Group* | *Positive* | *Negative* | *Control* |
|---|---|---|---|
| 18- to 20-month-olds (*n* = 20) | 15% | 20% | 0% |
| 24- to 26-month-olds (*n* = 20) | 5% | 5% | 5% |
| 30- to 32-month-olds (*n* = 18) | 0% | 6% | 17% |

We fitted binary logistic regression models to analyse whether partial-responders (outcome variable: partial-response = 1, no response/response = 0) differed as a function of condition or age. Models with either age or condition (or both) as predictors did not improve on a null model with only participant ($n = 58$) as a random effect and no predictors. Therefore partial responders did not differ as a function of either condition or age group.

**Pretending Not to Hear**

Of the participants who gave no explicit response some engaged in switching behaviour. Six instances of switching behaviour were coded from participants who had initially made eye contact, and table 8 shows the percentage of these switchers as a function of condition and age.

Table 8

*Percentage of Switchers Who Made Eye Contact as a Function of Condition and Age*

| *Age Group* | *Positive* | *Negative* | *Control* |
|---|---|---|---|
| 18- to 20-month-olds (*n* = 20) | 10% | 5% | 0% |
| 24- to 26-month-olds (*n* = 20) | 5% | 0% | 0% |
| 30- to 32-month-olds (*n* = 18) | 0% | 6% | 6% |

We fitted binary logistic regression models to analyse whether these switchers (outcome variable: switched behaviour = 1, other = 0) differed as a function of condition or age. Models with either age or condition (or both) as predictors did not improve on a null model with only participant (*n* = 58) as a random effect and no predictors. Therefore these switchers who made eye contact did not differ as a function of either condition or age group.

## Discussion

A number of participants did not explicitly respond to the request, despite making eye contact, and went back to playing, when asked to do something that they did not want to do. This behaviour did not significantly differ as a function of age, and also was not significantly more likely to occur in any condition. Likewise, switching behaviours that we originally took as indicative of pretence were not more likely to occur in any condition.

This re-analysis highlights the need in to ensure caregivers do not prompt their children during the experiment before they explicitly respond. A set up whereby caregivers are distracted immediately after making the request, putting the onus on toddlers to chose to explicitly respond or not, and not to be forced to do so as a result of having made eye contact (which due to a lack of inhibitory control, they may not be accomplished at doing) would be particularly illuminating.

## General Discussion

Caregivers report that toddlers as young as 18 months pretend not to hear them when they ask them to do things that the toddler does not want to do. This would suggest that children engage in a deceptive behaviour that demonstrates an understanding of the intentional structure of communication. In Experiments 1 and 2, we attempted to assess whether this behaviour could be observed under controlled experimental conditions. In Experiment 1 children appeared to selectively not respond to negative requests. Experiment 2 sought to control for alternative explanations of this behaviour but resulted in children responding to negative requests at a very high rate. Observations of potential pretending not to hear behaviour (indicated by switching behaviours) in response to these requests were rare in both experiments. In addition, in Experiment 2, toddlers were far more likely to engage in the switching behaviors when there was irrelevant speech, or no child-directed speech at all. As a result it is difficult to conclude whether children this age genuinely pretend not to hear, or selectively avoid ostensively recognizing communicative intentions. It is therefore also unclear whether they understand the intentional structure of human communication.

Paradoxically, perhaps the most convincing evidence that toddlers can pretend not to hear their caregivers comes from cases where they fail convincingly to do so, in cases where they make eye contact with their caregiver. By doing so, they demonstrate that they've heard what their caregiver has said, but frequently did not explicitly respond to them. A small number of children also went back to playing and engaged in the switching behaviours we associated with pretending not to hear. In contrast to full pretending behaviour, this behaviour was not significantly more likely to occur in response to irrelevant speech in the control condition and potentially provides a further avenue for research in tandem with pretending not to hear behaviour where children do *not* make eye contact with their caregivers first. However this behaviour, much like the switching behaviour without eye contact, was infrequent, and could be more likely to be observed in more naturalistic experiments.

On the basis of caregiver report, toddlers from 18 months pretend not to hear requests to do things that they dislike doing, however we were unable to demonstrate this experimentally. While this provides some evidence that toddlers *may* understand the relation between the different layers of intention in communication through their ability to manipulate them, without unambiguous observations of such behaviour under controlled conditions (where it can be verified that children have both *heard* and *understood* the request), it would be premature to claim that toddlers do understand the intentional structure of communication. While toddlers' engagement in this pretence may have indicated that they do understand the intentional structure of communication prior to turning 3, we can further seek to test this hypothesis by examining their responses to the communication of others when such structure is violated, in cases of communication breakdown. This is the focus of Chapter 3.

# 3. Do Toddlers Understand the Intentional Structure of Communication Between Others?

**Abstract**

In this chapter we sought to investigate whether children younger than 3 years could demonstrate an understanding of the relation between levels of intention in communication through their response to communicative breakdowns between others. In Experiment 1, toddlers were third-party observers to a game between two researchers who requested balls from each other. In Experiment 2 they were actively involved in completing a puzzle with one researcher who requested shapes from another researcher. Both experiments involved two types of communicative breakdown between the researchers, one in which communication failed because a researcher did not notice the other's communication, and one where they ignored it. Participant's communicative repair and looking behaviour was coded in response to this communicative breakdown. We did not observe any differences in these measures that would suggest that toddlers are sensitive to the intentional structure of communication. Potential further avenues for research are discussed.

## Introduction

Studies on informative intentions have provided evidence that infants infer what a communicator *means* (i.e., their informative intention) by a signal (be it a gesture or a word) well before their second birthday. For example, children as young as 12 to 14 months can locate hidden objects when adults point to their location (Behne et al., 2005, 2012; Kristin Liebal et al., 2009). Some authors argue that this means that the infants inferred that the signaller's informative intention was to let them know the location of the hidden object (Behne et al., 2012, p. 360). Others suggest that infants make associations between pointing and specific objects of interest, without understanding an intentional relation between pointer and object (Sodian & Thoermer, 2004). It does seem that by 18 months however, toddlers *are* able to comprehend something about how communicative acts convey information in terms of the intentional states of the signaller. 18-month-olds seem to infer that two adults intend different things when pointing at the same objects, depending on the context in which they had previously interacted (Kristin Liebal et al., 2009) and 14- to 20-month-olds seem to infer that two adults intended to ask about different balls even when making the same utterance ('where's the ball?'), depending on the context in which they had previously interacted (Saylor & Ganea, 2007).

Studies on communicative intentions have also demonstrated that toddlers are sensitive to the difference between actions that are produced with or without communicative intent. In accidental versus intentional paradigms, adults perform actions that accidentally inform children of the location of an object (e.g., absent-mindedly pointing), or do so in an ostensive way (using behaviours to signal that they have communicative intent). 14-month-olds performed at chance (in identifying a hidden referent) when such actions were accidental, but not when there were markers of ostension (Behne et al., 2005). More persuasive evidence comes from paradigms where 2- and 3-year-olds can distinguish novel actions (e.g., pressing a button that lights up a location in the room) performed with communicative intent (but without gaze or gesture) from accidental actions (R. Moore et al., 2013, 2015).

While these studies demonstrate that toddlers are sensitive to communicative and informative intentions; they do not show that they understand the relation between the two. Indeed there remains doubt that children under 3-years-old do understand this relation (Tomasello, 2008). In the previous chapter, we suggested that toddlers may understand the relation between communicative and informative intentions much younger than 3, by their engagement in pretending not to hear behaviour. However, tentative conclusions from caregiver report were not substantiated by observations under experimental conditions.

The current chapter seeks to complement Chapter 2 with two novel comprehension experiments designed to determine what toddlers understand about the intentional structure of communication when that communication is between others. Thus in these experiments, children observe communication between two adults as an onlooker. We seek to test whether toddlers are sensitive to the idea that when someone detects another communicating, they should engage with what that person is communicating about. Specifically, to determine this, we are interested in how toddlers respond to two different types of communication breakdown. In one situation, an adult visibly detects another adult communicating, but then does not engage at all with the content of the communication, returning to what they were doing. We refer to this as *ignoring*. In the other, an adult does not detect another adult communicating, and so does not engage at all with the content of the communication, continuing what they are doing. We refer to this simply as *not noticing*.

The two situations represent two different reasons why someone might not respond to the intentions of a signaller. In the first, the recipient visibly detects the signaller's communicative intention, but crucially does not engage with their informative intention (which they could easily have inferred on detection of the communicative intention). If someone understands the relationship between intentions, this would be unusual, indeed uncooperative recipient behaviour: if you detect *that* someone is communicating, you normally engage with *what* they are communicating about. In the second situation, the recipient does not detect the signaller's communicative intention, and does not engage with their informative intention. This is reasonable recipient behaviour: if you don't detect *that* someone is communicating, then you cannot engage with *what* they are communicating about.

If toddlers are sensitive to a difference between the two situations, it suggests that they understand that the cases are different, and provides evidence that they understand something about the structure of intentional communication, i.e., that detection of someone's communicative intention normally leads to engagement with their informative intention.

The purpose of the studies in this chapter is two-fold. Firstly, we seek to determine whether we can demonstrate that toddlers understand communicative breakdowns differently. Secondly, we seek to determine whether this understanding differs across age groups. In Experiment 1, children passively observe two people passing balls back and forth to each other. In Experiment 2, the child is more actively involved in helping one person complete a large floor puzzle while this person requests pieces from another person.

## Experiment 1

In this study, 18- and 24-month-olds observed two researchers rolling balls across a table to each other upon request (by pointing). The researchers were also engaged intently with instructions on a clipboard such that it was plausible they would sometimes miss each other's requests. Thus, in the *Doesn't Notice* condition, one researcher did not detect the other pointing to request a ball, and so did not respond by passing it. In the *Ignore* condition, one researcher detected the other pointing to request a ball (looking up, and then back down) but did not pass it.

Our original intention with this study was to measure the children's repair behaviour upon viewing the two types of communication breakdown. For example, it they selectively attempted to attract the attention of the recipient in the situation where he didn't detect that the signaller was communicating, this would imply that they understood communication had broken down due to non-detection of the signal. To have the strongest possible test of understanding, toddlers were third-party to these communicative interactions (Gräfenhain et al., 2009) and thus had no expectation of material or affective reward if they were to intervene to repair the misunderstandings. Possibly for this reason, no repair was observed in this study. This experiment thus focuses on the secondary measure of children's looking times.

There are several ways in which looking times could plausibly differ in reaction to the two breakdown types. Children might look more to the recipient when they believed that they had transgressed conventions of communication. However, it is equally plausible that they would do so if they were waiting for them to respond. Children might look more to the signaller when they expected them to signal again, but again this could be because the recipient was uncooperative, or because the recipient was unobservant, and looking time would not distinguish between the two.  It could also be that children are more attentive overall in the case of uncooperative behaviour. We therefore did not make make strong directional predictions about where children would look more in this preliminary study. Rather, we were interested to discover whether the manipulation had any effect, so that we could narrow down the possible hypothesis space, allowing future testing to pinpoint how children's reactions to these communicative settings differed.

# Method

## Participants

Fifteen 18-month-olds (7 girls, 8 boys, *M* = 18 months and 10 days, *SD* = 18 days), and thirteen 24-month-olds (6 girls, 7 boys, *M* = 24 months and 6 days, *SD* = 11 days) were included in the final sample.

A further eight children took part, but were excluded due to researcher error (*n* = 2), bilingualism (*n* = 2) not sufficiently engaging with the task (*n* = 1) and during coding for failure to attend to trials (*n* = 3).

## Apparatus and Materials

We used two Sony CX-280 video cameras to record testing sessions. One camera was mounted in the top corner of the room to provide a scene view, while another was hidden behind a curtain on a tripod facing the participant. A small wall (50cm) was constructed out of cardboard boxes, and covered in curtains to separate participants from the researchers playing the communicative game.

The communicative game was played on a table, in front of a white curtain (hiding the participant facing camera). The walls in the testing room were blank. A microphone stand (covered in white masking tape) with a white fabric sling inserted through the curtain allowed a ball to be suspended behind the area where the researchers played, for the purpose of a warm-up exercise.  Materials for the communicative game played by researchers were 7 x 80mm, 220g rubbery balls in red and blue (http://www.oddballs.co.uk/juggle-dream-stage-contact-ball-80mm-p-3133.html), and four rolls of masking tape, which acted as stands for the balls, to hold them still.

All video data was coded in ELAN (Slotjes & Wittenburg, 2008).

## Procedure

Prior to taking part in the experiment, participants played in a warm up room with the researchers for 10-20 minutes. After this, participants were brought into the testing room along with their caregiver. The participant was encouraged to sit on a chair, while the researchers took their place for the experiment on the other side of the dividing wall.

Prior to beginning the habituation trials, the participant was encouraged to point at the ball suspended through the curtain behind the researchers, and point to the location of a box of balls under the table. The purpose of these exercises was to ensure that participant felt comfortable communicating with the researchers in this novel environment, and to introduce them to the room setup as well as confirming that they could observe the researchers interacting, and the materials they were using.

For the habituation trials and the experiment, the two researchers knelt in front of a table facing each other, playing a communicative game where they requested balls from one another by pointing (Figure 5). A standard trial consisted of a start position whereby both researchers were looking away from each other and in the opposite direction to the toddler, looking at clipboards, with one hand on the clipboard, as if reading off of it. One researcher (the signaller) would then look towards the other researcher (the recipient), whilst pointing at a ball. The recipient would look up from their clipboard, say, ‘Oh!’, and roll the ball towards the signaller. On receipt of the ball, the signaller would say, ‘Yes!’, and both researchers would return to the start position ready for the next trial. Trials operated in both directions, with both researchers acting as signallers or recipients. At no point did the interaction involve the participant, the participant was third-party to the game, able to see what was going on, but behind the dividing wall not allowing them access to the game itself.

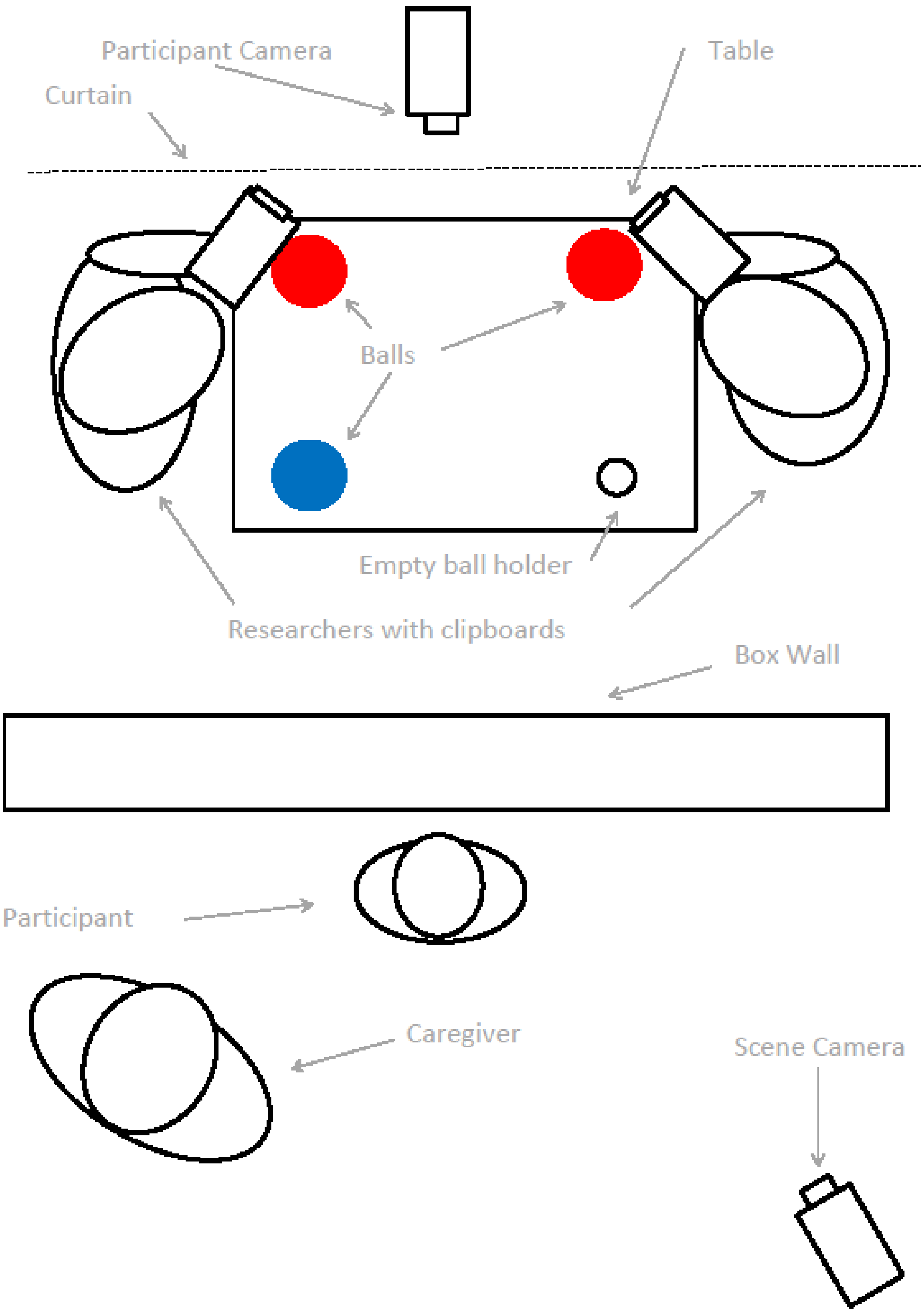


*Figure 5.* Room set up for comprehension task

The familiarisation phase consisted of six standard trials. After these trials, there was a break for a habituation exercise. This involved each researcher requesting a ball that was hidden on the participant's side of the wall. The participant was encouraged to pass the ball to the researcher that had requested it to ensure that participant felt comfortable interacting with the researchers.

The experimental phase of the study consisted of fourteen trials, both standard, and two types of experimental trial (Table 9). After each experimental trial, there would be two standard trials. Each type of experimental trial was repeated twice, meaning that there were four experimental trials in total, separated by two standard trials, with four standard trials prior to the first experimental trial.

Table 9

*Order of trials for Experiment 1*

| *Trial Number* | *Trial Type* |
|---|---|
| 1-6 | Standard |
| … | *Break* |
| 7-10 | Standard |
| 11 | Experimental A |
| 12-13 | Standard |
| 14 | Experimental B |
| 15-16 | Standard |
| 17 | Experimental A |
| 18-19 | Standard |
| 20 | Experimental B |

The two types of experimental trial were the *Doesn't Notice* trial, and the *Ignore* trial. In the Doesn't Notice trial, the signaller pointed to the ball in front of the recipient, but the recipient did not look up, so they did not pass the ball. In the Ignore trial, the signaller pointed to the ball in front of the recipient, the recipient looked towards the signaller, but then looked back towards their clipboard, without passing the ball.

In both trials, there was a response period of 7s in which the signaller maintained pointing and the recipient either looked or did not look. In both cases, after 7s, the situation was resolved when the signaller shouted ‘Hey!’, causing the recipient to look up, say ‘Oh!’, and pass the ball as in the standard trials (the signaller also responding by saying, ‘Yes!’ as normal).

The order in which participants saw the Doesn’t Notice or Ignore trials was counterbalanced across participants, as were the researchers playing the roles. For all but two participants, the two researchers were both male. For the remaining two participants one of the researchers was female.

After the trials were completed, we performed a preference test, to see if the participants had a preference for either researcher. The researchers moved to the wall and at the same time held up stickers, while caregivers read out a prompt to the child, saying, ‘They’ve got a sticker for you! Can you ask them for a sticker?’. The researchers maintained a neutral expression, kneeling equidistantly from the participant holding up identical stickers, and participants were allowed to freely choose. After 10s, if participants had not requested a sticker, one of the researchers would give them a sticker, and this participant was not included for this part of the analysis. Since few children responded to this test without substantial encouragement that lead to loss of experimental control, the results of this test are not reported.

**Coding**

Response windows to the two experimental trials were determined from the moment that the signaller’s arm reached full extension to either a maximum of 7 seconds later, or to the moment where they said ‘hey’ to resolve the trial (in cases where this happened slightly earlier).

We had two presentations of both experimental types, but only included one of each for the analysis. By default we took our measures from the first presentation of each of the two experimental trials, and would only take the measures from the second presentation if the first was invalid.

During the Ignore trials, it was absolutely crucial that participants had attended to the experimental manipulation (that the recipient had looked at the signaller and then looked away). Therefore, if any participants did not see all or part of the head turns (i.e., were not looking towards the appropriate side of the scene at any point during the head turn) during the first presentation of the Ignore trial, this trial was excluded. If, and only if, the first trial was excluded, we coded the second trial. If participants did not look at the head turn in either presentation of the Ignore trial, they were excluded (see Participants).

If participants looked at the researchers or their caregivers for less than 1 second in the 7 second response window, that trial was also excluded. One participant did not attend for more than 1 second in either Ignore trial, and so was excluded (see Participants).

For coding eye-movements, we used a camera focused on the participant and positioned at their eye level, which was broadly in line with the chests of the researchers. We divided the screen into four portions (halving it vertically and horizontally). Looks by the participant to the edge of the top right or top left portions were judged to be looks to faces of a researcher (on the child's right or left respectively). Looks to the caregiver (normally sat behind the participant) were coded when the participant clearly turned towards their caregiver. For every trial, we thus had three targets: signaller, recipient and caregiver. We coded whether or not participants looked to the signaller, recipient or their caregiver as a binary measure. We also coded both the *duration* and *frequency* of looks at the signaller, the recipient and the participant's caregiver.

We also intended to code participant's communicative repair behaviour, however no instances of repair behaviour occurred.

## Results

On experimental trials, where the recipient either didn't see or ignored the signaller's gestures, participants frequently looked to the recipient and the signaller, and less frequently looked to their caregiver. Table 10 shows the percentage of participants who looked to the signaller, recipient and caregiver as a function of age and condition.

Table 10

*Percentage of Participants Looking At The Signaller, Recipient and Caregiver as a Function of Age and Condition*

| | | *Condition* | |
|---|---|---|---|
| *Age* | *Looked to* | *Doesn't Notice* | *Ignore* |
| | *Signaller* | 100% | 100% |
| 18-month-olds (*n* = 15) | *Recipient* | 82% | 94% |
| | *Caregiver* | 29% | 29% |
| | *Signaller* | 85% | 100% |
| 24-month-olds (*n* = 13) | *Recipient* | 92% | 69% |
| | *Caregiver* | 23% | 31% |

We fitted binary logistic regression models to analyse whether participants' looking to the signaller, recipient or their caregiver (outcome variable: participant looked = 1, did not look = 0) differed as a function of condition or age. We fitted a separate regression model for each target (signaller, recipient and caregiver). Participant was included as a random effect for all models (*N* = 28). For all targets, no predictors (either age group or condition) improved on the null model (with no predictors), therefore participants looking to the separate targets did not differ as a function of age or condition.

Figure 6 shows the median frequency of looks to the signaller, recipient and caregiver as a function of age and condition.

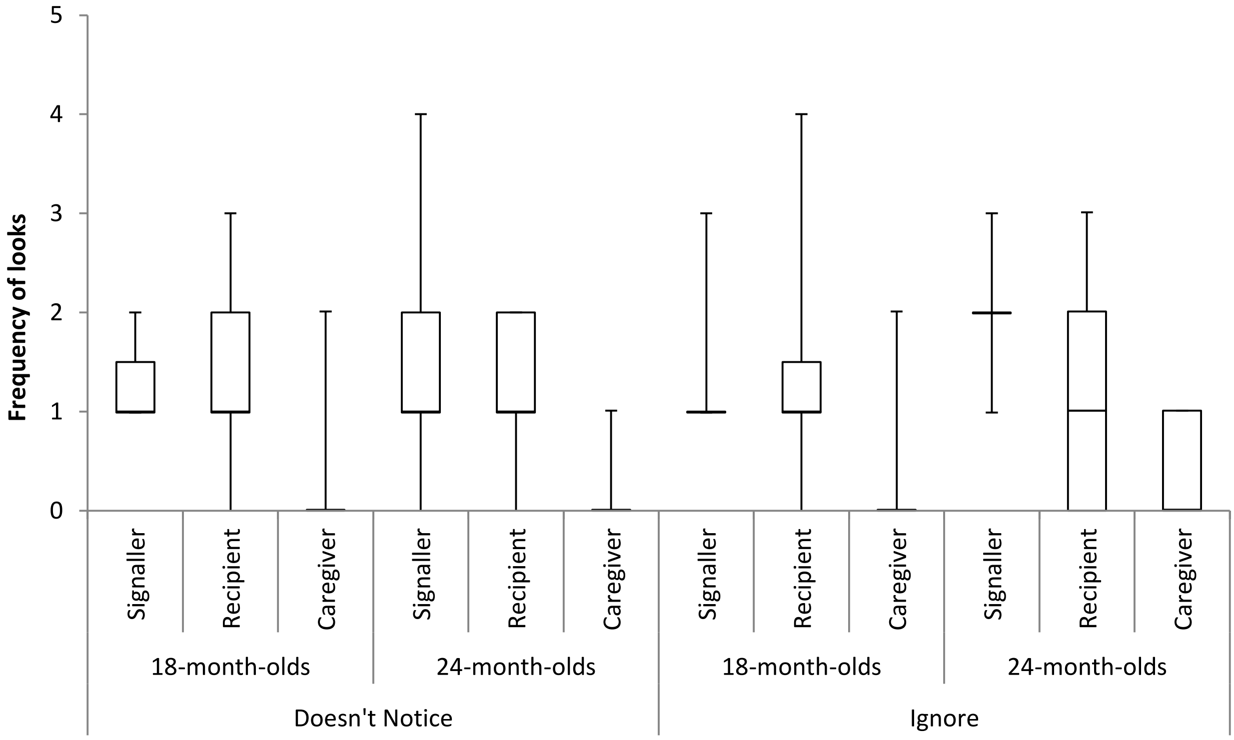


*Figure 6.* Frequency of looks towards the signaller, recipient and caregiver as a function of age and condition. Boxes represent inter-quartile range about the median and error bars represent minimum/maximum frequency of looks

Separate Wilcoxon's signed-rank tests for each age group, revealed no significant difference in frequency of looks to *signaller*, *recipient* or *caregiver* as a function of condition.

Wilcoxon's rank-sum test revealed that in the Ignore condition, 24-month-olds (*Mdn* = 2) looked to the signaller significantly more frequently than the 18-month-olds (*Mdn* = 1), $W = 35.5$, $p = .002$, $r = .57$. There was no significant difference in the Doesn't Notice condition. Wilcoxon's rank-sum tests did not reveal any significant difference in frequency of looks to *recipient* as a function of age in either condition, nor in frequency of looks to *caregiver* as a function of age in either condition.

We also analysed the total number of looks per child to all other people in the room in order to test whether gaze alternating was more likely in either condition. Wilcoxon's signed-rank test for each age group revealed no significant difference in looks a function of condition, while Wilcoxon's rank-sum test revealed no difference as a function of age in either condition.

Figure 7 shows the mean duration of looks to the signaller, recipient and caregiver as a function of age and condition.

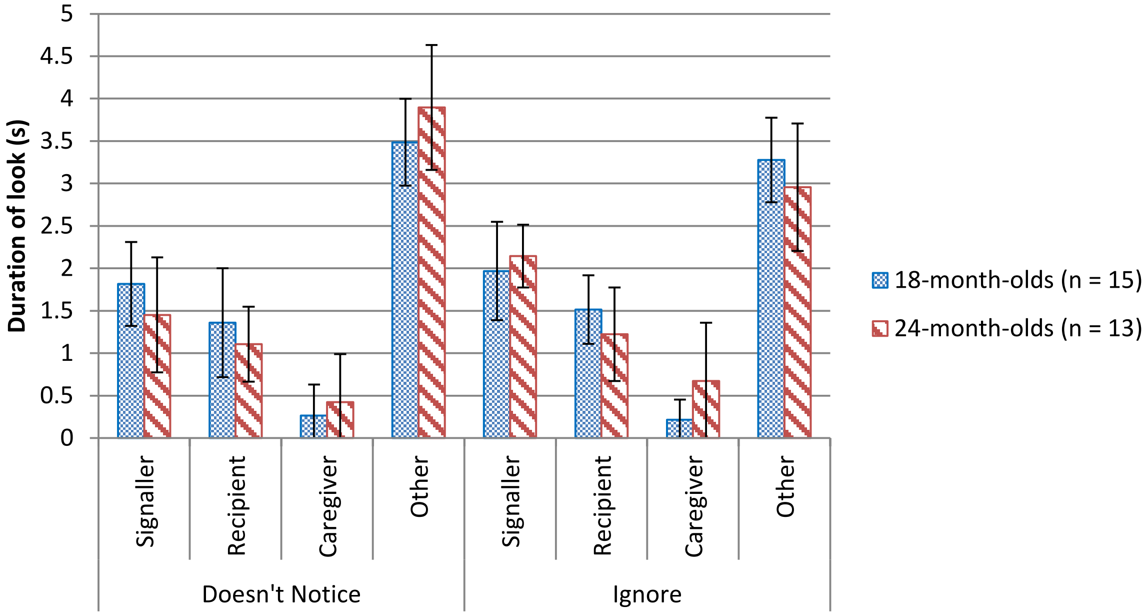


*Figure 7.* Mean duration of looks to the signaller, recipient caregiver and other as a function of age and condition. Error bars represent 95% confidence intervals about the mean

Since looking times to each target were not independent of each other, we focused statistical analysis on looks to signaller and the recipient. We thus calculated the proportion of these looks that were to the signaller or to the recipient (denominator = looks to signaller + looks to recipient + looks to caregiver + looks to other). A 2 (age) x 2 (condition) ANOVA demonstrated that there was no significant effect of age or condition nor any significant interaction for proportion of time spent looking to the signaller.

## Discussion

While the children did not engage in any repair behaviour, they attended to the signaller and the recipient during these games and we were thus able to analyse their looking times. Looking behaviour was largely stable across ages and conditions. Children generally looked once each to the recipient and the signaller. The only case in which this differed as a function of condition was for the 24-month-olds in the ignoring condition who tended to look twice to the signaller. Generally children spent slightly more time looking to the signaller than the recipient but there were no statistically reliable differences in looking times.

The significant difference in frequency of looks suggests that the older children differentiate the two types of communicative breakdown and pay special attention to the case of ignoring.  This pattern of results is repeated with the looking duration measure but is not statistically significant. A larger sample would be needed to have any confidence in this conclusion. Further study is also needed to explore why this difference occurs. Previous research involving 14-month-olds passively watching two researchers point to request objects suggested that longer looking times indicate an expectation of an appropriate response (Thorgrimsson, Fawcett, & Liszkowski, 2014). In line with this, we hypothesise that 24-month-olds looked to the signaller more because they expected them to act in order to resolve the problem of an uncooperative partner.

Any further study would also ideally elicit repair behaviour as differences in who this was directed at and how it was achieved would provide stronger evidence for an ability to tell the two breakdowns apart. In previous studies of children's repair where they experience communication breakdown first person, only around half of the participants engaged in repair behaviour (Grosse et al., 2010). In the current study, since the children were not involved in the passing game, and indeed were separated from it by barrier, they most likely didn't feel it was their place to intervene, or if they did they were not motivated to.  While 18- to 24-month-olds engage in communicative repair, when their own requests are misunderstood or seemingly not heard, repeating and elaborating their communicative efforts (Marcos, 1991), it is perfectly plausible that  they would not do so in third-party interactions. The aim of having children in a third-party role was to avoid the problem of rewarding toddlers for specific behaviours (thus potentially confounding attempts to evaluate their understanding of communication). However, this appears to have been too ambitious, and thus  future variants on this design should seek to involve them in the game, and motivate them to intervene in cases of communication breakdown.

In addition, by attempting to obtain measures of where children were looking, a compromise was made to an imperfect setup for coding looks to either signaller or researcher (coding looks to either side of the screen, though fairly robust could potentially have led to inaccuracies). Given these criticisms, a new study was designed with improved camera angles and with a set up that included the child more, to encourage repair behaviours.

## Experiment 2

The aim of Experiment 2 was to provide better evidence that toddlers are sensitive to the intentional structure of communication. Specifically, we sought to increase child involvement in the paradigm, as we hypothesised that this would motivate them to engage in repair. In this study, toddlers were involved in a making a giant floor puzzle with one of the researchers who then requested pieces from another. In addition, we decided to focus analysis of looking behaviour of participants towards the signaller only. This was because in Experiment 1, participants looked much more to signaller than recipient or caregiver, and this was the measure where we saw most interesting variance across age groups and condition. Therefore we made the setup as conducive to collecting this measure as possible (i.e., having both participant and signaller in the same camera angle, to allow for more accurate coding of looking behaviour).

We aimed to test whether there were differences in children's understanding (as revealed by their repair or looking behaviour) as a function of condition and age group. Our working hypothesis, from the findings in Experiment 1, was that 24-month-olds are more likely to expect the signaller to act next in cases where the recipient is ignoring them, than in cases where the recipient fails to notice them signalling, but that 18-month-olds are not (as they looked equally to the signaller in both conditions). Therefore, we predicted that 24- to 26-, and 30- to 32-month-olds are more likely to look to the signaller, and will look more frequently and longer to the signaller in cases of communicative breakdown where the recipient has ignored them, than when they don't notice them. We did not expect 18- to 20-month-olds to look more frequently or longer to the signaller in either case. As we did not observe any repair behaviour in Experiment 1, it was harder to generate a strong prediction for this experiment. However, following the hypothesis that 24-month-olds may have expected signallers to act in the *Ignore* condition more than in the *Doesn't Notice* condition, we predicted that 24- to 26- and 30- to 32-

month-olds are more likely to repair in the *Ignore* condition when they are effectively working together with the signaller.

## Method

### Participants

Twenty-three 18- to 20-month-olds (9 girls, 14 boys, *M* = 18 months and 28 days, *SD* = 29 days), twenty-four 24- to 26-month-olds (12 girls, 12 boys, *M* = 25 months and 10 days, *SD* = 25 days) and twenty-four 30- to 32-month-olds (11 girls, 13 boys, *M* = 30 months and 26 days, *SD* = 28 days) were included in the final sample.

A further thirteen children took part, but were excluded for caregiver interference (*n* = 3), fussiness (*n* = 4), sibling interference (*n* = 2), during coding for not attending to the trials (*n* = 3), and for bilingualism (*n* = 1).

### Apparatus and Materials

We used two Sony CX-280 video cameras to record testing sessions. One camera was mounted in the top corner of the room to provide a scene view, while another was hidden in a fabric covered box on a tripod facing the participant.

Materials for the game were 3 sets of 4 foam tiles, each with 4 animal cutout shapes that could be slotted in (http://www.softfloorkids.co.uk/jungle-mats-4pack.shtml). For habituation we used a set with dinosaur shapes and one with animals (lion, monkey, elephant and giraffe). For the experimental trials we used an animal set with the same animals, but in different colours. Each set of foam tiles was 120 x 120 cm approximately, and were placed on the floor during the game, or stacked against the wall when not in use.

For part of the experiment, one researcher sat behind a plain custom-made wooden bench (50cm high), that separated off a portion of the room.

All data was coded in ELAN (Slotjes & Wittenburg, 2008).

**Procedure**

Prior to taking part in this experiment, participants had played in a warm up room with the researchers for 10-20 minutes. Participants were brought into the room with their caregiver and at the same time as the researchers.

As participants entered the room, there was a foam floor jigsaw on the floor with 4 spaces for dinosaur cut-out shapes. Three dinosaur shapes that were designed to be slotted into the jigsaw were scattered around on top of the mat. The final dinosaur shape, and four animal shapes (two monkeys, two giraffes) were on a shelf out of reach of the toddler and above the chair where Researcher 2 would be reading later in the experiment. Four more shapes were hidden in the room (two elephants behind the wall, and two lions under chairs). The space where Researcher 2 would be reading (underneath the shelf) was behind a wall that participants could not easily get over.

The researchers played with the participant and encouraged them to slot the three dinosaur shapes into the foam tiles. After the first dinosaur shapes were filled in, there was the first habituation trial to demonstrate that Researcher 2 could be recruited to get shapes from the shelf. For this first habituation trial, Researcher 2 was with the participant and Researcher 1 in the main part of the room with the foam tiles. During this trial, the participant would be encouraged to look at the final dinosaur on the shelf (if they had not noticed it already), and Researcher 1 would ask Researcher 2 to retrieve it. Researcher 2 would climb over the wall to the shelf to retrieve the final dinosaur. After this habituation trial, Researcher 2 sat reading a book underneath the shelf while Researcher 1 and the participant continued with the game. This was the setup for all other trials (Figure 8).

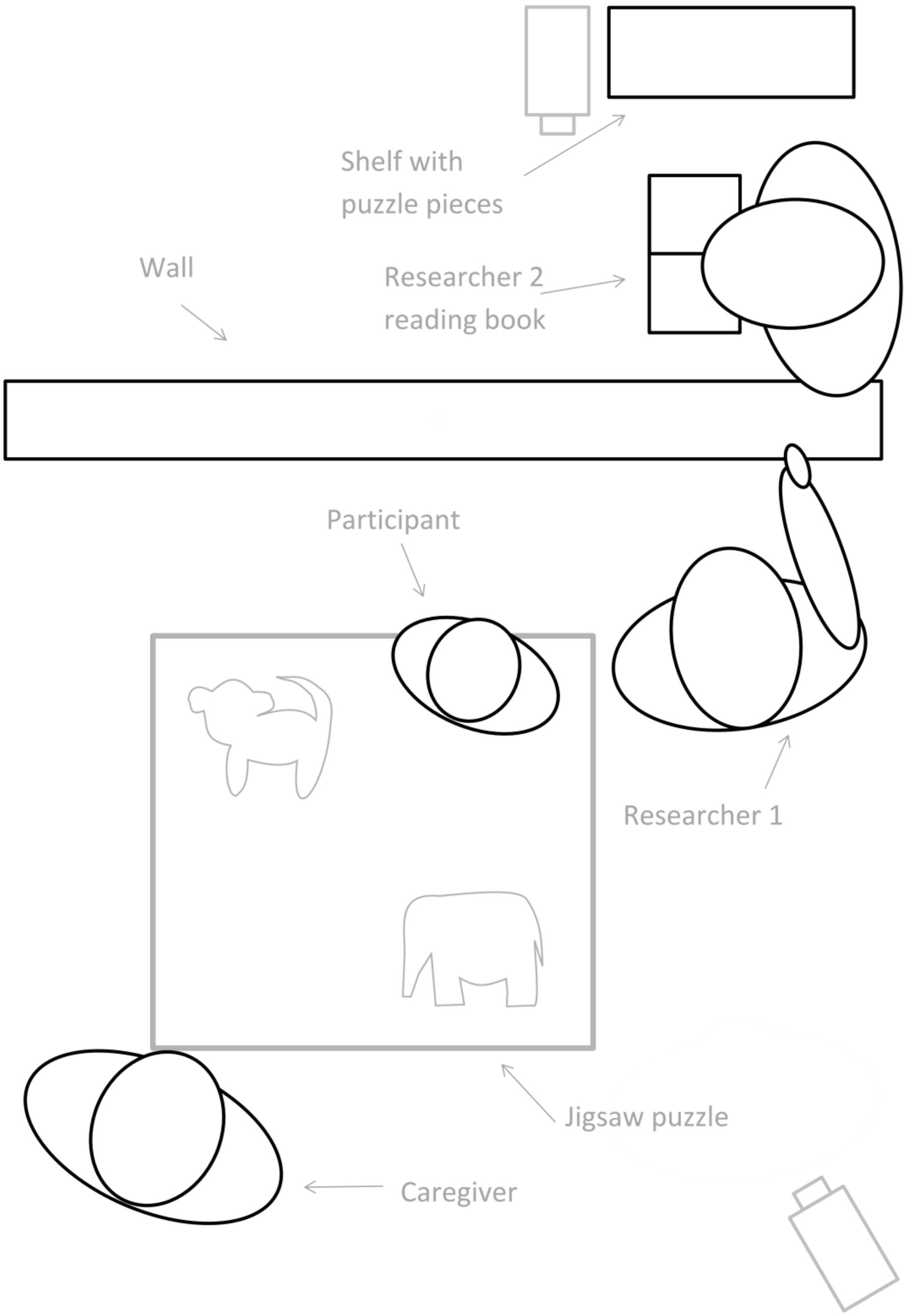


*Figure 8.* The layout of the room after warm up and the first habituation trial – for the other habituation trials, and experimental trials

Following this, there were two more habituation trials, and two experimental trials, interspersed with four hidden piece trials, which served as a break between experimental and habituation trials. The order of trials is set out in Table 11.

Table 11

*Order of trials for Experiment 2*

| *Trial Type* |
|---|
| *Warm-up (Three Dinosaurs)* |
| Habituation Trial (1) (Final Dinosaur) |
| *Hidden Piece* |
| Habituation Trial (2) |
| *Hidden Piece* |
| Habituation Trial (2) |
| *Hidden Piece* |
| Doesn't Notice / Ignore Trial |
| *Hidden Piece* |
| Ignore / Doesn't Notice Trial |

**Habituation trial (1).** After determining that the participant had seen the dinosaur on the shelf, Researcher 1 said, '{Researcher 2}, can you get me the {shape name} from the shelf?' whilst pointing to the shelf, and looking towards Researcher 2. Researcher 2 climbed over the wall, reached up to the shelf, brought down a piece and passed it to the participant. Then Researcher 2 sat down under the shelf and began to read a book.

**Habituation trial (2).** Researcher 1 said, 'Oh, I know where another shape is. {Researcher 2}, can you get me the {shape name} from the shelf?' whilst pointing to the shelf, and looking towards Researcher 2. Researcher 2 looked up from the book, said, 'Oh!', and reached up to the shelf, bringing down a piece and passing it to the participant. The participant was encouraged to put it in the correct tile.

**Hidden Piece Trial.** Researcher 1 indicates the location of a hidden piece to the participant (either an elephant hidden behind the wall, or a lion hidden under a caregiver's chair), and they put that into place.

**Doesn't Notice trial.** Researcher 1 said, '{Researcher 2}, can you get me the {shape name} from the shelf?' whilst pointing to the shelf, and looking towards Researcher 2. Researcher 2 did not look up and carried on reading. After 5 seconds, Researcher 1 said 'hmm', and stopped pointing. After a further 5 seconds Researcher 1 said, 'Nevermind, I know where another one is' and moved on to the hidden piece trial. Alternatively if this was the final experimental trial, Researcher 1 would prompt Researcher 2 for the piece(s) a second time, to which Researcher 2 would respond.

**Ignore trial.** Researcher 1 said, '{Researcher 2}, can you get me the {shape name} from the shelf?' whilst pointing to the shelf, and looking towards Researcher 2. Researcher 2 briefly looked up (and ensured that the participant had seen her do so), but then looked down and carried on reading. After 5 seconds, Researcher 1 said 'hmm', and stopped pointing. After a further 5 seconds Researcher 1 said, 'Nevermind, I know where another one is' and moved on to the hidden piece trial. Alternatively, if this was the final experimental trial, Researcher 1 would prompt Researcher 2 for the piece(s) a second time, to which Researcher 2 would respond.

For the two experimental trials, and the habituation trials, a specific spatial setup was achieved to make it as likely as possible that all age groups understood the interaction (see Figure 8). Researcher 1 would tell the participant that the next shape was up on the shelf, and make sure they had understood before beginning the trial (by pointing and requesting the shape from Researcher 2). Equally, the test was designed in such a way as there would be no other option but to request a shape from the shelf (as all other shapes were either already in the puzzle or hidden from the participant). Researcher 1 pointed and made the request to Researcher 2, with the participant to their left. The aim was that the participant would always see Researcher 1 pointing, Researcher 2's reaction and the shelf in the same direction, to reduce the comprehension demand of the trials.

**Coding**

Videos from the two camera angles were synchronised and coded using ELAN. A response window was set for each participant in each experimental trial, and all coding was restricted to these windows. A window began from the frame of the video when Researcher 1 began to make the request to Researcher 2 (the verbal part of the request). It lasted until 10 seconds had elapsed or until Researcher 1 turned back to the participant to carry on with the game (if this happened slightly earlier). Therefore the maximum a trial could last was 10 seconds.

For Ignore trials, it was crucial that participants had seen the manipulation. Participants were excluded during coding when they were judged not to have been looking when Researcher 2 looked up in the Ignore condition ($n$ = 3).

**Repair behaviour.** In the literature, communicative repair behaviour has a number of key characteristics. It is a communicative act e.g., a verbal utterance, vocalisation, pointing, reaching or showing (Golinkoff, 1986; Grosse et al., 2010). It is an expression of dissatisfaction (not confirmation that someone has done the right thing) i.e., it is clearly not affirmative (Grosse et al., 2010) and can involve repetition, augmentation or substitution of the signal (Golinkoff, 1986). It can redirect attention to establish shared reference (Liszkowski et al., 2007) or be directed at a person; either looking towards that person or contingently responding to their communication (Grosse et al., 2010).

We coded repair behaviour as any communicative act (vocalisation, gesture or banging) that did not express some sort of affirmative response, but seemed to be imperative, or norm enforcing (expressing dissatisfaction with the current state of affairs). This could either be directed at Researcher 2 (the recipient) to attract their attention, or towards the object to attempt to redirect Researcher 2's attention to it. If the participant imitated all or part of the communicative act that Researcher 1 (the signaller) had used already – this counted as repair only if it followed Researcher 1's attempt. Behaviour was not repair behaviour if it coincided with Researcher 1's attempt. If the participant selected a different type of communicative behaviour, or elaborated on the original message to convey the same message as Researcher 1, this was repair behaviour. Additionally if they repeated or elaborated (selecting a different type of communicative behaviour, or elaborating on their original attempt using the same type of behaviour) their own signal, this counted as repair, even if it was non-contingent on Researcher 1's communicative attempts (but it still had to follow Researcher 1's attempt).

A participant could engage in more than one attempt at repair behaviour in the same trial (repair was considered separate attempts when there was a clear break between efforts), and the behaviour type of each instance was coded as to whether it involved vocalisation, gesture or banging. For each participant, they had a binary outcome for the response window on four categories – whether they repaired or not, and whether their repair involved gesture, vocalisation or banging.

**Looks to the signaller.** We coded whether or not participants looked to the signaller as a binary measure, and both the look *duration* and *frequency*. In cases where the participant was already looking at the signaller at the beginning of the trial, looks to the signaller were not counted (for the binary outcome or for duration of frequency) until after the participant looked up in the direction of the recipient. If participants did not look in the direction of the recipient during the response window, they would not have been included in the analysis. However this did not occur: all participants looked in the direction of the recipient.

## Results

### Repair Behaviour

A number of participants engaged in communicative repair during the response windows in experimental trials. The overwhelming majority of repairs (around 75%) involved pointing to the shape, or saying the shape name. Table 12 shows the percentage of participants that engaged in repair behaviour as a function of condition and age.

Table 12

*Percentage of Participants Engaging in Repair as a Function of Condition and Age*

| *Age* | *Doesn't Notice* | *Ignore* |
|---|---|---|
| 18- to 20-month-olds (*n* = 23) | 26% | 26% |
| 24- to 26-month olds (*n* = 24) | 25% | 25% |
| 30- to 32-month-olds (*n* = 24) | 25% | 25% |

We fitted binary logistic regression models to analyse whether the participants engaging in repair (outcome variable: participant repaired = 1, no repair = 0) differed as a function of condition or age. Models with either age or condition (or both) as predictors did not improve on a null model that contained only participant (*N* = 71) as a random effect and no predictors. Therefore repair behaviour did not differ as a function of either condition or age group.

Participants' repair behaviour involved vocalisations, gestures and banging and on occasion combined these. Table 13 shows the percentage of participants engaging in repair using vocalisations, gestures or banging as a function of age and condition.

Table 13

*Percentage of Participants Engaging in Repair of Each Type as a Function of Age and Condition*

| *Age* | *Type* | *Doesn't Notice* | *Ignore* |
|---|---|---|---|
| 18- to 20-month-olds | *No repair* | 74% | 74% |
| | *Vocal* | 22% | 26% |
| | *Gesture* | 26% | 22% |
| | *Banging* | 4% | 0% |
| 24- to 26-month-olds | *No repair* | 75% | 75% |
| | *Vocal* | 25% | 25% |
| | *Gesture* | 17% | 17% |
| | *Banging* | 4% | 4% |
| 30- to 32-month-olds | *No repair* | 75% | 75% |
| | *Vocal* | 21% | 21% |
| | *Gesture* | 8% | 13% |
| | *Banging* | 4% | 4% |

We fitted binary logistic regression models to the data to analyse whether the type of behaviour participants used to repair varied as a function of age and condition. We conducted separate regression models for each type of behaviour (vocalisations, gestures and banging), with a binary outcome variable (e.g., for vocalisation models, 1 = repair involved vocalisation, 0 = repair did not involve vocalisation/no repair). Participant ($N$ = 71) was included as a random effect for all models. For all types of behaviour, predictors (either condition or age group) did not improve on the null models (with no predictors), therefore the likelihood that participants would vocalise, gesture, or bang did not vary as a function of age and condition.

**Looking to the Signaller**

Table 14 shows the percentage of participants that looked to the signaller as a function of age and condition.

Table 14

*Percentage of Participants Looking to the Signaller as a Function of Age and Condition*

| *Age* | *Doesn't Notice* | *Ignore* |
|---|---|---|
| 18- to 20-month-olds (*n* = 23) | 83% | 91% |
| 24- to 26-month-olds (*n* = 24) | 71% | 75% |
| 30- to 32-month-olds (*n* = 24) | 79% | 83% |

We fitted binary logistic regression models to analyse whether participant looking to the signaller (outcome variable: participant looked = 1, did not look = 0) differed as a function of condition or age. Models with either age or condition (or both) as predictors did not improve on a null model with participant (*N* = 71) as a random effect and no predictors. Therefore looking behaviour did not differ as a function of condition or age.

Figure 9 shows the frequency of looks participants made to the signaller as a function of condition and age.

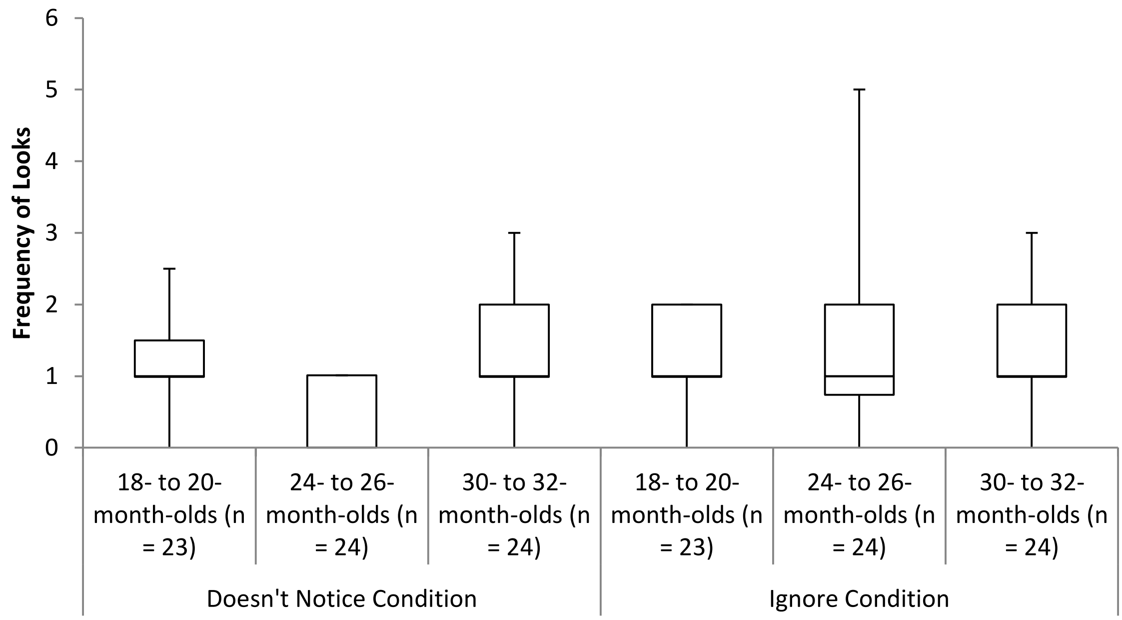


*Figure 9.* Frequency of looks to the signaller as a function of condition and age. Boxes represent inter-quartile range, error bars represent maximum and minimum frequency of looks to signaller

Separate Wilcoxon signed-rank tests revealed no significant difference in the frequency of looks to the signaller for any age group as a function of condition. Separate Kruskal-wallis tests revealed no difference on the frequency of looks in either condition as a function of age.

Figure 10 shows how long participants looked towards the signaller as a function of condition and age.

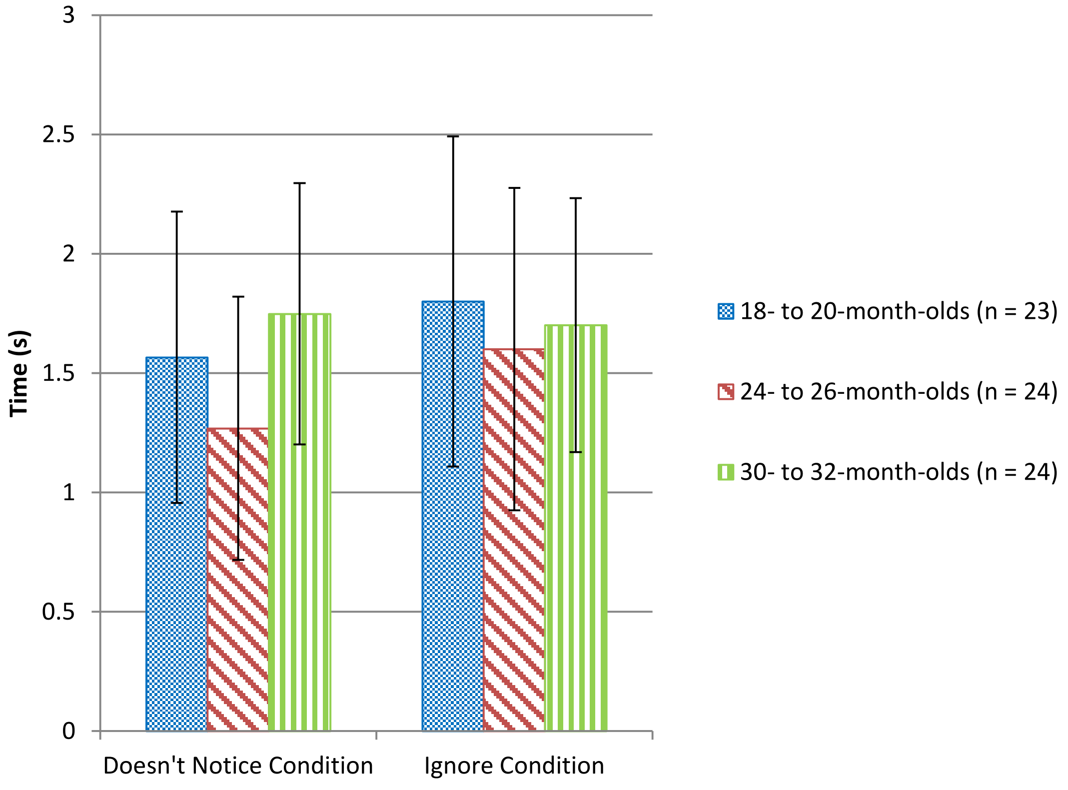


*Figure 10.* Mean duration of looks to the signaller as a function of condition and age. Error bars represent the 95% confidence intervals about the mean

The proportion of the time spent looking at the signaller was calculated for each participant. A 3 (age) x 2 (condition) ANOVA found no significant effects of age on the proportion of time participants spent looking at the signaller.

**Length of Response Windows**

As response windows were determined by Researcher 1 counting to 10 seconds in their head, there was some variation in their exact length. Out of 144 trials (72 Doesn't Notice, 72

Ignore), 35 (49.6%) Doesn't Notice trials and 43 (59.7%) Ignore trials lasted to the cut-off point of 10s. The average length of the Doesn't Notice trials was 9.5 s (SD = 0.74), and the average length of the Ignore trials was 9.7s (SD = 0.47).

## Discussion

Around 25% of toddlers engaged in communicative repair and around 80% looked to the signaller during the response windows in experimental trials. Our predictions were not supported by the data, the two older age groups were not more likely to look to the signaller, and did not look longer or more frequently at the signaller in the *Ignore* condition than in the *Doesn't Notice* condition. In addition, the two older groups were not more likely to engage in repair behaviour in the *Ignore* condition than in the *Doesn't Notice* condition.

There are three possible explanations of these null findings. Firstly, Tomasello (2008) argues that it is unlikely that children do understand the intentional structure of communication prior to 3-years-old. The current findings are in line with this hypothesis.

A second explanation is that while participants understand the intentional structure of communication from a first person perspective, they do not apply this knowledge to make sense of third-party interactions. Other studies have found a developmental gap between, for example, following pointing gestures in first-person contexts, which infants can follow at 12 to 14 months (Behne et al., 2005), and in third-party contexts (Gräfenhain et al., 2009). So it is plausible that a similar developmental lag exists here. We could conclude that while toddlers are capable of understanding and manipulating intentional structure in first-party contexts, they do not understand intentional structure holistically in the communication of others, and so do not recognize when this is manipulated. Future research should look at whether toddlers understand the difference between the two situations in first-party contexts (i.e., by having researchers ignore or not notice requests from toddlers).

A third possibility is that children do have some understanding of the difference between the communicative breakdowns but this methodology was not able to detect them. For example, it is not clear whether, in the *Doesn't Notice* condition, children understand that the experimenter hasn't noticed, or in the *Ignore* condition that they had noticed but were being uncooperative. There are several ways to fix these methodological problems including making the current conditions more distinctive. For example, if children can demonstrate auditory perspective-taking behaviour when observing a researcher wearing headphones, the two cases (not noticing, and ignoring) could be distinguished by Researcher 2 wearing headphones, and not by Researcher 2 looking up. Equally, the motivations of Researcher 2 *to* ignore communication could be made clearer, for example conveying to participants that they were very busy and not happy to be disturbed (although this would possibly make them difficult to approach). Additionally, it is possible that by actively involving children in the game, this means that the subtleties of the communicative breakdown may be lost, as children's primary goal is to get hold of the shape. The third-party setup from Experiment 1 may be preferable in order to measure children's understanding of the situation in strictly communication terms in terms of their looking behaviours.

However, with hindsight, we have a more serious concern with this experimental design. Similar patterns of behaviour across conditions might underpinned by different understanding of the situation, and different patterns of behaviour might not necessarily rely on a different social-cognitive understanding of those situations. Although we predicted that older children would be more likely to repair in the *Ignore* condition than the *Doesn't Notice* condition, there are equally plausible reasons to think that they would be more likely to repair and look in the *Doesn't Notice* condition than the *Ignore* condition, or in both, if they understood the intentional structure of communication. For example, it is perfectly plausible that participants would repair in the *Doesn't Notice* condition because the recipient hasn't detected the communication, which is easy to repair. In contrast in the *Ignore* condition, the recipient has detected the communication and appears to being uncooperative so the repair is not as straightforward for a toddler. Equally, participants might repair in both, but for different reasons (in one because they think that the researcher hasn't noticed, but in the other because they have understood that they are being uncooperative). The same criticisms could easily be levelled at data on looking behaviour. The lack of well-motivated directional predictions as a result of children's understanding of intentional structure was a major weakness of our design. This study was in essence exploratory: a significant tendency for older children to intervene in one condition but not the other, whilst younger children were equally likely to repair or not in both, might have given us good reason to conclude that older children understand the situation differently, and would have provided fruitful observations with which to base future predictions on. Given that there was no such significant tendency (we suspect due to a plethora of underlying motivations governing children's communicative repair) we are sceptical that this experimental design could ultimately provide conclusive insight into children's understanding of intentional communication in its current form.

## General Discussion

In this chapter, we sought to determine whether toddlers would distinguish two situations involving communication breakdown that differed subtly according to whether or not a recipient had apparently detected a communicative intention. In Experiment 1, children did not produce repairs, however there was some tentative evidence from looking times that the older children looked more to the signaller in the ignoring cases. However, a larger sample size was needed to have confidence in this. In Experiment 2 we had a larger sample size and a set up that was more conducive to child repair but in this study neither repair behaviour nor looking times differed as a function of condition.

Combined with the lack of substantial evidence that toddlers pretended not to hear in Chapter 2, a reasonable conclusion is that we have not demonstrated that toddlers understand the relation between the intentions in communication prior to 3 years of age. This is in line with the literature (Tomasello, 2008). While we have created a novel method that could potentially shed light on children's understanding of the intentional structure of communication, it is clear that it needs substantial work in order to show beyond doubt whether toddlers understand the intentional structure of communication or not.

The challenge is that an experimental study would need to maintain ecological validity across a wide age group (from as young as 12 to 15 months, right up until 4 years) while using a manipulation that would be understandable to all ages. There is almost certainly a lower limit on such a study, as there is doubt that 12-month-olds can take the visual perspective of others (Gaunet & Massioui, 2014), however they are seemingly able to understand give-and-take interactions even when they are third-party to them (Elsner, Bakker, Rohlfing, & Gredeback, 2014). There would also have to be a reasonable expectation that children could engage in behaviour that would indicate their understanding at all ages, and there would have to be clear directional predictions as to what sort of looking behaviour would imply that they understood the situation in the way that indicated they understood the intentional structure of communication. We argue that a design where participants *passively* observe the communicative breakdowns (such as Experiment 1) would yield more meaningful data on their understanding of the communicative interaction in terms of intentional structure, as a set-up more conducive to repair (Experiment 2) could distract children from the subtlety of the manipulation. However, of greater concern, we also do not have clear directional predictions of what would entail an understanding of the intentional structure of communication, and were unable to generate these in order to pursue this line of enquiry further.

The communicative breakdowns of the type that we used for this experiment represent subtle ways in which the intentional structure of communication can be manipulated, which could provide evidence of toddlers' understanding of the relation between communicative and informative intentions. However, they are undoubtedly rare in the normal communicative environment of young children, compared to occasions whereby caregivers ensure communication *doesn't* break down, by semantically and temporally responding to children's communicative efforts (McGillion et al., 2013), and engaging in ostensive child-directed speech (Csibra, 2010). A more fruitful avenue of research therefore maybe in more typically occurring contexts, whereby early communication produced naturally by infants is investigated for indications that they are using the intentional structure of communication.

We have not provided unambiguous evidence to suggest that infants understand the intentional structure of communication at 18 to 30 months. Indeed, the earliest age that such an understanding can be demonstrated remains around 3 years (Grosse et al., 2013; Tomasello, 2008). Prior to this point, as reviewed in Chapter 1, infants seem to engage in intentional communication, and must do so without understanding its intentional structure. In the second half of the thesis, we explore whether infants around their first birthday produce behaviour that is intentionally communicative in terms of directing the attention of others. The following chapter is a second literature review, building on work in Chapter 1, to provide the theoretical and empirical backdrop for the second half of this thesis.

# 4. Prelinguistic Intentional Communication

In the previous chapters, we asked whether toddlers might understand the intentional structure of communication, focusing on manipulations of the two levels of intentions involved when we communicate. The following chapters pose a different question. Specifically, before they produce language, do infants vocalise or gesture in a way that indicates that they intend for these actions to influence the behaviour of others? As a potential indicator of such intent, we will focus on infants' coordination of gaze to the faces of others with their vocalisations and gestures, and will ascertain the *relative* rates of gaze coordination for vocalisations and gestures. We will test whether infants' preverbal vocalisations and gestures (intentional or otherwise) predict later expressive vocabulary and whether any predictive links are explained by how caregivers respond. As such, while the previous chapters focused on both the *communicative intention* (that you are intending that another recognises your actions as communicative) and the *informative intention* (what it is that you intend to communicate), the following chapters would be better conceptualized as an investigation of the informative intention only.

This chapter first reviews the literature on how instances of intentional communication can be identified in infancy. Secondly, we briefly describe the specific vocalisation and gesture types that are produced by infants at the end of the first year of life and review the literature on gaze coordination with these behaviours, which we take as a marker of communicative intent. Thirdly, we give an account of what is known about the links between prelinguistic vocalisations and gestures and later language. Fourth, we review the literature on caregiver responsiveness to infant vocalisations and gestures, and how this might impact on later language. Finally, we detail a number of methodological issues that arise from this review, and that will be addressed throughout the following chapters.

## Measuring Intentional Communication in Infancy

Intending that your interlocutor does something (i.e., intending that they give you something, look at something, understand something etc.) is perhaps obviously, a necessary condition for intentional communication. It is the content of our communication, and what we intend to communicate; and can be termed our *informative intention*. It has been demonstrated that infants' prelinguistic vocalisations and gestures influence the behaviour of interlocutors. For example, prior to 12 months, infants' vocalisations can be used to request or express discontent (Esteve-Gibert & Prieto, 2013; Oller et al., 2013; Papaeliou et al., 2002; Papaeliou & Trevarthen, 2006). Infants' gestures can be used to draw caregivers' attention to objects, which might also result in caregivers performing actions with the objects (Cameron-Faulkner et al., 2015; Carpenter et al., 1998). At the beginning of the second year of life, the pointing gesture is used by infants to direct adult attention and interest (E. Bates et al., 1975; Liszkowski et al., 2004, 2006). However, while there is evidence that infants' prelinguistic vocalisations and gestures influence others' behaviour, whether or not infants *intend* for their actions to affect others' behaviour represents the empirical challenge for the second half of this thesis. In the following chapters, we investigate whether infants gaze to their caregiver's face whilst vocalising or gesturing, in order to determine if infants intend for their actions to influence their caregiver's behaviour. Gaze to an interlocutor's face has been used as a marker of intention by numerous researchers investigating intentionality in both *human* infants, and in non-human animals (Balog & Brentari, 2008; E. Bates et al., 1975; Bourjade, Meguerditchian, Maille, Gaunet, & Vauclair, 2013; Camaioni et al., 2004; Franco & Butterworth, 1996; Harding, 1982; Harding & Golinkoff, 1979; Hostetter, Cantero, & Hopkins, 2001; Krause & Fouts, 1997; Leavens & Hopkins, 1998, 1999; Leavens, Hopkins, & Thomas, 2004; Leavens, Russell, & Hopkins, 2005; Katja Liebal, Pika, & Tomasello, 2006; Masur, 1983; Meunier, Prieur, & Vauclair, 2013; O'Connell & Farran, 1982; A. I. Roberts, Vick, & Buchanan-Smith, 2013; Schel, Machanda, Townsend, Zuberbühler, & Slocombe, 2013; Schel, Townsend, et al., 2013; Snyder, 1978; Tomasello, Call, Warren, et al., 1997; Tomasello, Kruger, George, Farrar, & Evans, 1985; Vauclair, 2004; Zinober & Martlew, 1985).

When an infant gazes to an adult's face whilst gesturing or vocalising, we argue that at the very least this indicates that the infant has an informative intention, and thus their actions are intentionally communicative. This is because the infant, by gazing to their caregiver's face whilst vocalising or gesturing, demonstrates awareness that when these actions are attended to by others, they function effectively in achieving what they want to happen (i.e., in achieving their informative intention). That is to say that when they gaze to their caregiver's face, infants are monitoring the attention of their caregiver, which demonstrates that they understand that in order for communication to function, others need to attend to their actions. We acknowledge that it is also possible to make a richer cognitive interpretation from gaze coordination of this type. As adults, when we communicate we also have a *communicative intention* that others recognize our actions as communicative, and should infer our *informative intentions* (Gómez, 1994; R. Moore, 2014; Sperber & Wilson, 1986; Tomasello, 2008). It is possible that gaze coordination could be interpreted as evidence of the adult-like communicative intention, in that infants are signalling to their caregivers that their vocalisations or gestures *are communicative*. However, this position is too mentalistic a conclusion to draw from gaze coordination alone. The simpler explanation that gaze coordination involves a monitoring of adult attention is more parsimonious. Furthermore, eye contact is not a *necessary* condition for intentional communication (Cochet & Vauclair, 2010; Crais et al., 2004; Gros-Louis, West, & King, 2014; Messinger & Fogel, 1998; Olson & Masur, 2013; Paavola, Kemppinen, Kumpulainen, Moilanen, & Ebeling, 2006; Paavola, Kunnari, & Moilanen, 2005; Tomasello et al., 2007). While we opt for the cognitively leaner explanation, we argue that by monitoring the attention of others, and demonstrating communicative intent (on the level of *informative intentions*) this represents an important developmental step towards the communicative intention proper.

In the following chapters, we opt not to use alternative behavioural indicators that infants' prelinguistic vocalisations or gestures are intentionally communicative. We believe that this requires justification, as these alternative indicators are often used in the literature. Specifically, an alternative way of determining if behaviours are intentionally communicative is to apply '*Brunerian*' criteria to them (Bruner, 1973, 1975). As reviewed in Chapter 1, a behaviour must demonstrate response-waiting, persistence or elaboration to be considered intentional. For example, Golinkoff (1986) investigated repair behaviour of infants during frustration episodes, to determine if infants persisted and elaborated on behaviours if they were denied a specific outcome (see also Liszkowski et al., 2007). Studies with great apes make greater use of measures of persistence and elaboration as behavioural markers of intentional communication (Cartmill & Byrne, 2010; Genty, Breuer, Hobaiter, & Byrne, 2009; Leavens et al., 2005; Pika, Liebal, & Tomasello, 2003, 2005; Poss, Kuhar, Stoinski, & Hopkins, 2006; A. I.

Roberts et al., 2013; C. Russell, Bard, & Adamson, 1997; J. L. Russell et al., 2005; Schel, Townsend, et al., 2013). However, in order to determine if an infant is persisting in an attempt to achieve some sort of goal we have to determine what their goal might be, i.e., we have to accurately identify the pragmatic function of their behaviour. The majority of the studies using *Brunerian* criteria either have good theoretical reasons to believe that a specific act has inferable goal (such as the act being produced is met with a reliable response), or have participants in a context where intent is contrived as a result of the experimental setup (i.e., in putting a desired object out of reach of an infant, all communicative bids towards the object will be treated as having an imperative function).

For the current study, the latter option is not possible, as we want to determine naturalistic frequencies of potentially intentionally communicative behaviour. With regards to the former, if we knew the goal of infant's vocalisations or gestures then of course we could apply *Brunerian* criteria. Naturalistic studies of great apes in the wild might be considered as an example whereby applying *Brunerian* criteria is possible, for example, researchers studying wild chimpanzee gestures (e.g., Hobaiter & Byrne, 2011). Implicit in these studies is an assumption that the function or 'goal' of chimpanzee gestures can be inferred by trained observers. For example, the form of the gesture can indicate that some sort of specific response from a conspecific might be the goal (i.e., begging for food, by extending an upturned hand towards another individual's mouth which is full of food, the goal most intuitively might be to obtain food in the upturned hand from the other individual). Secondly, in these studies there is a possibility that chimpanzees might demonstrate persistence towards this inferred goal. Only around 50% of wild chimpanzee gestures (or gesture-vocal combinations) are responded to by conspecifics that are close by (Wilke et al., 2017), meaning that there is a higher likelihood of individuals demonstrating persistence. However, how this approach could be applied to both human infant vocalisations and gestures in naturalistic settings is not clear. Firstly, regarding goal attribution, the goal of vocalisations is often impossible to ascertain, as vocalisations do not clearly indicate *what response* might be the infant's goal, though it is potentially plausible that specific gestures or specific vocalisations (e.g., words) are linked to specific outcomes (e.g., requesting gestures, or naming of objects leading to objects being given). Secondly, regarding persistence, caregivers are highly responsive to their infant's behaviours during play, rarely denying infants things that they think they want (Baumwell, Tamis-LeMonda, & Bornstein, 1997) and without manipulation we would observe low frequencies of 'frustration episodes' (Golinkoff, 1986) where such behaviours are repeated, that might lead to demonstrations of persistence. Finally, a more worrying theoretical concern is that to pre-suppose some specific goal of individual vocalisations or gestures begs the question as to whether such actions are intentionally communicative. Hence, we opt not to use these alternative criteria for intentional communication, and instead focus solely on gaze coordination.

As gaze coordination is our measure of whether infants are intentionally communicating, the following sections of this chapter focus on what evidence there is that infants' vocalisations and gestures are gaze-coordinated around the end of the first year of life. However, prior to that, we briefly describe the repertoire of vocalisations and gestures that infants produce at this age.

**Prelinguistic Infants' Repertoire of Vocalisations and Gestures**

As reported in Chapter 1, by the end of the first year of life, infants typically produce speech-like babbling, consisting of repeated syllables involving both vowels and consonants (Oller, 2000; Stark, 1980; Vihman, 1996). Additionally, they produce other vocalisations that are not so speech-like, such as grunts, and a variety of idiosyncratic, but non-vegetative, vocalisations (e.g., McCune, Roug-Hellichius, Vihman, Delery, & Gogate, 1996). Infants also produce a number of gestures, namely, extended arm gestures, such as gives, shows and reaching, which typically precede the onset of pointing (Masur, 1983).

There is evidence that a developmental trajectory exists for the onset of gestures. A large scale cross-sectional study using caregiver report suggests that reaching emerges first at 8 months, followed by showing at 10 months, followed by giving and pointing at 11 months (Caselli, Rinaldi, Stefanini, & Volterra, 2012, p. 535, Figure 4, age of emergence defined as age when 50% or more infants are reported to perform the gesture). These results are supported by further studies using caregiver report (Sansavini et al., 2010) and from longitudinal studies using gesture-eliciting paradigms (Cameron-Faulkner et al., 2015; Carpenter et al., 1998). It is argued that reaching, giving and showing plausibly emerge earlier as they are more tractable for young infants in terms of communicating intent, as sign and referent are in the same place, i.e., at the end of your arm which is gesturing, unlike pointing, where the referent is distal (Rodríguez, Moreno-Núñez, Basilio, & Sosa, 2015).

It is worth noting that there is a discrepancy in the literature over age of pointing onset, mainly attributable to how pointing is measured. Pointing onset, conceived of as the age where 50% of infants point with their index-finger, has been reported at 11 months (Caselli et al., 2012), 12 months (Desrochers et al., 1995) and estimated at 12.56 months (Leung & Rheingold, 1981). Mean average age of pointing onset has been reported at 11.3 months (Butterworth & Morissette, 1996) and 12.3 months (Carpenter et al., 1998). On a large study of naturalistic observations, median onset of pointing was 12.6 months, with a range of 9.3-18.3 months (McGillion, Herbert, et al., 2017). Elsewhere, Leung and Rhinegold (1981) reported 12.56 months as their estimated age of onset (when 50% of infants pointed) with a 95% range from 11.16-13.47 months. Thus, it is fair to say that pointing can emerge as early as 9 to 10 months in precocious infants, but is likely to emerge between 11 to 13 months for a majority of infants, and could emerge much later for others.

Infants frequently combine gestures and vocalisations. The onset of the specific gesture-vocal combination of pointing and vocalising has been the most extensively studied. At 12 months, around half of infant pointing is produced in combination with a vocalisation during pointing eliciting experiments (Aureli et al., 2017; Igualada, Bosch, & Prieto, 2015), although elsewhere this is reported to be as high as 75% (Franco & Butterworth, 1996). Note however, that studies used different criteria for combinations, with two studies requiring that pointing and vocalisations had to occur within 2 seconds of each other (Aureli et al., 2017; Franco & Butterworth, 1996), and the other where pointing and vocalisations had to overlap (Igualada et al., 2015). Nevertheless, at 12 months, we therefore expect infants to be engaging in gesture-vocal combinations.

Therefore, we can characterize the repertoire of infants around the end of the first year of life as follows. We expect infants to be producing speech-like babble and using a variety of non-babble vocalisations (such as grunting or other idiosyncratic vocalisations). Furthermore, we expect that they will be producing giving and showing gestures at a high frequency, and some infants will be producing index-finger pointing. Finally, we expect that infants will be producing gesture-vocal combinations.

## Gaze Coordination

How much infants engage in gaze coordination (i.e., gazing to their caregiver's face) whilst producing vocalisations or gestures (or combinations of both), and whether this could indicate that the infant is engaging in intentional communication presents different challenges depending on whether you are focusing on vocalisations, gestures or both. As such, our review of the current evidence will focus on vocalisations produced alone, gestures produce alone, and studies of both.

### Vocalisations

There are two major challenges in investigating how much gaze coordination occurs with infants' vocalisations that do not apply (or are less pertinent) when considering gestures. Firstly, vocalisations are frequently produced, so there is a high chance of these co-occurring with gaze to caregivers' face by chance. In a recent naturalistic study, 12-month-olds produced eighteen times more vocalisations (excluding vegetative sounds, crying, fussing and laughing) than gestures (including pointing, reaching, showing and other conventional gestures; Miller & Lossia, 2013). Co-occurrence of vocalisations with looks to faces are therefore far more likely to occur *by chance* than looks to faces whilst gesturing. Thus appropriate statistical controls are necessary when comparing between vocalisations and gestures (see D'Odorico et al., 1997; D'Odorico & Cassibba, 1995 for an example of such a control looking at vocalisations alone). Secondly, whereas gestures such as pointing (if intentionally communicative) require the interlocutor to visually attend to whatever is pointed at, vocalisations do not necessarily require this response. One solution to this problem is to focus on vocalisations that are more likely to have a similar communicative function to pointing. One class of such vocalisations are object-directed vocalisations, where non-cry vocalisations are given when an infant is attending to an object that they are either holding or that is within reach (Goldstein, Schwade, Briesch, & Syal, 2010). However, while studies have successfully looked at these vocalisations, restricting the types of vocalisations in this way is not appropriate for the current question, as it is circular logic to attempt to look for potential instances of intentional communication, by only selecting behaviours to code that we think *a priori* might be communicative (i.e., 'about' an object – see above discussion of this in the primate literature). Additionally, this may be a useful distinction to use in the lab, when the number of objects in the environment are controlled, however in the home, infants are surrounded by within-reach objects, thus making this coding unlikely to reduce the number of vocalisations substantially. We thus consider all vocalisations while being mindful of the potential pitfalls of cross-modality comparison here.

Coordination between vocalisations and gaze to caregiver's face in infants around 12 months has rarely been studied in naturalistic settings, and less frequently still, to our knowledge, in participant's homes. An early attempt to study vocalisations in semi-naturalistic settings (i.e., observational sessions in the laboratory, with an experimenter present and occasionally instructing the interactions) found that 61% of 12-month-olds looked to their caregiver whilst vocalising, a level of co-occurrence above that expected by chance (D'Odorico et al., 1997). One recent study that observed caregiver and infant dyads under semi-naturalistic settings (again in a laboratory), found that, on average, 13% of 12-month-olds non-vegetative vocalisations were given whilst the infant made eye contact with the caregiver, compared to 75% which appeared to be object-directed, and 12% which were not directed towards objects or caregiver (Miller & Lossia, 2013). Finally, another recent paper looking at semi-naturalistic settings in the lab reported that 12-month-olds did gaze check caregivers *during* object-directed vocalisations, but did not report the relative frequency, as the primary focus of the study was caregiver labelling responses, and these did not vary according to whether infants gaze-checked or not (Wu & Gros-Louis, 2015).

The rate at which a specific type of vocalisation (e.g., 'grunting') is gaze-coordinated has been observed under naturalistic conditions (McCune et al., 1996). This study used gaze coordination as an indication of communicative intent when infants grunt, however it was part of a longer list of criteria for communicative intent including combinations with gestures (such as 'giving') and tactile contact with the mother (McCune et al., 1996). Using this criteria, this small (*n*=5), but intensive longitudinal study (home visits every month between 8 and 24 months of age) showed that in the main, infants grunting around 8 to 12 months was not communicative, and grunts coordinated with extravocal behaviours that the authors argued demonstrated communicative intent only emerged around 13 months.

As studies of vocalisations with gaze coordination in naturalistic contexts are lacking, experimental paradigms are one source of information about the rate at which such gaze coordination might occur. For example, it has been demonstrated in a lab eliciting context, that 45% of 8- to 14-month-olds (mean age of 10 months) alternated gaze between caregiver's eyes and an object whilst vocalising when an object was put out of their reach, while 30% alternated gaze between caregiver's eyes and hands, and 33% alternated gaze between all three whilst vocalising (calculated from Harding & Golinkoff, 1979, p. 38, Table 2).

A key study for this thesis demonstrates how appropriate statistical controls can be used when investigating gaze coordination and vocalisations (see above), and provides tentative evidence that vocalisations are intentionally communicative prior to the end of the first year of life (D'Odorico & Cassibba, 1995). In this experimental paradigm, infants sat in a high chair at 90 degrees to their caregiver whilst being given toys. 10-month-olds looked to their caregiver's faces prior to, or during vocalising more than would have been expected by chance, whereas 4-, 6- and 8-month-old infants did not coordinate gaze and vocalisations above chance levels. This finding is important, as it shows that before the onset of pointing infants already coordinate vocalisations with gaze to caregivers' face. This therefore suggests that infants intentionally communicate well prior to the onset of index-finger pointing (thought by some to herald the onset of intentional communication), and so intentionally communicative communication using vocalisations is ontogenetically prior to gestures. There are a number of reasons to be cautious of this interpretation however. Firstly, the sample was small, with only eight infants studied at 10 months. Secondly, this study *only* looked at vocalisations and did not consider gestures (e.g., giving or showing), so we do not know what relative rates are to earlier gestures. Gaze to the caregiver and vocalisations were coded in 500ms windows (as technology allowed), so co-occurrence judgements were not fine-grained (with vocalisations and gaze having to cover 300ms of a 500ms window). There were large individual differences in the number of vocalisations combined with gaze. Six infants (75%) produced gaze-coordinated vocalisations 7 times or less in 10 minutes, while the remaining two produced 19 and 22 gaze-coordinated vocalisations, over twice the mean, and four times the median. This suggests that the data at 10 months was driven by more precocious infants, and that this is not representative of all 10-month-olds. However, this study provides highly suggestive evidence that intentionally communicative vocalisations might be ontogentically prior to, or concurrent with intentionally communicative gestures.

**Gestures**

Naturalistic observation has demonstrated that infants' gestures (including pointing, reaching, giving and showing) are coordinated with gaze at around 12 months. However, such studies are often limited to the investigation of one gesture (usually pointing), and few have large sample sizes and/or are observed in naturalistic conditions like the home. In an early longitudinal study in the homes of four infants, by 12 months, three of the infants' reaching and extension gestures (e.g., showing/giving) were often accompanied by gaze to their caregiver, and of the two infants observed to point by 12 months, one of these had been gaze checking from 9 months (Masur, 1983). Another early study using caregiver report revealed that 34% of 10-month-olds, and 69% of 13-month-olds reportedly gaze check whilst pointing to redirect attention (Bretherton, McNew, & Beeghly-Smith, 1981). More recent studies have involved a naturalistic environment, but set up in a laboratory for ease of observation. In these, no significant difference was found between rates of gaze-checking before, during or after index-finger pointing and open-hand pointing in 12-month-olds, however exact rates were not reported (Gros-Louis & Wu, 2012). 38% of 12-month-olds (eighteen out of forty-seven) were observed to point and gaze-check their caregiver, with around 40.78% of points occurring with gaze checking (Wu & Gros-Louis, 2015). Focusing on giving and showing gestures, it was found that 10- to 13-month-old infants alternated gaze to their caregiver with gaze to the toy around 64% of the time when performing giving and showing gestures, but that gaze to caregiver's face was more likely to occur for showing (Boundy, Cameron-Faulkner, & Theakston, 2016). There is mixed evidence comparing rates of gestures observed in naturalistic settings and those observed through experimental paradigms. Some findings seem to match rates observed in naturalistic settings (Franco & Butterworth, 1996; Kutsuki, Ogura, Egami, Itakura, & Children's Study Group, 2009), however others report much lower rates (Desrochers et al., 1995; Lempers, 1979; Murphy & Messer, 1977).

**Vocalisations and Gestures**

Studying vocalisations and gestures in isolation of each other provides a number of problems for interpreting exactly what modalities children might intentionally communicate in. Given that even *within* the same modality, varying methodologies provide wildly different rates of the behaviours and different rates of their co-occurrence with gaze, comparisons between studies looking at different modalities is almost certainly not a valid way to evaluate this relationship. Since this has not been the central goal of prior research, gestures have often been the primary focus of studies, with vocalisations coded only when they co-occur with gestures, in the context of elaborating on gestures, or modifying the pragmatic content of the gesture (Aureli et al., 2017; Gros-Louis & Wu, 2012; Grünloh & Liszkowski, 2015; Legerstee & Barillas, 2003; Wu & Gros-Louis, 2014). The opposite is also true, with some researchers only coding gestures when they co-occur with vocalisations (Esteve-Gibert & Prieto, 2013). Similarly, where gestures are the primary focus, both gaze and vocalisations are coded, but only when they co-occur with gestures (Masur, 1983; Messinger & Fogel, 1998). Few studies have looked at *all* instances of both behaviours in the same children in the same observation period (Igualada et al., 2015; Jones & Hong, 2001; Murillo & Capilla, 2016; Wetherby, Cain, Yonclas, & Walker, 1988; Zinober & Martlew, 1985), but these studies did not investigate gaze checking behaviour for both. To our knowledge, few have investigated all instances of both behaviours and related gaze checking in the same children around the end of the first year of life. This work is crucial to ascertain the *relative* rates of gaze-checking with prelinguistic vocalisations and gestures. However, of these studies, four did not report the relative rates of gaze coordination for both vocalisations and gestures (Harding & Golinkoff, 1979; Murillo & Belinchón, 2012; Spencer, 1993; Wu & Gros-Louis, 2015), and the final paper reported the relative rates that infants looked to their caregiver whilst vocalising or gesturing, but this category was not restricted to gaze at the caregiver's face (instead including gaze to caregiver's body and hands), which we argue would more unambiguously demonstrate that infants are monitoring attention (Miller & Lossia, 2013).

In sum, around the end of the first year of life, the relative rates of gaze coordination with infants' vocalisations and gestures produced under naturalistic conditions is still not known. There is suggestive evidence that at 10 months, infants vocalise with gaze coordination above the level predicted by chance (D'Odorico & Cassibba, 1995) however such evidence is limited to a small sample in an experimental paradigm. There is some evidence that giving and showing gestures, produced ontogenetically earlier than index-finger pointing, are highly likely to be gaze-coordinated (Boundy et al., 2016), and that index-finger pointing by contrast is less likely to be gaze-coordinated and rare (Wu & Gros-Louis, 2015), however these studies did not consider relative rates for vocalisations. The following section considers what is known about the relationship between prelinguistic vocalisations and gestures (intentional or otherwise) and infants' later language abilities.

## Predicting Later Language

The view that there is continuity between prelinguistic communication and later language ability is relatively uncontroversial, however, the focus of much recent work on prelinguistic communication has been on infants gestures, as gestures have been seen by some as more likely precursors to language than prelinguistic vocalisations (the gesture-first view). In support of this view, a large body of evidence demonstrates early gesture use correlating with or predicting later language abilities. Onset of gestures, frequency and number of gesture-types used around children's first birthday have all been correlated with an array of measures of later expressive and receptive vocabulary (E. Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bavin et al., 2008; Blake, Vitale, Osborne, & Olshansky, 2005; Kraljević, Cepanec, & Šimleša, 2014; Laakso, Poikkeus, Katajamaki, & Lyytinen, 1999; Murillo & Belinchón, 2012; Olson & Masur, 2015; Paavola et al., 2006, 2005; Rowe & Goldin-Meadow, 2009a; Rowe, Özçalışkan, & Goldin-Meadow, 2008). It has been noted that the onset or early frequency of *pointing* especially is found to correlate with later measures of expressive and receptive vocabulary (Camaioni, Castelli, Longobardi, & Volterra, 1991; Desrochers et al., 1995; Harris, Barlow-Brown, & Chasin, 1995; Igualada et al., 2015; Murillo & Belinchón, 2012; Wu & Gros-Louis, 2014). A relatively recent meta-analysis supports this view (Colonnesi, Stams, Koster, & Noom, 2010), although it is worth bearing in mind that that perhaps only specific types of pointing (i.e., declarative not imperative pointing) are predictors of later language (Cochet & Byrne, 2016). As a result of these findings, it is claimed that pointing represents a 'royal road (but not the only route) to language' (Butterworth, 2003). To further reinforce the dominant position of index-finger pointing in language development over vocalisations, it is also sometimes claimed that prelinguistic vocalisations *are not* related to later language based on negative or null findings, but this stronger claim is rare (Camaioni et al., 1991).

However, the evidence does not unambiguously support gesture-first accounts. For example, gesture-first accounts are undermined by studies that *do* show predictive links and correlations between prelinguistic vocalisations and language (D'Odorico et al., 1997; D'Odorico, Salerni, Cassibba, & Jacob, 1999; McCune et al., 1996; McCune & Vihman, 2001; Menyuk, Liebergott, & Shultz, 1986; Murillo & Belinchón, 2012; Rome-Flanders & Cronk, 1995; Watt, Wetherby, & Shumway, 2006). Additionally, some measures of early gestures (including pointing) *do not* correlate with or predict later language ability even in studies where other positive findings are reported (Blake, Osborne, Cabral, & Gluck, 2003; Brooks & Meltzoff, 2008; Butterworth & Morissette, 1996; Colonnesi, Rieffe, Koops, & Perucchini, 2008; Harris et al., 1995; Mumford & Kita, 2016).

Methodological differences between researchers almost certainly contribute to the seemingly contradictory results. Many only study *either* vocalisations *or* gestures, or one in relation only to the other (e.g., vocalisations *only* during pointing). Some studies use eliciting tasks (which are not designed to get natural frequencies of communicative behaviours rather demonstrate potential competence), while others use naturalistic observations and caregiver report (which are more likely to give natural frequencies, but may not detect *rare* instances of behaviours of interest that would not be demonstrated without experimental prompting). The key distinction here is that while demonstrating an infant's potential competence through eliciting tasks might reveal an infant's readiness for later communicative acts (e.g., checking that infants can produce certain phoneme, or a have the motor control required for specific gestures) it doesn't necessarily speak to how that infant might *learn* to communicate. In contrast, determining the natural frequencies or onset of behaviour through naturalistic observation might reveal the how much an infant experiences communicative-like events in their normal environment that might scaffold learning about communication (e.g., when an infant first produces specific behaviours like vocalising or gesturing that receive communicative responses from adults, and what the frequency of such events is).

Addressing at least some of these issues, one recent longitudinal study investigated the link between early pointing *and* babble and later language, where both babble, pointing and words were transcribed from naturalistic recordings of play interactions between child and caregiver in the home between 9 and 18 months of age (McGillion, Herbert, et al., 2017). It was found that the onset of babbling (when children were observed to have mastered two consonants), but *not* the onset of pointing (where children were observed to produce at least one index-finger point) predicted the age at which children were first observed to produce 4 different words (i.e., the very beginnings of expressive vocabulary). Additionally, 13% of the sample (6 children) had already mastered 4 words *prior* to being observed to point, suggesting the transition to producing conventional linguistic forms can be made without first passing through a pointing stage. Receptive and expressive vocabulary scores at 18 months were obtained from caregiver report on the Oxford Communicative Development Inventory (OCDI), a UK-modified version of the MacArthur Bates Communicative Development Inventory (MCDI). Children's expressive vocabulary at 18 months was predicted by babble, receptive vocabulary was predicted by pointing and maternal education. This highlights the importance of studying *both* prelinguistic vocalisations and gestures in tandem, as *both* are linked to different aspects of later language.

One limitation of the McGillion et al. (2017) study, acknowledged by the authors, was that there was no investigation into the *intentionality* of early pointing or vocalisations. Prelinguistic babble constitutes the building blocks of spoken language, thus there is phonological continuity between individual infants' babble and first words (Fagan, 2009; Majorano & D'Odorico, 2011; McCune & Vihman, 2001; Menyuk et al., 1986; Stoel-Gammon, 2011; Vihman, 2014; Vihman & Ferguson, 1986; Vihman, Macken, Miller, Simmons, & Miller, 1985). Therefore, it is not clear if babble predicts later language because it represents instances of prelinguistic intentional communication, or if it may do so solely because it is an indicator of phonological development necessary for speech. Similarly, the predictive relationship between gesture and later language might not be because early gesture represents the first steps of prelinguistic intentional communication, it may just be a marker of general socio-cognitive or motoric development that would necessarily correlate with later language under a Piagetian 'deep homology' model (E. Bates et al., 1979, p. 69).

Vocalisations involving both a consonant and vowel (CV vocalisations) have been found to predict later language explicitly (Menyuk et al., 1986; Stoel-Gammon, 1992), and implicitly, as necessary components of canonical babbling (D'Odorico et al., 1999), and as part of vocal-motor scheme measures (McCune & Vihman, 2001; McGillion, Herbert, et al., 2017). CV vocalisations (unlike gestures) are produced well in advance of 12 months, with canonical babbling onset around 6 to 8 months (Oller, 2000). Mastery of these type of vocalisations are undoubtedly a necessary phonological prerequisite for speech, and so necessarily predict expressive language abilities, but a central empirical question of the following chapters is whether they are also predictive of later language due to their intentionally communicative use (if and when that occurs). Furthermore, it is possible that around 11 months there could be a shift in their use, from vocal play towards more communicative purposes, as one study suggests that it is only from around this age that CV vocalisations more clearly begin to resemble adult language in the environment, with the most frequently occurring syllables in the adult language being produced more frequently (Majorano & D'Odorico, 2011). Therefore thorough investigation of the potential intentionally communicative nature of CV vocalisations at the end of the first year of life is warranted.

Additionally, non-CV vocalisations could also be used to intentionally communicate (and equally predict later language). A previous longitudinal study using a lab-based communication-eliciting task (involving out of reach objects), investigated infants CV and non-CV vocalisations directed towards their caregivers or elsewhere (Marchand, Ricard, & Gouin Décarie, 1994). This study demonstrated that CV vocalisation production, and the proportion of infant's vocalisations (both CV and non-CV) directed towards their caregiver increased between 6 and 15 months. However, infants did not seem to preferentially direct CV vocalisations over non-CV vocalisations towards their caregiver at any age. This finding suggests that there might not be a privileged prelinguistic communicative role for CV vocalisations over non-CV vocalisations. Thus, there remains an empirical question as to whether non-CV vocalisations (like gestures) might be intentionally communicative and predict later language ability despite not necessarily being key phonological components of spoken language. This hypothesis has been advanced especially in the case of grunt vocalisations (McCune & Vihman, 2001).

It is possible that a unique relationship exists between combinations of gestures and vocalisations, and later language that does not exist for either vocalisations or gestures produced alone. In studies designed to elicit infants communicative behaviours, the frequency with which infants produce combinations involving vocalisations and pointing at 12 months predicts infants expressive vocabulary at 15 months (Murillo & Belinchón, 2012) and at 18 months (Igualada et al., 2015). Moreover, the frequency with which infants produced gestures (including pointing, reaching, giving, showing and others) combined with vocalisations at 12 months was correlated with infants' expressive vocabulary at 15 months, and a stronger relationship was observed when the frequency of gesture-vocal combinations *with gaze to caregiver's face* was considered (Murillo & Belinchón, 2012). A privileged role in language development for gesture-vocal combinations is intuitively possible. For example, a predictive link between the frequency of gesture-vocal combinations produced by older infants (18 months) and later language, was hypothesised to be because infants can express communicative meanings beyond what they can produce with speech alone at that age, by using a concurrent gesture (Rowe & Goldin-Meadow, 2009b). Furthermore, recent findings suggest that the mechanism by which this is predictive is that caregivers' specific type of responses to gesture-vocal combinations at 18 months exposes infants to more complex language input, i.e., by repeating infants' simple utterances with more complexity added (Fasolo & D'Odorico, 2012). Finally, even prior to 18 months, when infants use non-linguistic vocalisations in these gesture-vocal combinations, caregivers may respond differently to combinations (compared to gestures or vocalisations produced alone) as the prosodic cues of these vocalisations combined with the referential nature of gestures allow caregivers to more reliably infer what infants are trying to communicate about (Balog & Brentari, 2008).

In sum, the production of both some gestures and vocalisations, and combinations of the two around the end of the first year of life predicts later language abilities (e.g., McGillion, Herbert, et al., 2017). However, what is not currently known is whether different types of vocalisations and gestures, or specific gesture-vocal combinations produced by the same infants under naturalistic conditions predict their later language abilities. Furthermore, it is not known whether the frequency of gaze-coordinated gestures, vocalisations or combinations predict later language abilities because they are early instances of intentional communication.

## Caregiver Responsiveness

Caregiver responsiveness is the mechanism by which some hypothesise that prelinguistic vocalisations and gestures lead to language. For example, the argument that pointing represents the 'royal road' to language is based on the assumption that because pointed-to objects provide a focus for caregiver's speech, this best facilitates infants' mapping of names to objects (Butterworth, 2003).

Support for this begins with findings that caregiver responsiveness, defined as 'the prompt, contingent, and appropriate responding by mothers to children's exploratory and communicative overtures' (Baumwell et al., 1997, p. 247) is related to infant's later language abilities. For example, a longitudinal study of 40 caregiver-infant dyads, observed during naturalistic play in the home revealed that caregivers responsiveness (appropriate verbal responses produced within 5 seconds of infants' communicative behaviours, object exploration, and playing behaviours) at 9 months, predicted infants receptive vocabulary at 13 months (Baumwell et al., 1997). Furthermore, in a similar study involving 30 caregiver-infant dyads, caregivers' responsiveness at 13 months was  predictive of infants expressive vocabulary at 20 months (Bornstein, Tamis-LeMonda, & Haynes, 1999).

A specific kind of caregiver responsiveness has been shown to be predictive of later language outcomes. One traditional candidate has been object-labelling, as suggested by Butteworth (2003). Around the first birthday (between 10 and 14 months) object-labelling is known to increase infants attention to the object that is being labelled (Baldwin & Markman, 1989), and naturalistic observations with older infants (15 months) have demonstrated that caregivers labelling of objects that infants were attending to predicts their expressive vocabulary at 21 months of age (Tomasello & Farrar, 1986). However, object-labelling is quite a narrow definition of a type of caregiver speech that is plausibly useful for infants to learn language (beyond learning labels for objects), and this narrowness might exclude other types of speech that could plausibly scaffold infants' language development. Researchers have instead focused on *semantically contingent talk* (McGillion et al., 2013), also referred to as *following in* behaviour (Carpenter et al., 1998). This behaviour involves caregivers talking about what is in their child's current focus of attention (which can include, but is not limited to, labelling of objects). Semantically contingent talk is thought to promote language development as it helps infants to understand the function of caregiver speech, which aids word learning.

Semantically contingent speech by caregivers has been shown to contribute to infants' language learning. In a small scale study (with 11 caregiver-infant dyads) the amount of speech that caregivers provided to 9-month-olds that was semantically contingent on what the infant was attending to (during 10 minutes of interaction where they were observed playing with toys in the lab with their caregiver) predicted infants' receptive vocabulary at 12 and 18 months, and expressive vocabulary at 30 months (Rollins, 2003). Semantically contingent talk at the end of the first year of life has also recently been demonstrated to predict infants' later vocabulary in a large randomized control trial (McGillion, Pine, Herbert, & Matthews, 2017). In this study, half of 142 caregiver-infant dyads were assigned to an intervention condition where they were asked to engage in semantically contingent talk with their infant for 15 minutes a day for a month (a control condition involved a dental health intervention) when their infant was between 11 and 12 months of age. Specifically caregivers were prompted to notice what their child was attending to, and then talk to them about it. At 15 and 18 months, infants of low-socio-economic status caregivers in the intervention condition had higher expressive vocabularies than those in the control condition (although this effect did not persist until 24 months). Finally, a longitudinal study of 40 infants revealed that semantically contingent responses that were also temporally contingent on infants actions (including vocalisations, object exploration and play) during play at 9 months predicted the age at which infants produced their first words (Tamis-LeMonda, Bornstein, & Baumwell, 2001).

Results are mixed however, as in a naturalistic study of 20 caregiver-infant dyads, caregiver's speech that was descriptive of what infants were attending to (which could be thought of as semantically contingent speech) at 10 months did not predict infants' vocabularies at 13 months, while the same kind of responsiveness at 15 months was a marginal predictor of 17-month vocabulary, and at 17 months was a significant predictor of 19-month vocabulary (Masur, Flynn, & Eichorst, 2005). The authors suggested that the effect of caregiver's semantically contingent talk may not be effective in scaffolding language development prior to the end of the first year of life, and would be more effective in the second half of the second year of life. This demonstrates that caregiver responsiveness may be related differently to later language outcomes (or different types of outcomes, i.e., expressive vocabulary counts versus onset of language milestones) at different ages.

Caregivers' responses that are specifically given in response to infants' vocalisations and gestures might also promote language development in the manner suggested by Butterworth (2003). In naturalistic studies of caregiver-infant dyads' play, caregiver responses (provided within 5 seconds – regardless of semantic content) to infants' vocalisations at 9 months predicted the onset of infants' first words, and responsiveness to vocalisations at both 9 and 13 months, was predictive of the age at which infants attained a lexicon of 50 words (Tamis-LeMonda et al., 2001). However, the amount of purely temporally contingent responses to infants vocalisations at 9.5 months was not found to be predictive of expressive vocabulary at 18 months (McGillion et al., 2013). In this latter study however, the proportion of caregiver utterances that were semantically contingent *and* temporally contingent (within 2 seconds) of infants' vocalisations was a significant predictor of infants' 18 month expressive vocabulary. Again, this highlights that while caregivers' responses to vocalisations (either regardless of content, or semantically contingent to infants focus of attention) are predictive of infants' later language, the type of responsiveness may be predictive of measures of infants' language at different ages, or of different types.

These findings do suggest however that responses (and specifically semantically contingent responses) to infants' vocalisations at the end of the first year of life predict later expressive language. However, neither study distinguished between different types of vocalisations, so we don't know whether responsiveness to only those vocalisations that resemble later speech (e.g., babbling) or those that do not (e.g., grunting) differs in relation to language outcomes. This is important because it has been argued that grunting and other non-speech like vocalisations may function communicatively and help scaffold language development through their responses from caregivers (McCune et al., 1996).

Similar research into gestures has focused mainly on object-labelling as opposed to the broader concept of semantically contingent talk (see discussion above). Infants' communicative bids involving gestures receive caregiver responses that include object labelling, and this has been demonstrated in both experimental settings (Olson & Masur, 2011) and through naturalistic observation (Goldin-Meadow, Goodrich, Sauer, & Iverson, 2007; Masur, 1982). Considering different gesture types, while caregivers were more likely to respond to object extension gestures (i.e., showing/giving) than pointing or reaching, index-finger pointing was proportionally more likely than the other gestures to receive object-labelling responses (Masur, 1982). In experimental studies (designed to encourage infants communicative bids), research has demonstrated that labelling responses are more likely to be given for communicative bids involving gestures at 13 months than those not involving gestures (Olson & Masur, 2013, 2015). Moreover, it has been found that caregivers in naturalistic play are significantly more likely to provide object-labelling responses to infants pointing at 12 months than to object-directed vocalisations, however such responses are significantly more frequently provided to object-directed vocalisations, as these were more frequently produced by infants (Wu & Gros-Louis, 2015).

In terms of the relationship with later language, a recent study has demonstrated that the predictive relationship between frequency of gestures at 13 months and expressive vocabulary at the same age was fully mediated by the frequency with which these gestural bids received responses involving object labelling (Olson & Masur, 2015). Furthermore, the proportion of gestures receiving an object labelling response at 13 months (along with infants 13 month expressive vocabulary) predicted 17-month vocabulary. This study was conducted using a gesture eliciting paradigm, involving opportunities for infants to vocalise or gesture when toys were revealed either within reach and required adults to activate (termed proto-imperative contexts) or inaccessible, and distal (termed proto-declarative contexts). Only the relationship between gestures (including index-finger and open-hand pointing) produced in the proto-declarative context, and caregiver responses to these behaviours were predictive of later language. While this provided good evidence that caregiver responses to some infant gestures predicts later vocabulary (and could be the mechanism by which gestures predict later vocabulary), it provided no evidence that caregiver responses to non-gestural bids, i.e., bids involving only vocalisations (or indeed non-gestural bids themselves) were predictive of later expressive vocabulary (Olson & Masur, 2015). This contrasts with evidence from the naturalistic studies discussed above that showed predictive links between semantically contingent responses to early vocalisations and later expressive vocabulary (McGillion et al., 2013). Whether this discrepancy is due to the difference in methodology (experimental paradigm versus naturalistic) or the type of caregiver response considered (object-labelling versus the broader category of semantic contingency), or the ages of observation and outcome (13 months to 17-month vocabulary versus 9.5 months to 18-month vocabulary) is not clear. It does however again highlight the need to study both infants' vocalisations and gestures (and responses to these) in the same infants at the same age under naturalistic conditions, using the same criteria for responsiveness (as with intentionality, discussed above).

One final issue is whether infants *expect* that their vocalisations and gestures will receive responses, and whether these responses have any real-time effect. This is important to consider when speculating about the mechanisms by which infants learn language when they produce vocalisations and gestures. From as early as 5 months, infants seem to expect responses when they produce vocalisations. This is demonstrated by studies using a still-face paradigm, whereby when caregivers stop responding to infant vocalisations, infants demonstrate an extinction burst, vocalising much more frequently (Goldstein et al., 2009). Furthermore, the magnitude of the extinction burst (taken as a proxy for how strongly infants expect responses from caregivers) predicts infants' expressive vocabularies at 13 months. Another experiment has demonstrated that if 9-month-olds received a temporally contingent response to their babble (within 2 seconds), they responded by matching their vocalisations to the phonological qualities of the caregiver speech that they had just heard (Goldstein & Schwade, 2008). Finally, in a study that manipulated caregivers attentional focus (either available and responsive or unavailable because they were reading a book and ignoring the infant) 10 month-olds seem to vocalise more when their caregivers were available than when they were not (Wu & Gros-Louis, 2017). These studies suggest strongly that infants expect responses to their vocalisations, can modify their behaviour as a result of the responses that they receive, and vocalise more when it is possible that responses from caregivers are forthcoming.

In contrast, evidence for this in early gestures is not so clear. In the study described above, infants did not gesture (or produce gesture-vocal combinations) more or less depending on caregiver's attentional availability (Wu & Gros-Louis, 2017). These infants were only 10 months old, which is admittedly prior to the onset of most gestures, and studies with infants when gestures are more frequent (12 to 16 months) suggest that infants are sensitive to the attention of their caregivers when producing gestures, pointing more when adults are attending to them (C. Moore & D'Entremont, 2001). However, recent findings suggest that at the onset of pointing, 12-month-olds do not map object labels provided by caregivers to objects when they point to them (Lucca & Wilbourn, 2016). As discussed above, object-labelling is a narrow form of semantically contingent caregiver response, so this finding is especially pertinent. While 18-month-olds seem to privilege label-learning when they have pointed to an object (i.e., they were better at learning labels given for objects that they had pointed at compared to objects they had reached towards or only gazed at prior to their labelling), 12-month-olds did not (Lucca & Wilbourn, 2016). The authors of this study explained these findings by suggesting that while 18-month-olds have learnt to expect caregiver labelling as they gain experience of pointing and labelling (between 12 and 18 months), 12-month-olds do not have such an expectation. Therefore, it may be the case that infants are not seeking out semantically contingent caregiver responses through pointing at its onset, and so caregiver responses to early pointing may not scaffold language development in the way that it does with older infants' pointing.

In sum, there is evidence that caregiver responsiveness, specifically providing semantically contingent infant-directed speech is predictive of infants later language abilities (e.g., McGillion, Pine, et al., 2017; Tamis-LeMonda et al., 2001). Similarly, semantically contingent responses to infants vocalisations and gestures is predictive of later expressive language outcomes (e.g., McGillion et al., 2013; Olson & Masur, 2015). Furthermore, there is evidence that the relationship between gestures and later expressive vocabulary is fully mediated by caregiver's semantically contingent responses (Olson & Masur, 2015), however this has only been demonstrated through an experimental paradigm, and with no developmental time lag. What has not yet been demonstrated is whether caregivers' responses to *both* vocalisations and gestures in the same infants under naturalistic conditions predict later language, and whether this mediates any relationship between infants' production of vocalisations and gestures and later language. Furthermore, the role of responsiveness with regards to whether infants' behaviours were intentionally communicative (i.e., gaze-coordinated) and later language has not been investigated.

## Methodology

Throughout this literature review, we have identified a number of key methodological issues that we address in the following chapters. These are dealt with in turn below.

### Gesture-vocal Combinations

In the following chapters, we opt to treat gesture-vocal combinations as a separate category of behaviour from vocalisations and gestures produced alone. As discussed previously, it is possible that these behaviours have a unique relationship to later language, and that caregivers' responses to these behaviours may differ from when vocalisations or gestures are produced alone in ways that differently scaffold language development. In addition, recent research has suggested that the vocalisations produced in combination with gestures are qualitatively different from those produced alone (Murillo & Capilla, 2016), suggesting that there is a compelling case for treating instances of gesture-vocal combinations separately from analyses of gestures and vocalisations produced in isolation. Infants vocalisations produced in combination with pointing are different in terms of intonation between pointing to inform and pointing to request (Grünloh & Liszkowski, 2015), suggesting that the combination of vocalisations and gestures provide unique events quite different from vocalisations or gestures produced in isolation. Finally, in terms of predicting language, if the aim is to determine whether the production of early vocalisations *or* gestures predict later language, then this question cannot be answered satisfactorily if categories where they co-occur are included. For example, if the frequency of vocalisations that co-occur with gestures are included with the frequency of vocalisations produced alone, and this frequency predicts later language, it is possible that (on the assumption that more gestural infants will produce more gesture-vocal combinations too) this predictive relationship may result from a high frequency of gestures, not vocalisations. We realise that this approach is uncommon in the developmental literature, and might mean our results are difficult to compare to those that have preceded it, so we opt to run analyses (included in Appendixes) for each chapter where vocalisations and gestures produced in combinations are treated as both vocalisations and gestures respectively

**Gaze Coordination**

As discussed above, appropriate statistical controls are required for high frequency behaviours (such as vocalisations and gazing to caregiver's face) to determine if gaze coordination occurs more than expected by chance. In the following chapters, we opt to determine chance rates for all behaviours (vocalisations, gestures and combinations) in how much they co-occur with gaze to caregiver's face, to determine if this exceeds chance rates. If infants gaze to their caregiver's face whilst engaging in vocalisations, gestures or gesture-vocal combinations above the level predicted by chance then it is likely that they are *deliberately* coordinating gaze with these behaviours. In these cases, it is possible that instances of such coordination might be taken as evidence of communicative intent. If infants do not gaze to their caregiver's face whilst engaging in these behaviours above the amount that chance alone would predict, it is not possible to determine if infants are deliberately coordinating gaze with these behaviours. In these cases, we should be cautious of inferring that these behaviours are intentionally communicative, in the absence of other compelling evidence.

### Frequency and Proportion Measures

Traditionally, researchers have looked at proportions or frequencies (but not both) of vocalisations or gestures produced by the infant that are coordinated with gaze. However, this is problematic. Vocalisations are much more frequently produced than gestures and gesture-vocal combinations in naturalistic contexts at this age (Miller & Lossia, 2013; Spencer, 1993). Vocalisations are produced for a higher variety of non-social reasons (i.e., vocal play, emotional arousal, involuntary grunting such as during exertion) than gestures; the production of gestures (and therefore gesture-vocal combinations) is more tightly limited to social interactions. If only a subset of vocalisations are used in social interactions, then the proportion of vocalisations with gaze coordination seems an inappropriate measure to compare with gestures and gesture-vocal combinations (as this will necessarily be lower for vocalisations than for gestures and gesture-vocal combinations). Instead, the *frequency* of vocalisations with gaze coordination may be a more appropriate measure. However, because gestures and gesture-vocal combinations are more infrequent, the frequency of these with gaze coordination may not be the most appropriate measure for determining if it is more likely that vocalisations, gestures and gesture-vocal combinations are produced intentionally. Additionally, in order for proportions to be calculated at least one instance of that behaviour has to be observed. The danger is that by only including infants who produce at least one instance of a specific behaviour of interest might select for some sort of precociousness.

Throughout the following chapters, we opt to investigate both frequency and proportion of vocalisations and gestures that are produced with gaze coordination (or responded to by caregivers), to determine whether either approach yields substantially different answers to whether early intentional communication is more likely to be gestural, vocal, or gestural-vocal.

### Bouts

In Chapter 5, we will also determine whether the results of our analyses are radically different if we consider individual vocalisations, gestures or gesture-vocal combinations or 'bouts' of these behaviours. Vocalisations are more frequent, and are often repeated in close temporal synchrony, with distinct vocalisations given within a short time of a previous vocalisation ending. If a child engages in a number of these vocalisations whilst engaged in a sustained look to their caregiver's face, we would report that a number of vocalisations occurred with gaze coordination in these instances. This is not necessarily problematic (or untruthful), but it is as coherent to argue that really there has been one 'event' or bout of vocalisations with gaze coordination as it is to say that a specific number of vocalisations were given with gaze coordination. Furthermore, infants might gaze to their caregiver's face prior to a vocalisation or gesture in order to establish that their caregiver is attending to their actions, and then continue to produce more vocalisations or gestures *without* monitoring their caregiver's attention further. Again in this case, we might want to argue that the entire sequence of vocalisations or gestures has been one *bout* of behaviours with gaze coordination, not one vocalisation or gesture with gaze coordination, and more without. Depending on whether bouts or individual behaviours are used, the frequency and proportion of behaviours that have gaze coordination will be affected.

In the following chapter, we will explore both traditional analyses that consider individual behaviours and also consider bouts of behaviours, which can be solely vocalisations or solely gestures (unimodal bouts), or involve both gestures and vocalisations (multimodal bouts - necessarily gesture-vocal combinations would be included in this category). This will inform us whether either approach yields substantially different answers to whether early intentional communication is more likely to be gestural, vocal, or gestural-vocal.

**Caregiver Responsiveness**

As demonstrated above, different infant behaviours are responded to in different ways, and this may in turn affect language learning. In Chapter 7 we opt to investigate caregiver responses that are both temporally contingent on infant's vocalisations, gestures and combinations *and* semantically contingent on what the child is attending to, as these have been demonstrated to be especially key to language learning.

## 5. Are Infants' Vocalisations, Gestures and Gesture-vocal Combinations Intentionally Communicative Around Their First Birthday?

**Abstract**

Prelinguistic vocalisations and gestures are predictive of infants' early language abilities. This could be because they are early instances of intentional communication. In this chapter, we sought to determine whether prelinguistic gestures and vocalisations were produced in an intentionally communicative way (produced with gaze to their caregiver's face) at 11 and 12 months. We addressed two main questions. Firstly, we considered whether infants' vocalisations, gestures or gesture-vocal combinations are coordinated with gaze to caregiver's face. We found that at 11 months infant's vocalisations, gestures and combinations in general co-occurred with gaze to caregiver's face significantly above the levels predicted by chance (although fine-grained analysis of gesture and vocalisation types revealed that this was not the case for some individual vocalisation and gesture types). At 12 months however, infant's vocalisations did not co-occur with gaze to caregiver's face above chance levels, while their gestures and combinations did. Secondly, we considered whether gaze coordination is more likely to occur with vocalisations or gestures. We found that infants more *frequently* produced vocalisations rather than gestures or combinations that were gaze-coordinated. However, a higher *proportion* of infant gestures and combinations than vocalisations were gaze-coordinated. We found no differences in whether vocalisations and gestures are conceptualized as *bouts* or individual behaviours, and no major differences in whether combinations are treated as mutually exclusive category from either gestures or vocalisations (produced alone) or included within those categories. Finally, we found no evidence to suggest that the age at which index-finger pointing emerges (12 months) heralds an increase in potential instances of intentional communication.

### Introduction

As the previous review chapter shows, there is some evidence to suggest that around one year of age, infants are using both vocalisations and gestures in potentially intentionally communicative ways. However, there are a number of shortcomings in previous research.

Firstly, there is a shortage of studies that consider both vocalisations *and* gestures (and indeed combinations of the two), as often they are studied in isolation. Secondly, there is a lack of *naturalistic* studies involving *large samples* investigating prelinguistic vocalisations, gestures and gesture-vocal combinations. Finally, few studies use the same measure for indicating communicative intent (coordination with gaze) for vocalisations, gestures and gesture-vocal combinations. The current study aimed to investigate infants' vocalisations, gestures and gesture-vocal combinations around the end of the first year of life to determine if they are intentionally communicative. In doing so, we seek to provide an indication of when infants might be making a transition to intentional communication.

The current study addresses limitations of previous research, by using the same token (coordination of gaze to the caregiver's face) to measure potential instances of intentional communication across vocalisations, gestures and combinations. We focus on a number of infant gestures including index-finger pointing, open-hand pointing, showing, giving and conventional gestures. We distinguish between two kinds of prelinguistic vocalisations, those that include both a consonant (C) and vowel (V) and are thus more speech-like, and those that do not (non-CV vocalisations). We also opt to treat gesture-vocal combinations as mutually exclusive behaviours to vocalisations and gestures. To provide as ecologically valid picture as possible, we focus on such behaviours in a naturalistic environment (in the home with the primary caregiver) to determine naturalistic frequencies of vocalisations, gestures and gesture-vocal combinations, and the frequency and proportion of these behaviours that are coordinated with gaze.

This study specifically aims to address two main questions:

(1) Do infants coordinate their vocalisations, gestures or combinations with gaze to caregiver's face above the rate expected by chance?

(2) Is coordination of gaze to caregiver's face more likely with vocalisations, gestures or combinations?

Regarding the first question, we aim to determine whether it is likely that infants deliberately coordinate their gestures or vocalisations with gaze to their caregiver's face. While coordination above chance levels does not necessarily indicate that such coordination is intentional, it provides stronger evidence that *when this coordination occurs* the infant might be communicating with the intention of affecting others' attention. Conversely coordination at, or below chance does not necessarily indicate that when such coordination occurs, it is not an attempt to intentionally communicate, but it provides reason to be cautious in overinterpretation of these instances as intentional communication.

It is plausible that infants' gesture-vocal combinations may be intentionally communicative in the absence of gaze coordination, as infants might use vocalisations to draw caregivers' attention to their gestures. It is also plausible however, that such combinations could happen by chance if the frequency of vocalisations and gestures are high enough. We will therefore provide a similar analysis for gesture-vocal combinations as with gaze coordination, to evaluate whether these too occur above the level predicted by chance. This will allow us to evaluate whether infants may be communicating with the intention of affecting others' attention when they produce combinations more globally.

Regarding the second question, we aim to determine whether vocalisations, gestures or combinations are coordinated with gaze to caregiver's face differently, and whether one form is privileged in terms of intentional communication. As we have seen, there are different ways to characterise rates of coordination. In prior developmental research, researchers have measured either the *proportion* of all instances of a behaviour that are coordinated with gaze, or at the *frequency* of all instances of a behaviour that are coordinated with gaze (but not both in the same study). Since both are informative in different ways, we will compare analyses considering frequency and proportion. There are also different levels of granularity at which raw behaviours are quantified. In some developmental research (e.g., Gros-Louis et al., 2006), but more commonly in primatological research (e.g., Schel, Machanda, et al., 2013), *bouts* have been the chosen unit of measurement: e.g., vocalisations repeated in close temporal synchrony followed by a pause are considered a single bout. One can thus consider frequency and proportional measures for bouts with gaze coordination. Since vocalisations are much more frequently produced than gestures and combinations in naturalistic contexts at this age (Miller & Lossia, 2013; Spencer, 1993), and may be given in close temporal synchrony, using different measures (e.g., frequency versus proportion, individual behaviours versus bouts) will potentially paint quite different pictures of coordination. We will therefore consider all approaches to check whether they yield substantially different answers.

We predict that vocalisations (or unimodal vocal bouts) with gaze coordination will be much more frequent than gestures and combinations (or unimodal gestural and multimodal bouts) with gaze coordination. Conversely, we predict that gestures and combinations (or unimodal gestural and multimodal bouts) will be proportionally more likely to be gaze-coordinated than vocalisations (or unimodal vocal bouts).

In the current study we opt to look at infants' behaviour at 11 and 12 months. There are a number of prevailing theories that suggest that the onset of index-finger pointing around 12 months heralds the onset of intentional communication in infancy. This could be because infants produce intentionally communicative index-finger pointing as a result of a pragmatic realisation about intentional communication, namely that they can direct the attention of others in order to achieve their own goals. However, as reviewed in the previous chapter, other gestures such as showing and giving emerge prior to pointing (around 10 to 12 months), and infants produce babbling vocalisations from much earlier (around 6 to 10 months), both of which could also be produced in an intentionally communicatively way. In order to assess whether intentional communication could plausibly arise prior to 12 months and the onset of index-finger pointing, in infants' earlier gestures and vocalisations, we opt to look at 11 months as this is an age where we can expect an appreciable frequency of these earlier gestures. By looking at development from 11 to 12 months, we can also see whether there is a marked increase in gaze coordination for infants' vocalisations and gestures, which coincides with the onset of index-finger pointing.

## Method

### Participants

Participants were drawn from a larger sample of families ($N$ = 140) who had previously participated in a longitudinal randomised control study (McGillion, Pine, et al., 2017). We selected participants at two age points from this sample; 11 and 12 months.

**11 months.** The 11 month sample included *n* = 134 caregiver-infant dyads (70 female infants, 64 male) for whom we had naturalistic videos when the infants were 11 months of age (*M* = 334 days, *SD* = 4 days). Two families were excluded because the play interaction was not dyadic (i.e., there was a third individual present), and a further 4 dyads were excluded for being in shot for less than 7 minutes.

**12 months.** The 12 month sample included *n* = 64 caregiver-infant dyads (35 female infants, 29 male) for whom we had naturalistic videos of when the infants were 12 months of age (*M* = 365 days, *SD* = 5 days). This included 3 participants for whom we did not have naturalistic data at 11 months. Only families included in the control condition of the original study were included at 12 months, as there was an effect of the intervention on the frequency of infants' vocalisations produced at 12 months in the intervention condition. Thus, we excluded 72 individuals who were part of the language intervention condition. Two dyads did not take part in the study at 12 months and a further 2 dyads were excluded for being in shot for less than 7 minutes.

**Procedure**

Participants were filmed from two camera angles in an unstructured play session in their home lasting from 10 to 15 minutes (McGillion, Pine, et al., 2017). The researcher who set up the recording was not present during the play session, having left the room after setting up the cameras. Only the caregiver-infant dyad was present for the duration of the video. Coding of the naturalistic videos was undertaken in ELAN (Sloetjes & Wittenburg, 2008). At 11 months 100% of videos were coded by the first author (with 10% double-coded by a research assistant in order to obtain reliabilities). For the videos at 12 months, the research assistant and the first author each coded 55% of the videos, thus 10% of videos were double-coded in order to obtain reliabilities.

**Coding**

**Observation period**. From the videoed session, 10 minutes were selected for coding. This period began from the moment the researcher left the room until 10 minutes later (excluding time offshot), or until the experimenter returned (if this was prior to 10 minutes being reached). For 9 participants, observation time was below 10 minutes (but above 7 minutes), so adjusted frequencies of behaviours were used throughout the analyses. These were calculated as follows:

$$frequency \times \left(\frac{10\ minutes}{duration\ of\ observation\ time}\right)$$

**Gaze to caregiver's face.** All instances where the infant looked to the caregiver's face were coded. These were marked from the frame that was judged to be the beginning of the look, to the last frame where the infant was judged to be looking at their caregiver's face.

**Vocalisations.** Initially, all infant vocalisations were coded except crying vocalisations, vegetative noises, and fussing noises. This is line with much of the existing literature that looks at the co-occurrence between infant gaze and vocalisations (D'Odorico & Cassibba, 1995; D'Odorico & Franco, 1991; Harding & Golinkoff, 1979; Legerstee & Barillas, 2003; Messinger & Fogel, 1998; Murillo & Belinchón, 2012; Wu & Gros-Louis, 2014, 2015). For all vocalisations, the beginning of the vocalisation was marked at the frame where the vocalisation began, and the end was marked at the last frame where the vocalisation was still audible.

In the literature, criteria for defining *separate* vocalisations differ widely, with most authors agreeing some sort of 'perceivable silence' (Wu & Gros-Louis, 2015) is necessary to delineate distinct vocalisations. How this silence can be quantified varies, with some choosing 50ms (Esteve-Gibert & Prieto, 2013; Papaeliou & Trevarthen, 2006), 300ms (D'Odorico & Franco, 1991), 500ms (Grünloh & Liszkowski, 2015), up to 1000ms periods of silence (or with audible/visible breath) to separate vocalisations (Fagan, 2009; Marchand et al., 1994; Murillo & Belinchón, 2012). When more detailed analysis of infant phonology are carried out, vocalisations can be delineated by breath control, timing and a unified intonation contour (Vihman & Ferguson, 1986; Vihman et al., 1985), or by syllables (Goldstein et al., 2010; Majorano & D'Odorico, 2011), but this is too fine-grained for our purposes. Therefore we coded vocalisations as separate when there was a perceivable silence of 200ms or more between them.

***CV Vocalisations***: We coded all vocalisations for whether they involved a syllable containing at least one consonant (C) and vowel (V). This is consistent with many studies in the literature that have examined the coordination of gaze and vocalisations (Grünloh & Liszkowski, 2015; Miller & Lossia, 2013; Spencer, 1993), and importantly with those who look at predictive links between early vocalisations and later language (D'Odorico et al., 1999; Menyuk et al., 1986). In line with McCune and Vihman (2001), we code only supraglottal consonants as consonants, excluding glides and glottals. All vocalisations that did not contain a CV syllable were coded as *non-CV vocalisations*.

**Gestures.** All occurrences of pointing and holdout gestures known to be used by infants around this age were coded (*index finger point, whole hand point, show* and *give*). Additionally, we coded a number of *conventional gestures*, (including *arm up, wave, all gone,* and instances of *baby sign*). While not an exhaustive set of infant gestures, any remaining types were so rare as to not warrant coding. For all these gestures, the beginning of the gesture was marked at the frame where arm reached maximum extension, and the end is marked at the frame where retraction of the arm began. To create continuity with the vocalisation coding scheme, if the arm was extended within 200ms of the previous arm retraction, this is counted as the same gesture. We used the following definitions for each gesture type.

***Pointing***. While looking at an object or event of interest, the infant extended left or right hand (or both). The arm(s) had to be extended, the hand(s) for the gesture had to be empty, and the child was not leaning forward and did not touch what was being pointed at (Matthews et al., 2012; McGillion, Herbert, et al., 2017). For *index-finger points*, the index finger(s) was clearly and visibly separate from the other fingers, which were partially or entirely curled back, and the index finger extended in the direction of the object or event being looked at. For *open-hand points*, a majority of fingers were extended in the direction of the object or event being looked at.

***Holdout***. While holding an object with one or both hands, the infant held out an object with their arm (or arms) extended towards the caregiver. For a *show*, the object was held up towards the caregiver's face, while for a *give* the object was extended in the direction of the caregiver's hands, or extended in a way so as to deliver the object into the vicinity of the caregiver (Cameron-Faulkner et al., 2015; Carpenter et al., 1998).

***Conventional gestures.*** Due to low frequencies, all the following gestures are included as *conventional gestures*. For *arm up*, the infant raised both arms in order to initiate being picked up. For *wave,* the infant waved with palm vertical (or close to vertical) and moving side to side. For *all gone*, the infant shrugged with palm of hand facing up, similar to adults asking, 'where?'. All instances of *baby sign* were also coded.

**Offshot and availability of modalities**. In order to be able to code the same length of video for all infants and to be able to calculate proportions, it was important to mark any sections of video where it was not possible to code the infant's behaviour. We coded the following three off-shot measures.

***Offshot.*** Where participants were completely out of shot.

***Unable to code gaze to caregiver's face***. This code was used for all the time that it was not possible to tell if the infant was looking to the caregiver's face. A few examples of this are (1) when infant's eyes were not in shot, (2) position of caregiver's face was not known (3) the infant was looking in the direction of caregiver's face, but there was some partial occluder between caregiver and infant that made it impossible to tell if the infant was looking to the caregiver's face (i.e., unclear whether they had a direct line of sight). This category was exclusively for when it was *possible* that the infant could have looked to the caregiver's face, but it *was not possible* to conclusively determine if they were or not. For instance, in some cases, if infant's eyes were not in shot (i.e., they were looking straight down, with head bent), this category was not marked, as it was clear that the infant was not capable of gazing to the caregiver's face. Equally, if the participant was facing away from the camera (i.e., only see the back of their head could be seen), but the caregiver was behind them, this category was not marked, as the participant could not possibly have gazed to their caregiver's face.

***Unable to code gestures***. This code was used for all the time where the infant's arms (or one arm) were not visible, and it was possible that they could have gestured, i.e., if the infant had gestured, it would not have been seen.

For any period with these categories marked, we took no data. This is because if gestures were unavailable for a period, we could not rule out the possibility that any vocalisations produced during that period were not gesture-vocal combinations. Additionally, if we did not know whether a gesture, vocalisation or combination was accompanied by gaze to the caregiver or not, this was not useful data.

**Measures**

**Gesture-vocal combinations.** When all or part of a vocalisation and gesture overlapped in time, this was considered a gesture-vocal combination (see also Igualada et al., 2015). Combinations could involve either CV or non-CV vocalisations. In cases where they involved both CV and non-CV vocalisations, they were counted as involving CV vocalisations. Combinations could also involve any of the five gesture types. No instances of combinations involving two different gesture types was observed. This gave us 10 types of gesture-vocal combination (2 vocalisation types x 5 gesture types).

**Bouts**. For the main analyses only (but not fine-grained analyses) concerning frequency and proportion of vocalisations, gestures and combinations that were coordinated with gaze, we consider bouts. These are when vocalisations, gestures or combinations are given with 1000ms of another vocalisation, gesture or combination ending. We classified three kinds of bouts.

***Unimodal vocal bouts.*** These only contained one or more vocalisations.

***Unimodal gestural bouts.*** These only contained one or more gestures.

***Multimodal bouts***. These contained both one or more gestures and one or more vocalisations (combinations are necessarily included in this category).

**Temporal window of co-occurrence of gaze to caregiver's face.** Studies looking at gaze coordination with vocalisations, gestures or combinations normally focus on the period co-occurring exactly with the start and end of the behaviour of interest, and a short temporal window before or after it. This varies across studies. Many researchers choose 1000ms (Desrochers et al., 1995; Matthews et al., 2012; Murillo & Belinchón, 2012; Wu & Gros-Louis, 2014), but others use 2000ms (D'Odorico & Franco, 1991; Franco & Butterworth, 1996; Legerstee & Barillas, 2003; Messinger & Fogel, 1998; Spencer, 1993; Wu & Gros-Louis, 2015). Alternatively, D'Odorico & Cassibba (1995) code gaze checking in two 500ms periods (if gaze to the face occurred for a minimum of 300ms of that interval) following or preceding a 500ms period containing a vocalisation (where it occupied a minimum of 300ms of that period), theoretically including gaze checks up to 900ms after the vocalisation. However, this solution was born out of the limits of the available technology, and we can now code frame-by-frame with relative ease. For this study, we choose a 1000ms temporal window prior to, or following each vocalisation, gesture or gesture-vocal combination in which to code 'co-occurrence' of gaze to caregiver's face. Only vocalisations, gestures or combinations with the full temporal windows were included for the analysis.

**Expected and observed co-occurrence of gaze to caregiver's face with vocalisations, gestures and combinations.** For the analysis relating to whether vocalisations, gestures or combinations co-occurred with gaze to caregiver's face above the rate that would be expected by chance, first we needed to calculate the *expected* (chance) rate of co-occurrence for each type of behaviour. To take the example of vocalisations, we calculated the proportion of time each infant spent vocalising and the proportion of time they spent gazing to the caregiver and then multiplied these to obtain the expected rate of co-occurrence. A slight modification to this procedure was necessary due to the fact that we counted gaze to caregiver's face as co-occurring if they happened within a 1 second window of the vocalisation. Thus the proportion of time spent vocalising was taken to be the time spent vocalising plus the 1 second windows before and after the vocalisations. The 1 second windows sometimes overlapped with another vocalisation, or the window of another vocalisation (see Figure 11), and so it was important not to double count this overlapping time. In these cases, we counted the intervening time between the behaviours only once.

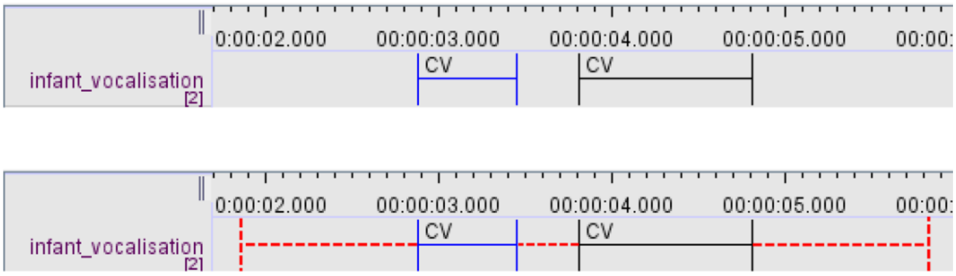


*Figure 11*. Top: Two vocalisations coded on the timeline (separated by *less* than 2s). Bottom: Annotated with temporal windows (indicated with red dotted line).

To calculate the *observed* rate of co-occurrence of gaze with vocalisations, we simply took the total duration for which gaze to caregiver's face co-occurred with vocalisations and the 1 second windows around them (see Figure 12) and divided this by the total time. Again, we made sure not to double count any time due to overlapping (see Appendix C for exhaustive scenarios). The same procedure was repeated for gestures and combinations.

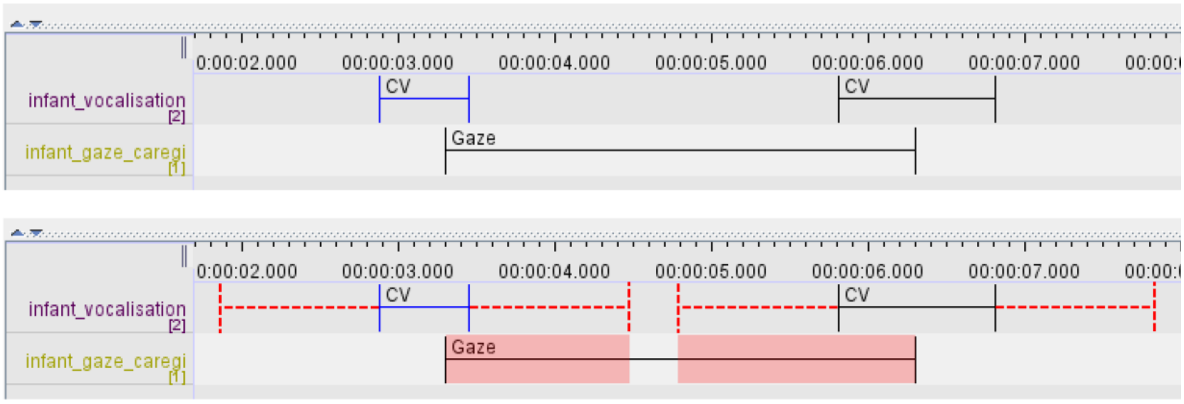


*Figure 12*. Top: Two vocalisations and one gaze to caregiver's face coded on the timeline. Bottom: Annotated with temporal windows (indicated by red dotted line). Area shaded red is counted as gaze to caregiver's face co-occurring with vocalisations.

**Expected and observed co-occurrence of vocalisations with gestures (combinations).** The same procedure for calculating expected and observed rates was used to obtain measures of the co-occurrence of gestures and vocalisations. Since no 1 second temporal windows were used for this analysis, we did not need to take into account potential overlap.

**Frequency and proportion of vocalisations, gestures and combinations with gaze coordination.** For the second part of the analyses, we calculated the frequency of each type of vocalisation, gesture or combination (or bouts) with gaze coordination and then the proportion of their vocalisations, gestures or combinations (or bouts) that had gaze coordination for each participant.

**Analyses**

All analyses were conducted in R (R Core Team, 2015).

**Comparing chance to observed rates.** To compare the expected duration of co-occurrence of two behaviours to the observed co-occurrence of those behaviours, we used a t-test.

**Analysing frequency of occurrence**. To determine if there were differences in the frequency with which types of behaviours were produced, we constructed linear regression models using "lme" function in "nlme" package (Pinheiro, Bates, Debroy, Sarkar, & Team, 2015). Initially, we constructed a null model, and then added behaviour type (e.g., was the behaviour a vocalisation, gesture or combination) as a predictor. We report if there was a significant improvement on the null model when adding behaviour type as a predictor, determined using "anova" function, part of the "stats" package (R Core Team, 2015). Post-hoc comparisons were conducted using "lsmeans" function in "lsmeans" package (Lenth, 2016).

When there were only two behaviour types to compare (e.g., vocalisation types), we conducted t-tests.

For models that investigated developmental change in frequencies of behaviours, the method was the same, except that age (11 or 12 months) was also added as a predictor. We first added behaviour type, then age, then the interaction term (behaviour type * age) and tested for improvement of fit.

**Analysing proportions.** To determine if there were differences in the proportion that types of behaviours were produced with gaze coordination, we constructed logistic regression models using "glmer" function in "lme4" (D. Bates, Mächler, Bolker, & Walker, 2015). For each behaviour (e.g each vocalisation, gesture or combination), the outcome variable was whether it was gaze-coordinated (1= gaze-coordinated, 0 = not). We included behaviour type (e.g., was the behaviour a vocalisation, gesture or combination) as a fixed effect. Infant was included as a random effect on the intercept and on the slope (behaviour type), except when model comparison established that this was unnecessary. Initially, we constructed a null model, and then added behaviour type as a predictor. Again, we report if there was a significant improvement on the null model when adding behaviour type as a predictor, determined using the "anova" function. Post-hoc comparisons were again conducted using "lsmeans" function.

For models that investigated developmental change in proportions of behaviours that were gaze-coordinated, the method was the same, except that both behaviour type and age (11 or 12 months) were fixed effects. Infant was included as a random effect on the intercept and on the slope (behaviour type * age) except when model comparison deemed that this was unnecessary. We first added behaviour type, then age, then the interaction term (behaviour type * age) and tested for improvement of fit. Note, we lacked the required number of observations to test developmental change for analyses of sub types of gestures or combinations.

**Reliabilities**

For 11 month reliabilities, a trained research assistant blind to the aims of the study coded gaze to caregiver's face, vocalisations and gestures for 10% of participants ($n$ = 14). Reliabilities for 12 month olds were calculated from a 10% (of the full sample, $n$ = 13) overlap in coding between the first author and the research assistant. Agreement on the frequency of these behaviours was high at both 11 months (for gaze to caregiver's face, $r$ = .95; for vocalisations, $r$ = .99; for gestures, $r$ = .82; for combinations, $r$ = .93) and 12 months (for gaze to caregiver's face, $r$ = .95; for vocalisations, $r$ = .98; for gestures, $r$ = .97; for combinations, $r$ = .95).

Additionally, we tested whether the frequency of vocalisations, gestures and gesture-vocal combinations with gaze coordination was reliable, and again, agreement was high at both 11 months (for vocalisations, $r$ = .95; for gestures, $r$ = .89; for combinations, $r$ = .94) and 12 months (for vocalisations, $r$ = .96; for gestures, $r$ = .96; for combinations, $r$ = .94).

For agreed vocalisations, gestures and combinations, Cohen's kappa was calculated for gaze coordination (was the behaviour coordinated with gaze or not), and indicated high levels of agreement at both 11 and 12 months. At 11 months, Cohen's kappa was high for vocalisations, $\kappa = .82$, $p < .001$ (agreement on coding of 96%); for gestures, $\kappa = .86$, $p < .001$ (93%); and for combinations, $\kappa = .77$, $p = .013$ (89%). At 12 months, Cohen's kappa was high for vocalisations, $\kappa = .85$, $p < .001$ (96%); for gestures, $\kappa = .97$, $p < .001$ (98%); and for combinations, $\kappa = .92$, $p < .001$ (96%).

In terms of gesture type coding, we intended to calculate kappas on gesture type (whether gestures were classified as index-finger pointing, open-hand pointing, giving, showing or conventional gestures) on agreed gestures, however there was 100% agreement at 11 months. At 12 months, Cohen's kappa for gestures was, $\kappa = 0.85$, $p < .001$ (agreement on coding of 90%), indicating excellent agreement.

Finally, for vocalisation type coding (whether they were classified as CV or non-CV), a separate phonologically trained researcher (MMG) independently classified vocalisations for 10% of the sample. Cohen's kappa for vocalisations at 11 months indicated excellent agreement, $\kappa = .80$, $p < .001$ (agreement on coding of 91%), as did Cohen's kappa at 12 months, $\kappa = .81$, (agreement on coding of 91%).

## Results

At 11 months, all infants gazed to their caregiver's face, all produced non-CV vocalisations, and 97% ($n = 130$) produced at least one CV vocalisation. Fewer infants produced gestures (either alone or in a gesture-vocal combination), with 67% ($n = 90$) producing one or more gestures. Most commonly, infants produced give gestures (produced by 36% of infants, $n = 48$), but a number also produced show gestures (22%, $n = 30$), index finger pointing (21%, $n = 28$), open hand pointing (20%, $n = 27$) and conventional gestures (19%, $n = 25$). Forty percent ($n = 53$) produced combinations.

At 12 months, again all infants gazed to their caregiver's face and all produced non-CV vocalisations, while 98% ($n = 63$) produced at least one CV vocalisation. Infants also produced gestures, with 81% ($n$ = 52) producing one or more gestures. Most commonly, infants produced give gestures (produced by 52% of infants, $n$ = 33), and index finger pointing (34%, $n$ = 22), but a number also produced show gestures (25%, $n$ = 16), conventional gestures (17%, $n$ = 11), and open hand pointing (17%, $n$ = 11). Fifty-six percent ($n$ = 36) of infants produced combinations.

Descriptive statistics for infant behaviours at 11 and 12 months are presented in Table 15 (combinations by both gesture *and* vocalisation type are reported in full in Appendix D, Table 39). For the following analyses, we treated combinations as a separate, mutually exclusive category from gesture and vocalisations. However, to check that this decision did not substantially affect the results, we also conducted inclusive analyses with only vocalisations and gestures (making no distinction for those involved in combinations). These inclusive analyses are reported in full in Appendix E and we note below whenever they differ from the reported exclusive analyses.

Table 15

*Average Frequency of Vocalisations, Gestures and Combinations from 10 minutes Observation at 11 (n = 134) and 12 months (n = 64)*

| | *11 months (n = 134)* | | | | *12 months (n = 64)* | | | |
|---|---|---|---|---|---|---|---|---|
| *Behaviour* | *M* | *SD* | *Median* | *Range* | *M* | *SD* | *Median* | *Range* |
| **Gaze to Caregiver's Face** | 22.21 | 12.11 | 19 | 1-53 | 22.18 | 13.68 | 19 | 1-63 |
| **Vocalisations (without gesture)** | 47.50 | 28.82 | 41.5 | 4-172 | 39.28 | 24.82 | 37 | 7-128 |
| CV | 18.45 | 17.94 | 14 | 0-108 | 16.30 | 12.74 | 12.5 | 0-50 |
| Non-CV | 29.05 | 17.49 | 25.5 | 2-82 | 22.98 | 18.80 | 17.5 | 2-123 |
| **Gestures (without vocalisation)** | 1.64 | 2.26 | 1 | 0-11 | 2.79 | 4.07 | 1 | 0-20 |
| Index-finger point | 0.22 | 0.74 | 0 | 0-7 | 0.53 | 1.44 | 0 | 0-9 |
| Open-hand point | 0.20 | 0.58 | 0 | 0-4 | 0.05 | 0.21 | 0 | 0-1 |
| Give | 0.68 | 1.41 | 0 | 0-8 | 1.88 | 3.32 | 0 | 0-15 |
| Show | 0.29 | 0.77 | 0 | 0-4 | 0.20 | 0.51 | 0 | 0-2 |
| Conventional gesture | 0.25 | 0.80 | 0 | 0-6 | 0.13 | 0.41 | 0 | 0-2 |
| **Gesture-vocal combinations** | 1.10 | 2.13 | 0 | 0-12 | 1.78 | 2.91 | 1 | 0-14 |
| CV | 0.63 | 1.45 | 0 | 0-9 | 0.94 | 1.96 | 0 | 0-12 |
| Non-CV | 0.47 | 1.15 | 0 | 0-8 | 0.84 | 1.86 | 0 | 0-13 |
| Index-finger point | 0.31 | 1.32 | 0 | 0-12 | 0.48 | 1.63 | 0 | 0-12 |
| Open-hand point | 0.19 | 0.96 | 0 | 0-10 | 0.19 | 0.50 | 0 | 0-2 |
| Give | 0.34 | 0.99 | 0 | 0-8 | 0.83 | 1.84 | 0 | 0-12 |
| Show | 0.16 | 0.58 | 0 | 0-3 | 0.20 | 0.47 | 0 | 0-2 |
| Conventional gesture | 0.10 | 0.36 | 0 | 0-2 | 0.08 | 0.27 | 0 | 0-1 |
| **Unimodal vocal bouts** | 35.94 | 18.47 | 33.5 | 4-94 | 30.01 | 16.89 | 28.5 | 6-76 |
| **Unimodal gestural bouts** | 1.33 | 2.01 | 1 | 0-11 | 2.29 | 3.55 | 1 | 0-18 |
| **Multimodal bouts** | 1.37 | 2.33 | 0 | 0-11 | 2.00 | 2.90 | 1 | 0-14 |

It is noteworthy that the frequency of different behaviours is remarkably stable across the two measurement points. There were however some instances of significant developmental change. For the infants whose measures are reported at both time points, regression analyses and follow up pair-wise comparisons revealed that the frequency of vocalisations significantly *decreased* between 11 and 12 months ($b = -7.62$, $t(240) = -3.23$, $p = .030$, Appendix F, Table 56 for model) whereas the frequency of giving significantly *increased* between 11 and 12 months ($b = 1.05$, $t(300) = 4.74$, $p < .001$, Appendix F, Table 57 for model). Furthermore, the frequency of unimodal vocal bouts also significantly *decreased* between 11 and 12 months ($b = -5.75$, $t(180) = -3.69$, $p = .004$, Appendix F, Table 58 for model). It is worth noting that when gesture-vocal combinations were not treated as a separate category from vocalisations (as in Appendix E), there was no significant decrease in vocalisations between 11 and 12 months. This suggests instead that the rate that infants produce vocalisations alone (i.e., not in combination with a gesture) decreases, but the overall rate of vocalisations does not (i.e., more vocalisations are produced in gesture-vocal combinations).

### Do Infants' Vocalisations, Gestures and Combinations Co-occur with Gaze to Caregiver's Face Above the Rate Expected by Chance?

Figure 13 shows the expected and observed duration that vocalisations, gestures and combinations co-occurred with gaze to caregiver's face at both 11 and 12 months.

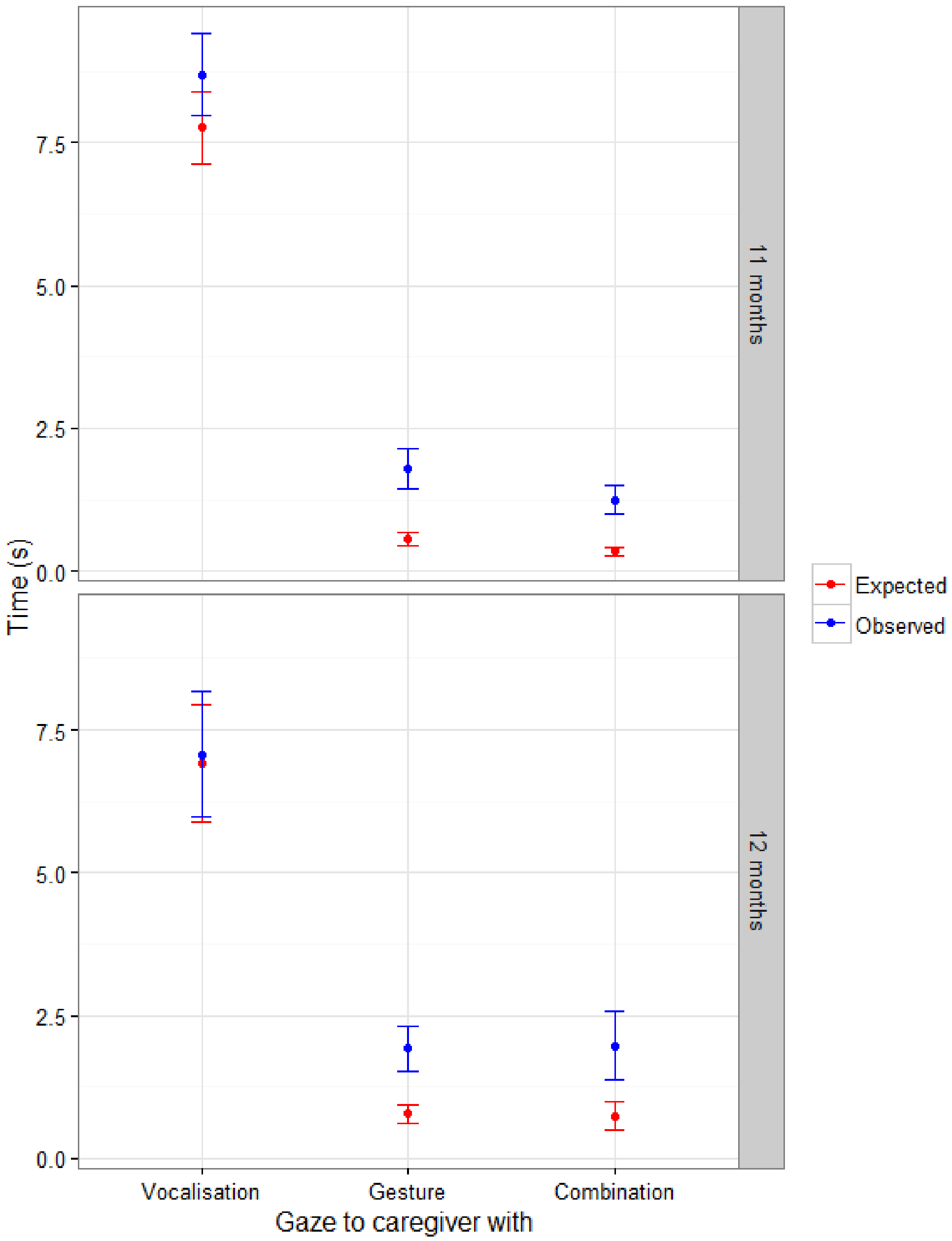


*Figure 13*. Mean expected and observed co-occurrence of vocalisations, gestures and gesture-vocal combinations with gaze to caregiver's face at 11 ($n$ = 134) and 12 months ($n$ = 64). Error bars represent standard error about the mean.

At 11 months, vocalisations, gestures and combinations co-occurred with gaze significantly above the level predicted by chance (vocalisations, $t(133) = 2.36$, $p = .020$, gestures, $t(133) = 4.85$, $p <.001$, combinations, $t(133) = 4.32$, $p <.001$). At 12 months however, only gestures and combinations co-occurred with gaze significantly above the level predicted by chance (gestures, $t(63) = 3.90$, $p <.001$, combinations, $t(63) = 3.02$, $p = .004$), while vocalisations did not. Inclusive analyses revealed the same picture except that vocalisations co-occurred with gaze significantly above chance at 12 months (see Appendix E, Figure 20). It should be noted that when we considered only data at 11 months from infants whom we had observations at both time points ($n$ = 61), the picture was the same as with the 12 month data, in that gestures and combinations, but not vocalisations, co-occurred with gaze significantly above the level predicted by chance (gestures, $t(60) = 2.97$, $p$ =.004, combinations, $t(60) = 2.98$, $p$ = .004). The same was true when we collapsed across age (see Appendix G, Figure 24). This suggests that this discrepancy between 11 and 12 months was as a result of sample size, not a genuine developmental effect.

Figures 14 and 15 shows the expected and observed duration that individual vocalisation and gesture types (both produced alone and as part of combinations) co-occurred with gaze to caregiver's face at 11 and 12 months.

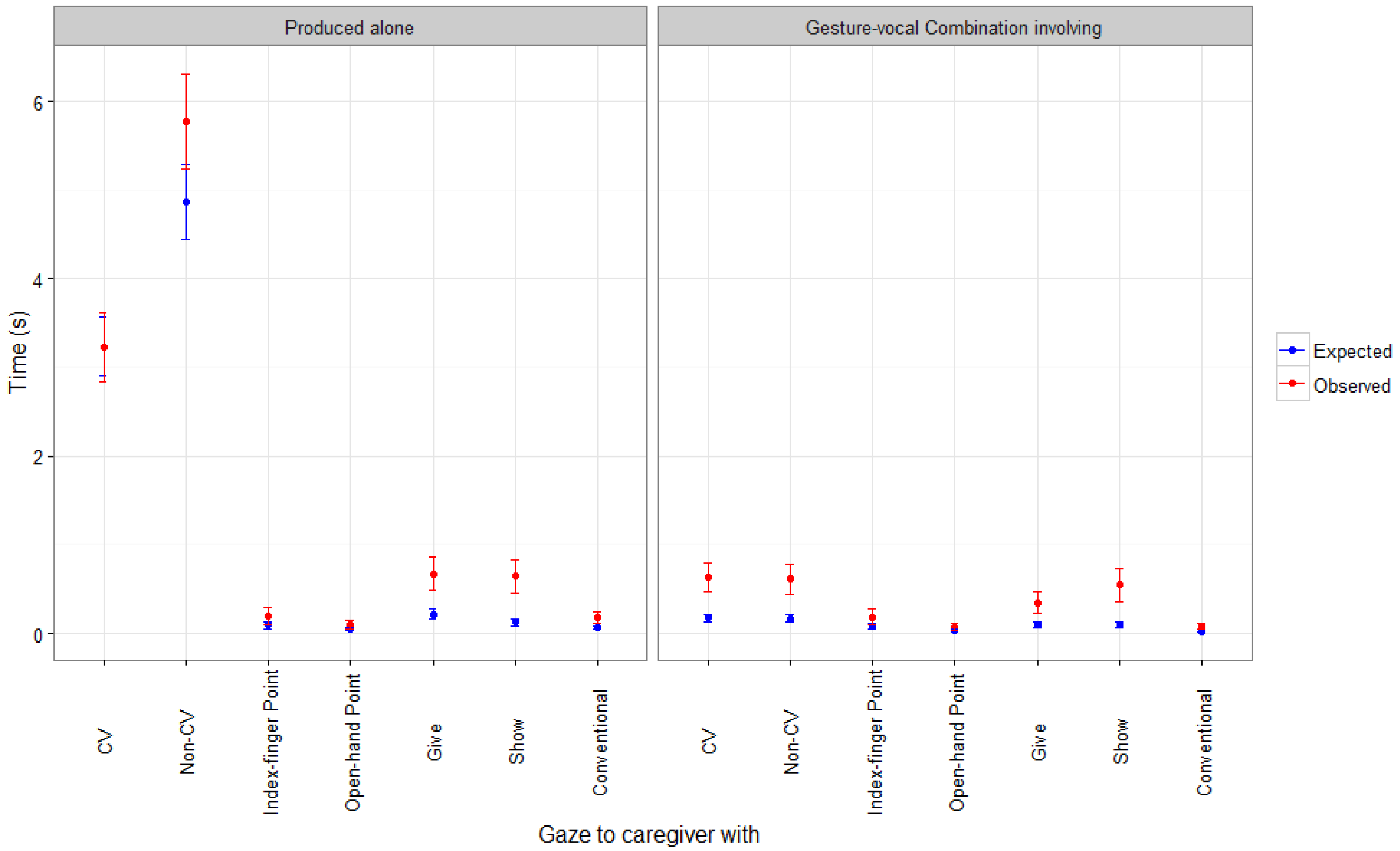


*Figure 14*. Mean expected and observed co-occurrence of individual vocalisation and gesture types (both produced alone and as part of gesture-vocal combinations) co-occurred with gaze to caregiver's face at 11 months ($n$ = 134). Error bars represent standard error about the mean

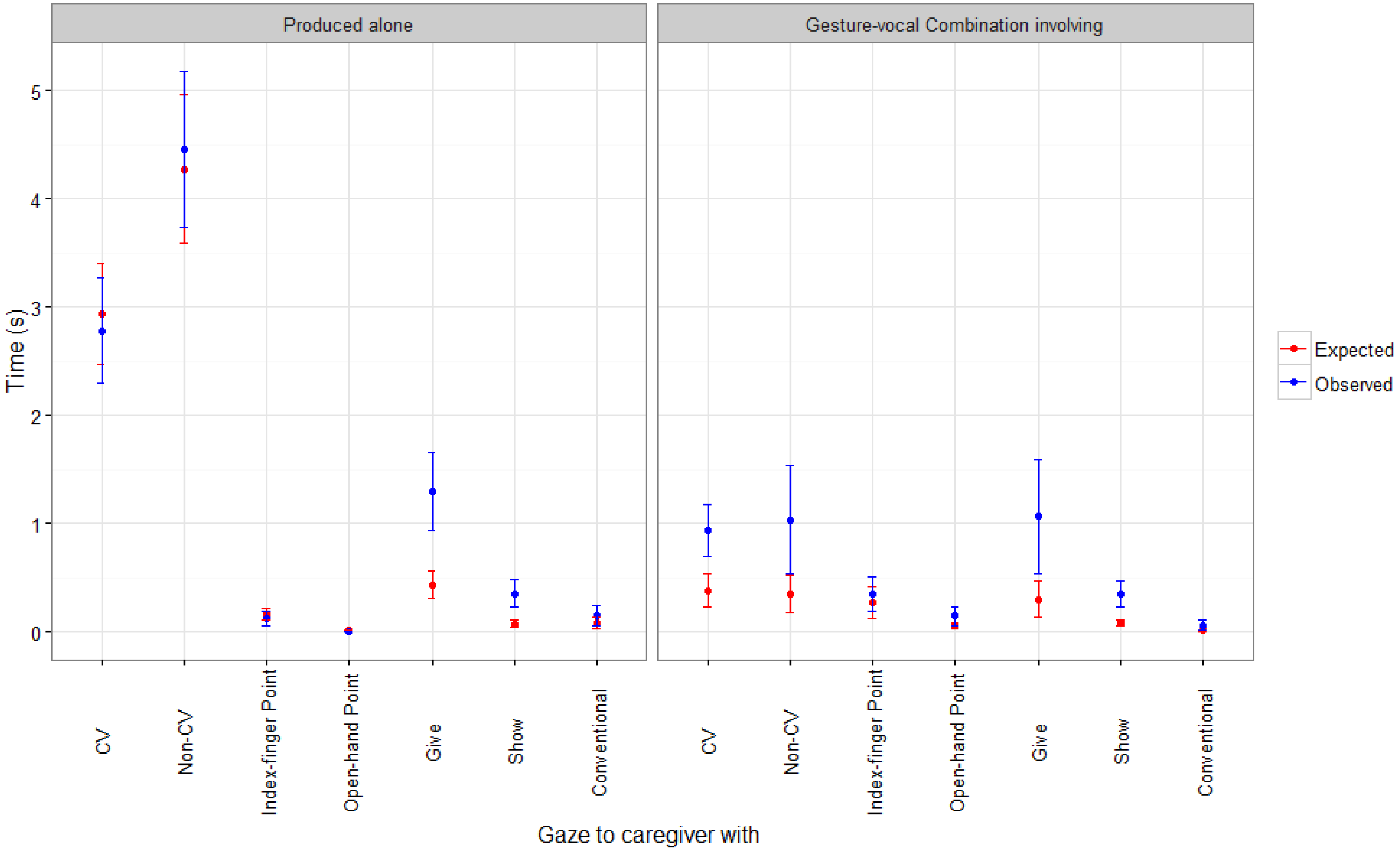


*Figure 15*. Mean expected and observed co-occurrence of individual vocalisation and gesture types (both produced alone and as part of gesture-vocal combinations) co-occurred with gaze to caregiver's face at 12 months ($n$ = 64). Error bars represent standard error about the mean

**Vocalisation type.** At 11 months, non-CV vocalisations co-occurred with gaze significantly above the level predicted by chance, $t(133) = 2.94$, $p = .004$, while CV vocalisations did not. At 12 months, however, neither CV nor non-CV vocalisations co-occurred with gaze significantly above the level predicted by chance. Again, when we considered only data at 11 months from infants whom we had observations at both time points ($n = 61$), the picture was the same as with the 12 month data in that neither CV or non-CV vocalisations co-occurred with gaze significantly above the level predicted by chance. The same was true when collapsed across age (see Appendix G, Figure 25). This suggests that this discrepancy between 11 and 12 months was as a result of sample size, not a genuine developmental effect. At both 11 and 12 months, combinations involving CV and non-CV vocalisations co-occurred with gaze above the level predicted by chance (11 months: combinations involving CV vocalisation(s), $t(133) = 3.22$, $p = .002$; involving only non-CV vocalisations, $t(133) = 3.17$, $p = .002$. 12 months: involving CV vocalisation(s), $t(63) = 2.99$, $p = .004$; involving only non-CV vocalisations, $t(129) = 2.09$, $p = .041$).

**Gesture type.** At 11 months giving, showing and conventional gestures co-occurred with gaze significantly above the level predicted by chance (giving, $t(133) = 3.20$, $p =.002$; showing, $t(133) = 3.44$, $p <.001$; conventional gestures, $t(133) = 2.27$, $p = .025$). At 12 months, only giving and showing co-occurred with gaze significantly above the level predicted by chance (giving, $t(63) = 3.14$, $p = .003$; showing, $t(63) = 2.71$, $p = .009$). When we considered only data at 11 months from infants whom we had observations at both time points ($n = 61$), the picture remained the same as when we considered the full sample at 11 months. This suggests a developmental change in that conventional gestures were likely to co-occur significantly above chance at 11, but not 12 months. However, the frequency of conventional gestures at both 11 and 12 months was low. At neither age did index-finger or open-hand pointing co-occur with gaze above the level predicted by chance. However, this differed for the inclusive analyses, where index-finger pointing at 11 months co-occurred with gaze above chance (see Appendix E, Figure 22). Additionally, it should be noted that when we collapsed across age, giving did not co-occur with gaze significantly above the level predicted by chance (though it did separately at

11 and 12 months), suggesting that the higher frequency of giving as a result of collapsing age brings this co-occurrence closer to chance (Appendix G, Figure 25).

The same pattern of results occurred for the type of gestures included in combinations. At 11 months, combinations involving giving, showing and conventional gestures co-occurred with gaze above the level predicted by chance (combinations involving giving, $t(133) = 2.59$, $p = .011$; showing, $t(133) = 2.95$, $p = .004$; conventional gestures, $t(133) = 2.42$, $p = .017$). At 12 months, only giving and showing co-occurred with gaze significantly above the level predicted by chance (giving, $t(63) = 2.08$, $p = .042$; showing, $t(63) = 2.81$, $p = .007$). When we considered only data at 11 months from infants whom we had observations at both time points ($n = 61$), only combinations involving giving co-occurred with gaze significantly above the level predicted by chance ($t(60) = 2.01$, $p = .048$). At neither age did combinations involving index-finger or open-hand pointing co-occur with gaze above the level predicted by chance. In general, low frequencies of combinations at 11 or 12 months might account for these discrepancies in results, as when we collapsed across age, combinations involving all types of gestures co-occurred with gaze significantly above the level predicted by chance (Appendix G, Figure 25).

### Do Infants' Vocalisations and Gestures Co-occur with Each Other Above the Rate Expected by Chance?

At both ages vocalisations and gestures co-occurred significantly above the level predicted by chance (11 months: $t(133) = 3.60$, $p < .001$; 12 months: $t(63) = 4.10$, $p < .001$).

**Vocalisation type.** At both ages, both types of vocalisations co-occurred with gestures significantly above the level predicted by chance (11 months: CV, $t(133) = 2.97$, $p = .004$; non-CV, $t(133) = 3.25$, $p = .001$; 12 months: CV, $t(63) = 3.36$, $p = .001$; non-CV, $t(63) = 2.64$, $p = .010$).

**Gesture type.** At 11 months only open-hand pointing and giving co-occurred with vocalisations significantly above the level predicted by chance (open hand pointing, $t(133) = 2.05$, $p = .043$; giving, $t(133) = 2.21$, $p = .029$). Index-finger pointing, showing and conventional gestures did not co-occur with vocalisations above the level predicted by chance. However, at

12 months all gestures (with the exception of conventional gestures) co-occurred with vocalisations significantly above the level predicted by chance (index finger pointing, $t(63) = 2.15$, $p = .035$; open hand pointing, $t(63) = 2.05$, $p = .027$; giving, $t(63) = 2.44$, $p = .017$; showing, $t(63) = 2.62$, $p = .011$).

**Vocalisation and gesture type.** At 11 months only open hand pointing with CV vocalisations, and giving with non-CV vocalisations co-occurred significantly above the level predicted by chance (open hand pointing and CV, $t(133) = 2.08$, $p = .039$; giving and non-CV, $t(133) = 2.19$, $p = .030$), all other combinations did not co-occur above the level predicted by chance. At 12 months only giving with CV vocalisations, and showing with non-CV vocalisations co-occurred significantly above the level predicted by chance (giving and CV, $t(63) = 2.35$, $p = .022$; showing and non-CV, $t(63) = 2.01$, $p = .049$), all other combinations did not co-occur above the level predicted by chance.

## Is Intentional Communication More Likely to be Vocal, Gestural or Part of a Combination?

Our second question concerned whether gaze coordination, taken as an index of intentional communication, was more likely to occur with vocalisations, gestures and combinations. Table 16 shows descriptive statistics including the frequency and proportion of vocalisations, gestures and combinations with gaze coordination (combinations by both gesture *and* vocalisation type with gaze coordination are reported in full in Appendix D, Table 40).

Table 16

*Average Frequency and Proportion of Vocalisations, Gestures and Gesture-vocal Combinations With Gaze Coordination at 11 (n = 134) and 12 months (n = 64)*

| | *11 months (n = 134)* | | | | | *12 months (n = 64)* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 8.48 | 7.53 | 7 | 0-36 | .18 | 7.01 | 6.78 | 5 | 0-31 | .19 |
| CV | 3.16 | 4.32 | 2 | 0-23 | .16 | 2.63 | 3.04 | 2 | 0-13 | .19 |
| Non-CV | 5.33 | 4.85 | 4 | 0-22 | .19 | 4.38 | 4.54 | 3 | 0-18 | .19 |
| **Gestures** | 0.99 | 1.73 | 0 | 0-9 | .56 | 1.22 | 2.10 | 0 | 0-11 | .46 |
| Index-finger point | 0.10 | 0.51 | 0 | 0-5 | .43 | 0.14 | 0.58 | 0 | 0-4 | .27 |
| Open-hand point | 0.07 | 0.34 | 0 | 0-3 | .29 | 0.00 | N/A | 0 | 0 | N/A |
| Give | 0.41 | 1.07 | 0 | 0-7 | .61 | 0.86 | 1.77 | 0 | 0-8 | .47 |
| Show | 0.28 | 0.76 | 0 | 0-4 | .93 | 0.16 | 0.44 | 0 | 0-2 | .75 |
| Conventional gesture | 0.13 | 0.45 | 0 | 0-3 | .52 | 0.06 | 0.30 | 0 | 0-2 | .42 |
| **Gesture-vocal combinations** | 0.62 | 1.20 | 0 | 0-6 | .59 | 1.07 | 2.04 | 0 | 0-12 | .54 |
| CV | 0.32 | 0.78 | 0 | 0-4 | .54 | 0.55 | 1.15 | 0 | 0-6 | .61 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Non-CV | 0.30 | 0.86 | 0 | 0-5 | .57 | 0.52 | 1.55 | 0 | 0-11 | .51 |
| Index-finger point | 0.11 | 0.50 | 0 | 0-4 | .37 | 0.27 | 0.91 | 0 | 0-6 | .50 |
| Open-hand point | 0.07 | 0.55 | 0 | 0-6 | .24 | 0.09 | 0.29 | 0 | 0-1 | .50 |
| Give | 0.21 | 0.64 | 0 | 0-4 | .61 | 0.47 | 1.38 | 0 | 0-10 | .51 |
| Show | 0.16 | 0.57 | 0 | 0-3 | .92 | 0.20 | 0.47 | 0 | 0-2 | 1.00 |
| Conventional gesture | 0.07 | 0.25 | 0 | 0-1 | .75 | 0.03 | 0.18 | 0 | 0-1 | .40 |
| **Unimodal vocal bouts** | 6.75 | 5.40 | 6 | 0-23 | .19 | 5.83 | 5.75 | 5 | 0-25 | .20 |
| **Unimodal gestural bouts** | 0.80 | 1.50 | 0 | 0-9 | .59 | 0.96 | 1.83 | 0 | 0-11 | .43 |
| **Multimodal bouts** | 0.81 | 1.51 | 0 | 0-8 | .54 | 1.11 | 1.74 | 0 | 0-7 | .54 |

As predicted, at both ages, vocalisations (or unimodal vocal bouts) were more frequently gaze-coordinated than gestures and combinations (or unimodal gestural and multimodal bouts). However, a far higher *proportion* of gestures and combinations (or unimodal gestural and multimodal bouts) were gaze-coordinated, compared to vocalisations (or unimodal vocal bouts). The following analyses focus on whether these predicted differences are significant, and also explores whether different vocalisation or gesture types are more frequently, or proportionally more likely to be gaze-coordinated at either age. We first consider analyses counting each vocalisation and gesture individually and then consider bouts.

**Individual behaviours.** We investigated whether single (i.e., not considering whether they were produced in a bout) vocalisations, gestures or combinations were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** At both ages, there was a significant improvement in model fit (see method for detailed description of models) when behaviour type (vocalisation, gesture or combination) was added as a predictor, (11 months: $\chi^2$ (2) = 200.63, $p < .001$ (Appendix F, Table 59); 12 months: $\chi^2$ (2) = 70.75, $p < .001$ (Appendix F, Table 60)). Planned contrasts revealed that at both ages, a significantly higher frequency of vocalisations were gaze-coordinated than both gestures and combinations (11 months: $b = 3.84$, $t(266) = 16.39$, $p < .001$; 12 months: $b = 2.93$, $t(126) = 9.60$, $p < .001$).

***Proportion.*** At both ages, there was a significant improvement in model fit (see method for detailed description of models) when behaviour type was added as a predictor, (11 months, $\chi^2$ (2) = 58.44, $p < .001$ (Appendix F, Table 61); 12 months, $\chi^2$ (2) = 36.21, $p < .001$(Appendix F, Table 62)). Planned contrasts revealed that at both ages, a significantly lower proportion of infants' vocalisations were gaze-coordinated than their gestures and combinations (11 months: $b = -1.04$, $z = -10.63$, $p < .001$; 12 months: $b = -0.89$, $z = -7.92$, $p < .001$).

**Bouts.** We investigated whether unimodal vocal, unimodal gestural or multimodal bouts were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** At both ages there was a significant improvement in model fit when bout type was added as a predictor (11 months: $\chi^2$ (2) = 213.85, $p < .001$ (Appendix F, Table 63); 12 months: $\chi^2$ (2) = 65.14, $p < .001$ (Appendix F, Table 64)). Planned contrasts revealed that at both ages, a significantly higher frequency of unimodal vocal bouts were gaze-coordinated than unimodal gestural or multimodal bouts (11 months: $b$ = 2.97, $t(266)$ = 17.07, $p$ <.001; 12 months, $b$ = 2.40, $t(126)$ = 9.07, $p$ < .001).

***Proportion.*** At both ages, there was a significant improvement in model fit when bout type was added as a predictor, (11 months, $\chi^2$ (2) = 56.12, $p$ < .001 (Appendix F, Table 65); 12 months, $\chi^2$ (2) = 36.23, $p$ < .001(Appendix F, Table 66)). Planned contrasts revealed that at both ages, a significantly lower proportion of infants' unimodal vocal bouts were gaze-coordinated than their unimodal gestural and multimodal bouts (11 months: $b$ = -1.02, $z$ = -10.21, $p$ < .001; 12 months: $b$ = -0.79, $z$ = -7.10, $p$ < .001).

**Vocalisation type**. We investigated whether CV or non-CV vocalisations (whether produced alone or part of a combination) were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** Regarding vocalisations produced alone, at both ages, dependent t-tests revealed that a significantly higher frequency of non-CV vocalisations were gaze-coordinated than CV vocalisations (11 months, $t(133)$ = 4.75, $p$ < .001; 12 months, $t(63)$ = 3.76, $p$ < .001). At neither age was gaze coordination significantly more or less frequent with combinations involving either vocalisation type.

***Proportion.*** At 12 months (but not 11 months), when a model was fitted to vocalisations (produced alone), adding vocalisation type improved model fit, $\chi^2$ (1) = 4.68, $p$ = .030 (Appendix F, Table 67). At 12 months, a significantly higher proportion of infants non-CV vocalisations were gaze-coordinated than their CV vocalisations, $b$ = 0.26, $z$ = 2.18, $p$ =.030. When similar models were fitted to datasets including combinations, adding vocalisation type did not improve model fit at either age. When inclusive analyses were run (Appendix E – where combinations were not considered as a mutually exclusive category) the picture was different, in that adding vocalisation type did not improve model fit at *either* age.

**Gesture type.** We investigated whether index-finger pointing, open-hand pointing, giving, showing or conventional gestures (whether produced alone or part of a combination) were more frequently gaze-coordinated (we lacked enough data to conduct proportional analyses).

***Frequency***. At both ages, there was a significant improvement in model fit when gesture type was added as a predictor, (11 months: $\chi^2$ (4) = 25.95, $p$ < .001 (Appendix F, Table 68); 12 months: $\chi^2$ (4) = 40.76, $p$ < .001 (Appendix F, Table 69)). Post-hoc pairwise tests revealed that at both ages, a significantly higher frequency of giving was gaze-coordinated than all other gestures, with the exception of showing only at 11 months (Table 17).

Table 17

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types (Produced Alone) with Gaze Coordination at 11 (n = 134) and 12 months (n = 64)*

| | *11 months* | | | | *12 months* | | | |
|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* |
| Index-finger point : Open-hand point | 0.38 | 532 | 0.48 | .989 | 0.14 | 252 | 0.93 | .884 |
| Index-finger point : Give | -0.31 | 532 | -3.87 | .001 | -0.72 | 252 | -4.79 | < .001 |
| Index-finger point : Show | -0.17 | 532 | -2.18 | .191 | -0.02 | 252 | -0.10 | 1.000 |
| Index-finger point : Conventional | -0.02 | 532 | -0.28 | .999 | -0.08 | 252 | 0.52 | .986 |
| Open-hand point : Give | -0.35 | 532 | -4.35 | < .001 | -0.86 | 252 | -5.73 | < .001 |
| Open-hand point : Show | -0.21 | 532 | -2.66 | .062 | -0.16 | 252 | -1.04 | .838 |
| Open-hand point : Conventional | -0.06 | 532 | -0.76 | .941 | -0.06 | 252 | -0.42 | .994 |
| Give : Show | 0.14 | 532 | 1.70 | .437 | 0.71 | 252 | 4.69 | < .001 |
| Give : Conventional | 0.29 | 532 | 3.59 | .003 | 0.80 | 252 | 5.31 | < .001 |
| Show : Conventional | 0.15 | 532 | 1.90 | .321 | 0.09 | 252 | 0.62 | .972 |

We constructed similar models for each age group to determine if there were significant differences in the frequency that combinations with different types of gestures were gaze-coordinated. At 12 months, but not 11 months, there was a significant improvement in model fit when gesture type was added as a predictor, $\chi^2$ (4) = 13.13, $p$ = .011 (Appendix F, Table 70).

Post-hoc tests revealed that, a significantly higher frequency of combinations involving giving were gaze-coordinated than open-hand pointing or conventional gestures (Table 18).

Table 18

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types When Produced as part of Gesture-vocal Combinations with Gaze Coordination at 12 months (n = 64)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Index-finger point : Open-hand point | 0.17 | 252 | 1.29 | .699 |
| Index-finger point : Give | -0.21 | 252 | -1.56 | .525 |
| Index-finger point : Show | 0.06 | 252 | 0.47 | .990 |
| Index-finger point : Conventional | 0.23 | 252 | 1.76 | .402 |
| Open-hand point : Give | -0.38 | 252 | -2.85 | .038 |
| Open-hand point : Show | -0.11 | 252 | -0.82 | .924 |
| Open-hand point : Conventional | 0.06 | 252 | 0.47 | .990 |
| Give : Show | 0.27 | 252 | 2.03 | .256 |
| Give : Conventional | 0.44 | 252 | 3.32 | .009 |
| Show : Conventional | 0.17 | 252 | 1.29 | .699 |

When combinations were not considered as a mutually exclusive category, the picture was identical in terms of pairwise comparisons to when gestures were produced alone (Appendix E).

**Developmental change.** We were interested in determining whether there were changes in the frequency, or proportion that behaviours were gaze-coordinated between 11 and 12 months. Again, it was noteworthy that the frequency of different behaviours was remarkably stable across the two measurement time points. There was only one instance of significant developmental change. For the infants whose measures are reported at both time points, regression analyses and follow up pair-wise comparisons revealed that the frequency of giving *with gaze coordination* significantly *increased* between 11 and 12 months ($b$ = 0.44, $t$(300) =

3.36, $p$ = .027, Appendix F, Table 71 for model). When combinations were not considered as a mutually exclusive category, the picture was identical, as the frequency of giving with gaze coordination significantly increased between 11 and 12 months (Appendix E, Table 55). However, the proportion of gaze coordination with gestures did not increase, so this phenomenon is most likely due to the increase in giving gestures between 11 and 12 months previously reported.

## Discussion

We investigated whether 11- and 12-month-old infants' vocalisations, gestures and gesture-vocal combinations co-occurred with gaze to their caregiver's face at a rate above chance and whether the vocalisations or gestures were more frequently or more likely to be gaze-coordinated. We found that at 11 months vocalisations, gestures and gesture-vocal combinations co-occurred with gaze to caregiver's face significantly above the levels predicted by chance, and therefore it is plausible that these were deliberately coordinated by the infant. We found that at 12 months, infants' gestures and combinations continued to co-occur with gaze to caregiver's face significantly above the levels predicted by chance, but vocalisations did not. Analysis of specific sub-types of vocalisation revealed that it was only non-CV vocalisations that were found to co-occur with gaze to caregiver's face above chance levels at 11 months. CV vocalisations did not co-occur with gaze to caregiver's face above chance levels at either age. Analysis of specific sub-types of gesture revealed that giving and showing co-occurred with gaze above chance levels at both ages, and conventional gestures co-occurred with gaze above chance levels only at 11 months. Index-finger and open-hand pointing did not co-occur with gaze to caregiver's face above chance levels at either age. At both ages vocalisations were also coordinated with gestures (forming gesture-vocal combinations) at above chance rates, again suggesting that infants at this age intend for their communicative acts to be perceived by their interlocutors. As predicted, we found that infants produced vocalisations with gaze coordination more frequently than gestures and combinations, but a higher *proportion* of gestures and combinations were coordinated with gaze. The same pattern held when we considered bouts of behaviours. Finally, we did not observe any wholesale increase in gaze coordination between 11 and 12 months. Only the frequency of giving with gaze coordination increased from 11 to 12 months, which was most likely due to the increase in giving between 11 and 12 months (regardless of gaze coordination) as the proportion of giving that was gaze-coordinated did not change.

Firstly, our results suggest that some vocalisations, gestures and combinations are deliberately coordinated with gaze to caregiver's face from 11 months of age. This is a necessary condition for intentional communication, though it is not a sufficient condition for demonstrating that intentional communication has occurred. This type of analysis has only previously been conducted on infant vocalisations. D'Odorico & Cassibba (1995) observed vocalisations to co-occur with gaze to a caregiver's face above the level predicted by chance at 10 months. Our research found the same was true at 11 months, and further demonstrated that it is non-CV vocalisations that drive this effect. Critically, we also considered gesture production in the same children and showed that some types of gestures and gesture-vocal combinations co-occur with gaze at above chance rates at 11 months (the earliest this has been studied). Therefore there is evidence in support of the hypothesis that infants produce both vocalisations and gestures with the intention of influencing their caregiver's behaviour at 11 months.

While overall these results support the hypothesis that infants have informative intentions at 11 months, there are several surprising observations that require further explanation. First, it is surprising that coordination between vocalisations and gaze was observed at 11, but not 12 months. However, this result is almost certainly as a result of the discrepancy in sample size between 11 and 12 months, as oppose to a genuine effect of development. When we only included data from infants at 11 months for whom we also had data at 12 months ($n$ = 61), the result was the same as at 12 months, in that only gestures and combinations co-occurred with gaze above chance levels, whilst vocalisations did not. Therefore, this change between 11 and 12 months is not likely to be a genuine developmental effect. It is a limitation of our study that our sample at 12 months was greatly reduced compared to the sample at 11 months. Future research should aim to maintain large samples at all ages in order to test for developmental effects.

Second, CV vocalisations, open-hand pointing, and index-finger pointing (if produced alone, i.e., not as part of a gesture-vocal combination) were not coordinated with gaze to caregiver's face above levels expected by chance alone at either age. Pointing (both index-finger and open-hand) was infrequently produced (29 times at 11 months and 34 times at 12 months) by a small number of the infants, thus providing a potential explanation why they did not exceed chance levels of co-occurrence with gaze. However, CV vocalisations were very frequent in the dataset. We suspect that as infants gaze to their caregiver's face, and produce CV vocalisations for a plethora of non-communicative reasons, such as checking for the reassuring presence of adults (in the case of gaze) and vocal play (Oller, 2000), these behaviours could naturally co-occur by chance. These frequent non-communicative vocalisations could obscure genuine instances of deliberate coordination. Additionally, there is arguably less need for infants to monitor their caregiver's attention when producing vocalisations, as they may function without being visually attended to, leading to less deliberate coordination. By contrast, showing involves an infant raising an object up to their caregiver's face, and so it is almost necessary that infants look to their caregiver's face when engaging in this behaviour.

Key to interpreting these findings is that while demonstrating that some vocalisations, gestures or combinations co-occurred with gaze above chance levels suggets that they are more likely to be deliberately coordinated, this does not mean that when they do not co-occur above chance levels that this is evidence that they are *not* deliberately coordinated. In sum, we have provided no evidence that infants *deliberately* coordinate their CV vocalisations, open-hand pointing and index-finger pointing (when produced alone) with gaze to their caregiver's face at 11 and 12 months, but we have not demonstrated conclusively that this is not the case.

Secondly, regarding whether intentional communication was more likely to be gestural, vocal, or gesture-vocal, as predicted, the results differed depending on whether we focussed on the frequency or proportion measures (but not on whether we focused on bouts or single behaviours not considering whether they were produced in bouts). In essence, vocalisations are by far the more frequent type of communicative behaviour and, in turn, the more frequent type of gaze-coordinated behaviour.  To the extent that we demonstrated that only non-CV vocalisations co-occur with gaze above chance rates (and even this was not always the case), we can say that infants vocalisations are the most frequent behaviour produced in a way that suggests communicative intent. Infants were far less likely to gesture, however when they engaged in giving and showing they very often did so in a way that suggests communicative intent.

One interesting next step would be to determine whether proportional or frequency measures are more valuable in the sense that they better predict later language outcomes. If the high frequency of gaze coordination with vocalisations is largely due to chance, then high levels of this coordination are not necessarily an indicator of developmental advance and therefore might not predict later language development. However, instances of such coordination might elicit caregiver responses that in turn affect language learning, and might therefore predict language. Equally, high frequencies of vocalisations might predict later language development, so instances of coordination may also predict these purely by virtue of the fact that a high frequency of vocalisations means a high frequency of vocalisations with gaze coordination. If, on the other hand, the high proportion of gaze coordination with some gestures (that are coordinated with gaze above chance levels) is indicative of developmental advance and readiness for conventional language, we might expect this to predict later language development. This is the focus of Chapter 6.

Infants more frequently produced non-CV vocalisations than CV vocalisations with gaze to their caregiver's face. However non-CV vocalisations were more frequent overall. Proportionally speaking, each type of vocalisation was coordinated with gaze at the same rate. This supports a previous finding from a vocalisation eliciting study, whereby 12-month-olds' 'person-directed' (not necessarily towards the face) non-CV vocalisations were more frequent than similarly directed CV vocalisations, but that neither type were proportionally more likely to be person-directed (Marchand et al., 1994). There are two plausible explanations of this finding. Firstly, as adult language users, we do not need to necessarily mark language with extra behaviours (i.e., eye contact) to demonstrate communicative intent, as using language (more often than not) indicates to others that we are communicating. It is possible that caregivers respond to infants CV vocalisations, which are more language-like, regardless of gaze coordination (as they might to uses of language in other adults), and so infants do not necessarily need to coordinate gaze in order to get responses from caregivers. This possibility is explored in Chapter 7.

Showing (either produced alone or in a combination) is proportionally more likely to be gaze-coordinated than other types of gestures. This supports previous findings that gaze coordination is more likely for showing than giving gestures (Boundy et al., 2016). Additionally, we found that both showing and giving gestures are proportionally more likely to be gaze-coordinated than index-finger pointing. To some extent, as noted before, it is almost necessary for infants to gaze to their caregiver's face whilst performing show gestures, and highly likely when they are performing give gestures, as they are extending objects towards their caregivers. Giving was more frequently gaze-coordinated than pointing, and gesture-vocal combinations involving giving were more likely at 11 months to have gaze coordination than those involving index-finger pointing. In terms of frequency, at both ages, giving was the most common potentially intentionally communicative gesture, whilst in terms of proportion, showing was the most was the most likely gesture to be potentially intentionally communicatively produced.

In some respects, this finding is surprising given the focus in recent research on index-finger pointing. There is no evidence, using the token of whether behaviours are coordinated with gaze to caregiver's face, to suggest that index-finger pointing is the gesture most likely to be intentionally communicative, or the most frequent intentionally communicative gesture at either 11 and 12 months. Less than half of index-finger pointing was produced with gaze coordination, as previously noted in this age group in similar naturalistic studies (Wu & Gros-Louis, 2014, 2015). This is surprisingly low for the gesture that has been thought to herald the onset of intentional communication. For most pointing gestures, there was nothing to suggest they were intending to direct the attention of a caregiver.

Infants often produced vocalisations and gestures together. These combinations were often coordinated with gaze, following a more similar pattern to gestures than vocalisations produced alone. There was no evidence that these combinations were more communicative in the sense that they were more likely to be coordinated with gaze. It is possible however that they would be more likely to elicit a responses from caregivers (e.g., because vocalisations attract attention to gestures). This possibility is explored in Chapter 7.

Considering the developmental transition from 11 to 12 months, we did not observe a watershed in gaze coordination with behaviours that might herald the onset of intentional communication. Some have claimed this happens when infants begin to point using their index-finger at around 12 months (Tomasello et al., 2007). Yet there was no increase in the proportion of gaze coordination for the behaviours that we studied between 11 and 12 months. There was only an increase in the frequency of giving gestures with gaze coordination, which could be in part due to the increase in raw frequency of giving gestures between 11 and 12 months. There are a number of explanations of this. Firstly, it could be that gaze coordination is given too much weight as a marker of intentional communication and that infants could determine that caregivers are attending to their behaviour without necessarily having to make eye contact (Akhtar & Gernsbacher, 2008; Cochet & Vauclair, 2010; Murphy, 1978). Perhaps greater expansion of the criteria for intentional communication (e.g., to include persistence, elaboration etc.) may demonstrate a more stark transition (Golinkoff, 1986; Spencer, 1993). This would account for why there is not an increase in gaze coordination, but there could still be a shift in the understanding of intentional communication that is undetected in this study. Secondly, it is possible that infants have already made the transition to intentional communication prior to 12 months (and the emergence of index-finger pointing) in their earlier vocalisations and gestures, and this is why we did not detect a change between 11 and 12 months. Equally, it is possible that infants have not made this transition yet, which is why such a change was not detected. We can test between these explanations by investigating whether the early potential instances of intentional communication (vocalisations and gestures coordinated with gaze) that we have observed here predict later language outcomes, indicating that they may be genuine instances of early intentional communication. We do this in the following chapter.

The main finding of this chapter is that some types of both vocalisations and gestures are coordinated with gaze, indicating that they might be early instances of intentional communication. Vocalisations are frequently gaze-coordinated, but while gestures and combinations are less frequent and less frequently gaze-coordinated, much higher proportions of gestures and combinations are gaze-coordinated than vocalisations. The following chapters will address issues for further study identified throughout this discussion. Namely, how do the measures of gaze coordination measured here predict infants' later expressive vocabulary? This may be one way of determining whether such early instances of gaze coordination indicate that a behaviour might be intentionally communicative. Additionally, this will also help us determine whether measures of frequency or proportion are better predictors of later language.

## 6. Do Infants' Vocalisations, Gestures and Gesture-vocal Combinations Predict Later Expressive Vocabulary?


### Abstract

Previous research suggests that some prelinguistic vocalisations and gestures are predictive of infants' early language abilities, however a complete developmental account is lacking. In this chapter, we first sought to test whether the *frequency* with which infants naturally produce vocalisations, gestures and gesture-vocal combinations (and specific types of these behaviours) at 11 & 12 months predicts later expressive vocabulary at 15 and 24 months. We found that a number of specific types of vocalisations and combinations predicted 15-month expressive vocabulary, while a number of gestures and combinations predicted 24-month expressive vocabulary. Plausibly, these could predict later vocabulary because they are early instances of intentional communication. To test this proposal, we next considered whether the frequency or proportion of vocalisations, gestures and combinations produced *with gaze coordination* were better predictors of later expressive vocabulary. We found that the *frequency* of CV vocalisations with gaze coordination, and the *proportion* of all vocalisations with gaze coordination were predictive of 15-month expressive vocabulary. Both gaze-coordinated measures were better predictors than when gaze coordination was not taken into account. This is the first demonstration that vocalisations, particularly CV vocalisations, *with gaze coordination* predict infants' later expressive vocabulary, suggesting that these could be predictive because they are early attempts to intentionally communicate. We found that showing (and combinations involving showing) with gaze coordination predicted 24-month expressive vocabulary. However, as showing is nearly always gaze-coordinated, we were unable to ascertain whether gaze-coordinated showing was a better predictor. This pattern of results gives us limited evidence that some behaviours predict later expressive vocabulary *because* they are early attempts to intentionally communicate.


### Introduction

As Chapter 4 demonstrated, a prevailing orthodoxy in the literature is that prelinguistic gestures are precursors to later language. However, this view is not unambiguously supported, as prelinguistic vocalisations also predict later language abilities, and it has not been unambiguously demonstrated that links between gestures and later language are dominant. An important methodological gap in the literature is that gestures and vocalisations are rarely studied simultaneously in naturalistic contexts. Going some of the way towards this, McGillion, et al., (2017) has demonstrated that the onset of both babble and pointing observed in naturalistic contexts are linked to later vocabulary (discussed in Chapter 4). The purpose of this chapter is to build on this finding, and investigate the relative contributions of prelinguistic infant vocalisations and a broader range of gestures to later expressive vocabulary.

There were two limitations of the McGillion et al. (2017) study. First, the study considered only babble and index-finger pointing. As we have demonstrated in the previous chapter, there are a number of other gestures produced by infants at this age that could plausibly be linked to later language abilities. Secondly, there was no investigation into the *intentionality* of early pointing or vocalisations. As discussed in Chapter 4, it is not clear if babble or pointing predict later language because they represent early instances of prelinguistic intentional communication, or because they are indicators of socio-cognitive or motoric development required for speech.

The current chapter aimed to address these two limitations. First, we tested whether the frequencies of early vocalisations, gestures and gesture-vocal combinations predict later expressive vocabulary, considering a wider range of vocalisations (not just babble involving consonants and vowels) and a wider range of gesture types (open-hand pointing, giving, showing and conventional gestures in addition to index-finger pointing). Second, we tested whether the frequency with which these behaviours were coordinated with gaze was especially predictive of language outcomes. This would be expected if the predictive relation holds because they are early instances of intentional communication.

In order to investigate these links fully, as in the previous chapter, we opt to run the analyses with gesture-vocal combinations treated as a mutually exclusive category from vocalisations or gestures, as this has implications for theories proposing that early intentional communication is either gestural or vocal, or a combination of the two. We also run the same analyses with gesture-vocal combinations not included as a mutually exclusive category to determine if this decision substantially affects the results.

Finally, we tested whether expressive vocabulary is best predicted by the *frequency* with which a given behaviour (vocalisations, gestures or combinations) is coordinated with gaze or by the *proportion* of occurrences of that behaviour that are coordinated with gaze. These two ways of quantifying the extent of gaze coordination are theoretically important for different reasons. When infants produce vocalisations or gestures with gaze coordination this provides potential learning opportunities about intentional communication, as infants receive information about their caregiver's attention and responses. On this basis, the more frequent the behaviour the more learning should occur. In contrast, some behaviours (gestures and combinations) were less frequently but proportionally more likely to be coordinated with gaze than others. These behaviours provide fewer learning opportunities, but the opportunities that do arise might be more likely to result in successful learning as there will be less noise.

## Method

### Participants

As for the previous chapter, participants were drawn from a larger sample of families who had previously participated in a longitudinal randomised control study (McGillion et al. 2017). We included dyads for whom we had naturalistic observations at both 11 and 12 months (and were from the control condition in the original study, see Chapter 5), and a measure of expressive vocabulary at either 15 or 24 months. For analyses of 15-month expressive vocabulary, 53 caregiver-infant dyads (30 female infants, 23 male) were included. For analyses of 24-month expressive vocabulary, 49 caregiver-infant dyads (28 female, 21 male) were included.

### Materials

Expressive vocabulary was assessed using the Lincoln Communicative Development Inventory (LCDI) Infant form at 15 months and the toddler form at 24 months (Meints & Woodward, 2011).

## Procedure

LCDIs were collected by post at 15 months and collected in person at 24 months by the researchers in the longitudinal study.

Infant behaviours at 11 and 12 months, as coded in Chapter 5, were collapsed across age. This involved summing the frequencies for every measure for each participant. We did this in order to maximise variance in the frequency of gesture particularly as gestures were produced at relatively low frequencies at each individual age. Furthermore, as became clear in the previous chapter, there were only isolated and modest changes in the rates that any behaviour was produced or produced with gaze coordination between 11 and 12 months.

## Analyses

All analyses consist of linear regression models, fit to either 15- or 24-month language outcomes using the “lm” function (part of the “stats” package) in R (R Core Team, 2015). Given the large number of predictors available, we initially tested those behaviours at the coarsest level of granularity (gaze, vocalisations, gestures and combinations). We then broke vocalisations and gestures into sub-types. First we broke combinations into finer subtypes (those that involved either CV or non-CV vocalisations regardless of gesture type, and by gesture type regardless of vocalisation type), and second, finest level subtypes (those that involved specific vocalisations and gestures). The different granular levels are depicted in Table 19. Testing predictors at different levels of granularity allowed us to determine what granularity of behaviours was warranted, and to reduce the number of predictors with which to build our final model (thus reducing the chance of type I errors). Note, that in the current chapter, we did not consider bouts as an alternate form of analyses since these did not differ from single behaviours in the previous chapter.

Table 19

*Granular levels of predictors*

| *Coarsest Level* | *Finer Level* | *Finest Level* |
|---|---|---|
| **Vocalisations** | CV | CV |
| | Non-CV | Non-CV |
| **Gestures** | Index-finger Point | Index-finger Point |
| | Open-hand Point | Open-hand Point |
| | Give | Give |
| | Show | Show |
| | Conventional | Conventional |
| **Combinations** | *By vocalisation (with any gesture)* | CV & Index-finger Point |
| | CV | CV & Open-hand Point |
| | Non-CV | CV & Give |
| | | CV & Show |
| | *By gesture (with any vocalisation)* | CV & Conventional |
| | Index-finger Point | Non-CV & Index-finger Point |
| | Open-hand Point | Non-CV & Open-hand Point |
| | Give | Non-CV & Give |
| | Show | Non-CV & Show |
| | Conventional | Non-CV & Conventional |

Once we had tested individual predictors, and identified those which improved on the null model (and therefore significantly predicted later expressive vocabulary), we sought to determine whether these predictors explained unique variance. We constructed a final model using only those predictors that had improved on the null model, systematically adding predictors and testing for improvement of fit. These final models included only predictors that were compatible, given that some might be represented at two or three levels of granularity. For example, we did not construct models with *frequency of vocalisations* and *frequency of CV* *vocalisations* as predictors in the same model as one predictor represents a subset of the data points of the other. If both behaviours had improved on the null model, we selected the predictor that explained more variance as indicated by the $R^2$ value.

Finally, we did not include behaviours as predictors in *any* model when less than 10% of infants had produced that specific type of behaviour (to see excluded predictors, see Appendix H, Tables 81 and 82).

**Proportional models.** When testing whether the proportions of behaviours that were gaze-coordinated at 11 & 12 months predicted language outcomes, only infants who produced the given behaviour were included (e.g., index-finger pointers). For proportional models, we did not test predictors when fewer than *n*=10 infants produced that type of behaviour. Furthermore, to ensure that there was variance in the proportional scores, 50% or more of infants had to produce more than 2 instances of the behaviour (to see excluded predictors, see Appendix H, Tables 81 and 82).

**Inclusive analyses.** For all analyses, we treated gesture-vocal combinations as a separate, mutually exclusive category from gesture and vocalisations. However, to check that this decision did not substantially affect the results, we also conducted inclusive analyses with only vocalisations and gestures (making no distinction for those involved in gesture-vocal combinations). These analyses are reported in full in Appendix I, here we only note when such inclusive analyses differ notably from the main exclusive analyses.

## Results

We focus first on whether the *raw frequency* of behaviours (e.g., vocalisations, gestures, gesture-vocal combinations, and individual vocalisation and gesture types), and the frequency of these behaviours *with gaze coordination* at 11 & 12 months predict expressive vocabulary at 15 months. We then focus on whether the *proportion* of behaviours with gaze coordination at 11 & 12 months predict expressive vocabulary at 15 months. Finally, all analyses are then repeated to with expressive vocabulary at 24 months as the outcome.

## Predicting Expressive Vocabulary at 15 Months

**Raw frequency.** We first explored whether raw frequencies of behaviours at 11 & 12 months (Appendix H, Table 81) predicted expressive vocabulary at 15 months. The mean number of words infants produced at 15 months was 20.49 (SD = 25.14). The median was 12 (range = 0-113).

We initially tested whether the frequency of gaze to caregiver's face, vocalisations, gestures or combinations individually predicted 15-month expressive vocabulary. Only the frequency of gaze to caregiver's face improved on the null model, $F(1,51) = 6.90$, $p = .011$, $R^2 = .12$.

We then explored whether the frequency of specific types of vocalisation, gesture or gesture-vocal combinations predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on a null model with no predictors.

***Vocalisation type.*** For models with frequency of CV and frequency of non-CV vocalisations as predictors, only the frequency of CV vocalisations improved on the null model, $F(1,51) = 8.77$, $p = .005$, $R^2 = .15$.

***Gesture type.*** No gestural predictors improved on the null model.

***Combination type.*** For models with the frequency of combinations involving CV or non-CV vocalisations as predictors, only the frequency of combinations involving CV vocalisations improved on the null model, $F(1,51) = 9.44$, $p = .003$, $R^2 = .16$. No predictors improved on the null model when we tested the frequency of combinations involving specific gestures as predictors.

Finally, for models with the frequency of specific vocalisation types combined with specific gesture types as predictors, the frequency of combinations involving index-finger pointing and CV vocalisations improved on the null model, $F(1,51) = 4.09$, $p = .049$, $R^2 = .07$, as did the frequency of combinations involving giving and CV vocalisations, $F(1,51) = 12.14$, $p = .001$, $R^2 = .19$.

***Final model.*** Given that a number of behaviours were predictive of later language, we next wanted to test whether they each explained separate variance. Four predictors were identified (see Method): gaze to caregiver's face, CV vocalisations, combinations involving index-finger pointing and CV vocalisations and combinations involving giving and CV vocalisations. It is worth noting that the latter two behaviours are not subsets of the CV vocalisation measure because gesture-vocal combinations are a separate category to vocalisations. However, the frequency of CV vocalisations was correlated with the frequency of combinations involving giving and CV vocalisations ($r$ = .53), and the frequency of gaze was correlated with the frequency of combinations involving index-finger pointing and CV vocalisations ($r$ = .32). We systematically tested for improvement in fit, combining the four predictors. The final model included all four predictors (suggesting that they contribute separate variance) and is presented in Table 20.

Table 20

*Regression Model fitting Frequency of Gazes to Caregiver's Face, CV Vocalisations, Combinations Involving Index-finger Pointing and CV Vocalisations, Combinations Involving Giving and CV vocalisations at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | -3.11 | 7.08 | -0.44 | .662 |
| Frequency of Gaze to Caregiver's Face | 0.27 | 0.14 | 2.03 | .048 |
| Frequency of CV Vocalisations | 0.15 | 0.13 | 1.15 | .255 |
| Frequency of Combinations involving Index-finger Pointing and CV vocalisations | 2.85 | 1.74 | 1.64 | .108 |
| Frequency of Combinations involving Giving and CV Vocalisations | 10.89 | 4.32 | 2.52 | .015 |

$R^2$ = .36, $F(4,48) = 6.64$, $p < .001$.

**Frequency of behaviours with gaze coordination.** We next investigated whether the frequency of behaviours *with gaze coordination* at 11 & 12 months predicted expressive vocabulary at 15 months.

Working at the coarser level of granularity first, we tested whether the frequency of vocalisations, gestures or combinations with gaze coordination individually predicted 15-month expressive vocabulary. Only the frequency of vocalisations with gaze coordination improved on the null model, $F(1,51) = 11.30$, $p = .001$, $R^2 = .18$.

We then explored whether the frequency of specific types of vocalisation, gesture or gesture-vocal combinations with gaze coordination predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** For models with the frequency of CV or non-CV vocalisations with gaze coordination as predictors, only the frequency of CV vocalisations with gaze coordination improved on the null model, $F(1,51) = 21.05$, $p < .001$, $R^2 = .29$.

***Gesture type.*** No predictors improved on the null model when we tested the frequency of specific gesture types with gaze coordination as predictors.

***Combination type.*** No predictors improved on the null model when we tested the frequency of combinations involving CV or non-CV vocalisations, or the individual gesture types with gaze coordination as predictors. No predictors improved on the null model when we tested the frequency of specific vocalisation types combined with specific gesture types with gaze coordination as predictors.

***Final model.*** The final model consists only of the frequency of CV vocalisations with gaze coordination, and is presented in Table 21.

Table 21

*Regression Model fitting Frequency of CV Vocalisations with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 7.07 | 4.14 | 1.71 | .094 |
| Frequency of CV Vocalisations with Gaze Coordination | 2.25 | 0.49 | 4.59 | < .001 |

$R^2 = .29$, $F(1,51) = 21.05$, $p < .001$.

**Proportion of behaviours with gaze coordination.** We next investigated whether the proportion of behaviours that were gaze-coordinated at 11 & 12 months (Appendix H, Table 81) predicted expressive language at 15 months.

We initially tested whether the proportion of vocalisations, gestures or combinations with gaze coordination individually predicted 15-month expressive vocabulary (on subsets where all participants had produced each type of behaviour – see Method). Only the proportion of vocalisations with gaze coordination improved on the null model $F(1,51) = 6.70$, $p = .013$, $R^2 = .12$.

We then explored whether the proportion of specific types of vocalisation or gestures (but not gesture-vocal combinations, as we lacked variance for any of these measures – see Method) with gaze coordination predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** For models with the proportion of CV or non-CV vocalisations with gaze coordination as predictors, the proportion of CV vocalisations with gaze coordination improved on the null model, $F(1,51) = 4.03$, $p = .050$, $R^2 = .07$, as did the proportion of non-CV vocalisations, $F(1,51) = 4.17$, $p = .046$, $R^2 = .08$.

***Gesture type.*** For the model with the proportion of giving with gaze coordination as a predictor (as this was the only gesture type it was possible to test – see Method), it did not improve on the null model.

***Final model.*** Adding the proportion of CV vocalisations with gaze coordination to a model with the proportion of non-CV vocalisations with gaze coordination as a predictor did not result in an improvement of fit. As such, the final model consists only of the proportion of vocalisations with gaze coordination (as this explained more variance than either other predictor), and is presented in Table 22.

Table 22

*Regression Model fitting Proportion of Vocalisations with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 6.64 | 6.27 | 1.06 | .295 |
| Proportion of Vocalisations with Gaze Coordination | 77.67 | 30.00 | 2.59 | .013 |

$R^2$ = .12, *F*(1,51) = 6.70, *p* = .013.

**Does Gaze-Coordinated Behaviour Better Predict Expressive Vocabulary at 15 Months than the Raw Frequency of Behaviours?**

We wanted to determine whether gaze-coordinated behaviours were better predictors of later expressive vocabulary than the frequency with which they were produced regardless of gaze coordination, as this would follow if they were early instances of intentional communication. A number of gaze-coordinated behaviours predicted expressive vocabulary at 15 months (as explored above), and so we consider this question for these behaviours. Table 23 shows the variance explained by predictors relating to vocalisations and specifically CV vocalisations and 15-month expressive vocabulary, as demonstrated by both $R^2$ and Akaike Information Criterion (AIC – a standardized measure of fit across models with a different number of predictors).

Table 23

*Summary of Regression Models ($R^2$ and AIC) fitting Predictors relating to Vocalisations and CV Vocalisations at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| *Predictor(s)* | $R^2$ | *AIC* |
|---|---|---|
| ***Predictors relating to Vocalisations*** | | |
| Frequency of Vocalisations (*non-significant predictor*) | .03 | 495.61 |
| Frequency of Vocalisations with Gaze Coordination | .18 | 486.59 |
| Frequency of Vocalisations + Frequency of Gaze | .13 | 491.99 |
| ***Predictors relating to CV Vocalisations*** | | |
| Frequency of CV Vocalisations | .15 | 488.79 |
| Frequency of CV Vocalisations with Gaze Coordination | .29 | 478.88 |
| Frequency of CV Vocalisations + Frequency of Gaze | .23 | 485.29 |

Table 23 shows that the frequency of CV vocalisations with gaze coordination was a better predictor of expressive vocabulary at 15 months than the frequency of CV vocalisations (as denoted by $R^2$) and also was a better predictor than a model with both the frequency of CV vocalisations and the frequency of gaze to caregiver's face as predictors (as denoted by AIC, full model in Appendix J, Table 92). The raw frequency of CV vocalisations is correlated with the frequency of CV vocalisations with gaze coordination ($r$ = .71). The same pattern was also true for vocalisations (regardless of sub-type), but the variance explained was far less.

**Inclusive Analyses**

There was a notable difference in the inclusive analyses in predicting 15-month expressive vocabulary (see Appendix I). When all instances of index-finger pointing (whether produced alone or in gesture-vocal combinations) were collapsed into one predictor, the frequency of index-finger pointing with gaze coordination was predictive of 15-month expressive vocabulary. However, index-finger pointing with gaze coordination was a non-significant predictor of 15-month vocabulary when combined in a model with CV vocalisations with gaze coordination. Furthermore, the frequency of index-finger pointing with gaze coordination was not a better predictor than the frequency of index-finger pointing. Overall, this suggests that gestures are only predictors of 15-month vocabulary when gestures are combined with gesture-vocal combinations.

## Predicting Expressive Vocabulary at 24 Months

**Raw frequency.** We tested whether raw frequencies of behaviours at 11 & 12 months (see Appendix H, Table 82) predicted expressive language at 24 months. The mean number of words infants produced at 24 months was 370.92 (SD = 172.02). The median was 371 (range = 0-689).

We initially tested whether the frequency of gaze to caregiver's face, vocalisations, gestures or combinations individually predicted 24-month expressive vocabulary. However, no predictors improved on the null model.

We then explored whether the frequency of specific types of vocalisation, gesture or gesture-vocal combinations predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on a null model with no predictors.

***Vocalisation type.*** Neither predictor improved on the null model when we tested the frequency of CV or non-CV vocalisations as predictors.

***Gesture type.*** For models with the frequency of specific gesture types as predictors, the frequency of index-finger pointing improved on the null model, $F(1,47) = 7.65$, $p = .008$, $R^2 = .14$, as did the frequency of showing, $F(1,47) = 5.42$, $p = .024$, $R^2 = .10$.

***Combination type.*** Neither predictor improved on the null model when we tested the frequency of combinations involving CV or non-CV vocalisations as predictors. For models with the frequency of combinations involving specific gesture types as predictors only the frequency of combinations involving showing improved on the null model, $F(1,47) = 5.11$, $p = .028$, $R^2 = .10$.

Finally, for models with the frequency of specific vocalisation types combined with specific gesture types as predictors, the frequency of combinations involving giving and CV vocalisations improved on the null model, $F(1,47) = 4.38$, $p = .042$, $R^2 = .09$, as did the frequency of combinations involving showing and non-CV vocalisations, $F(1,47) = 4.98$, $p = .030$, $R^2 = .10$.

***Final model.*** We next tested whether the compatible predictors explained separate variance. Four predictors (see Method) were identified: frequency of index-finger pointing, showing, combinations involving giving and CV vocalisations and combinations involving showing. We systematically tested for improvement in fit, combining the four predictors. The final model is presented in Table 24.

Table 24

*Regression Model fitting Frequency of Index-Finger Pointing, Showing and Giving Combined with CV Vocalisations at 11 & 12 months to Expressive Vocabulary at 24 months (n=49)*

| | *B* | *SE* | *T* | *p* |
|---|---|---|---|---|
| Intercept | 291.52 | 27.80 | 10.49 | < .001 |
| Frequency of Index-finger Pointing | 32.38 | 10.98 | 2.95 | .005 |
| Frequency of Showing | 52.54 | 21.82 | 2.41 | .020 |
| Frequency of Combinations involving Giving and CV Vocalisations | 48.70 | 23.80 | 2.05 | .047 |

$R^2 = .31$, $F(3,45) = 6.67$, $p = .001$.

**Frequency of behaviours with gaze coordination.** We next investigated whether the frequency of behaviours *with gaze coordination* at 11 & 12 months predicted expressive vocabulary at 24 months.

We initially tested whether the frequency of vocalisations, gestures or combinations with gaze coordination individually predicted 24-month expressive vocabulary. Only the frequency of gestures with gaze coordination improved on the null model, $F(1,47) = 4.57$, $p = .038$, $R^2 = .09$.

We then explored whether the frequency of specific types of vocalisation, gesture or gesture-vocal combinations with gaze coordination predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** Neither predictor improved on the null model when we tested the frequency of CV or non-CV vocalisations as predictors with gaze coordination.

***Gesture type.*** For models with the frequency of specific gesture types with gaze coordination as predictors, only the frequency of showing with gaze coordination improved on the null model, $F(1,47) = 4.50$, $p = .039$, $R^2 = .09$.

***Combination type.*** Neither predictor improved on the null model when we tested the frequency of combinations involving CV or non-CV vocalisations with gaze coordination as predictors. For models with the frequency of combinations involving specific gesture types with gaze coordination as predictors, only the frequency of combinations involving showing with gaze coordination improved on the null model, $F(1,47) = 5.11$, $p = .028$, $R^2 = .10$.

Finally, for models with the frequency of specific vocalisation types combined with specific gesture types as predictors, the frequency of combinations involving showing and non-CV vocalisations with gaze coordination improved on the null model, $F(1,47) = 4.98$, $p = .030$, $R^2 = .10$.

***Final model.*** We next tested whether compatible predictors explained separate variance. The final model included only the frequency of combinations involving showing with gaze coordination, and is presented in Table 25.

Table 25

*Regression Model fitting Frequency of Gestures with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n=49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 343.69 | 26.48 | 12.98 | < .001 |
| Frequency of Combinations involving Showing with Gaze Coordination | 74.07 | 32.75 | 2.26 | .028 |

$R^2 = .10$, $F(1,47) = 5.11$, $p = .028$.

**Proportion of behaviours with gaze coordination.** We next investigated whether the proportion of behaviours that were gaze-coordinated at 11 & 12 months (Appendix H, Table 82) predicted expressive language at 24 months.

We initially tested whether the proportion of vocalisations, gestures or combinations with gaze coordination individually predicted 24-month expressive vocabulary (on subsets where all participants had produced each type of behaviour– see Method). However, no predictors improved on the null model. We then explored whether the proportion of specific types of vocalisation or gestures (but not gesture-vocal combinations, as we lacked variance for any of these measures – see Method) with gaze coordination predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model, but again, no predictors improved on the null model.

## Does Gaze-Coordinated Behaviour Better Predict Expressive Vocabulary at 24 Months than the Raw Frequency of Behaviours?

We wanted to determine whether gaze-coordinated behaviours were better predictors of later expressive vocabulary than the frequency with which they were produced regardless of gaze coordination, as this would follow if they were early instances of intentional communication. A number of gaze-coordinated behaviours predicted expressive vocabulary at 24 months (as explored above), and so we consider this question for these behaviours. Table 26 shows the variance explained by predictors relating to gestures and 24-month expressive vocabulary, as demonstrated by both $R^2$ and AIC.

Table 26

*Summary of Regression Models ($R^2$ and AIC) fitting Predictors relating to Gestures at 11 & 12 months to Expressive Vocabulary at 24 months (n=49)*

| *Predictor(s)* | $R^2$ | *AIC* |
|---|---|---|
| Frequency of Gestures (*non-significant predictor*) | .05 | 645.98 |
| Frequency of Gestures with Gaze Coordination | .09 | 643.96 |
| Frequency of Gestures + Frequency of Gaze | .06 | 647.37 |

Table 26 shows that the frequency of gestures with gaze coordination was a better predictor of expressive vocabulary at 24 months than the frequency of gestures (as denoted by $R^2$) and also was a better predictor than a model with both the frequency of gestures and the frequency of gaze to caregiver's face as predictors (as denoted by AIC, full model in Appendix J, Table 94). The raw frequency of gestures is correlated with the frequency of gestures with gaze coordination ($r$ = .68).

For showing gestures we were unable to establish whether gaze coordination better predicted later expressive vocabulary. The frequency of showing, combinations involving showing, and showing combined with non-CV specifically were all predictive of 24-month expressive vocabulary, as were the same predictors with gaze coordination. However, we are unable to determine whether gaze coordination is a better predictor than the raw frequency, as showing (whether in combination or produced alone) is almost always produced with gaze coordination (see Appendix H, Table 82). As such these predictors are highly correlated (for showing, $r$ = .97; for both combinations involving showing, and showing combined with non-CV vocalisations, $r$ = 1).

## Discussion

The current study demonstrated that the frequency with which a number of specific types of early vocalisations, gestures and gesture-vocal combinations are produced at the end of the first year of life predicts expressive vocabulary at 15 and 24 months of age. Furthermore, we have demonstrated that the extent to which these behaviours are coordinated with gaze to the caregiver's face is predictive of later expressive vocabulary, but only in a limited number of cases did considering gaze coordination result in being able to predict more variance in outcomes. These results suggest that for the majority of behaviours that predict later expressive vocabulary (most notably for index-finger pointing), we have provided no evidence that their predictive value was because they are early attempts to intentionally communicate. However, there is some limited scope to conclude that certain behaviours (e.g., CV vocalisations) predict later language abilities *because* they are early instances of intentional communication, discussed below.

Some of our findings replicated previous findings. We found that the frequency with which infants produced CV vocalisations at the end of the first year of life was predictive of 15-month expressive vocabulary, in line with previous research that links babble to later expressive vocabularies (D'Odorico et al., 1999; McCune & Vihman, 2001; McGillion, Herbert, et al., 2017; Menyuk et al., 1986; Stoel-Gammon, 1992). We also found that the frequency with which infants produced index-finger pointing (either alone, or in combination with a vocalisation) is a predictor of 15- and 24-month expressive vocabulary, confirming links between index-finger pointing and later language (Camaioni et al., 1991; Desrochers et al., 1995). We found that the frequency of showing gestures (whether produced alone or in combination with vocalisations), like index-finger pointing, positively predicted 24-month expressive vocabulary. As showing is often conceptualized as an early declarative, the link to later language has been frequently hypothesized (Boundy et al., 2016). However, to our knowledge, this has only been empirically tested once, on a smaller sample (E. Bates et al., 1979). Finally, we found that infants that gaze to their caregiver's face more frequently have higher expressive vocabulary at 15 months. This finding is broadly consistent with literature that demonstrates the link between the infants gaze to caregivers' faces and later language measures. For example, gazing to a caregiver's face could be a conceived of as initiation of joint attention, which at 9 and 15 months is known to predict receptive vocabulary at 24-months, and at 18 months is known to predict expressive vocabulary (Mundy et al., 2007).

Other findings are novel, particularly in relation to how gaze-coordinated behaviours predict later language. We found that the frequency of CV vocalisations with gaze coordination was predictive of 15-month expressive language, and furthermore was a *better* predictor of language than the frequency that infants' produced CV vocalisations (regardless of gaze coordination). This could be taken as evidence that these vocalisations are predictors of later vocabulary because they were early instances of intentional communication, and thus demonstrate that the infant is ready to learn words. However, this interpretation should be considered in light of the fact that in the previous chapter, we provided no evidence that these behaviours were intentionally gaze-coordinated by the infant. Though this is not evidence of the absence of intentional communication, it is an absence of evidence of intentional communication. Therefore caution is warranted in the in interpretation of this relationship in terms of intentional communication. If such behaviours are not intentionally communicative, why are they predictive of later language? An intriguing alternative (but not mutually exclusive) possibility is that even if infants do not intentionally coordinate these behaviours at 11 months, caregivers respond to these speech-like sounds made with eye contact (in a way which resembles adult communication) in a way that promotes language development. This is a possibility that is explored in the next chapter.

We also found that the frequency of showing and combinations involving showing with gaze coordination (depending on inclusive and exclusive analyses), significantly predicted expressive vocabulary at 24-months. However, the vast majority of showing was gaze-coordinated, meaning it was a moot question as to whether gaze coordination increased the predictive power of showing for 24-month expressive vocabulary. It remains possible that due to the high proportion of gaze coordination, showing could be conceived of as always being intentionally communicative, and that is why it is predictive. However this is hard to unpick using our data. Showing is highly likely to be gaze-coordinated because of the physical configuration of the gesture (holding objects up to caregiver's face). It could be this physical configuration that allows infants' to attend to both an object of interest and the attention of their caregiver to that object, which could scaffold the transition to later intentional communication. It could also be that the unique configuration of the showing gesture elicits a specific kind of response from caregivers that scaffolds vocabulary learning, a possibility explored in the next chapter. While the frequency of gestures (in general) with gaze coordination was predictive of

24-month expressive vocabulary, it is likely that this effect is due to the effect of showing with gaze coordination, as the two are highly correlated and explain similar amounts of variance.

A further novel finding is that the frequency of combinations involving giving and CV vocalisations is a positive predictor of language at both 15 and 24 months. This finding highlights the need to investigate gesture-vocal combinations as a mutually exclusive category. Additionally, this was the only consistent predictor of both 15- and 24-month expressive vocabulary, suggesting that the combination of speech-like vocalisations (CV vocalisations) with gestures that invite some sort of caregiver response (giving an object) are an important means for infants to learn language. This is explored in the next chapter.

There are limitations with the current study. One set of limitations concerns the sample that we used. Firstly, it is not possible for us to generalize to other cultures outside of the UK, as our study only involved infants from UK households. Cross-cultural studies would be needed to determine if the findings observed here are generalizable to development across cultures. There is some evidence to suggest that infants could be on similar developmental paths in different cultures with regards to many early communicative behaviours (Lieven & Stoll, 2013; Liszkowski et al., 2012), however there might be differences in the frequencies with which infants produce gestures in different cultures (see Huttunen, Pine, Thurnham, & Khan, 2013 for a review). Our study only included infants brought up in a monolingual home, so it is possible that the study is not generalizable even within all UK homes, where more than one language is spoken at home. No differences have been found in the *quantity* of infant vocalisations produced in bilingual and monolingual homes (Karousou & López-Ornat, 2013; Oller, Eilers, Urbano, & Cobo-Lewis, 1997), although infants' vocalisations in bilingual homes are qualitatively different (Andruski, Casielles, & Nathan, 2014; Maneva & Genesee, 2002). There is some tentative evidence to suggest that older bilingual children (aged 4-6 years) gesture more than monolingual peers (Nicoladis, Pika, & Marentette, 2009), but more work is needed in infancy to ascertain if this is the same for prelinguistic gestures. Finally, all infants in the study were first-born, and birth order effects are known to effect expressive vocabulary and production of early gestures (Fenson et al., 1994). All these factors affect the generalizability of our results.

Concerning methodological implications, one further limitation is that we used caregiver reports of infants' expressive vocabulary as an outcome measure. This may be seen as a potentially unreliable indicator of infants' vocabulary size, as caregivers may overestimate or underestimate their children's abilities. However, recently studies have found that observational measures of infants' expressive vocabulary correlate highly with caregiver report at 18 months (McGillion, Pine, et al., 2017). Additionally, a meta-analyses from eighteen language intervention studies has demonstrated that there are no differences between expressive vocabulary measures from caregiver report and from observation, on studies with children from 18 to 60 months (M. Y. Roberts & Kaiser, 2011).

We have demonstrated that both prelinguistic vocalisations and gestures, and specific combinations of the two predicted later expressive vocabulary, thus rejecting accounts that suggest either modality is uniquely important for the development of language. It certainly seems the case that gestures, and specifically index-finger pointing are *not* the 'royal road' to language (Butterworth, 2003), though they are predictive of later language. They are predictive in tandem with vocalisations, other types of gestures and gazing to caregiver's face in general. Though we have provided no evidence that many of these behaviours (including index-finger pointing) predict later language *because* they represent early attempts to intentionally communicate, we have provided the first demonstration that vocalisations, and in particular CV vocalisations, with gaze coordination predict infants' later expressive vocabulary. This suggests that their predictive power is not exclusively because they represent the phonological building blocks of language, but also because of their social use. Whether or not infants deliberately coordinate their vocalisations with gaze remains an empirical question, and one that has implications for theories of how infants learn to intentionally communicate. On the current picture, if this coordination of gaze with vocalisations *is* as a result of chance, this could be the first chaotic steps in the learning process that leads children to their understanding of adult-like communication.

In sum, the frequency that infants produce a number of behaviours including vocalisations, gestures and specific gesture-vocal combinations, at the end of the first year of life predict later vocabulary. Some of these behaviours (e.g., CV vocalisations) are especially predictive when they are gaze-coordinated, suggesting that they may predict language as they are early instances of intentional communication. However, some may lead to later expressive vocabulary in part, or wholly because of the way in which caregivers' respond to them. This is the focus of the next chapter.

# 7. Do Caregivers Responses to Infants' Vocalisations, Gestures and Gesture-vocal Combinations Scaffold Infants' Vocabulary Learning?

**Abstract**

In the current chapter, we investigated whether the mechanism by which prelinguistic behaviours predict later language is by eliciting caregiver responses that are contingent on infants' focus of attention and thus provide material from which to learn language. In general, more frequently produced behaviours (such as vocalisations) at 11 & 12 months were more frequently responded to by caregivers than low frequency behaviours (such as gestures and combinations). However, we found that that a higher proportion of gestures and combinations were met with a response. Furthermore, a higher proportion of gaze-coordinated behaviours were met with responses than those that were not gaze-coordinated. We found no evidence of social shaping of infant behaviours through caregiver responsiveness between 11 and 12 months. We investigated the links between responses to infant behaviours produced at 11 & 12 months, and later expressive vocabulary. The best predictors of 15-month expressive vocabulary were responses to vocalisations. In particular, the frequency of gaze-coordinated CV vocalisations that were met with a response were the best predictors of 15-month expressive vocabulary (explaining 42% of variance). Furthermore, caregiver responsiveness mediated the relationship between these vocalisations and expressive vocabulary, suggesting that the predictive value of vocalisations is explained in large part because caregivers respond to them, providing material for infants to learn vocabulary from. Conversely, the best predictors of 24-month expressive vocabulary were gestural in nature. However, caregiver responsiveness to infant vocalisations, gestures or combinations was not a better predictor of expressive language at 24-months than the frequency with which infants produced these behaviours. Instead, the frequency with which infants produced index-finger pointing and showing, along with the frequency with which caregiver's produced semantically contingent infant-directed speech was predictive. This suggests that responsiveness to infants' gestures are not a main driver of infants' early vocabulary acquisition (at 15 months), but that there is good reason to expect that they may be later in development when they are produced more frequently (towards 24 months).

## Introduction

In the previous chapter, we demonstrated that a number of infant vocalisations, gestures and specific gesture-vocal combinations predicted later expressive vocabulary. These early behaviours may predict later vocabulary because they are a marker of infants' motoric or socio-cognitive readiness for speech. Furthermore, we highlighted a number of behaviours that were especially predictive when they were gaze-coordinated. Such behaviour could be predictive of later language because they are instances of intentional communication. It is also possible that infants vocalisations, gestures and combinations (either gaze-coordinated or not) may be predictive of later expressive vocabulary as they provoke a linguistic response from caregivers, effectively shaping the environment to provide exactly the experiences they need to help learn language. In this chapter, we investigate this final possibility, considering whether the reason why these early behaviours predict later language is because caregivers respond to them.

As discussed in Chapter 4, caregivers can respond in a way that is purely *temporally* contingent on infants' behaviour (i.e., they say something in quick temporal succession of a infant behaviour), or, in a way that is also *semantically* contingent (i.e., they say something in quick temporal succession of an infant behaviour that relates to the infant's focus of attention). In theory, semantically contingent caregiver responses would provide a better opportunity to learn vocabulary than just any response, as the infant will be in a better position to infer the function of the words the caregiver is using and thereby build up their lexicon (Butterworth, 2003; McGillion et al., 2013). Furthermore, if infants are intentionally communicating, and therefore expecting some kind of relevant caregiver response (i.e., one that is semantically contingent), this will be even more powerful for language learning. The current study therefore focuses only on semantically contingent responses to infants' behaviours.

Previous studies have established that the amount that caregivers respond in a semantically contingent manner to infant vocalisations, gestures or combinations predicts later expressive vocabulary. Of particular note are findings from two key studies, reviewed in Chapter 4. Firstly, the amount that caregivers responded to infants' vocalisations in a way that was semantically contingent on the infants' focus of attention at 9 months was predictive of expressive vocabulary at 18 months (McGillion et al., 2013). Secondly, the frequency with which caregivers responded to infants gestures by providing an object label (which is a type of semantically contingent response) at 13 months, predicted concurrent expressive vocabulary (Olson & Masur, 2015). Furthermore this 13-month expressive vocabulary, combined with the proportion of gestures that received an object-labelling response, predicted 17 month expressive vocabulary. Importantly, the frequency with which caregivers provided responses fully mediated the relationship between infants' gestures and expressive vocabulary. It should be noted, that in contrast to the finding by McGillion et al. (2013), Olson & Masur (2015) did not find a relationship between responses to infants communicative bids that did not involve gestures (and so involved vocalisations only), and later expressive vocabulary. However, responsiveness to gestures and vocalisations might not *differently* predict later language outcomes. McGillion et al. (2013) did not investigate whether responses to gestures were predictive of later outcomes along with vocalisations, and the differences observed by Olson & Masur (2015) could be an artefact of the way in which infant behaviours were elicited (through experimental paradigms) and might not truly reflect how responsiveness occurs in the home.

Given these contrasting findings, an important outstanding question remains as to whether responsiveness to gestures or vocalisations is especially predictive of later language. Furthermore, neither of these studies reviewed the contribution of different sub-types of vocalisations or gestures (or gesture-vocal combinations) produced by the infant, or considered whether the infant might have been intentionally communicating. The current study aims to address these limitations by providing a finer-grained analysis of a naturalistic dataset that considers responses to both vocalisations and gestures, and whether such behaviours are intentionally communicative.

Why might caregiver responses to some of infants' behaviours help them learn vocabulary, whilst responses to others do not? One simple explanation is that caregivers selectively respond to some behaviours more frequently, or are proportionally more likely to do so. This would create more learning opportunities from specific infant behaviours but not others. A related explanation is that infants have an active role in their vocabulary learning, in which they are motivated to more frequently produce specific vocalisations or gestures that elicit such responses. This could be demonstrated through social shaping of preverbal behaviours, whereby infants increase the amount that they produce specific behaviours (i.e., vocalisations or gestures) because they get a response, in turn allowing for more responses, and subsequently more opportunities for language learning.

A key focus of this chapter is the role of features of intentional communication in eliciting caregiver responses and facilitating language learning, which have previously not been studied. It is possible that behaviours that are gaze-coordinated (i.e., intentionally communicative) are more likely to be met with caregiver responses because caregivers are more likely to interpret these as attempts to communicate. Furthermore, infants' gaze-coordinated behaviours, if intentionally communicative, are likely to be attempts to elicit responses from their caregivers. It would be reasonable to hypothesise that responses to infants' intentionally communicative behaviours would better facilitate vocabulary learning as infants would be actively seeking out caregiver responses, and so more likely to learn from them. It is also possible that responses to specific gaze-coordinated behaviours are key predictors of later expressive vocabulary, suggesting that responses to specific intentionally communicative behaviours (i.e., specific vocalisation or gesture types) facilitate language.

In the following analyses it was important to control for the possibility that high levels of semantically contingent infant-directed speech (IDS) by caregivers would predict expressive vocabulary outcomes regardless of whether that speech was produced in response to infants' behaviours. High frequency of such speech could lead to high levels of this speech being produced 'in response to' infants more frequent behaviours (i.e., just after infants vocalise) due to chance. Therefore, it may be that focusing on semantically contingent speech in a temporal window after high frequency infant behaviours (e.g.., focussing on caregiver responsiveness to vocalisations) might create a proxy measure for how much IDS is semantically contingent on infant activity *in general*. The amount of semantically contingent IDS given by caregivers (not necessarily in response to infant behaviours) is known to predict later expressive language outcomes (McGillion, Pine, et al., 2017; Rollins, 2003). It is worth noting that McGillion et al. (2013) found that the proportion of caregiver IDS that was semantically contingent did not predict 18-month expressive vocabulary outcomes, while semantically contingent *responses* to infant vocalisations did. However, this study did not test whether the *frequency* with which IDS was semantically contingent predicted outcomes, which is especially important. This motivates the need to consider both the frequency and proportion of IDS that was semantically contingent, when assessing the role of responses to infant behaviours.

In the previous chapter, we highlighted a number of infant behaviours that predict later expressive vocabulary, some of which were especially predictive when gaze-coordinated. The final part of our analyses will focus on whether the predictive relationship between these behaviours and expressive vocabulary is mediated by the effect of caregiver responsiveness. Thus, we will investigate the mechanism by which infants' production of prelinguistic vocalisations and gestures may facilitate vocabulary learning.

In sum, in this chapter, we will explore six questions regarding the role of caregiver responsiveness in infants' language learning. The first two concern the interaction between infant behaviours and caregiver responses at 11 & 12 months:

1. Are certain types of infant behaviour (e.g., vocalisations, gestures or combinations, or sub-types), or gaze-coordinated behaviours frequently, or proportionally more likely to be met with a response?
2. Do responses to infant behaviours at 11 months cause them to be produced more frequently at 12 months?

The second group of questions concern the way in which such caregiver responsiveness to infant behaviours at 11 & 12 months might predict later expressive vocabulary outcomes. These analyses will be tested first in relation to 15-month, and then 24-month expressive vocabulary outcomes.

3. Does the frequency or proportion of infant behaviours that are met with a response predict later expressive vocabulary?
4. Furthermore, does the frequency or proportion of *gaze-coordinated* infant behaviours that are met with a response predict later expressive vocabulary?
5. Do relationships between infant behaviours met with a response and later expressive vocabulary (identified for question 3 & 4) hold when the amount of caregiver's semantically contingent IDS (not necessarily speech given as a response to infant behaviour) is controlled for?

6. Does caregiver responsiveness to infant behaviours mediate the relationship between infants' production of these behaviours and later expressive vocabulary (established in Chapter 6)?

## Method

### Participants

As in the previous chapters, participants were drawn from a large sample of families (*N* = 140) who had previously participated in a longitudinal randomised control study (McGillion, Pine, et al., 2017). This data comes from the same participants as in Chapter 5 (see Chapter 5, Participants). We also pooled data from 11 and 12 months for some analyses, resulting in 61 caregiver-infant dyads with data at both time points. For analyses of language outcomes at 15- and 24-months, participants were the same as in Chapter 6 (see Chapter 6, Participants).

### Coding

All IDS produced by the caregiver had been transcribed as part of the longitudinal study (McGillion, Pine, et al., 2017). This had also been categorised into speech that was *semantically contingent* to what the infant was attending to, and speech that was not semantically contingent. Semantically contingent speech is defined as an utterance that follows the infant's focus of attention. Caregiver's utterances were coded as semantically contingent or not, based on whether they followed the infant's focus of attention, based on the infant's focus of attention in a 5 second window prior to the onset of the caregiver's utterance. An infant's attention could be either towards an object or activity. The focus of an infant's attention on an object could be determined by whether the infant was looking at an object, manipulating an object (with hands or mouth), gesturing using an object (e.g., showing or giving) or towards an object (e.g., pointing), or naming an object. Note that the infant had to be manipulating, looking, or gesturing to the object, holding or touching was not considered to indicate their attention was focused on it. The focus of an infant's attention on an activity (with a caregiver, e.g., playing peek-a-boo) could be determined by whether the infant was gazing to the caregiver's face, vocalising in a related way to the activity (e.g., imitating the caregiver's vocalisations, laughing or shouting), and whether the infant was using actions, gestures or postural cues relevant to the activity (Appendix S3, McGillion, Pine, et al., 2017).

## Measures

**Infant behaviours.** Infant behaviours produced at 11 and 12 months from Chapter 5 (see Chapter 5, Coding, Measures and Results) were used in this chapter.

**Caregiver responses.** We extracted caregivers' IDS occurring after an infant began a vocalisation, gesture or combination, and within 1s of that infant behaviour ending (1s window also used in McGillion, Pine, et al., 2017). We considered only semantically contingent utterances, as this was our focus.

**Expressive vocabulary at 15- and 24-months.** Language outcomes from Chapter 6 (see Chapter 6, Materials and Results) were used for this chapter.

## Analyses

Throughout this chapter, we compare how often infant behaviours were met with a response. Analyses differ depending on whether we were considering frequency of responses or proportion of behaviours that met with a response.

***Frequency***. To determine if there were differences in the frequency with which behaviours were met with a response depending on whether they were gaze-coordinated, and/or on the behaviour type (either the main analyses comparing vocalisations, gestures and combinations, or across sub-types of these behaviours), we constructed linear regression models using "lme" function in "nlme" package (Pinheiro et al., 2015) in R (R Core Team, 2015). For each type of behaviour (e.g., vocalisation), for each infant there were two data points: the frequency with which the infant produced this behaviour and it was met with a response when 1) that behaviour was gaze-coordinated, and 2) that behaviour was not gaze-coordinated. We used the following random effect structure: (1|infant/behaviour/coordination). Initially, we constructed a null model, and then added gaze coordination as a predictor which allowed us to compare the frequency with which gaze-coordinated and non-gaze-coordinated behaviours were met with a response). Following this we added behaviour type as a predictor, and finally we added the interaction term (gaze coordination * behaviour type). We report if there was a significant improvement from adding either predictor or the interaction term, determined using the "anova" function, part of the "stats" package (R Core Team, 2015). Post-hoc comparisons were conducted using "lsmeans" function in "lsmeans" package (Lenth, 2016).

***Proportion*.** To determine if there were differences in the proportion of behaviours that were met with a response depending on whether they were gaze-coordinated, and/or on the behaviour type, we constructed logistic regression models using "glmer" function in "lme4" (D. Bates et al., 2015). These models were fitted to datasets including every behaviour as a separate data point. In the case of analyses of sub-types, e.g., types of vocalisations, the dataset only included instances of that behaviour, e.g., vocalisations. For each behaviour, we coded whether it was met with a response or not (1 = met with a response, 0 = not). This was the outcome variable. We also included whether it was gaze-coordinated (1= gaze-coordinated, 0 = not), and behaviour type as fixed effects. Infant was included as a random effect on the intercept and on the slope (gaze coordination * behaviour type), except when model comparison deemed that this was unnecessary. Initially, we constructed a null model, and then added gaze coordination as a predictor. Following this, we added behaviour type as a predictor, and finally we added the interaction term (gaze coordination * behaviour type). Again, we report if there was a significant improvement from adding behaviour type as a predictor, determined using "anova" function. Post-hoc comparisons were again conducted using "lsmeans" function.

**Social shaping.** To determine whether social shaping occurred between 11 and 12 months for any behaviour, we constructed linear regression models using "lm" function, part of the "stats" package (R Core Team, 2015). We were interested in whether either the frequency that the behaviour was met with a response or the proportion of times it was met with a response at 11 months predicted the amount the behaviour was produced at 12 months. Our first model included the frequency of the behaviour at 12 months as the outcome variable, and the frequency of the behaviour at 11 months as the predictor. We then added either the frequency that the behaviour was met with a caregiver response at 11 months, or the proportion that the behaviour was met with a caregiver response at 11 months and tested for improvement of fit.

We report if there was a significant improvement from adding either predictor, determined using the “anova” function.

We did not test for effects of social shaping when fewer than 20 infants had produced that specific type of behaviour at 11 months (given that the regression model required 2 predictors – applying the rule of thumb that $n$ = 10 are required per predictor). Additionally, to ensure that there was variance in the proportional scores, 50% or more of infants had to produce more than 2 instances of the behaviour in order for us to test whether the proportion that the behaviour was met with a caregiver response at 11 months predicted the frequency it was produced at 12 months. Finally, if the frequency that an infant produced a behaviour, and the frequency that the behaviour was met with a response were too highly correlated ($r$ > .80), we did not run this model. In practice, this meant that we could only test models involving frequency of behaviours for vocalisations (and only non-CV vocalisations), and frequency of gaze-coordinated behaviours for vocalisations (both CV and non-CV) and gestures (regardless of sub-type). We only tested models involving proportions of behaviours or gaze-coordinated behaviours for vocalisations (both CV and non-CV).

**Predicting language.** All analyses predicting expressive vocabulary consist of linear regression models, fit to either 15- or 24-month expressive vocabulary outcomes using the “lm” function.

For measures of whether caregiver responsiveness to specific behaviours was predictive of later language, as in the previous chapter, our analyses were complicated by the sheer number of potential predictors to test, and differing levels of granularity of infant behaviours. Our method is therefore identical to Chapter 6 (testing broad categories first, followed by specific vocalisation or gesture types, and finishing with a final model from only those predictors who had improved at previous levels of granularity).

Again, we did not include behaviours as predictors in *any* model when less than 10% of infants had produced that specific type of behaviour.

***Proportional analyses*.** Similarly, when testing whether the proportions of behaviours that were responded to by caregivers predict language outcomes, the procedure for testing was the same as in the previous chapter (for each behaviour, the dataset for the model only included those infants who produced the behaviour)

Again, for proportional models, we did not include behaviours when fewer than $n = 10$ infants produced that type of behaviour. Furthermore, to ensure that there was variance in the proportional scores, 50% or more of infants had to produce more than 2 instances of the behaviour.

**Controlling for rates of semantically contingent IDS.** Once final models were created (see above), we added a control variable to determine whether predictive relationships held when the amount of caregiver's semantically contingent IDS was controlled for. We separately added the frequency of IDS that was semantically contingent, and the proportion of IDS that was semantically contingent into models. For brevity, we only report these analyses where either of these predictors were significant predictors when included in the model, or if they changed another predictor to non-significance.

**Mediation analyses.** Mediation analyses were conducted in R, using "lm" and "mediate" function, part of the "mediation" package (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014)). Standardized regression coefficients between predictor, mediator and the outcome variable were calculated through linear regression (using "lm"). The average causal mediation effect (or indirect effect) and significance of this effect was tested by generating 95% confidence intervals using quasi-Bayesian Monte Carlo approximation (with 1000 resamples) in "mediate".

## Results

We answer the questions posed in the introduction in order. The first questions focus exclusively on behaviours at 11 and 12 months, whilst the second set of questions focus on expressive vocabulary outcomes at 15- and 24-months. As in the previous chapters, we present the inclusive analyses (where gesture-vocal combinations are not considered as a mutually exclusive category from gestures and vocalisations) in Appendix K.

**Are Certain Types of Infant Behaviour, or Gaze-coordinated Behaviours Frequently, or Proportionally More Likely to be Met with a Response?**

Descriptive statistics for caregiver responses at 11 and 12 months separately as well as 11 & 12 months combined, are presented in Tables 27-29. These show both the frequency and proportion with which infants produced behaviours that were met with a caregiver response regardless of gaze coordination. Additionally, they show both the frequency and proportion with which infants produced specifically gaze-coordinated behaviours that were met with a caregiver response. Exhaustive descriptives relating to sub-types of gesture-vocal combinations (e.g., index-pointing and CV vocalisations) are presented in Appendix L, Tables 125-127.

Table 27

*Average Frequency of Infant Behaviours, and Proportion of Infant Behaviours that were Met with a Response at 11 months (n = 134)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 10.67 | 9.00 | 8 | 0-52 | .22 | 3.90 | 4.08 | 3 | 0-20 | .27 |
| CV | 4.49 | 5.95 | 2 | 0-43 | .23 | 1.03 | 1.99 | 0 | 0-15 | .28 |
| Non-CV | 6.18 | 5.06 | 5 | 0-30 | .22 | 1.60 | 2.07 | 1 | 0-10 | .26 |
| **Gestures** | 1.22 | 1.87 | 0.5 | 0-10 | .74 | 0.78 | 1.45 | 0 | 0-8 | .78 |
| Index-finger point | 0.17 | 0.65 | 0 | 0-6 | .76 | 0.09 | 0.43 | 0 | 0-4 | .87 |
| Open-hand point | 0.11 | 0.39 | 0 | 0-2 | .57 | 0.04 | 0.28 | 0 | 0-2 | .67 |
| Give | 0.60 | 1.30 | 0 | 0-8 | .90 | 0.38 | 1.05 | 0 | 0-7 | .90 |
| Show | 0.20 | 0.60 | 0 | 0-3 | .68 | 0.20 | 0.59 | 0 | 0-3 | .69 |
| Conventional gesture | 0.13 | 0.53 | 0 | 0-5 | .56 | 0.07 | 0.25 | 0 | 0-1 | .61 |
| **Gesture-vocal combinations** | 0.81 | 1.65 | 0 | 0-10 | .73 | 0.49 | 0.96 | 0 | 0-5 | .80 |
| CV | 0.49 | 1.14 | 0 | 0-8 | .81 | 0.27 | 0.66 | 0 | 0-3 | .85 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Non-CV | 0.32 | 0.94 | 0 | 0-6 | .58 | 0.22 | 0.69 | 0 | 0-4 | .72 |
| Index-finger point | 0.24 | 1.11 | 0 | 0-10 | .74 | 0.09 | 0.41 | 0 | 0-3 | .81 |
| Open-hand point | 0.10 | 0.64 | 0 | 0-7 | .54 | 0.04 | 0.30 | 0 | 0-3 | .75 |
| Give | 0.29 | 0.81 | 0 | 0-5 | .87 | 0.20 | 0.61 | 0 | 0-4 | .98 |
| Show | 0.11 | 0.42 | 0 | 0-3 | .75 | 0.10 | 0.41 | 0 | 0-3 | .73 |
| Conventional gesture | 0.07 | 0.30 | 0 | 0-2 | .70 | 0.04 | 0.21 | 0 | 0-1 | .67 |

Table 28

*Average Frequency of Infant Behaviours, and Proportion of Infant Behaviours that were Met with a Response at 12 months (n = 64)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 8.54 | 7.74 | 7 | 0-36 | .21 | 1.91 | 2.24 | 1 | 0-9 | .28 |
| CV | 3.45 | 4.26 | 2 | 0-21 | .20 | 0.67 | 0.92 | 0 | 0-4 | .30 |
| Non-CV | 5.09 | 5.15 | 4 | 0-33 | .23 | 1.23 | 1.71 | 1 | 0-7 | .28 |
| **Gestures** | 1.86 | 3.05 | 1 | 0-15 | .69 | 0.82 | 1.50 | 0 | 0-7 | .71 |
| Index-finger point | 0.31 | 1.07 | 0 | 0-7 | .45 | 0.06 | 0.30 | 0 | 0-2 | .40 |
| Open-hand point | 0.03 | 0.17 | 0 | 0-1 | .67 | 0.00 | 0.00 | 0 | 0-0 | N/A |
| Give | 1.32 | 2.61 | 0 | 0-12 | .71 | 0.61 | 1.34 | 0 | 0-7 | .77 |
| Show | 0.13 | 0.38 | 0 | 0-2 | .65 | 0.08 | 0.27 | 0 | 0-1 | .56 |
| Conventional gesture | 0.06 | 0.30 | 0 | 0-2 | .42 | 0.06 | 0.30 | 0 | 0-2 | 1.00 |
| **Gesture-vocal combinations** | 1.12 | 1.91 | 0 | 0-8 | .60 | 0.71 | 1.37 | 0 | 0-7 | .67 |
| CV | 0.61 | 1.31 | 0 | 0-7 | .67 | 0.40 | 0.90 | 0 | 0-4 | .71 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Non-CV | 0.50 | 1.16 | 0 | 0-7 | .55 | 0.31 | 0.95 | 0 | 0-6 | .60 |
| Index-finger point | 0.27 | 0.97 | 0 | 0-7 | .60 | 0.13 | 0.57 | 0 | 0-4 | .40 |
| Open-hand point | 0.06 | 0.30 | 0 | 0-2 | .28 | 0.05 | 0.21 | 0 | 0-1 | .50 |
| Give | 0.57 | 1.27 | 0 | 0-7 | .72 | 0.35 | 0.93 | 0 | 0-6 | .78 |
| Show | 0.16 | 0.44 | 0 | 0-2 | .73 | 0.16 | 0.44 | 0 | 0-2 | .73 |
| Conventional gesture | 0.06 | 0.24 | 0 | 0-1 | .80 | 0.03 | 0.18 | 0 | 0-1 | 1.00 |

Table 29

*Average Frequency of Infant Behaviours, and Proportion of Infant Behaviours that were Met with a Response at 11 & 12 months (n = 61)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 18.78 | 14.09 | 14 | 0-66 | .21 | 4.45 | 3.96 | 4 | 0-19 | .28 |
| CV | 7.92 | 9.61 | 4 | 0-51 | .22 | 1.66 | 2.34 | 1 | 0-13 | .32 |
| Non-CV | 10.87 | 7.66 | 10 | 0-44 | .21 | 2.79 | 2.58 | 2 | 0-9 | .26 |
| **Gestures** | 3.10 | 3.81 | 2 | 0-16 | .72 | 1.55 | 2.01 | 1 | 0-8 | .74 |
| Index-finger point | 0.51 | 1.41 | 0 | 0-7 | .55 | 0.13 | 0.56 | 0 | 0-4 | .54 |
| Open-hand point | 0.13 | 0.49 | 0 | 0-3 | .47 | 0.06 | 0.38 | 0 | 0-3 | .50 |
| Give | 1.84 | 2.72 | 1 | 0-12 | .77 | 0.89 | 1.55 | 0 | 0-7 | .83 |
| Show | 0.36 | 0.71 | 0 | 0-3 | .77 | 0.30 | 0.64 | 0 | 0-3 | .74 |
| Conventional gesture | 0.26 | 0.75 | 0 | 0-5 | .57 | 0.16 | 0.42 | 0 | 0-2 | .69 |
| **Gesture-vocal combinations** | 1.86 | 2.44 | 1 | 0-9 | .71 | 1.15 | 1.72 | 0 | 0-8 | .73 |
| CV | 1.04 | 1.64 | 0 | 0-7 | .77 | 0.66 | 1.07 | 0 | 0-4 | .75 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Non-CV | 0.82 | 1.51 | 0 | 0-7 | .59 | 0.49 | 1.15 | 0 | 0-6 | .61 |
| Index-finger point | 0.43 | 1.33 | 0 | 0-7 | .65 | 0.13 | 0.59 | 0 | 0-4 | .30 |
| Open-hand point | 0.26 | 0.98 | 0 | 0-7 | .50 | 0.13 | 0.46 | 0 | 0-3 | .62 |
| Give | 0.78 | 1.52 | 0 | 0-7 | .78 | 0.51 | 1.15 | 0 | 0-6 | .86 |
| Show | 0.30 | 0.67 | 0 | 0-3 | .77 | 0.30 | 0.67 | 0 | 0-3 | .77 |
| Conventional gesture | 0.10 | 0.30 | 0 | 0-1 | .75 | 0.08 | 0.28 | 0 | 0-1 | .83 |

We first investigated whether the type of behaviour (either vocalisations, gestures or combinations) and whether or not such behaviour was gaze-coordinated meant that they were more frequently, or proportionally more likely to be met with a response.

***Frequency***. As was established in Chapter 5, infants vocalise more frequently than they gesture or produce combinations. It is therefore not surprising that responses to vocalisations are more frequent than responses to gestures or combinations. At both ages, there was a significant improvement in model fit (see Method) when gaze coordination (11 months: $\chi^2$ (1) = 50.00, $p$ < .001; 12 months: $\chi^2$ (1) = 19.26, $p$ < .001; 11 & 12 months: $\chi^2$ (1) = 27.60, $p$ < .001) and behaviour type (11 months: $\chi^2$ (2) = 217.75, $p$ < .001; 12 months: $\chi^2$ (2) = 83.09, $p$ < .001; 11 & 12 months: $\chi^2$ (2) = 112.88, $p$ < .001) were added as predictors. There was also an interaction between behaviour type and gaze coordination (11 months: $\chi^2$ (2) = 173.59, $p$ < .001; 12 months: $\chi^2$ (2) = 51.07, $p$ < .001; 11 & 12 months: $\chi^2$ (2) = 85.41, $p$ < .001 (Appendix M, Tables 128-130)). Post-hoc tests (Appendix M, Table 131) revealed that at 11 months (and when collapsed across age), a significantly higher frequency of vocalisations were met with a response than gestures or combinations (both gaze-coordinated and not gaze-coordinated). At 12 months, this was only the case for infant behaviours that were not gaze-coordinated (i.e., gaze-coordinated vocalisations were not met with a response more frequently than gaze-coordinated gestures or combinations). This is possibly due to the reduction in vocalisations, and increase in gaze-coordinated gestures (specifically giving) between 11 and 12 months reported in Chapter 5. At both ages (and when collapsed across age), a significantly higher frequency of vocalisations that were *not* gaze-coordinated were met with a response than those that were gaze-coordinated. This is most likely due to the higher frequency of vocalisations produced without gaze coordination than produced with gaze coordination (see Chapter 5).

***Proportion*.** At both ages (and when collapsed across age), there was a significant improvement in model fit when gaze coordination (11 months: $\chi^2$ (1) = 17.96, $p$ < .001; 12 months: $\chi^2$ (1) = 8.44, $p$ = .004; 11 & 12 months: $\chi^2$ (1) = 15.20, $p$ < .001) and behaviour type (11 months: $\chi^2$ (2) = 100.29, $p$ < .001; 12 months: $\chi^2$ (2) = 65.18, $p$ < .001; 11 & 12 months: $\chi^2$ (2) = 89.63, $p$ < .001 (Appendix M, Tables 132-134)) were added as predictors. There was no interaction between behaviour type and gaze coordination. At both ages (and when collapsed across age), a significantly higher proportion of gestures and combinations were met with a response than vocalisations. A significantly higher proportion of behaviours that were gaze-coordinated were met with a response than those that were not gaze-coordinated.

**Type of vocalisation**. We investigated whether the type of vocalisation (produced alone or in a combination; either CV or non-CV vocalisations) and whether or not such vocalisations or combinations were gaze-coordinated meant that they were more frequently, or proportionally more likely to be met with a response.

***Frequency.*** As was established in Chapter 5, non-CV vocalisations were more frequently produced than CV vocalisations, so again it is not surprising non-CV vocalisations are more frequently met with a response. Considering vocalisations produced alone (i.e., not as part of a combination), at both ages (and when collapsed across age), there was a significant improvement in model fit when gaze coordination, (11 months: $\chi^2$ (1) = 102.82, $p < .001$; 12 months: $\chi^2$ (1) = 38.30, $p < .001$; 11 & 12 months: $\chi^2$ (1) = 58.85, $p < .001$), and vocalisation type (11 months: $\chi^2$ (1) = 9.03, $p = .003$; 12 months: $\chi^2$ (1) = 5.18, $p = .023$; 11 & 12 months: $\chi^2$ (1) = 5.00, $p = .025$ (Appendix M, Tables 135-137)) were added as predictors, but these two factors did not interact. At both ages (and when collapsed across age), a significantly higher frequency of non-CV vocalisations were met with a response than CV vocalisations. At both ages (and when collapsed across age), a significantly higher frequency of vocalisations without gaze coordination were met with a response than those that were gaze-coordinated (as noted above). This is most likely due to the higher frequency of vocalisations without gaze coordination (80%) produced by infants than those with gaze coordination (see Chapter 5).

We conducted the same analyses for combinations. When we collapsed across age only, there was a significant improvement in model fit when gaze coordination was added as a predictor ($\chi^2$ (1) = 5.03, $p = .025$ (Appendix M, Table 138)), but not for vocalisation type. A significantly higher frequency of gaze-coordinated combinations were met with a response than combinations that were not gaze-coordinated.

***Proportion.*** Considering vocalisations produced alone, at both ages (and when collapsed across age) there was a significant improvement in model fit when gaze coordination was added as a predictor (11 months: $\chi^2$ (1) = 19.74, $p < .001$; 12 months: $\chi^2$ (1) = 7.00, $p =$

.008; 11 & 12 months: $\chi^2$ (1) = 21.85, $p$ < .001 (Appendix M, Tables 139-141)), but not vocalisation type. At both ages (and when collapsed across age), a significantly higher proportion of gaze-coordinated vocalisations were met with a response than those that were not gaze-coordinated.

We lacked data to properly investigate the effect of vocalisation type and gaze coordination for combinations at separate ages, however when we collapsed across age we found that there was no effect of either on the proportion of infants' combinations that were met with a response.

**Type of gesture.** We investigated whether the type of gesture (produced alone or in a combination; either index-finger pointing, open-hand pointing, giving, showing or conventional gestures) and whether or not such gestures or combinations were gaze-coordinated meant that they were more frequently, or proportionally more likely to be met with a response.

***Frequency.*** As was established in Chapter 5, giving is more frequently produced than other gestures, so again it is not surprising giving is more frequently met with a response. Considering gestures produced alone (i.e., not as part of a combination), at 11 months only, there was a significant improvement in model fit when gaze coordination was added as a predictor, $\chi^2$ (1) = 7.12, $p$ = .008. At both ages (and when we collapsed across age), there was a significant improvement in model fit when gesture type was added as a predictor (11 months: $\chi^2$ (4) = 38.49, $p$ < .001; 12 months: $\chi^2$ (4) = 43.76, $p$ < .001; 11 & 12 months: $\chi^2$ (4) = 55.01, $p$ < .001 (Appendix M, Table 143 and 144)). At 11 months only, there was also an interaction effect between behaviour type and gaze coordination, $\chi^2$ (4) = 13.25, $p$ = .010 (Appendix M, Table 142). Post-hoc tests (Appendix M, Tables 145 and 146) revealed that at 11 months, a significantly higher frequency of gaze-coordinated giving was met with a response than gaze-coordinated index-finger pointing, open-hand pointing and conventional gestures. At 12 months (and when collapsed across age), a significantly higher frequency of giving was met with a response than all other gestures (regardless of gaze coordination). This is most likely due to the higher frequency of giving (and giving with gaze coordination) produced by infants compared to other gestures (see Chapter 5). At 11 months, a significantly higher frequency of giving without gaze coordination was met with a response than showing without gaze coordination.

Further, a significantly higher frequency of gaze-coordinated showing was met with a response than showing without gaze coordination. This is most likely due to a low frequency of showing without gaze coordination produced by infants (see Chapter 5).

We conducted the same analysis for combinations. At 12 months (and when we collapsed across age) there was a significant improvement in model fit when gaze coordination was added as a predictor (12 months: $\chi^2$ (1) = 4.19, $p$ = .041; 11 & 12 months: $\chi^2$ (4) = 4.41, $p$ = .036). At both ages (and when we collapsed across age) there was a significant improvement in model fit when gesture type was added as a predictor (11 months: $\chi^2$ (4) = 9.87, $p$ = .043 (Appendix M, Table 147); 12 months: $\chi^2$ (4) = 20.24, $p$ < .001; 11 & 12 months: $\chi^2$ (4) = 14.34, $p$ = .006). At 11 months (and when we collapsed across age), there was an interaction effect between gaze coordination and gesture type, (11 months: $\chi^2$ (4) = 12.68, $p$ = .013; 11 & 12 months: $\chi^2$ (4) = 16.20, $p$ = .003 (Appendix M, Tables 148 and 149)). Post-hoc tests revealed that at 11 months no pairwise interactions were significant. However when we collapsed across age, a significantly higher frequency of gaze-coordinated combinations involving giving were met with a response than gaze-coordinated combinations involving index-finger pointing, open-hand pointing or conventional gestures (Appendix M, Table 151). At 12 months, a significantly higher frequency of combinations involving giving were met with a response than combinations involving open-hand pointing, showing, and conventional gestures (Appendix M, Table 150). This is most likely due to the higher frequency of combinations involving giving (and combinations involving giving with gaze coordination) produced by infants compared to combinations involving other gestures (see Chapter 5). When we collapsed across age a significantly higher frequency of gaze-coordinated combinations involving showing was met with a response than combinations involving showing without gaze coordination. This is most likely due to a very low frequency of combinations involving showing without gaze coordination produced by infants (see Chapter 5).

***Proportion.*** We lacked the data to properly investigate proportional differences for either gestures or combinations.

**Summary.** As has been a theme of the previous chapters, focusing on frequency or proportional measures of caregiver responsiveness affects what behaviours are more likely to be met with a caregiver response. Certain types of infant behaviour are more frequent and are therefore more frequently met with caregiver responses. Vocalisations (compared to gestures and combinations), non-CV vocalisations (compared to CV vocalisations), and giving gestures (compared to other gestures and combinations) were most frequently responded to. Similarly, regarding gaze coordination, vocalisations without gaze coordination were more frequent than those with gaze coordination, and equally were more frequently met with caregiver responses. However, higher proportions of gestures and combinations were met with a response than vocalisations, and higher proportions of gaze-coordinated vocalisations, gestures and combinations were met with a response than those that were not gaze-coordinated.

**Do Responses to Infant Behaviours at 11 Months Cause Infants to Produce These Behaviours More Frequently at 12 Months?**

If infants have an active role in their vocabulary learning, in where they are motivated to more frequently produce specific vocalisations or gestures to elicit specific responses, we could expect social shaping of preverbal behaviours, whereby infants increase the amount that they produce specific behaviours (i.e., vocalisations or gestures) between 11 and 12 months because they get a response at 11 months. We first investigated whether the frequency or proportion of infant behaviours that were met with a response at 11 months predicted the frequency that they were produced at 12 months, controlling for the frequency that they were produced at 11 months. Then, as we were particularly interested in whether there was social shaping of gaze-coordinated (i.e., intentionally communicative) behaviours, we conducted the same analyses, focusing on the frequency of gaze-coordinated infant behaviours. However, we found no evidence of social shaping for any behaviour (that it was possible to test, see Method).

**Predicting Expressive Language**

We posed a number of questions (questions 3 - 6, see Introduction) that relate to later expressive vocabulary outcomes. For analyses relating to these questions, we opted to collapse the 11- and 12-month data (as in Chapter 6), in order to provide enough variance in measures regarding infants’ gestures and combinations (see Table 29 for descriptives).

We first focus on the relationship between behaviours at 11 & 12 months and expressive vocabulary at 15 months. We initially focus on question 3, posed in the introduction: whether the frequency with which specific types of infant behaviour (vocalisations, gestures and combinations, and the sub-types) were met with a response at 11 & 12 months predicted 15-month expressive vocabulary. We then consider the same question focussing on the *proportion* of their behaviours that were met with a response. Then, as we are also interested in whether responsiveness to gaze-coordinated behaviours (i.e., intentionally communicative behaviours) is especially predictive of later expressive vocabulary, we then looked at predictors involving gaze-coordinated infant behaviours for both frequency and proportional measures (question 4). Next, we investigated whether predictive relationships still held when controlling for the frequency of IDS that was semantically contingent (question 5). Finally, we explored whether the relationship between infant behaviours at 11 & 12 months and expressive vocabulary at 15 months is mediated by caregiver responsiveness (question 6). Subsequently, all analyses are rerun (questions 3-6) focusing on 24-month vocabulary outcomes.

### Does the Frequency or Proportion of Infant Behaviours that are Met with a Response Predict 15-month Expressive Vocabulary?

**Frequency.** We first explored whether the frequency with which infant behaviours were produced and met with a response predicted 15-month expressive vocabulary. The method for testing predictors (see Method) was identical to the method also used in Chapter 6. However, unlike in Chapter 6 (see Chapter 6, Results), here we compress the results for brevity of reporting. That is to say, we initially tested behaviours at the coarsest level of granularity (vocalisations, gestures and combinations that were met with a response) before considering sub-types of these behaviours. Instead of presenting our results in subsections at separate granular levels (as in Chapter 6), we list all predictors that improved on the null model regardless of granular level, and then present the final model.

After determining which predictors improve on the null model, we also consider whether predictive relationships between behaviours held after controlling for semantically contingent speech (question 5). As a first step we checked whether the overall rate of semantically contingent IDS at 11 & 12 months predicted expressive vocabulary at 15 months regardless of whether it was produced in response to an infant behaviour. However neither the frequency of semantically contingent IDS nor the proportion of IDS that was semantically contingent was predictive of 15-month expressive vocabulary. We return to this point, after determining whether the frequency with which vocalisations, gestures and combinations are met with a caregiver response predict later expressive vocabulary (question 4).

The frequency with which a number of behaviours were produced by infants and were met with a response predicted 15-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,51) = 9.97$, $p = .003$, $R^2 = .16$), and specifically CV vocalisations ($F(1,51) = 17.84$, $p < .001$, $R^2 = .26$), combinations ($F(1,51) = 4.20$, $p = .046$, $R^2 = .08$) and specifically combinations involving CV vocalisations ($F(1,51) = 14.14$, $p < .001$, $R^2 = .22$), combinations involving index-finger pointing ($F(1,51) = 4.87$, $p = .032$, $R^2 = .09$), combinations involving index-finger pointing and CV vocalisations ($F(1,51) = 4.81$, $p = .033$, $R^2 = .09$), and combinations involving giving and CV vocalisations ($F(1,51) = 16.14$, $p < .001$, $R^2 = .24$). To check the extent to which these behaviours explained separate variance, we systematically tested for improvements in fit when combining appropriate predictors (see Method). The final model is presented in table 30.

Table 30

*Regression Model fitting Frequency of CV Vocalisations, Combinations involving Index-finger Pointing and CV Vocalisations, Combinations involving Giving and CV Vocalisations that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 8.55 | 3.68 | 2.32 | .025 |
| Frequency of CV Vocalisations that were met with a response | 0.80 | 0.34 | 2.37 | .022 |
| Frequency of Combinations involving Index-finger Pointing and CV Vocalisations that were met with a response | 5.98 | 2.62 | 2.28 | .027 |
| Frequency of Combinations involving Giving and CV Vocalisations that were met with a response | 9.96 | 4.59 | 2.17 | .035 |

$R^2$ = .38, $F(3,49) = 10.19$, $p < .001$.

**Proportion.** We next explored whether the proportion that infant behaviours were met with a response were predictors of 15-month expressive vocabulary. We used the same method as for frequency analyses, regarding granularity of predictors (i.e., testing main types, followed by sub-types) on subsets where all participants had produced each type of behaviour (see Method).

Only the proportion of vocalisations ($F(1,51) = 5.16$, $p = .027$, $R^2 = .09$) and specifically CV vocalisations ($F(1,51) = 5.92$, $p = .019$, $R^2 = .10$ (Table 31)) that were met with a response improved on the null model.

Table 31

*Regression Model fitting the Proportion of CV Vocalisations that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 9.23 | 5.69 | 1.62 | .111 |
| Proportion of CV Vocalisations that were met with a response | 53.59 | 22.03 | 2.43 | .019 |

$R^2$ = .10, $F(1,51) = 5.92$, $p = .019$.

**Does the Frequency or Proportion of *Gaze-coordinated* Infant Behaviours that are Met with a Response Predict 15-month Expressive Vocabulary?**

We were particularly interested in whether responses to gaze-coordinated (i.e., intentionally communicative) behaviours were especially predictive of later expressive vocabulary. We conducted the same analyses as above but considered whether the frequency or proportion of infant behaviours that were *gaze-coordinated* and met with a response were predictors of 15-month expressive vocabulary.

**Frequency.** The frequency with which a number of behaviours were produced with gaze coordination by infants and were met with a response predicted 15-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,51) = 23.45$, $p < .001$, $R^2 = .31$), and specifically CV vocalisations ($F(1,51) = 36.90$, $p < .001$, $R^2 = .42$), combinations involving CV vocalisations ($F(1,51) = 7.39$, $p = .009$, $R^2 = .13$), combinations involving index-finger pointing and CV vocalisations ($F(1,51) = 4.51$, $p = .039$, $R^2 = .08$) and combinations involving giving and CV vocalisations ($F(1,51) = 5.58$, $p = .022$, $R^2 = .10$). The final model was similar to when we considered responsiveness to these behaviours regardless of gaze coordination, except that it explained much more variance (Table 32).

Table 32

*Regression Model fitting Frequency of Gaze-coordinated CV Vocalisations, Combinations involving Index-finger Pointing and CV Vocalisations, Combinations involving Giving and CV Vocalisations that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 5.89 | 3.13 | 1.89 | .065 |
| Frequency of *Gaze-coordinated* CV Vocalisations that were met with a response | 6.00 | 1.00 | 6.02 | < .001 |
| Frequency of *Gaze-coordinated* Combinations involving Index-finger Pointing and CV Vocalisations that were met with a response | 11.53 | 4.03 | 2.86 | .006 |
| Frequency of *Gaze-coordinated* Combinations involving Giving and CV Vocalisations that were met with a response | 12.41 | 5.21 | 2.38 | .021 |

$R^2 = .54$, $F(3,49) = 19.07$, $p < .001$.

***Proportion.*** We conducted the same analyses but considered whether the proportion of specific infant behaviours that were *gaze-coordinated* and met with a response were predictors of 15-month expressive vocabulary.

Only the proportion of vocalisations ($F(1,51) = 11.71$, $p = .001$, $R^2 = .19$), specifically CV vocalisations ($F(1,51) = 9.22$, $p = .004$, $R^2 = .15$) and non-CV vocalisations ($F(1,51) = 4.53$, $p = .038$, $R^2 = .08$) that were *gaze-coordinated* and met with a response improved on the null model. Adding the two predictors regarding the sub-types of vocalisation resulted in an improvement of fit, and represents our final model (Table 33).

Table 33

*Regression Model fitting Proportion of Gaze-coordinated CV and Non-CV Vocalisations that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 9.06 | 4.90 | 1.85 | .070 |
| Proportion of CV Vocalisations that were *Gaze-coordinated* and met with a response | 144.71 | 58.20 | 2.49 | .016 |
| Proportion of Non-CV Vocalisations that were *Gaze-coordinated* and met with a response | 89.53 | 66.70 | 1.34 | .186 |

$R^2$ = .18, $F(2,50) = 5.58$, $p = .006$.

**Do Relationships Between Infant Behaviours Met with a Response and 15-month Expressive Vocabulary Hold when the Amount of Caregiver's Semantically Contingent IDS is Controlled For?**

We wanted to check whether the relationship between caregivers' responsiveness to specific infant behaviours and 15-month expressive vocabulary (as investigated above) was explained instead by rates of semantically contingent IDS by caregivers. However, when we added the frequency of semantically contingent IDS or the proportion of IDS that was semantically contingent as a predictor to the models in Tables 30-33, they were non-significant predictors. This suggests that the relationship between caregivers' responsiveness to specific infant behaviours and 15-month expressive vocabulary (as investigated above) is not due to these behaviours being a proxy for rates of semantically contingent IDS.

**Does Caregiver Responsiveness to Infant Behaviours Mediate the Relationship Between Infants' Production of these Behaviours and 15-month Expressive Vocabulary?**

In Chapter 6, we demonstrated that 15-month expressive vocabulary was predicted by the frequency with which infants produced a number of behaviours, and in some cases, the frequency or proportion with which these behaviours were gaze-coordinated. In the current chapter, we have demonstrated that responses to some of these behaviours also predicts expressive vocabulary at 15 months. The question arises, then, as to whether some infant behaviours predict outcomes *because* they are responded to. To recap results so far, Table 34 summarises the predictive relationships we have demonstrated throughout this chapter and Chapter 6 and expressive vocabulary at 15 months.

Table 34

*Predictors of 15-month expressive vocabulary (with $R^2$)*

| | *Vocalisations* | | | *Combinations* | | | | |
|---|---|---|---|---|---|---|---|---|
| *Predictors of 15-month expressive vocabulary* | *All* | *CV* | *Non-CV* | *All* | *CV* | *Index-finger Point* | *Index-finger Point + CV* | *Give + CV* |
| *Frequency* | | | | | | | | |
| Frequency produced* | | .15 | | | .16 | | .07 | .19 |
| Gaze-coordinated* | .18 | .29 | | | | | | |
| Met with a response | .16 | .26 | | .08 | .22 | .09 | .09 | .24 |
| Gaze-coordinated and met with a response | .31 | .42 | | | .13 | | .08 | .10 |
| *Proportion* | | | | | | | | |
| Gaze-coordinated* | .12 | .07 | .08 | | | | | |
| Met with a response | .09 | .10 | | | | | | |
| Gaze-coordinated and met with a response | .19 | .15 | .08 | | | | | |

** relationship established in Chapter 6*

For the behaviours in Table 34 where both the infant behaviour and the caregiver response predicted outcomes, we conducted mediation analyses to determine whether caregiver responsiveness was the mechanism by which these behaviours predicted 15-month vocabulary. Caregiver responsiveness did not mediate the relationship between non-CV vocalisations, combinations involving index-finger pointing and CV vocalisations or combinations involving giving and CV vocalisations, to 15-month expressive vocabulary. For other behaviours, we found mediating effects of caregiver responsiveness, discussed in detail below. Note, that we did not investigate the proportional measures for these, as the frequency measures explained more variance and thus were of greater interest.

***Vocalisations***. Figure 16 shows that the relationship between the frequency of gaze-coordinated vocalisations and 15-month expressive vocabulary was mediated by the frequency with which these vocalisations were met with a response, $b = 0.44$, $p < .01$.

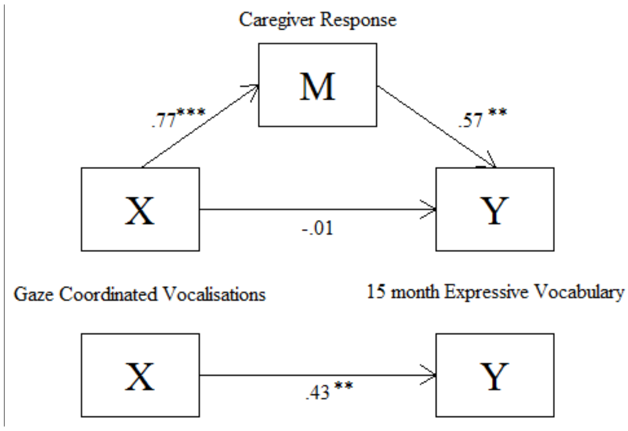


*Figure 16.* Standardized regression coefficients for the relationship between frequency of gaze-coordinated vocalisations and 15-month expressive vocabulary, mediated by the frequency with which these were met with a response. ** p < .01, *** p < .001

***CV Vocalisations.*** Figure 17 shows that the relationship between the frequency of CV vocalisations (and *gaze-coordinated* CV vocalisations) and 15-month expressive vocabulary was mediated by the frequency with which these vocalisations were met with a response, (regardless of gaze coordination: $b = 0.54$, $p < .01$; gaze-coordinated: $b = 0.43$, $p < .01$).

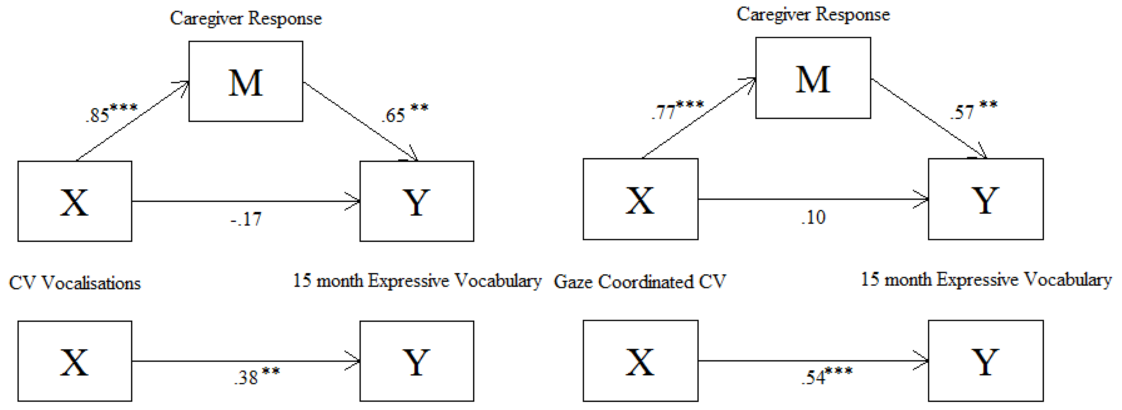


*Figure 17.* Standardized regression coefficients for the relationship between frequency of CV vocalisations (left) or gaze-coordinated CV vocalisations (right) and 15-month expressive vocabulary, mediated by the frequency with which these were met with a response. ** p < .01, *** p < .001

***Combinations involving a CV vocalisation.*** Figure 18 shows that the relationship between the frequency of combinations involving CV vocalisations and 15-month expressive vocabulary was mediated by the frequency that these were met with a response, $b = 0.94$, $p = .03$.

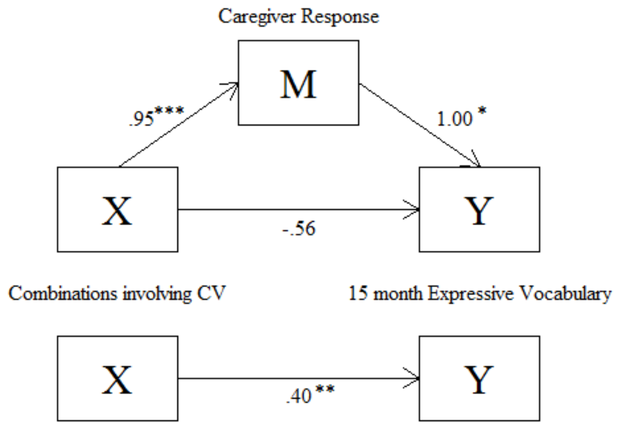


*Figure 18.* Standardized regression coefficients for the relationship between frequency combinations involving CV vocalisations and 15-month expressive vocabulary, mediated by the frequency with which these were met with a response. * p < .05, ** p < .01, *** p < .001

**Does the Frequency or Proportion of Infant Behaviours that are Met with a Response Predict 24-month Expressive Vocabulary?**

Having investigated how caregiver responses to infant behaviours predicts 15-month outcomes, we then turned our attention to 24-month outcomes. Again, we focus on questions 3-6 posed in the introduction.

We initially tested whether responses to specific types of infant behaviour (vocalisations, gestures and combinations, and the sub-types) at 11 & 12 months predicted 24-month expressive vocabulary. As before, we first considered the frequency with which infants produced behaviours that were responded to, and then proportional measures.

As for the 15-month analyses, as a first step, we checked whether the overall rate of semantically contingent IDS at 11 & 12 months predicted expressive vocabulary at 24 months regardless of whether it was produced in response to an infant behaviour. The frequency of IDS that was semantically contingent at 11 & 12 months was a significant predictor of 24-month expressive vocabulary, $F(1,47) = 8.40$, $p = .006$, $R^2 = .15$, however the proportion of IDS that was semantically contingent did not. Unlike for 15-month expressive vocabulary, controlling for the frequency of semantically contingent IDS also affected some final models. We have highlighted where this is the case.

**Frequency.** We explored whether the frequency of specific infant behaviours that were met with a response were predictors of 24-month expressive vocabulary. We used the same method as for predicting 15-month expressive vocabulary (regarding granularity of predictors to be tested).

The frequency with which a number of behaviours were produced by infants and were responded to predicted 24-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,47) = 4.91$, $p = .032$, $R^2 = .09$), and specifically CV vocalisations ($F(1,47) = 4.35$, $p = .042$, $R^2 = .08$), index-finger pointing ($F(1,47) = 7.70$, $p = .008$, $R^2 = .14$), showing ($F(1,47) = 6.80$, $p = .012$, $R^2 = .13$), combinations involving CV vocalisations ($F(1,47) = 4.45$, $p = .040$, $R^2 = .09$), and giving combined with a CV vocalisation ($F(1,47) = 4.90$, $p = .032$, $R^2 = .09$).

We systematically tested for improvements in fit when combining appropriate predictors (see Method). The final model is presented in Table 35, Model 1. However, additionally, when we added the frequency of semantically contingent IDS as a predictor to this model, though it was a non-significant predictor it reduced the role of combinations involving giving and CV vocalisations that were met with a response to non-significant (Table 35, Model 2). This suggests that the frequency of combinations involving giving and CV vocalisations that were met with a response at 11 & 12 months did not predict expressive vocabulary at 24 months when we controlled for the frequency of semantically contingent IDS, however the same was not true for showing or index-finger pointing.

Table 35

*Regression Model fitting Frequency of Index-finger Pointing, Showing, and Combinations involving Giving and CV Vocalisations that were Met with a Caregiver Response (model 1) and Frequency of Semantically Contingent IDS (model 2) at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| **Model 1** | | | | |
| Intercept | 287.53 | 26.36 | 10.91 | < .001 |
| Frequency of Index-finger Pointing that was met with a response | 40.22 | 13.19 | 3.05 | .004 |
| Frequency of Showing that was met with a response | 82.36 | 27.02 | 3.05 | .004 |
| Frequency of Combinations involving Giving and CV Vocalisations that were met with a response | 63.53 | 24.13 | 2.63 | .012 |
| $R^2 = .36, F(3,45) = 8.51, p < .001.$ | | | | |
| **Model 2** | | | | |
| Intercept | 205.93 | 64.42 | 3.20 | .003 |
| Frequency of Index-finger Pointing that was met with a response | 35.80 | 13.44 | 2.66 | .011 |
| Frequency of Showing that was met with a response | 76.87 | 27.04 | 2.84 | .007 |
| Frequency of Combinations involving Giving and CV Vocalisations that were met with a response | 49.64 | 25.90 | 1.92 | .062 |
| Frequency of Semantically Contingent IDS | 0.50 | 0.36 | 1.39 | .173 |
| $R^2 = .39, F(4,44) = 6.99, p < .001.$ | | | | |

**Proportion.** We next explored whether the proportion with which specific infant behaviours were met with a response were predictors of 24-month expressive vocabulary. We used the same method as for frequency analyses, regarding granularity of predictors (i.e., testing main types, followed by sub-types) on subsets where all participants had produced each type of behaviour (see Method).

The proportion of vocalisations ($F$(1,47) = 10.08, $p$ = .003, $R^2$ = .18), specifically non-CV vocalisations ($F$(1,47) = 11.12, $p$ = .002, $R^2$ = .19), gestures ($F$(1,47) = 4.25, $p$ = .046, $R^2$ = .09) and specifically giving ($F$(1,30) = 4.83, $p$ = .036, $R^2$ = .14) that were met with a response improved on null models.  Combining the two compatible predictors (non-CV vocalisations and giving) did not improve fit, suggesting that they explained similar variance. Additionally, when we controlled for the frequency of semantically contingent IDS by caregivers, neither significantly predicted expressive language at 24 months.

**Does the Frequency or Proportion of Gaze-coordinated Infant Behaviours that are Met with a Response Predict 24-month Expressive Vocabulary?**

As we were particularly interested in whether responses to gaze-coordinated (i.e., intentionally communicative) behaviours were especially predictive of later expressive vocabulary. We conducted the same analyses as above but considered whether the frequency or proportion of infant behaviours that were *gaze-coordinated* and met with a response were predictors of 24-month expressive vocabulary.

**Frequency.** The frequency with which a number of behaviours were produced with gaze coordination by infants and were met with a response predicted 24-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F$(1,47) = 5.69, $p$ = .021, $R^2$ = .11), and specifically non-CV vocalisations ($F$(1,47) = 4.39, $p$ = .041, $R^2$ = .09), showing ($F$(1,47) = 5.57, $p$ = .022, $R^2$ = .11) and combinations involving CV vocalisations ($F$(1,47) = 4.66, $p$ = .036, $R^2$ = .09).

The final model included the frequency of gaze-coordinated vocalisations, and showing as significant predictors of 24-month expressive vocabulary (Table 36, model 1). However, when we added the frequency of semantically contingent IDS as a predictor, it was a significant predictor, and furthermore reduced the role of the frequency of gaze-coordinated vocalisations that were met with a response to non-significant (Table 36, model 2). This suggests that the frequency of gaze-coordinated vocalisations that were met with a response at 11 & 12 months did not predict expressive vocabulary at 24 months when we controlled for the frequency of semantically contingent IDS, however the same was not true for gaze-coordinated showing.

Table 36

*Regression Model fitting Frequency of Gaze-coordinated Vocalisations and Showing (model 1) and Frequency of Semantically Contingent IDS (model 2) at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| **Model 1** | | | | |
| Intercept | 285.15 | 35.55 | 8.02 | < .001 |
| Frequency of *Gaze-coordinated* Vocalisations that were met with a response | 14.35 | 6.15 | 2.33 | .024 |
| Frequency of *Gaze-coordinated* Showing that was met with a response | 76.18 | 32.99 | 2.31 | .026 |
| *$R^2$ = .20, F(2,46) = 5.77, p = .006.* | | | | |
| **Model 2** | | | | |
| Intercept | 158.08 | 68.14 | 2.32 | .025 |
| Frequency of *Gaze-coordinated* Vocalisations that were met with a response | 8.22 | 6.57 | 1.25 | .217 |
| Frequency of *Gaze-coordinated* Showing that was met with a response | 76.08 | 31.76 | 2.40 | .021 |
| Frequency of Semantically Contingent IDS | 0.83 | 0.39 | 2.16 | .036 |
| *$R^2$ = .28, F(3,45) = 5.70, p = .002.* | | | | |

**Proportion.** Finally, we conducted the same analyses but considered whether the proportion of specific infant behaviours that were *gaze-coordinated* and met with a response were predictors of 24-month expressive vocabulary.

Only the proportion of vocalisations ($F(1,47) = 5.30$, $p = .026$, $R^2 = .10$) and specifically non-CV vocalisations ($F(1,47) = 6.10$, $p = .017$, $R^2 = .11$) that were gaze-coordinated and met with a response improved on the null model. However, when we controlled for the frequency of semantically contingent IDS by caregivers, the proportion of gaze-coordinated non-CV vocalisations that were met with a respond did not predict expressive language at 24 months.

**Does Caregiver Responsiveness to Infant Behaviours Mediate the Relationship Between Infants' Production of these Behaviours and 24-month Expressive Vocabulary?**

Through this chapter, and Chapter 6, we have demonstrated that frequency and proportional measures of infant behaviours, their gaze coordination and the extent to which they are responded to by caregivers predict expressive vocabulary at 24 months. Table 37 demonstrates what predictive relationships we have demonstrated throughout this chapter and Chapter 6 between and expressive vocabulary at 24 months.

Table 37

*Predictors of 24-month expressive vocabulary (with $R^2$)*

| *Predictors of 24-month expressive vocabulary* | *Vocalisations* | | | *Gestures* | | | | *Combinations* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *All* | *CV* | *Non-CV* | *All* | *Index-finger Pointing* | *Giving* | *Showing* | *CV* | *Showing* | *Giving + CV* | *Showing + Non-CV* |
| *Frequency* | | | | | | | | | | | |
| Frequency produced* | | | | | .14 | | .10 | | .10 | .09 | .10 |
| Gaze-coordinated* | | | | .09 | | | .09 | | .10 | | .10 |
| Met with a response | .09 | .08 | | | .14 | | .13 | .09 | | .09 | |
| Gaze-coordinated and met with a response | .11 | | .09 | | | | .11 | .09 | | | |
| *Proportion* | | | | | | | | | | | |
| Gaze-coordinated* | | | | | | | | | | | |
| Met with a response | .18 | | .19 | .09 | | .14 | | | | | |
| Gaze-coordinated and met with a response | .10 | | .11 | | | | | | | | |

** relationship established in Chapter 6*

For all of the behaviours in Table 37 where both the infant behaviour and the caregiver response predicted outcomes,we conducted mediation analyses to determine whether caregiver responsiveness was the mechanism by which these behaviours predicted 24-month vocabulary. However, caregiver responsiveness did not mediate the relationship between any of these behaviours and 24-month expressive vocabulary.

## Discussion

We investigated whether infant's vocalisations, gestures and combinations at 11 & 12 months were met with responses from caregivers that were contingent on the infants' focus of attention, and whether this responsiveness predicted expressive vocabulary at both 15 and 24 months. We found that infants' vocalisations were more frequently met with responses than gestures or combinations were, but that higher proportions of infants' gestures and combinations were met with caregiver responses (with minor differences when considering 12-month-olds alone). Furthermore, gaze-coordinated vocalisations, gestures and combinations were proportionally more likely to meet with a response than those that were not gaze-coordinated. We provided no evidence that responses to behaviours at 11 months cause these to be produced more at 12 months, although we demonstrated that the frequency of responses given to gestures at 11 months was negatively related to the frequency gestures were produced at 12 months.

We investigated the links between responses to infant behaviours produced at 11 & 12 months, and later expressive vocabulary. There were large contrasts in which behaviours were predictive of either 15- or 24-month vocabulary outcomes. Regarding 15-month expressive vocabulary, the best predictors tended to be vocal in nature. We demonstrated that the frequency with which infants' CV vocalisations, combinations involving index-finger pointing and CV vocalisations, and combinations involving giving and CV vocalisations were met with a response, predicted 15-month expressive vocabulary, as did the proportion of CV vocalisations that were met with a response. The relationship between these behaviours and 15-month expressive vocabulary were even stronger when we considered only gaze-coordinated vocalisations. The single best predictor of 15-month language outcomes was the frequency with which gaze-coordinated CV vocalisations were met with a response (which singularly explained

42% of variance). These relationships held when we controlled for the amount of semantically contingent IDS produced by caregivers, suggesting that it is specifically semantically contingent speech given *in response to* these infant behaviours that is especially predictive of language. Furthermore, we demonstrated that the relationship between CV vocalisations (and gaze-coordinated CV vocalisations) and 15-month vocabulary was mediated by the frequency with which they were met with a caregiver response. We also observed mediation effects for vocalisations in general, and combinations involving CV vocalisations, but these explained far less variance in later expressive vocabulary.

Regarding 24-month expressive vocabulary, the best predictors tended to be gestural in nature. We demonstrated that the frequency with which infants' index-finger pointing and showing gestures (which was nearly always gaze-coordinated) were met with a response predicted later expressive vocabulary. Other relationships did not hold when we controlled for the frequency of semantically contingent IDS produced by caregivers, which in itself was predictive of 24-month expressive vocabulary. However, caregiver responsiveness did not mediate the relationship between these behaviours (index-finger pointing and showing) and 24-month expressive vocabulary. These findings are discussed below.

Firstly, by far the most important finding of this chapter is that, during 20 minutes of naturalistic interaction at 11 and 12 months, the frequency with which infants produce gaze-coordinated CV vocalisations that are met with a response is highly predictive of expressive vocabulary at 15-months, explaining 42% of the variance in caregivers' report of vocabulary. This suggests that both the frequency of learning events, and infant communicative intent matter for learning. Gaze-coordinated CV vocalisations are produced more frequently than gestures (though less frequently than non-CV vocalisations; as we demonstrated in Chapter 5). As a result of being produced frequently, they are more frequently met with a response from caregivers that is semantically contingent to what they are attending to. Gaze-coordinated CV vocalisations are also more likely to be responded to than non-gaze-coordinated CV vocalisations. Thus, the key driver of early vocabulary acquisition is frequency of caregivers' prompt responses, contingent to the focus of an infant's attention, to gaze-coordinated, speech-like vocalisations. Crucially, this explains more variance than other behaviours, including gestures, which are often thought of representing a developmental royal road to language (Tomasello et al., 2007). Specifically, it has been hypothesised that caregivers' responses to infants' gestures, and specifically index-finger pointing, are the key means by which infants acquire early vocabulary (Butterworth, 2003). Our findings do not support this view. However, rehabilitation of this view bears discussion (as below), especially in relation to our other findings.

When we looked at 24-month language outcomes, the picture changed substantially. In the previous chapter we found that the frequency of two early declarative gestures, namely showing and index-finger pointing, best predicted 24-month expressive vocabulary. Importantly, here we have shown the relationship between the frequency of these behaviours and later language was not due to caregiver responsiveness. Given that we demonstrated so strongly the link between prompt, relevant caregiver responses to infants' gaze-coordinated CV vocalisations and 15-month vocabulary, one might ask why caregiver responses to infants' behaviours are not linked so strongly to 24-month vocabulary. One explanation is that a year is a long time developmentally speaking, *especially* given that during that time infants begin to use more gestures and some words, which will in turn shapes the way in which their caregivers respond. This means that developmental relationships between specific behaviours, caregiver responses and language development do not remain static, and so characterising them at 11 & 12 months does not necessarily characterize the interactions the child will experience in the following year (Masur et al., 2005).

Furthermore, behaviours that best predict 24-month outcomes, might give us an insight to the sort of interactions that will be crucial for vocabulary development during the second year of life. Higher frequencies of showing and index-finger pointing along with high frequencies of semantically contingent IDS predict 24-month expressive vocabulary, suggesting that the infants who are gesturing more at the end of the first year of life, and whose caregivers who are especially responsive to the focus of their attention will be best placed in the intervening year to learn vocabulary. As gestures become more frequent in the second year of life, responses to them might become more crucial to language development. This is line with findings that suggest that responsiveness to gestures aids transitions from one-word to two-word stage during the end of second of year of life (Goldin-Meadow & Butcher, 2003). In this way, gestures can provide referential indicators with which caregivers who are sensitive to the focus of infants' attention can respond to with more complex linguistic responses (Balog & Brentari, 2008; Fasolo & D'Odorico, 2012; Rowe & Goldin-Meadow, 2009b). It should be noted that pointing is not unique in directing caregivers attention towards a referent, as caregivers are fairly accomplished at recognising the referent of infants' vocalisations even in the absence of any visual information (Kersken, Zuberbühler, & Gomez, 2017). However, this pattern of results is in line with findings that infants do not privilege information given in response to pointing at 12 months, whereas they do at 18 months (Lucca & Wilbourn, 2016). This warrants further investigation to determine whether caregivers' responses to gestures during the second year of life (for example, around 18 months) is predictive of later language outcomes. For the moment we conclude that, while it is certainly not the case that responses to infants' gestures at 11 & 12 months are the main driver of vocabulary acquisition, there is good reason to predict that they will become increasingly important once they begin to be produced more frequently.

These results give reason to think that intentional communication may also play an important role in infants' vocabulary development. That is to say that caregivers respond to behaviours that appear intentionally communicative (i.e., are gaze-coordinated) in a way that scaffolds infants' vocabulary development. However, it is worth noting that this does not constitute evidence that such behaviours are intentionally communicative. Higher proportions of gaze-coordinated behaviours are responded to in a semantically contingent manner by caregivers. Furthermore the best predictor of expressive vocabulary at 15 months was the frequency with which *gaze-coordinated* CV vocalisations were responded to (as discussed above). In addition, predictors of 24-month expressive vocabulary included responses to showing (and combinations involving showing), which (as discussed in Chapter 5) are highly likely to be gaze-coordinated. However, for both of these behaviours, the caveats discussed in Chapter 6 (regarding whether infants intentionally coordinate gaze to their caregiver's face with CV vocalisations or showing) still remain. It is worth noting that it is hard to tease apart from our findings if communicative intent is important because it indicates that infants wanted to learn, or because it prompted a response that aided infants' learning.

We did not demonstrate evidence of social shaping, thus we did not provide support for the view that infants have an active role in their vocabulary learning in the sense that they engage more in behaviours (over time) that get caregiver responses. Given that vocalisations are the most frequently responded to behaviours, this is perhaps unsurprising, and social shaping may already have occurred prior to the end of the first year of life, as evidenced by high rates of vocalisations, and caregiver responses. Equally, prelinguistic vocalisations may no longer be infants 'target' behaviour, and we may see social shaping effects instead with their early words.

We demonstrated the importance of controlling for the frequency of semantically contingent IDS when determining whether responsiveness to specific behaviours predicts later language. Though this was not a factor in predicting expressive vocabulary at 15 months, a number of predictors of expressive vocabulary at 24 months were not predictive when this was controlled for. This is in line with other findings that higher amounts of semantically contingent IDS are predictive of later expressive vocabulary (McGillion, Herbert, et al., 2017; Rollins, 2003).

Our findings also highlight an important methodological point for further studies. In general, measures of frequency of infant and caregiver behaviours were better predictors of later language than proportional ones (as in the previous chapter). This highlights the need to focus on natural frequencies of behaviours produced by infants and caregivers. Underlining this point, we found that caregiver responsiveness to both infants' vocalisations and gestures predict later language, in contrast to findings from experimental paradigms that demonstrated only a link between caregiver responses to gestures and later language (Olson & Masur, 2015). This is almost certainly due to our focus on behaviours in naturalistic settings. As we have demonstrated here, overwhelmingly predictors involving *frequencies* of behaviours that are met with response are predictive of later language, it is crucial therefore to observe natural frequencies of behaviours and how they relate to later language in order to provide a complete picture of what facilitates vocabulary development.

In sum, this chapter has demonstrated that caregiver responses to infants' behaviours, specifically gaze-coordinated speech-like vocalisations, are crucial for early language development. The previous chapter demonstrated that the frequency of CV vocalisations seemed to be predictors of language, and this could be explained in terms of vocal readiness. However, it is clear from this chapter that their predictive value is explained in large part because caregivers respond to them, and this provides material for them to learn from. It seems likely that both explanations are true. In terms of predicting more distant language outcomes at 2 years, we did not demonstrate that responsiveness to specific infant behaviours was crucial, however we argue that there are indications that responsiveness, and in particular responsiveness to infants' gestures will play more of a role during the second year of life. Overall, around the first birthday, prompt, contingent responses to infants speech-like (and potentially intentionally communicative) vocalisations are the most important facilitators of language acquisition while over the second year of life this picture is likely to change.

# 8. General Discussion

In this thesis we have explored two broad questions concerning the development of intentional communication in the first years of life. The first two empirical chapters focused on whether toddlers might demonstrate an understanding of the intentional structure of communication earlier than has previously been demonstrated. The second half of the thesis focused on whether infants produced intentionally communicative behaviour at the end of the first year of life and how this links to infants' later expressive vocabulary.

In this final discussion chapter, we briefly summarize the main results from the empirical chapters in this thesis. Then, our discussion focuses on four areas. Firstly, we discuss our findings from the first half of the thesis in terms of whether it is possible to determine if young children understand the intentional structure of communication. Secondly, we discuss our findings from the second half of the thesis with respect to whether infants CV vocalisations are intentional around their first birthday. Thirdly, we discuss whether early infant behaviours are indicators of communicative readiness. Fourthly, we discuss our findings from the second half of the thesis in the context of theories of language evolution. For each area, we consider the theoretical implications of our results, the limitations of our work and discuss future directions. Finally, we provide some concluding remarks.

## Summary of Main Results

In the first half of the thesis (Chapters 2-3), we investigated whether toddlers understand the intentional structure of communication in terms of communicative and informative intentions. The current evidence, based on hidden authorship studies, suggests that such an understanding is present at 3 years of age (Grosse et al., 2013), however in Chapters 2 and 3 we attempted to test whether it could be demonstrated earlier. We investigated whether 18- to 30-month-olds could demonstrate that they understood the intentional structure of communication, by manipulating it or recognising when it had been manipulated. Specifically, in Chapter 2, we investigated whether infants pretended not to hear the communication of others, selectively avoiding ostensively recognizing the communicative intention of others, so as to avoid complying with their informative intention. Despite caregiver report that such behaviour is present for at least some children as young as 18-month-olds, we were unable to demonstrate that infants pretended not to hear their caregivers under controlled conditions. In Chapter 3, we investigated whether toddlers' at the same age would recognise the difference between a scenario in which communication broke down between two adults because one did not notice the other communicating, and another scenario in which communication broke down because one adult ignored the communication of the other. We did not observe any differences in any measures of toddlers' responses to these scenarios that would suggest that toddlers are sensitive to the intentional structure of communication. These chapters collectively did not give us grounds to claim that an understanding of the intentional structure of communication is present prior to 3 years of age.

In the second half of the thesis (Chapters 5-7), we investigated whether infants produced intentionally communicative vocalisations, gestures or gesture-vocal combinations at the end of the first year of life, and how these behaviours predicted later expressive vocabulary. To do so, we focused on whether these behaviours were coordinated with gaze to caregiver's face, which we argued (in Chapter 4) demonstrates that they have informative intentions (they intend to direct their interlocutor's attention). We demonstrated that infants do coordinate gaze with vocalisations, gestures and combinations, making it plausible that they engage in intentional communication around the end of the first year of life (Chapter 5). We then asked whether infants coordinated their vocalisations, gestures or combinations with gaze to caregiver's face above the rate expected by chance. We found that on the whole vocalisations, gestures and combinations were coordinated significantly above chance. However, fine-grained analyses of vocalisation and gesture types revealed that a number of behaviours were not coordinated above chance, most notably CV vocalisations and index-finger pointing. We asked whether gaze coordination was more likely for any behaviours, and found that vocalisations were more frequently gaze-coordinated than gestures or combinations, but that higher proportions of gestures and combinations were gaze-coordinated.

In Chapter 6, we investigated whether the behaviours investigated in Chapter 5 were predictive of later expressive vocabulary at 15 and 24 months, and whether those behaviours that were gaze-coordinated were especially predictive of later vocabulary. The frequency with which CV vocalisations were coordinated with gaze was especially predictive of 15-month expressive vocabulary (more so than the frequency with which they were produced regardless of gaze coordination). The frequency with which infants produced index-finger pointing was predictive of 24-month expressive vocabulary, but it was not more predictive when gaze-coordinated. The frequency with which infants produced showing (which was nearly always gaze-coordinated) was especially predictive of 24-month expressive vocabulary.

Finally, in Chapter 7 we investigated caregivers' semantically contingent responses to infant behaviours at 11 & 12 months, and how these responses might predict later language, or mediate predictive links observed in Chapter 6. Our main finding was that the frequency with which infants produced gaze-coordinated CV vocalisations that were responded to by caregivers predicted 42% of variance in 15-month expressive vocabulary outcomes. Caregiver responsiveness also mediated the relationship between gaze-coordinated CV vocalisations and expressive language observed in Chapter 5. In contrast, caregivers' responses to infants' vocalisations, gestures or combinations was not especially predictive of 24-month outcomes, over and above the predictive links established in the previous chapter based on the infants' behaviour alone.

## The Intentional Structure of Communication

In Chapters 2 and 3, we attempted to determine whether toddlers understood the intentional structure of communication, in terms of communicative intentions and informative intentions and the relationship between these levels. While there is good reason to believe that pretending not to hear provides a demonstration of this, we were unable to demonstrate unambiguously that toddlers pretend not to hear their caregivers under experimental conditions. However, this is by no means evidence that it is impossible to demonstrate if toddlers pretend not to hear, and that evidence that they understand the intentional structure of communication is not unobtainable in principle. One option, as we mentioned in our discussion in Chapter 2, would be to investigate this in more naturalistic settings. Specifically, it is likely that toddlers would not pretend not to hear every caregiver request that they don't want to comply with, but do so only when they have a good chance of successfully "pulling off" the pretence. For example, when caregivers are distracted after giving the request, so the onus is on the toddler to reengage with their caregiver, or continue with their activity having pretended they hadn't heard. Whether it is plausible to set this up experimentally is debatable. This is much more likely to happen in the home during play. Given this, the most promising avenue for future work would be to video naturalistic play sessions in the home (much as in the second half of thesis) to find naturalistic observations of this behaviour and potentially validate caregiver report. However, we do not know how frequent this behaviour is, and just how much observation would be required.

Likewise, we provided no evidence that infants noticed a difference between our two types of communicative breakdown in Chapter 3 (not noticing versus ignoring). Potential reasons for this have been discussed in Chapter 3. Thesis include the fact that our dependent variables (e.g. looking times or repair) may not be appropriate to measure infants' understanding in these situations, because of the lack of clear hypotheses for why infants might look at one researcher over the other, or repair one situation over the other. A possible avenue would be an experimental setup whereby infants apportion rewards to either researcher based on what they have witnessed. The clear hypothesis here being that the researcher who actively ignores the communication of the other would receive less than the researcher who doesn't notice the communication of the other.

In sum, it is no easy task to investigate whether toddlers' understand the intentional structure of communication, and it should be noted that only one published study has attempted this (Grosse et al., 2013). This previous study investigated cases of hidden authorship, whilst we focused on pretending not to hear. Certainly, the clear logic is that instances where toddlers have to manipulate the intentional structure of communication, or recognise that is had been manipulated provides the best opportunity to test whether they understand this structure. Our attempt here represents the first steps towards a new way of answering this question that warrants further study.

**Are Infants' CV Vocalisations Intentionally Communicative Around Their First birthday?**

In Chapters 5-7, we demonstrated that the frequency with which infants produced CV vocalisations was especially predictive of later 15-month expressive vocabulary when these CV vocalisations were accompanied with gaze to the caregiver's face, and when their caregiver responded in a semantically contingent way. This provides evidence that infants may be producing CV vocalisations in an intentionally communicative manner, and the responses to which facilitate language learning. However, there is a key issue that warrants discussion regarding this finding, namely that in Chapter 5 we demonstrated that CV vocalisations were not coordinated with gaze above the level expected by chance alone. This therefore provides no evidence that infants' are deliberately coordinating their CV vocalisations with gaze, and so these may not be intentionally communicative.

There are therefore two possible interpretations at this juncture. First, it is possible that we can conclude that infants are not communicating intentionally because their gaze coordination with CV vocalisations happened as a result of chance. However, caregivers interpret their infants' actions as intentionally communicative (even though they were not), and as such their responses still facilitate language development. On this account, infants' would be passive learners, who are not deliberately attempting to create opportunities to learn language, but learn from them when they occur. Alternatively, we could conclude that regardless of the chance analyses, infants are communicating intentionally because instances of gaze coordination with CV vocalisations predict later expressive vocabulary, a process facilitated by caregivers' semantically contingent responses. On this account, infants are more active learners, deliberately attempting to communicate, and learning from the responses of their caregivers. As noted in Chapter 7, absence of evidence of intentional communication is not evidence of absence of intentional communication. Though there is an *absence* of evidence of intentional communication in terms of whether CV vocalisations are coordinated with gaze above chance levels, there *is* suggestive evidence of intentional communication in terms of the predictive link between instances of this coordination and later language (specifically, the frequency with which this coordination occurs is predictive of later expressive vocabulary). Additionally, it is worth noting that while it would follow that frequencies of intentionally communicative behaviour would predict later expressive vocabulary, the predictive relation is not necessarily proof that such early behaviours are intentionally communicative. This is undoubtedly a limitation of our study, and further evidence that CV vocalisations are intentionally communicative at 11 & 12 months is warranted. This highlights the need for further study to provide more evidence that infants' behaviour at 11 & 12 months might be intentionally communicative.

A possible avenue for further research would be to address whether infants expect a response that is informative (and from which they could learn) when they produce vocalisations, gestures, or combinations. This could be tested in two ways. In an experimental setting, we could determine if infants are in some way dissatisfied when they do not receive a semantically contingent response from their caregiver when they produce intentionally communicative behaviour. Caregivers could be instructed to give a response that is temporally contingent, but not semantically contingent to their infant's focus of attention or give a temporally and semantically contingent response. This would allow us to test whether infants expect a response that is informative to the situation from which they could learn from. Secondly, this could be tested naturalistically, looking for evidence of social shaping of infant behaviours when they receive a temporally and semantically contingent response, and determining if this is stronger than for when they get only temporally contingent responses (similar to our analyses on social shaping conducted in Chapter 7). If we can demonstrate that infants expect semantically contingent responses that are informative to their vocalisations, gestures and combinations, this would provide stronger evidence that such behaviours are intentionally communicative. This is therefore a crucial question for further study.

In sum, our findings show that the frequency with which infants produce gaze-coordinated CV vocalisations that are met with a response from caregivers, predicts their later expressive vocabulary.  Thus, it provides some evidence that such CV vocalisations are intentionally communicative. However, given that this coordination does not occur above chance levels, to support this claim, further evidence is needed that such coordination indicates that these vocalisations are intentionally communicative, such as determining whether infants expect informative responses when they produce them.

## Communicative Readiness

The results presented in Chapters 6 and 7 suggest that the predictive link between CV vocalisations and later expressive language is not solely because they are indicators of readiness for speech. Specifically, one claim is that CV vocalisations necessarily predict later language because they represent the phonological building blocks of speech, without which speech can't be produced (see Chapter 4). We demonstrated that when infants produce CV vocalisations whilst gazing to the face of a caregiver, who responds promptly by talking about the focus of the infant's attention, this scaffolds early language learning. This suggests strongly that CV vocalisations do not solely predict later language because they are indicators of readiness for speech; this suggests that they may be genuine instances of early intentional communication (though see above), where infants learn vocabulary from caregivers' responses. In contrast, we found that the frequency with which infants produced index-finger pointing, and showing (thought to be early declaratives) at the end of the first year of life was predictive of 24-month outcomes, but crucially the frequency with which they were responded to or gaze-coordinated was not especially predictive (in the case of showing we were unable to tell whether gaze coordination was especially predictive).

As we argued in Chapter 7, it is likely that this indicates that caregivers' responses to infants' gestures may become important during the second year of life. It is plausible that, as gestures become more common in infants' play in the home, caregivers' responses to these gestures begin to aid infants' language learning. This would fit with the literature that suggests that gestures elicit more linguistically complex responses from caregivers (especially when the infant produces a concurrent word) as they provide caregivers with more information about what infants are trying to communicate about (Balog & Brentari, 2008; Fasolo & D'Odorico, 2012; Rowe & Goldin-Meadow, 2009b). Though this is plausible, it isn't borne out of our findings, although we have demonstrated that high proportions of gestures and gesture-vocal combinations are met with a response by caregivers, suggesting that when they are more frequently produced, they will reliably get responses from caregivers. Our study is limited on this front, and further study around 18 months is needed to establish whether it is caregiver responses to these gestures that scaffold language learning in the home at this time. In any case, on the current evidence it is appropriate to conclude that index-finger pointing and showing at the end of the first year of life predict later vocabulary at the end of the second year of life because these are indicators of communicative readiness. Specifically, they indicate that infants are able to produce gestures, responses to which will scaffold language learning during the second year. However, we provided no evidence that they predict later vocabulary because they are intentionally communicative, or because they are responded to by caregivers at the end of the first year of life.

One limitation of our study was the length of observation time. This was especially a problem for gestures, as these were so infrequently produced. Longer observation time would have meant that we could observe higher frequencies of gestures, and future studies in this age group that seek to investigate the role of gestures in language development would need to involve longer periods of observation. However, there is an important and simple point that comes out of this research and that is that gestures are fleetingly rare at this age for the majority of children when compared to vocalisations. It is therefore unlikely that they are the main driver of early vocabulary development.

Another limitation was that we were unable to properly evaluate whether showing could be demonstrated to be intentionally communicative. This is because it was always gaze-coordinated, so we couldn't evaluate whether such cases of gaze-coordinated showing better predicted expressive language than cases where they were not gaze-coordinated. One tempting option is to claim that because of high levels of gaze coordination, showing is intentionally communicative, and that is why it predicts later language. However, we don't subscribe to this view, firstly because of the confound that arises from infants holding an object that they are attending to in front of their caregivers face, so necessarily attending to their caregiver's face (as discussed in Chapter 6). Secondly, showing only explains variance in 24-month expressive vocabulary, not 15-month expressive vocabulary. Given that, it seems unlikely that they are indicative of any advanced intentional communication at 11 & 12 months which would aid word learning in the early part of the second year. Instead, they are predictive of a more distant vocabulary outcome (24-month) in the absence of closer outcomes (15-months) suggesting that they are predictive because they are indicative of interactions that are yet to occur throughout the second year of life (as discussed previously). Finally, caregivers' responses to showing don't mediate the relationship between the frequency with which infants produce showing and 24-month expressive vocabulary. This suggests that if they are intentionally communicative, we have no evidence that infants' find responses to these informative. Unpicking precisely whether showing is predictive of later language because it is intentionally communicative is not possible with our measure of gaze coordination. Further study could seek to test whether showing is intentionally communicative based on other behavioural markers, for example whether infants persist in showing if caregivers do not seem to show an interest in the object. Similarly, investigation into whether infants find caregivers' responses informative (as proposed for CV vocalisations above) might be illuminating. However, this is unlikely to be possible at 11 & 12 months in naturalistic settings, as showing (like other gestures) is rare.

A further limitation of this study is that we only focussed on expressive vocabulary outcomes, not receptive vocabulary. After all, infants have to understand when others use language, in order to be able to communicate linguistically. This is perhaps more pertinent because gestures may better predict infants' receptive vocabulary than expressive vocabulary (McGillion, Herbert, et al., 2017). However, receptive vocabulary is problematic to test, as caregiver report of this may not be accurate; as judging whether your child understands word is no easy task when compared with judging whether your child has produced a word (see Tomasello & Mervis commentary on Fenson et al., 1994). Further study could consider infants' receptive vocabulary scores, perhaps using looking while listening paradigms (Bergelson & Swingley, 2012). However, as discussed in Chapter 1, it is unclear whether sound-object associations are the same thing as understanding a word in the truer sense (Bannard & Tomasello, 2012). If these methodological constraints could be overcome, the relative value of early vocalisations and gestures in promoting infants' receptive vocabulary could be better assessed, and would be informative in tandem with the work presented here.

## Language Evolution

The results in the second half of this thesis also pertain to debates about language evolution. On one prominent recent theory, the human capacity for language evolved in the gestural domain, that is to say that our evolutionary ancestors first began to use gestural language before spoken language (Tomasello, 2008). This theory relies upon two claims. Firstly, the claim that in the development of human infants, gestures are the primary means by which infants begin to acquire language. Secondly, the theory relies on the claim that the gestures of extant primates are much more likely to be intentionally communicative than their vocalisations (which are thought to be genetically fixed, unlearned repertoires tightly tied to emotions) suggesting that the last common ancestors of humans and great apes communicated primarily using gestures. Our data goes some way towards challenging the first claim. In Chapter 8, we discussed how our results were not compatible with the claim that gestures, specifically index-finger pointing, are the key means by which infants acquire language (Butterworth, 2003; Tomasello et al., 2007), instead demonstrating that infants' vocalisations are crucial early on. That is not to say that gestures do not have a role during the second year of life in promoting language acquisition, in fact, from our data, we propose that this is likely. However, the claim that gestures are an ontogenetically prior means by which infants acquire language is not supported here, and hence the theory that language developed in the gestural domain is also unsupported.

Furthermore, a large body of recent work demonstrates that chimpanzee's vocalisations are not genetically fixed, display evidence of vocal learning and are produced sensitively to the presence of specific audiences (Fedurek, Schel, & Slocombe, 2013; Fedurek & Slocombe, 2013; Mitani & Gros-Louis, 1998; Schel, Machanda, et al., 2013; Slocombe et al., 2010; Slocombe & Zuberbühler, 2007; Townsend, Deschner, & Zuberbühler, 2008). Chimpanzee vocalisations demonstrate key markers of intentional communication such as those 'Brunerian' criteria discussed in Chapter 1 and 4 (Schel, Townsend, et al., 2013). Our research fits with this literature in providing evidence challenging the theory that language evolved in the gestural domain. Moreover, it motivates further investigation into the prelinguistic vocalisations of young children, to determine how these scaffold later language, and strongly suggests that further investigation into the vocalisations of great apes may be informative in understanding precursors to human language.

## Concluding Remarks

This thesis has investigated intentional communication in the early years of life, addressing whether toddlers understand the intentional structure of communication, and whether infants produce intentional communication prior to the onset of language. The studies presented here demonstrate that it is possible to investigate whether toddlers understand the intentional structure of communication earlier than has previously been investigated, although our findings do not conclusively demonstrate that toddlers do understand the intentional structure of communication. The studies presented here also demonstrate that infants at the end of the first year of life might produce intentionally communicative behaviours that, when responded to by caregivers, scaffold their acquisition of language. The studies contained within this thesis therefore demonstrate ways in which toddlers' understanding of the intentional structure of communication could be measured, and provides demonstrations that infants might produce intentionally communicative behaviours prior to the onset of language.

## References


Akhtar, N., & Gernsbacher, M. A. (2008). On Privileging the Role of Gaze in Infant Social Cognition. *Child Development Perspectives*, *2*, 59–65. http://doi.org/10.1111/j.1750-8606.2008.00044.x

Ambridge, B., & Lieven, E. V. M. (2011). *Child Language Acquisition*. Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511975073

Andruski, J. E., Casielles, E., & Nathan, G. (2014). Is bilingual babbling language-specific? Some evidence from a case study of Spanish–English dual acquisition. *Bilingualism: Language and Cognition*, *17*(3), 660–672. http://doi.org/10.1017/S1366728913000655

Aureli, T., Perucchini, P., & Genco, M. (2009). Children's understanding of communicative intentions in the middle of the second year of life. *Cognitive Development*, *24*(1), 1–12. http://doi.org/10.1016/j.cogdev.2008.07.003

Aureli, T., Spinelli, M., Fasolo, M., Garito, M. C., Perucchini, P., & D'Odorico, L. (2017). The Pointing-Vocal Coupling Progression in the First Half of the Second Year of Life. *Infancy*, (I), 1–18. http://doi.org/10.1111/infa.12181

Baldwin, D. A., & Markman, E. M. (1989). Establishing Word-Object Relations: A First Step. *Child Development*, *60*(2), 381–398.

Balog, H. L., & Brentari, D. (2008). The relationship between early gestures and intonation. *First Language*, *28*(2), 141–163. http://doi.org/10.1177/0142723708088722

Bannard, C., & Tomasello, M. (2012). Can we dissociate contingency learning from social learning in word acquisition by 24-month-olds? *PloS ONE*, *7*(11), e49881. http://doi.org/10.1371/journal.pone.0049881

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). http://doi.org/10.18637/jss.v067.i01

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). Cognition and Communication from Nine to Thirteen Months: Correlational Findings. In E. Bates

(Ed.), *The Emergence of Symbols: Cognition and Communication in Infancy* (pp. 69–140). New York; San Francisco; London: Academic Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The Acquisition of Performatives Prior to Speech. *Merrill-Palmer Quarterly of Behavior and Development, 21*(3), 205–226.

Baumwell, L., Tamis-LeMonda, C. S., & Bornstein, M. H. (1997). Maternal verbal sensitivity and child language comprehension. *Infant Behavior and Development*, *20*(2), 247–258. http://doi.org/10.1016/S0163-6383(97)90026-6

Bavin, E. L., Prior, M., Reilly, S., Bretherton, L., Williams, J., Eadie, P., … Ukoumunne, O. C. (2008). The Early Language in Victoria Study: predicting vocabulary at age one and two years from gesture and object use. *Journal of Child Language, 35*(3), 687–701. http://doi.org/10.1017/S0305000908008726

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science, 8*(6), 492–499. http://doi.org/10.1111/j.1467-7687.2005.00440.x

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *The British Journal of Developmental Psychology*, *30,* 359–375. http://doi.org/10.1111/j.2044-835X.2011.02043.x

Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. http://doi.org/10.1073/pnas.1113380109

Blake, J., Osborne, P., Cabral, M., & Gluck, P. (2003). The Development of Communicative Gestures in Japanese Infants. *First Language, 23*(1), 3–20. http://doi.org/10.1177/0142723703023001001

Blake, J., Vitale, G., Osborne, P., & Olshansky, E. (2005). A cross-cultural comparison of communicative gestures in human infants during the transition to language. *Gesture, 5*(1/2), 201–217.

Bloom, K. (1975). Social Elicitation of Infant Vocal Behaviour. *Journal of Experimental Child Psychology*, *20,* 51–58.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *14*(2), 211–227. http://doi.org/10.1017/S0305000900012897

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, O. M. (1999). First Words in the Second Year : Continuity , Stability , and Models of Concurrent and Predictive Correspondence in Vocabulary and Verbal Responsiveness Across Age and Context. *Infant Behavior and Development*, *22*(1), 65–85.

Boundy, L., Cameron-Faulkner, T., & Theakston, A. (2016). Infant Behavior and Development Exploring early communicative behaviours: A fine-grained analysis of infant shows and gives. *Infant Behavior and Development*, *44*, 86–97. http://doi.org/10.1016/j.infbeh.2016.06.005

Bourjade, M., Meguerditchian, A., Maille, A., Gaunet, F., & Vauclair, J. (2013). Olive baboons, Papio anubis, adjust their visual and auditory intentional gestures to the visual attention of others. *Animal Behaviour*, 1–8. http://doi.org/10.1016/j.anbehav.2013.10.019

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early Person Knowledge as Expressed in Gestural and Verbal Communication: When Do Infants Acquire A "Theory of Mind"? In M. E. Lamb & L. R. Sherrod (Eds.), *Infant Social Cognition: Empirical and Theoretical Considerations* (pp. 333–374). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. *Journal of Child Language*, *35*(1), 207–220. http://doi.org/10.1017/S030500090700829X

Bruner, J. S. (1973). Organization of Early Skilled Action. *Child Development*, *44*(1), 1–11.

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, *2*(1), 1–19. http://doi.org/10.1017/S0305000900000866

Bruner, J. S. (1976). From communication to language - A psychological perspective. *Cognition*, *3*(3), 255–287.

Bruner, J. S., Roy, & Ratner. (1982). The Beginnings of Request. In K. E. Nelson (Ed.), *Children's Language, Volume 3* (pp. 91-96). Lawrence Erlbaum Associates.

Butterworth, G. (2003). Pointing Is the Royal Road to Language for Babies. In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* (pp. 9–34). Lawrence Erlbaum Associates.

Butterworth, G., & Grover, L. (1988). The origins of referential commmunication in human infancy. In L. Weiskrantz (Ed.), *Thought Without Language* (pp. 5–24). Oxford University Press.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. http://doi.org/10.1111/j.2044-835X.1991.tb00862.x

Butterworth, G., & Morissette, P. (1996). Onset of pointing and the acquisition of language in infancy. *Journal of Reproductive and Infant Psychology*, *14*(3), 219–231. http://doi.org/10.1080/02646839608404519

Camaioni, L., Castelli, M. C., Longobardi, E., & Volterra, V. (1991). A parent report instrument for early language assessment. *First Language*, *11*(33), 345–358. http://doi.org/10.1177/014272379101103303

Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnesi, C. (2004). The Role of Declarative Pointing in Developing a Theory of Mind. *Infancy*, *5*(3), 291–308.

Cameron-Faulkner, T., Theakston, A., Lieven, E., & Tomasello, M. (2015). The Relationship Between Infant Holdout and Gives, and Pointing. *Infancy*, *20*(5), 576-586. http://doi.org/10.1111/infa.12085

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The Role of Inhibitory Processes in Young Children's Difficulties with Deception and False Belief. *Child Development*, *69*(3), 672. http://doi.org/10.2307/1132197

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

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind : The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*, 79–151.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1-174.

Cartmill, E. A., & Byrne, R. W. (2010). Semantics of primate gestures: intentional meanings of orangutan gestures. *Animal Cognition*, *13*(6), 793–804. http://doi.org/10.1007/s10071-010-0328-7

Casby, M. W., & Cumpata, J. F. (1986). A Protocol for the Assessment of Prelinguistic Intentional Communication. *Journal of Communication Disorders*, *19*, 251–260.

Caselli, M. C., Rinaldi, P., Stefanini, S., & Volterra, V. (2012). Early Action and Gesture "Vocabulary" and Its Relation With Word Comprehension and Production. *Child Development*, *83*(2), 526–542. http://doi.org/10.1111/j.1467-8624.2011.01727.x

Cochet, H., & Byrne, R. W. (2016). Communication in the second and third year of life: Relationships between nonverbal social skills and language. *Infant Behavior and Development*, *44*, 189–198. http://doi.org/10.1016/j.infbeh.2016.07.003

Cochet, H., & Vauclair, J. (2010). Features of spontaneous pointing gestures in toddlers. *Gesture*, *10*(1), 86–107. http://doi.org/10.1075/gest.10.1.05coc

Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology*, *26*(4), 561–577. http://doi.org/10.1348/026151008X285660

Colonnesi, C., Stams, G. J. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352–366. http://doi.org/10.1016/j.dr.2010.10.001

Colonnesi, C., Zijlstra, B. J. H., van der Zande, A., & Bögels, S. M. (2012). Coordination of gaze, facial expressions and vocalizations of early infant communication with mother and father. *Infant Behavior and Development*, *35*(3), 523–532. http://doi.org/10.1016/j.infbeh.2012.02.004

Cooperrider, K., & Núñez, R. (2012). Nose-pointing: Notes on a facial gesture of Papua New Guinea. *Gesture*, *12*(2), 103–129. http://doi.org/10.1075/gest.12.2.01coo

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*(1), 28–38.

Crais, E., Douglas, D. D., & Campbell, C. C. (2004). The Intersection of the Development of Gestures and Intentionality. *Journal of Speech, Language, and Hearing Research*, *47*(June), 678–694.

Csibra, G. (2010). Recognizing Communicative Intentions in Infancy. *Mind & Language*, *25*(2), 141–168. http://doi.org/10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. http://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B, Biological Sciences*, *366*(1567), 1149–1157. http://doi.org/10.1098/rstb.2010.0319

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A Demonstration of Gaze Following in 3- to 6-Month-Olds. *Infant Behavior and Development*, *20*(4), 569–572.

D'Odorico, L., & Cassibba, R. (1995). Cross-Sectional Study of Coordination Between Infants' Gaze and Vocalizations Towards Their Mothers. *Early Development and Parenting*, *4*(1), 11–19.

D'Odorico, L., Cassibba, R., & Salerni, N. (1997). Temporal relationships between gaze and vocal behavior in prelinguistic and linguistic communication. *Journal of Psycholinguistic Research*, *26*(5), 539–556.

D'Odorico, L., & Franco, F. (1991). Selective production of vocalization types in different communication contexts. *Journal of Child Language*, *18*(3), 475–499.

D'Odorico, L., Salerni, N., Cassibba, R., & Jacob, V. (1999). Stability and change of maternal speech to Italian infants from 7 to 21 months of age: a longitudinal study of its influence on early stages of language acquisition. *First Language*, *19*(57), 313–346. http://doi.org/10.1177/014272379901905702 de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychological Research*, *73*(4), 463–476. http://doi.org/10.1007/s00426-009-0230-6

Delgado, B., Gómez, J.-C., & Sarriá, E. (2011). Pointing gestures as a cognitive tool in young children: experimental evidence. *Journal of Experimental Child Psychology*, *110*(3), 299–312. http://doi.org/10.1016/j.jecp.2011.04.010

Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences, 6*, 343–390. http://doi.org/10.1017/S0140525X00016393

Desrochers, S., Morissette, P., & Ricard, M. (1995). Two perspectives on pointing in infancy. In P. J. Dunham & C. Moore (Eds.), *Joint Attention: Its origins and role in development* (pp. 85–101). Hillsdale, NJ: Lawrence Erlbaum.

Elsner, C., Bakker, M., Rohlfing, K., & Gredeback, G. (2014). Infants' online perception of give-and-take interactions. *Journal of Experimental Child Psychology*, *126*, 280–294. http://doi.org/10.1016/j.jecp.2014.05.007

Enfield, N. J. (2001). "Lip-pointing": A discussion of form and function with reference to data from Laos. *Gesture*, *1*(2), 185–211. http://doi.org/10.1075/gest.1.2.06enf

Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: evidence from Catalan-babbling infants. *Journal of Child Language*, *40*(5), 919–944. http://doi.org/10.1017/S0305000912000359

Fagan, M. K. (2009). Mean Length of Utterance before words and grammar: Longitudinal trends and developmental implications of infant vocalizations. *Journal of Child Language*, *36*(3), 495. http://doi.org/10.1017/S0305000908009070

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. http://doi.org/10.1073/pnas.152159999

Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology*, *95*(4), 298–308. http://doi.org/10.1016/j.jecp.2006.08.001

Fasolo, M., & D'Odorico, L. (2012). Gesture-plus-word combinations, transitional forms, and language development. *Gesture*, *12*(1), 1–15. http://doi.org/10.1075/gest.12.1.01fas

Fedurek, P., Schel, A. M., & Slocombe, K. E. (2013). The acoustic structure of chimpanzee pant-hooting facilitates chorusing. *Behavioral Ecology and Sociobiology*. http://doi.org/10.1007/s00265-013-1585-7

Fedurek, P., & Slocombe, K. E. (2013). The social function of food-associated calls in male chimpanzees. *American Journal of Primatology*, *75*(7), 726–739. http://doi.org/10.1002/ajp.22122

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i. http://doi.org/10.2307/1166093

Fogel, A., & Hannan, T. E. (1985). Manual Actions of Nine- to Fifteen-Week-Old Human Infants during Face-to-Face Interaction with Their Mothers. *Child Development*, *56*, 1271–1279.

Franco, F., & Butterworth, G. (1996). Pointing and social awareness: declaring and requesting in the second year. *Journal of Child Language*, *23*(2), 307–336. http://doi.org/10.1017/S0305000900008813

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170. http://doi.org/10.1016/j.cognition.2008.11.010

Gaunet, F., & Massioui, F. El. (2014). Marked referential communicative behaviours, but no differentiation of the "knowledge state" of humans in untrained pet dogs versus 1-year-old infants. *Animal Cognition*, 1137–1147. http://doi.org/10.1007/s10071-014-0746-z

Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (Gorilla gorilla): repertoire, intentionality and possible origins. *Animal Cognition*, *12*(3), 527–546. http://doi.org/10.1007/s10071-009-0213-4

Goldin-Meadow, S., & Butcher, C. (2003). Pointing Toward Two-Word Speech in Young Children. In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* (pp. 85–108). Lawrence Erlbaum Associates.

Goldin-Meadow, S., Goodrich, W., Sauer, E., & Iverson, J. (2007). Young children use their hands to tell their mothers what to say. *Developmental Science*, *10*(6), 778–785. http://doi.org/10.1111/j.1467-7687.2007.00636.x

Goldstein, M. H., & Schwade, J. A. (2008). Social Feedback to Infants' Babbling Facilitates Rapid Phonological Learning. *Psychological Science*, *19*(5), 515–523. http://doi.org/10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development*, *80*(3), 636–644. http://doi.org/10.1111/j.1467-8624.2009.01287.x

Goldstein, M. H., Schwade, J., Briesch, J., & Syal, S. (2010). Learning While Babbling: Prelinguistic Object-Directed Vocalizations Indicate a Readiness to Learn. *Infancy*, *15*(4), 362–391. http://doi.org/10.1111/j.1532-7078.2009.00020.x

Golinkoff, R. M. (1986). "I beg your pardon?": the preverbal negotiation of failed messages. *Journal of Child Language*, *13*(3), 455–476. http://doi.org/10.1017/S0305000900006826

Golinkoff, R. M. (1993). When is communication a "meeting of minds"? *Journal of Child Language*, *20*, 199–207.

Gómez, J.-C. (1994). Mutual awareness in primate communication: A Gricean approach. In S. T. Parker, R. W. Mitchell, & M. L. Boccia (Eds.), *Self-Awareness in Animals and Humans* (pp. 61–80). Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511565526

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*, 544–549.

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). One-year-olds' understanding of nonverbal gestures directed to a third person. *Cognitive Development*, *24*(1), 23–33. http://doi.org/10.1016/j.cogdev.2008.10.001

Greenfield, P. M., & Dent, C. H. (1979). A Developmental Study of the Communication of Meaning: The Role of Uncertainty and Information. In P. French (Ed.), *The Development of Meaning* (pp. 563–598). Bunka Hyoron Press.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*(3), 377–388.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*(6), 509–516. http://doi.org/10.1177/0165025406071914

Gros-Louis, J., West, M. J., & King, A. P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. *Infancy*, *19*(4), 385–408. http://doi.org/10.1111/infa.12054

Gros-Louis, J., & Wu, Z. (2012). Twelve-month-olds' vocal production during pointing in naturalistic interactions: sensitivity to parents' attention and responses. *Infant Behavior and Development*, *35*(4), 773–778. http://doi.org/10.1016/j.infbeh.2012.07.016

Grosse, G., Behne, T., Carpenter, M., & Tomasello, M. (2010). Infants communicate in order to be understood. *Developmental Psychology*, *46*(6), 1710–22. http://doi.org/10.1037/a0020727

Grosse, G., Scott-Phillips, T. C., & Tomasello, M. (2013). Three-Year-Olds Hide Their Communicative Intentions in Appropriate Contexts. *Developmental Psychology*. http://doi.org/10.1037/a0032017

Grünloh, T., & Liszkowski, U. (2015). Prelinguistic vocalizations distinguish pointing acts. *Journal of Child Language*, *42*(6), 1312–1336. http://doi.org/10.1017/S0305000914000816

Harding, C. G. (1982). Development of the Intention to Communicate. *Human Development*, *25*, 140–151.

Harding, C. G. (1984). Acting With Intention: A Framework for Examining the Development of the Intention to Communicate. In L. Feagans, C. Garvey, & R. M. Golinkoff (Eds.), *The Origins and Growth of Communication* (pp. 123–135). Norwood, New Jersey: Ablex Publishing Corporation.

Harding, C. G., & Golinkoff, R. M. (1979). The origins of intentional vocalizations in prelinguistic infants. *Child Development*, *50*(1), 33–40.

Harris, M., Barlow-Brown, F., & Chasin, J. (1995). The emergence of referential understanding: Pointing and the comprehension of object names. *First Language*, *15*(43), 19–34.

Heyes, C., & Dickinson, A. (1990). The Intentionality of Animal Action. *Mind & Language*, *5*(1), 87–103. http://doi.org/10.1111/j.1468-0017.1990.tb00154.x

Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. http://doi.org/10.1007/s10071-011-0409-2

Hostetter, A. B., Cantero, M., & Hopkins, W. D. (2001). Differential Use of Vocal and Gestural Communication by Chimpanzees in Response to the Attentional Status of a Human. *Journal of Comparative Psychology*, *115*(4), 337–343.

Hsu, H.-C., Fogel, A., & Messinger, D. S. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences. *Infant Behavior and Development*, *24*, 107–128.

Huttunen, K. H., Pine, K. J., Thurnham, A. J., & Khan, C. (2013). The Changing Role of Gesture in Linguistic Development: A Developmental Trajectory and a Cross-Cultural Comparison Between British and Finnish Children. *Journal of Psycholinguistic Research*, *42*(1), 81–101. http://doi.org/10.1007/s10936-012-9205-7

Igualada, A., Bosch, L., & Prieto, P. (2015). Language development at 18 months is related to multimodal communicative strategies at 12 months. *Infant Behavior and Development*, *39*, 42–52. http://doi.org/10.1016/j.infbeh.2015.02.004

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1651), 20130298–20130298. http://doi.org/10.1098/rstb.2013.0298

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound Symbolism Facilitates Word Learning in 14-Month-Olds. *PLOS ONE*, *10*(2), e0116494. http://doi.org/10.1371/journal.pone.0116494

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19.

Johnson, S. C., Ok, S.-J., & Luo, Y. (2007). The attribution of attention: 9-month-olds' interpretation of gaze as goal-directed action. *Developmental Science*, *10*(5), 530–537. http://doi.org/10.1111/j.1467-7687.2007.00606.x

Jones, S. S., & Hong, H.-W. (2001). Onset of Voluntary Communication: Smiling Looks to Mother. *Infancy*, *2*(3), 353–370. http://doi.org/10.1207/S15327078IN0203_4

Karousou, A., & López-Ornat, S. (2013). Prespeech vocalizations and the emergence of speech: a study of 1005 Spanish children. *The Spanish Journal of Psychology*, *16*(e32), 1–21. http://doi.org/10.1017/sjp.2013.27

Kaye, K., & Fogel, A. (1980). The Temporal Structure of Face-to-Face Communication Between Mothers and Infants. *Developmental Psychology*, *16*(5), 454–464.

Kersken, V., Zuberbühler, K., & Gomez, J. (2017). Listeners can extract meaning from non-linguistic infant vocalisations cross-culturally. *Scientific Reports*, *7*(January), 41016. http://doi.org/10.1038/srep41016

Klossek, U. M. H., Russell, J., & Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Child Psychology: General*, *137*(1), 39–51. http://doi.org/10.1037/0096-3445.137.1.39

Kraljević, J. K., Cepanec, M., & Šimleša, S. (2014). Gestural development and its relation to a child's early vocabulary. *Infant Behavior and Development*, *37*(2), 192–202. http://doi.org/10.1016/j.infbeh.2014.01.004

Krause, M. A., & Fouts, R. S. (1997). Chimpanzee (Pan troglodytes) pointing: hand shapes, accuracy, and the role of eye gaze. *Journal of Comparative Psychology*, *111*(4), 330–336.

Krehm, M., Onishi, K. H., & Vouloumanos, A. (2014). I See Your Point: Infants Under 12 Months Understand That Pointing Is Communicative. *Journal of Cognition and Development*, *15*(4), 527–538. http://doi.org/10.1080/15248372.2012.736112

Kutsuki, A., Ogura, T., Egami, S., Itakura, S., & Children's Study Group, J. (2009). Development of Infants' Request Expressions from 11 to 15 Months. *Psychological Reports*, *105*(3), 865–878. http://doi.org/10.2466/PR0.105.3.865-878

Laakso, M. L., Poikkeus, A. M., Katajamaki, J., & Lyytinen, P. (1999). Early intentional communication as a predictor of language development in young toddlers. *First Language*, *19*(56), 207–231. http://doi.org/10.1177/014272379901905604

Leavens, D. A., & Hopkins, W. D. (1998). Intentional Communication by Chimpanzees: A Cross-Sectional Study of the Use of Referential Gestures. *Developmental Psychology*, *34*(5), 813–822.

Leavens, D. A., & Hopkins, W. D. (1999). The Whole-Hand Point: The Structure and Function of Pointing From a Comparative Perspective. *Journal of Comparative Psychology*, *113*(4), 417–425.

Leavens, D. A., Hopkins, W. D., & Thomas, R. K. (2004). Referential communication by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *118*(1), 48–57. http://doi.org/10.1037/0735-7036.118.1.48

Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (Pan troglodytes). *Child Development*, *76*(1), 291–306. http://doi.org/10.1111/j.1467-8624.2005.00845.x

Legerstee, M., & Barillas, Y. (2003). Sharing attention and pointing to objects at 12 months: Is the intentional stance implied? *Cognitive Development*, *18*(1), 91–110. http://doi.org/10.1016/S0885-2014(02)00165-X

Lempers, J. D. (1979). Young children's production and comprehension of nonverbal deictic behaviors. *The Journal of Genetic Psychology*, *135*, 93–102. http://doi.org/10.1080/00221325.1979.10533420

Lenth, R. V. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *69*(1). http://doi.org/10.18637/jss.v069.i01

Leslie, A. M. (1987). Pretense and Representation - the Origins of Theory of Mind. *Psychological Review*, *94*(4), 412–426. http://doi.org/10.1037/0033-295x.94.4.412

Leung, E. H. L., & Rheingold, H. L. (1981). Development of Pointing as a Social Gesture. *Developmental Psychology*, *17*(2), 215–220.

Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science*, *12*(2), 264–271. http://doi.org/10.1111/j.1467-7687.2008.00758.x

Liebal, K., Pika, S., & Tomasello, M. (2006). Gestural communication of orangutans (Pongo pygmaeus). *Gesture*, *6*(1), 1–38.

Lieven, E., & Stoll, S. (2013). Early Communicative Development in Two Cultures: A Comparison of the Communicative Environments of Children from Two Cultures. *Human Development*, *56*(3), 178–206. http://doi.org/10.1159/000351073

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & de Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cognitive Science*, *36*(4), 698–713. http://doi.org/10.1111/j.1551-6709.2011.01228.x

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*(3), 297–307.

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. http://doi.org/10.1207/s15327647jcd0702_2

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

Lucca, K., & Wilbourn, M. P. (2016). Communicating to Learn: Infants' Pointing Gestures Result in Optimal Learning. *Child Development*, *0*(0), 1–20. http://doi.org/10.1111/cdev.12707

Majorano, M., & D'Odorico, L. (2011). The transition into ambient language: A longitudinal study of babbling and first word production of Italian children. *First Language*, *31*(1), 47–66. http://doi.org/10.1177/0142723709359239

Maljaars, J., Noens, I., Jansen, R., Scholte, E., & van Berckelaer-Onnes, I. (2011). Intentional communication in nonverbal and verbal low-functioning children with autism. *Journal* *of Communication Disorders*, *44*(6), 601–614. http://doi.org/10.1016/j.jcomdis.2011.07.004

Maneva, B., & Genesee, F. (2002). Bilingual babbling: Evidence for language differentiation in dual language acquisition. In S. Fish & A. H.-J. Do (Eds.), *Proceedings of the 26th Boston University Conference on Language Development* (pp. 383–392). Somerville, MA: Cascadilla Press.

Marchand, N., Ricard, M., & Gouin Décarie, T. (1994). Vocalisations et communication chez le nourrisson de 6 á 15 mois. *International Journal of Psychology, 29*(1), 71–84.

Marcos, H. (1991). Reformulating requests at 18 months: Gestures, vocalizations and words. *First Language*, *11*(33, Pt 3), 361–375. http://doi.org/http://dx.doi.org/10.1177/014272379101103304

Marschik, P. B., Prechtl, H. F. R., Prayer, D., Peyton, C., & Einspieler, C. (2013). An antecedent of later developing communicative functions: the fetal index finger. *BMJ*, *347*, f7232. http://doi.org/10.1136/bmj.f7232

Masur, E. F. (1982). Mothers’ responses to infants’ object-related gestures: influences on lexical development. *Journal of Child Language*, *9*(1), 23–30. http://doi.org/10.1017/S0305000900003585

Masur, E. F. (1983). Gestural development, dual-directional signaling, and the transition to words. *Journal of Psycholinguistic Research*, *12*(2), 93–109. http://doi.org/10.1007/BF01067406

Masur, E. F., Flynn, V., & Eichorst, D. L. (2005). Maternal responsive and directive behaviours and utterances as predictors of children’s lexical development. *Journal of Child Language*, *32*(1), 63–91. http://doi.org/10.1017/S0305000904006634

Matthews, D., Behne, T., Lieven, E., & Tomasello, M. (2012). Origins of the human pointing gesture: a training study. *Developmental Science*, *15*(6), 817–829. http://doi.org/10.1111/j.1467-7687.2012.01181.x

McCune, L., Roug-Hellichius, L., Vihman, M. M., Delery, D. B., & Gogate, L. (1996). Grunt Communication in Human Infants (Homo sapiens), *110*(1).

McCune, L., & Vihman, M. M. (2001). Early Phonetic and Lexical Development: A Productivity Approach. *Journal of Speech, Language and Hearing Research*, *44*(3), 670–684. http://doi.org/10.1044/1092-4388(2001/054)

McGillion, M. L., Herbert, J. S., Pine, J. M., Keren-portnoy, T., Vihman, M. M., & Matthews, D. E. (2013). Supporting early vocabulary development: What sort of responsiveness matters?

McGillion, M. L., Herbert, J. S., Pine, J., Vihman, M. M., DePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017). What Paves the Way to Conventional Language? The Predictive Value of Babble, Pointing, and Socioeconomic Status. *Child Development*, *88*(1), 156–166. http://doi.org/10.1111/cdev.12671

McGillion, M. L., Pine, J. M., Herbert, J., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse SES backgrounds. *Journal of Child Psychology and Psychiatry*, (June).

Meints, K., & Woodward, A. Electronic vocabulary database Lincoln Lincoln UK-CDI Infants and/or the Lincoln UK-CDI Toddlers. (2011). WWW document: Computer software & manual].

Menyuk, P., Liebergott, J., & Shultz, M. (1986). Predicting Phonological Development. In B. Lindblom & R. Zetterstrom (Eds.), *Precursors of Early Speech* (pp. 79–94). New York, NY: Stockton Press.

Messinger, D. S., & Fogel, A. (1998). Give and take: the development of conventional infant gestures. *Merrill-Palmer Quarterly*, *44*(4), 566–590.

Meunier, H., Prieur, J., & Vauclair, J. (2013). Olive baboons communicate intentionally by pointing. *Animal Cognition*, *16*(2), 155–63. http://doi.org/10.1007/s10071-012-0558-y

Miller, J. L., & Lossia, A. K. (2013). Prelinguistic infants' communicative system: Role of caregiver social feedback. *First Language*, *33*(5), 524–544. http://doi.org/10.1177/0142723713503147

Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and Call Convergence in Chimpanzees: Tests of Three Hypotheses. *Behaviour*, *135*(8), 1041–1064.

Moore, C., Angelopoulos, M., & Bennett, P. (1997). The role of movement in the development of joint visual attention. *Infant Behavior and Development*, *20*(1), 83–92. http://doi.org/10.1016/S0163-6383(97)90063-1

Moore, C., & Corkum, V. (1994). Social Understanding at the End of the First Year of Life. *Developmental Review*, *14*, 349–372.

Moore, C., & D'Entremont, B. (2001). Developmental Changes in Pointing as a Function of Attentional Focus. *Journal of Cognition and Development*, *2*(2), 109–129. http://doi.org/10.1207/S15327647JCD0202

Moore, R. (2014). Enacting and Understanding Communicative Intent.

Moore, R. (2016). Gricean Communication and Cognitive Development. *The Philosophical Quarterly*, *67*(267), 303–326. http://doi.org/10.13140/RG.2.1.1875.6000

Moore, R., Liebal, K., & Tomasello, M. (2013). Three-year-olds understand communicative intentions without language, gestures, or gaze. *Interaction Studies*, *14*(1), 62–80. http://doi.org/10.1075/is.14.1.05moo

Moore, R., Mueller, B., Kaminski, J., & Tomasello, M. (2015). Two-year-old children but not domestic dogs understand communicative intentions without language, gestures, or gaze. *Developmental Science*, *18*(2), 232–242. http://doi.org/10.1111/desc.12206

Morissette, P., Ricard, M., & Décarie, T. G. (1995). Joint visual attention and pointing in infancy: A longitudinal study of comprehension. *British Journal of Developmental Psychology*, *13*(2), 163–175. http://doi.org/10.1111/j.2044-835X.1995.tb00671.x

Mumford, K. H., & Kita, S. (2016). At 10-12 Months, Pointing Gesture Handedness Predicts the Size of Receptive Vocabularies. *Infancy*, *21*(6), 751–765. http://doi.org/10.1111/infa.12138

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual Differences and the Development of Joint Attention in Infancy. *Child Development*, *78*(3), 938–954. http://doi.org/10.1111/j.1467-8624.2007.01042.x

Murillo, E., & Belinchón, M. (2012). Gestural-vocal coordination: Longitudinal changes and predictive value on early lexical development. *Gesture*, *12*(2012), 16–39. http://doi.org/10.1075/gest.12.1.02mur

Murillo, E., & Capilla, A. (2016). Properties of vocalization- and gesture-combinations in the transition to first words. *Journal of Child Language*, *43*(4), 890–913. http://doi.org/10.1017/S0305000915000343

Murphy, C. M. (1978). Pointing in the Context of a Shared Activity. *Child Development*, *49*(2), 371–380.

Murphy, C. M., & Messer, D. J. (1977). Mothers, Infants and Pointing: A Study of a Gesture. In H. R. Schaffer (Ed.), *Studies in Mother-Infant Interaction* (pp. 325–354). London: Academic Press.

Nicoladis, E., Pika, S., & Marentette, P. (2009). Do French-English bilingual children gesture more than monolingual children? *Journal of Psycholinguistic Research*, *38*(6), 573–585. http://doi.org/10.1007/s10936-009-9121-7

O'Connell, J. C., & Farran, D. C. (1982). Effects of day-care experience on the use of intentional communicative behaviors in a sample of socioeconomically depressed infants. *Developmental Psychology*, *18*(1), 22–29. http://doi.org/10.1037//0012-1649.18.1.22

Oller, D. K. (2000). *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6318–6323. http://doi.org/10.1073/pnas.1300337110

Oller, D. K., Eilers, R. E., Urbano, R., & Cobo-Lewis, A. B. (1997). Development of precursors to speech in infants exposed to two languages. *Journal of Child Language*, *24*(2), 407–425. http://doi.org/10.1017/S0305000997003097

Olson, J., & Masur, E. F. (2011). Infants' gestures influence mothers' provision of object, action and internal state labels. *Journal of Child Language*, *38*(5), 1028–1054. http://doi.org/10.1017/S0305000910000565

Olson, J., & Masur, E. F. (2013). Mothers respond differently to infants' gestural versus nongestural communicative bids. *First Language*, *33*(4), 372–387. http://doi.org/10.1177/0142723713493346

Olson, J., & Masur, E. F. (2015). Mothers' labeling responses to infants' gestures predict vocabulary outcomes. *Journal of Child Language*, *42*(6), 1289–1311. http://doi.org/10.1017/S0305000914000828

Paavola, L., Kemppinen, K., Kumpulainen, K., Moilanen, I., & Ebeling, H. (2006). Maternal sensitivity, infant co-operation and early linguistic development: Some predictive relations. *European Journal of Developmental Psychology, 3*(1), 13–30. http://doi.org/10.1080/17405620500317789

Paavola, L., Kunnari, S., & Moilanen, I. (2005). Maternal responsiveness and infant intentional communication: implications for the early communicative and linguistic development. *Child: Care, Health and Development*, *31*(6), 727–735. http://doi.org/10.1111/j.1365-2214.2005.00566.x

Papaeliou, C. F., Minadakis, G., & Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech, Language, and Hearing Research, 45*(2), 311–317.

Papaeliou, C. F., & Trevarthen, C. (2006). Prelinguistic pitch patterns expressing "communication" and "apprehension." *Journal of Child Language*, *33,* 163–178. http://doi.org/10.1017/S0305000905007300

Pfandler, E., Lakatos, G., & Miklósi, Á. (2013). Eighteen-month-old human infants show intensive development in comprehension of different types of pointing gestures. *Animal Cognition*, *16*(5), 711–719. http://doi.org/10.1007/s10071-013-0606-2

Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (Gorilla gorilla): gestural repertoire, learning, and use. *American Journal of Primatology*, *60*(3), 95–111. http://doi.org/10.1002/ajp.10097

Pika, S., Liebal, K., & Tomasello, M. (2005). Gestural communication in subadult bonobos (Pan paniscus): repertoire and use. *American Journal of Primatology, 65*(1), 39–61. http://doi.org/10.1002/ajp.20096

Pinheiro, J., Bates, D., Debroy, S., Sarkar, D., & Team, R. C. (2015). {nlme}: Linear and Nonlinear Mixed Effects Models. Retrieved from http://cran.r-project.org/package=nlme

Poss, S. R., Kuhar, C., Stoinski, T. S., & Hopkins, W. D. (2006). Differential use of attentional and visual communicative signaling by orangutans (Pongo pygmaeus) and gorillas (Gorilla gorilla) in response to the attentional status of a human. *American Journal of Primatology*, *68*(10), 978–992. http://doi.org/10.1002/ajp.20304

Povinelli, D. J., & O'Neill, D. K. (2000). Do chimpanzees use their gestures to instruct each other? In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (2nd ed., pp. 459–487). Oxford University Press.

Povinelli, D. J., Reaux, J. E., Bierschwale, D. T., Allain, A. D., & Simon, B. B. (1997). Exploitation of Pointing As A Referential Gesture in Young Children, But Not Adolescent Chimpanzees. *Cognitive Development*, *12*, 423–461.

Prizant, B. M., & Wetherby, A. M. (1987). Communicative Intent: A Framework for Understanding Social-Communicative Behavior in Autism. *Journal of American Academy of Child and Adolescent Psychiatry*, *26*(4), 472–479.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognitive Systems Research*, *7*(2–3), 113–127. http://doi.org/10.1016/j.cogsys.2005.11.008

Rakoczy, H. (2008). Taking fiction seriously: young children understand the normative structure of joint pretence games. *Developmental Psychology*, *44*(4), 1195–1201. http://doi.org/10.1037/0012-1649.44.4.1195

Ramsdell, H. L., Oller, D. K., Buder, E. H., Ethington, C. A., & Chorna, L. (2012). Identification of Prelinguistic Phonological Categories. *Journal of Speech, Language, and Hearing Research*, *55*, 1626–1639. http://doi.org/10.1044/1092-4388(2012/11-0250)b

Robert Thompson, J. (2014). Meaning and Mindreading. *Mind & Language*, *29*(2), 167–200. http://doi.org/10.1111/mila.12046

Roberts, A. I., Vick, S.-J., & Buchanan-Smith, H. M. (2013). Communicative intentions in wild chimpanzees: persistence and elaboration in gestural signalling. *Animal Cognition*, *16*(2), 187–196. http://doi.org/10.1007/s10071-012-0563-1

Roberts, M. Y., & Kaiser, A. P. (2011). The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis. *American Journal of Speech-Language Pathology*, *20*(3), 180. http://doi.org/10.1044/1058-0360(2011/10-0055)

Rodríguez, C., Moreno-Núñez, A., Basilio, M., & Sosa, N. (2015). Ostensive gestures come first: their role in the beginning of shared reference. *Cognitive Development*, *36*, 142–149. http://doi.org/10.1016/j.cogdev.2015.09.005

Rollins, P. R. (2003). Caregivers' contingent comments to 9-month-old infants: Relationships with later language. *Applied Psycholinguistics*, *24*(May 2017), 221–234. http://doi.org/10.1017.S0142716403000110

Rome-Flanders, T., & Cronk, C. (1995). A longitudinal study of infant vocalizations during mother-infant games. *Journal of Child Language*, *22*(2), 259–274. http://doi.org/10.1017/S0305000900009788

Rowe, M. L., & Goldin-Meadow, S. (2009a). Differences in Early Gesture Explain SES Disparities in Child Vocabulary Size at School Entry. *Science*, *323*(5916), 951–953. http://doi.org/10.1126/science.1167025

Rowe, M. L., & Goldin-Meadow, S. (2009b). Early gesture selectively predicts later language learning. *Developmental Science*, *12*(1), 182–187. http://doi.org/10.1111/j.1467-7687.2008.00764.x

Rowe, M. L., Özçalışkan, Ş., & Goldin-Meadow, S. (2008). Learning words by hand: Gesture's role in predicting vocabulary development. *First Language*, *28*(2), 182–199. http://doi.org/10.1177/0142723707088310

Russell, C., Bard, K. A., & Adamson, L. B. (1997). Social Referencing by Young Chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *111*(2), 185–193.

Russell, J. L., Braccini, S., Buehler, N., Kachin, M. J., Schapiro, S. J., & Hopkins, W. D. (2005). Chimpanzee (Pan troglodytes) intentional communication is not contingent upon food. *Animal Cognition*, *8*(4), 263–272. http://doi.org/10.1007/s10071-005-0253-3

Salomo, D., & Liszkowski, U. (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child Development*, *84*(4), 1296–1307. http://doi.org/10.1111/cdev.12026

Sansavini, A., Bello, A., Guarini, A., Savini, S., Stefanini, S., & Caselli, M. C. (2010). Early development of gestures, object-related-actions, word comprehension and word production, and their relationships in Italian infants: A longitudinal study. *Gesture*, *10*(1), 52–85. http://doi.org/10.1075/gest.10.1.04san

Saylor, M. M., & Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, *43*(3), 696–704. http://doi.org/10.1037/0012-1649.43.3.696

Schel, A. M., Machanda, Z., Townsend, S. W., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee food calls are directed at specific individuals. *Animal Behaviour*, *86*(5), 955–965. http://doi.org/10.1016/j.anbehav.2013.08.013

Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PloS One*, *8*(10), e76674. http://doi.org/10.1371/journal.pone.0076674

Schulze, C., & Tomasello, M. (2015). 18-month-olds comprehend indirect communicative acts. *Cognition*, *136*, 91–98. http://doi.org/10.1016/j.cognition.2014.11.036

Scoville, R. (1984). Development of the Intention to Communicate: The Eye of the Beholder. In L. Feagans, C. Garvey, & R. M. Golinkoff (Eds.), *The Origins and Growth of Communication* (pp. 109–122). Norwood, New Jersey: Ablex Publishing Corporation.

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: infants' preference for object-directed gaze. *Cognition*, *108*(2), 303–319. http://doi.org/10.1016/j.cognition.2008.02.009

Shatz, M. (1981). Communication. In P. H. Mussen (Ed.), *Handbook of Child Psychology, Volume III: Cognitive Development* (pp. 841–889). John Wiley & Sons.

Sherzer, J. (1973). Verbal and nonverbal deixis: the pointed lip gesture among the San Blas Cuna. *Language in Society*, *2*, 117–131.

Slocombe, K. E., Kaller, T., Turman, L., Townsend, S. W., Papworth, S., Squibbs, P., & Zuberbühler, K. (2010). Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behavioral Ecology and Sociobiology*, *64*(12), 1959–1966. http://doi.org/10.1007/s00265-010-1006-0

Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, *104*(43), 17228–17233.

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

Snyder, L. S. (1978). Communicative and Cognitive Abilities and Disabilities in the Sensorimotor Period. *Merrill-Palmer Quarterly of Behavior and Development*, *24*(3), 161–180.

Sodian, B., & Thoermer, C. (2004). Infants’ Understanding of Looking, Pointing, and Reaching as Cues to Goal-Directed Action. *Journal of Cognition and Development*, *5*(3), 289–316. http://doi.org/10.1207/s15327647jcd0503_1

Spencer, P. E. (1993). Communication Behaviors of Infants With Hearing Loss and Their Hearing Mothers. *Journal of Speech Language and Hearing Research*, *36*(2), 311–321. http://doi.org/10.1044/jshr.3602.311

Sperber, D. (2000). Metarepresentations in an evolutionary perspective. In *Metarepresentations: A Multidisciplinary Perspective* (pp. 117–137). Oxford University Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.

Stark, R. (1980). Stages of speech development in the first year of life. In *Child Phonology, Volume 1: Production* (Vol. 1, pp. 73–92). New York, NY: Academic Press.

Stoel-Gammon, C. (1992). Prelinguistic Vocal Development: Measurement and Predictions. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, Implications* (pp. 439–456). Timonium, Maryland: York Press.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*(1), 1–34. http://doi.org/10.1017/S0305000910000425

Sugarman, S. (1984). The Development of Preverbal Communication: Its Contribution and Limits in Promoting the Development of Language. In R. L. Schiefelbusch & J. Pickar (Eds.), *The Acquisition of Communicative Competence* (pp. 23–67). Baltimore: University Park Press.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal Responsiveness and Children's Achievement of Language Milestones. *Child Development*, *72*(3), 748–767. http://doi.org/10.1111/1467-8624.00313

Thorgrimsson, G. B., Fawcett, C., & Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Frontiers in Psychology*, *5*(April), 321. http://doi.org/10.3389/fpsyg.2014.00321

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5), 1–38. Retrieved from http://www.jstatsoft.org/v59/i05/

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

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, *10*(2), 201–224. http://doi.org/10.1016/0885-2014(95)90009-8

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of Novel Communicative Signs by Apes and Human Children. *Child Development*, *68*(6), 1067–1080.

Tomasello, M., Call, J., Warren, J., Frost, G. T., Carpenter, M., & Nagell, K. (1997). The Ontogeny of Chimpanzee Gestural Signals: A Comparison Across Groups and Generations. *Evolution of Communication*, *1*(2), 223–259. http://doi.org/10.1075/eoc.1.2.04tom

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

Tomasello, M., & Farrar, M. J. (1986). Joint Attention and Early Language. *Child Development*, *57*(6), 1454–1463.

Tomasello, M., Kruger, A. C., George, B. L., Farrar, M. J., & Evans, A. (1985). The Development of Gestural Communication in Young Chimpanzees. *Journal of Human Evolution*, *14*, 175–186.

Townsend, S. W., Deschner, T., & Zuberbühler, K. (2008). Female Chimpanzees Use Copulation Calls Flexibly to Prevent Social Competition. *PloS ONE*, *3*(6), e2431. http://doi.org/10.1371/journal.pone.0002431

Trevarthen, C., & Hubley, P. (1978). Secondary Intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, Gesture and Symbol: The Emergence of Language* (pp. 183–229). Academic Press.

Triesch, J., Teuscher, C., Deák, G. O., & Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, *9*(2), 125–147. http://doi.org/10.1111/j.1467-7687.2006.00470.x

Vauclair, J. (2004). Lateralization of communicative signals in nonhuman primates and the hypothesis of the gestural origin of language. *Interaction Studies*, *5*(3), 365–386. http://doi.org/10.1075/is.5.3.04vau

Vihman, M. M. (1996). *Phonological Development: The Origins of Language in the Child*. Cambridge, MA: Blackwell.

Vihman, M. M. (2014). *Phonological Development: The First Two Years* (2nd ed.). Chichester: Wiley Blackwell.

Vihman, M. M., & Ferguson, C. A. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, *7*(1), 3–40. http://doi.org/10.1017/S0142716400007165

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From Babbling to Speech: A Re-Assessment of the Continuity Issue. *Language*, *61*(2), 397. http://doi.org/10.2307/414151

Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, *17*(6), 872–879. http://doi.org/10.1111/desc.12170

Vouloumanos, A., Onishi, K. H., & Pogue, A. (2012). Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences*, *109*(32), 12933–12937. http://doi.org/10.1073/pnas.1121057109

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic Predictors of Language Outcome at 3 Years of Age. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1224–1237. http://doi.org/10.1044/1092-4388(2006/088)

Wermke, K., Mende, W., Manfredi, C., & Bruscaglioni, P. (2002). Developmental aspects of infant's cry melody and formants. *Medical Engineering & Physics*, *24*(7–8), 501–514. http://doi.org/10.1016/S1350-4533(02)00061-9

Wetherby, A. M., Cain, D. H., Yonclas, D. G., & Walker, V. G. (1988). Analysis of intentional communication of normal children from the prelinguistic to the multiword stage. *Journal of Speech and Hearing Research*, *31*, 240–252.

Wilke, C., Kavanagh, E., Donnellan, E., Waller, B. M., Machanda, Z. P., & Slocombe, K. E. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, Pan troglodytes schweinfurthii. *Animal Behaviour*, *123*, 305–316. http://doi.org/10.1016/j.anbehav.2016.10.024

Wilkins, D. (2003). Why Pointing With the Index Finger Is Not a Universal (in Sociocultural and Semiotic Terms). In S. Kita (Ed.), *Pointing: Where Language, Culture and Cognition Meet* (pp. 171–215). Lawrence Erlbaum Associates.

Wu, Z., & Gros-Louis, J. (2014). Infants' prelinguistic communicative acts and maternal responses: Relations to linguistic development. *First Language*, *34*(1), 72–90. http://doi.org/10.1177/0142723714521925

Wu, Z., & Gros-Louis, J. (2015). Caregivers provide more labeling responses to infants' pointing than to infants' object-directed vocalizations. *Journal of Child Language*, *42*(3), 538–561. http://doi.org/10.1017/S0305000914000221

Wu, Z., & Gros-Louis, J. (2017). The Value of Vocalizing: 10-Month-Olds' Vocal Usage Relates to Language Outcomes at 15 Months. *Infancy*, *22*(1), 23–41. http://doi.org/10.1111/infa.12150

Zinober, B., & Martlew, M. (1985). Developmental changes in four types of gesture in relation to acts and vocalizations from 10 to 21 months. *British Journal of Developmental Psychology*, *3*, 293–306.

**Appendix A**

Questionnaire for Caregivers: Pretending Not To Hear

We are interested in child development in the 18- to 36-month age range. Specifically, we are investigating what toddlers do when you ask them to do something that they don't like doing. For example, what do they do when you say it's time for a nappy change while they are playing with their toys.

Toddlers might react in different ways when you ask them to do something that they don't like to do. This questionnaire will ask about some of these reactions. Some things may sound familiar, but others may not. We are looking at a wide range of ages, and looking at a large number of behaviours that toddlers may or may not do.

**How to fill out this questionnaire**

*Please think about occasions over the past month when you have asked your child to do something that they don't like doing whilst they were playing*. As you read each item, please indicate how often your child did this in the last month by circling a number from 1 – 5.

| Number | 1 | 2 | 3 | 4 | 5 | NA |
|---|---|---|---|---|---|---|
| How often you have noticed the behaviour described | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |

If you have not noticed anything about the behaviour described, please circle NA.

Some of the questions may not apply. For example if you answer that you have never noticed some behaviour (circling a "1"), this may mean that some other questions that ask about this behaviour are not applicable. In these cases, please follow the instructions in the grey boxes.

*For these questions, please think about occasions where your child is playing when you ask them to do something that they don't like doing.*

**Question 1**

Think about how your child responds to you. By responding to you, we mean, do they give you any indication that they have heard what you have said? When your child is playing and you ask them to do something that they don't like doing, how often do they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| respond to you? | 1 | 2 | 3 | 4 | 5 | NA |

If you have circled "1" for this question, please go to question 3.

If you have circled "2", "3", or "4", for this question, please go to question 2.

If you have circled "5" for this question, please only answer question 2, and then go straight to question 9.

If you have circled "NA" for this question, please go to question 9.

**Question 2**

Think about when your child *responds* to you on these occasions. How often do they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| make eye contact with you? | 1 | 2 | 3 | 4 | 5 | NA |
| say "no" or shake their head? | 1 | 2 | 3 | 4 | 5 | NA |
| say something else? | 1 | 2 | 3 | 4 | 5 | NA |
| get angry? | 1 | 2 | 3 | 4 | 5 | NA |
| start laughing or smiling? | 1 | 2 | 3 | 4 | 5 | NA |

**Question 3**

Now think about when your child *does not respond* to you on these occasions. How often do they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| make eye contact with you? | 1 | 2 | 3 | 4 | 5 | NA |
| say “no” or shake their head? | 1 | 2 | 3 | 4 | 5 | NA |
| say something else? | 1 | 2 | 3 | 4 | 5 | NA |
| get angry? | 1 | 2 | 3 | 4 | 5 | NA |
| start laughing or smiling? | 1 | 2 | 3 | 4 | 5 | NA |

**Question 4**

Again, thinking about when your child *does not respond* to you, how often do you feel that they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| avoid making eye contact with you? | 1 | 2 | 3 | 4 | 5 | NA |
| avoid saying anything? | 1 | 2 | 3 | 4 | 5 | NA |

**Question 5**

Again, thinking about when your child *does not respond* to you, how often do they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| carry on playing? | 1 | 2 | 3 | 4 | 5 | NA |

If you have circled “1” or “NA” for this question, please go to question 9.

If you have circled “2”, “3”, “4” or “5” for this question, please go to question 6.

**Question 6**

Think about when your child *does not respond, and carries on playing* when you ask them do something that they don't like doing. How often do they…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| carry on playing in the same way? | 1 | 2 | 3 | 4 | 5 | NA |
| start playing more intently? | 1 | 2 | 3 | 4 | 5 | NA |

**Question 7**

Again, think about when your child *does not respond, and carries on playing* when you ask them do something that they don't like doing. How often do you feel that this is because they are…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| too engaged in play to hear you? | 1 | 2 | 3 | 4 | 5 | NA |
| deliberately not responding to you? | 1 | 2 | 3 | 4 | 5 | NA |

If you have circled "1" or "NA" for the item 'deliberately not responding to you', please go to question 9.

If you have circled, "2", "3", "4" or "5" for the item 'deliberately not responding to you', please go to question 8.

**Question 8**

Think about when your child *deliberately does not respond, and carries on playing* when you ask them do something that they don't like doing. How often do you feel that they are…

| | Never | Less Than Half the Time | About Half the Time | More Than Half the Time | Always | Does not apply |
|---|---|---|---|---|---|---|
| ignoring you and carrying on with what they are doing? | 1 | 2 | 3 | 4 | 5 | NA |
| actively pretending that they haven't heard you (by pretending to be very busy)? | 1 | 2 | 3 | 4 | 5 | NA |

**Question 9**

Does your child have any older siblings?

☐ Yes

☐ No

**Question 10**

Are there any other languages other than English regularly spoken at home?

☐ Yes - Please state ______________________

☐ No

Thanks so much for your help! If you have any further comments on the study, or any other information that you think is important, please write it here (you can continue on the back)!

## Appendix B

Descriptives for Caregiver Report in Chapter 2, Experiment 2

Table 38

*Responses to Questionnaire in Appendix A: Number of Responses, Median and Interquartile Range*

| | *18- to 20-months* | | | *24- to 26-months* | | | *30- to 32-months* | | |
|---|---|---|---|---|---|---|---|---|---|
| *Question* | *n* | *Mdn* | *IQR* | *n* | *Mdn* | *IQR* | *n* | *Mdn* | *IQR* |
| 1 | 29 | 4 | 1.00 | 34 | 4 | 1.00 | 24 | 4 | 1.00 |
| 2.1 | 29 | 4 | 2.00 | 34 | 3 | 2.00 | 24 | 3 | 1.00 |
| 2.2 | 29 | 4 | 1.00 | 34 | 4 | 1.00 | 24 | 3 | 1.00 |
| 2.3 | 29 | 2 | 1.00 | 31 | 2 | 1.00 | 23 | 2 | 1.00 |
| 2.4 | 28 | 2 | 1.25 | 34 | 2 | 0.00 | 24 | 2 | 0.25 |
| 2.5 | 29 | 2 | 1.00 | 34 | 2 | 1.75 | 24 | 2 | 1.00 |
| 3.1 | 28 | 2 | 1.00 | 32 | 2 | 1.00 | 23 | 2 | 1.00 |
| 3.2 | 28 | 2 | 1.00 | 32 | 2 | 1.25 | 23 | 2 | 2.00 |
| 3.3 | 28 | 1 | 1.00 | 31 | 2 | 1.00 | 23 | 2 | 1.00 |
| 3.4 | 28 | 1 | 1.00 | 32 | 1 | 1.00 | 22 | 1 | 1.00 |
| 3.5 | 28 | 1 | 1.00 | 31 | 1 | 1.00 | 22 | 1 | 1.00 |
| 4.1 | 29 | 4 | 3.00 | 33 | 3 | 2.00 | 23 | 3 | 2.00 |
| 4.2 | 29 | 3 | 2.00 | 32 | 3 | 2.00 | 23 | 3 | 2.00 |
| 5 | 29 | 4 | 1.00 | 33 | 4 | 1.00 | 23 | 4 | 1.00 |
| 6.1 | 29 | 4 | 1.00 | 32 | 4 | 1.00 | 23 | 4 | 1.00 |
| 6.2 | 29 | 2 | 1.00 | 32 | 3 | 2.00 | 23 | 2 | 1.50 |
| 7.1 | 28 | 2 | 1.25 | 33 | 3 | 1.00 | 23 | 3 | 1.50 |
| 7.2 | 29 | 4 | 1.00 | 33 | 3 | 1.00 | 23 | 3 | 1.00 |
| 8.1 | 26 | 4 | 1.00 | 31 | 3 | 1.00 | 23 | 3 | 1.00 |
| 8.2 | 25 | 3 | 2.00 | 31 | 3 | 2.00 | 23 | 3 | 2.00 |

**Summary of Missing Data from Caregiver Report in Chapter 2, Experiment 2**

Caregivers that didn't respond to items are not included in the analysis for those items, and *N* is adjusted accordingly. These are detailed below:

- Two caregivers (of two 24- to 26-month-olds) left question 2.3 blank, while one caregiver (of a 24- to 26-month-old) ticked NA.
- Two caregivers (of one 18- to 20-month-old and of one 30- to 32-month-old) left question 2.4 blank.
- Two caregivers (of one 24- to 26-month-old, and of one 30- to 32-month-old) did not answer questions 3-8, as they answered that their child always responded to them (5) on question 1.
- Two caregivers (of one 18- to 20-month-old and of one 24- to 26-month-old) answered NA for all items on question 3.
- One caregiver (of a 24- to 26-month-old) left question 3.5 blank.
- One caregiver (of a 30-to 32- month-old) answered NA for questions 3.4 and 3.5.
- One caregiver (of a 24- to 26-month-old) answered NA for question 4.1.
- One caregiver (of a 24- to 26-month-old) did not fill out 6.1.
- One caregiver (of a 24- to 26-month-old) did not fill out 6.2.
- One caregiver (of an 18- to 20-month-old) did not fill out 7.1.
- Three caregivers (of two 18- to 20-month-olds, and of a 24- to 26-month-old) did not answer more questions, as they answered that their child never deliberately didn't respond to them (selecting option 1) on question 7.2.
- One caregiver (of an 18- to 20-month-old) answered NA to both 8.1 and 8.2
- One caregiver (of a 24- to 26-month-old) did not answer either 8.1 or 8.2.
- Finally, one caregiver (of an 18- to 20-month-old) answered 8.1, but not 8.2.

## Appendix C

Procedure to Avoid Double Counting (Appendix to Chapter 5)

In order to calculate a measure of the observed duration where vocalisations, gestures and gesture-vocal combinations have co-occurred with gaze to caregiver's face, data is extracted from ELAN and put into excel. The output from ELAN tells us a where a behaviour has co-occurred with a gaze to caregiver's face, and excel formulas (designed by the author) tells us the *duration* that the gaze has co-occurred with the behaviour and associated temporal windows. However, sometimes there are two behaviours (i.e., vocalisations) associated with the same gaze, and there is a risk that such duration of co-occurrence may be double counted. They can be organized in 25 different ways (Figure 19).

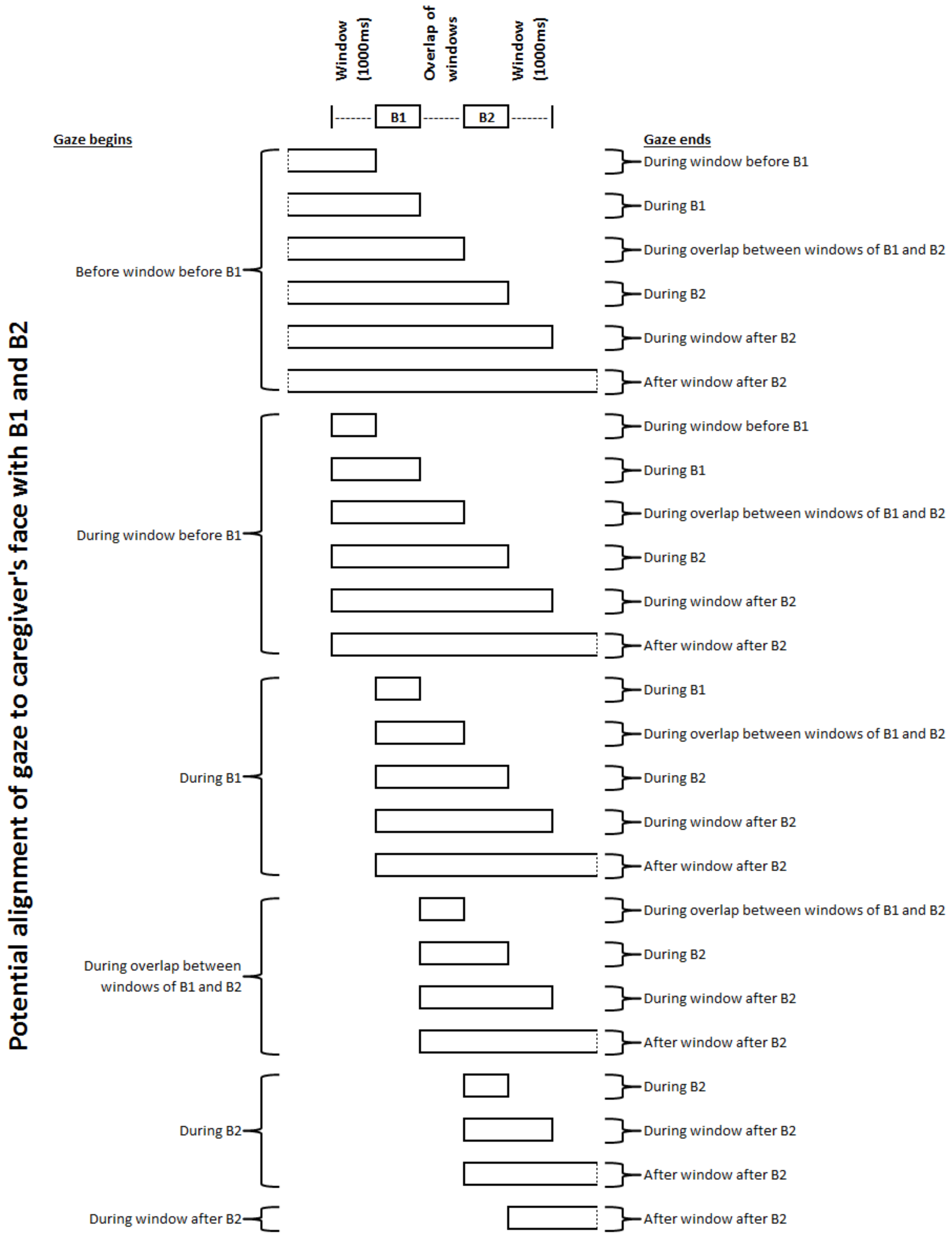


*Figure 19.* Potential alignment of gaze to caregiver's face with B1 and B2

In these cases, there will be two rows of data. They both relate to a different behaviour, but both count a portion (or the entirety) of the same gaze to caregiver's face (hereafter referred to as *gaze*), and give a duration for this co-occurrence. For the rest of this appendix, the duration of co-occurrence between gaze and the first behaviour (B1) and gaze and the second behaviour (B2) will be given as $_{dco}$B1 and $_{dco}$B2 respectively. Depending on the organisation of the gaze with the behaviour, these need to be checked and amended.

We devised exhaustive logical rules for the 25 cases where overlap could occur, as to what to do with the duration measures. These rules are hierarchical. The rules only apply to cases where it is possible that we have duplicated duration of co-occurrence data. There is firstly a quick heuristic to run if two behaviours have coded the same gaze, which negate the need to use the rules below. If the two behaviours (B1 and B2 – ordered temporally) are separated by 2000ms or more, then the 1000ms window after B1, and the 1000ms window prior to B2 (which are both in the gap between the two behaviours) will not capture the same part of the gaze. They therefore *cannot* code the same portion of the gaze, and so $_{dco}$B1 and $_{dco}$B2 are both kept.

The following rules all rely on the premise that B1 and B2 are ***not*** separated by 2000ms or more, and therefore some *same* portion of the gaze is captured by both $_{dco}$B1 and $_{dco}$B2. Both values are to be deleted, and instructions are given for the value of $_{dco}$B*x*, which will replace them with one single duration value. As these values are summed for each participant, it does not matter that two values are replaced by one as this measure is not used for generating data on *frequency* of co-occurrence, purely *total duration* of co-occurrence per participant.

1. If either $_{dco}$B1 or $_{dco}$B2 is equal to the full duration of the gaze, then we give the full duration of the gaze for $_{dco}$B*x*.
2. If the gaze *offset* (i.e., the end of the gaze) occurs within the temporal window of B1, then the duration of co-occurrence is represented by $_{d}$B1, so give this value for $_{dco}$B*x*.
3. If gaze *onset* (i.e., the beginning of the gaze) occurs within the temporal window of B2, then the duration of co-occurrence is captured by $_{d}$B2, so give this value for $_{dco}$B*x*.
4. If gaze onset occurs within the temporal window of B1, and offset occurs within the temporal window of B2, then $_{dco}$B*x* is the full duration of the gaze.

5. If gaze onset is prior to the temporal window of B1, and the offset is after the temporal window of B1 (but contained by B2), $_{dco}Bx$ is the total time from beginning of the temporal window of B1 to the gaze offset.
6. If gaze offset is after the temporal window of B2, and the onset is before the temporal window of B2 (but contained by B1). $_{dco}Bx$ is the total time from gaze onset to the end of the temporal window of B2.
7. If gaze onset is before the temporal window of B1, and the offset is after the temporal window of B2, $_{dco}Bx$ is the total time from the beginning of the temporal window before B1 to the end of the temporal window after B2.

**Multiple behaviours associated with the same gaze**

When there are more than one behaviours associated with the same gaze, similar rules apply. If any two behaviours in the 'chain' of behaviours (say B3 and B4) are separated by 2000ms or more, then the 1000ms window after B3, and the 1000ms window prior to B4 (which are both in the gap between the two behaviours) will not capture the same part of the gaze. They therefore cannot code the same portion of the gaze. This defines the length of the chains of behaviours associated with the same gaze that need dealing with (and highlight that you can have more than one chain relating to one gaze, if there is a gap in the initial chain of more than 2000ms between some behaviours). The following rules assume that you have more than two behaviours associated with the same gaze, and that all behaviours are within 2000ms of the next (i.e., code the same portion of the gaze). The final behaviour in a chain of this kind, is referred to as B*f*.

1. If any $_{dco}B$ in the chain is equal to the full duration of the gaze, then we give the full duration of the gaze for $_{dco}Bx$.
2. If the gaze *offset* (i.e., the end of the gaze) occurs within the temporal window of B1, then the duration of co-occurrence is represented by $_{d}B1$, so give this value for $_{dco}Bx$.
3. If gaze *onset* (i.e., the beginning of the gaze) occurs within the temporal window of B*f*, then the duration of co-occurrence is captured by $_{d}Bf$, so give this value for $_{dco}Bx$.

4. If gaze onset occurs within the temporal window of B1, and offset occurs within the temporal window of any other behaviour in the chain, then ${}_{dco}Bx$ is the full duration of the gaze
5. If gaze onset is prior to the temporal window of B1, and the offset is after the temporal window of B1 (but contained by another behaviour in the chain), ${}_{dco}Bx$ is the total time from beginning of the temporal window of B1 to the gaze offset.
6. If gaze offset is after the temporal window of B*f*, and the onset is before the temporal window of B*f* (but contained by any other behaviour in the chain), ${}_{dco}Bx$ is the total time from gaze onset to the end of the temporal window of B*f*.
7. If gaze onset is before the temporal window of B1, and the offset is after the temporal window of B*f*, ${}_{dco}Bx$ is the total time from the beginning of the temporal window before B1 to the end of the temporal window after B*f*.

# Appendix D

Extra Gesture-vocal Combination Descriptives (Appendix to Chapter 5)

Table 39 shows the average frequency with which specific types of gesture-vocal combinations were produced by infants at 11 & 12 months, and Table 40 shows the frequency and proportion with which these were gaze coordinated.

Table 39

*Average Frequency of Gesture-vocal Combinations (by Both Gesture and Vocalisation Type) from 10 minutes Observation at 11 (n=134) and 12 months (n=64)*

| | *11 months (n=134)* | | | | *12 months (n=64)* | | | |
|---|---|---|---|---|---|---|---|---|
| *Behaviour* | *M* | *SD* | *Median* | *Range* | *M* | *SD* | *Median* | *Range* |
| Index-finger point & CV | 0.24 | 0.99 | 0 | 0-9 | 0.36 | 1.56 | 0 | 0-12 |
| Index-finger point & Non-CV | 0.07 | 0.39 | 0 | 0-3 | 0.13 | 0.38 | 0 | 0-2 |
| Open-hand point & CV | 0.10 | 0.45 | 0 | 0-4 | 0.14 | 0.43 | 0 | 0-2 |
| Open-hand point & Non-CV | 0.09 | 0.71 | 0 | 0-8 | 0.05 | 0.21 | 0 | 0-1 |
| Give & CV | 0.14 | 0.51 | 0 | 0-4 | 0.38 | 0.83 | 0 | 0-4 |
| Give & Non-CV | 0.20 | 0.70 | 0 | 0-4 | 0.45 | 1.47 | 0 | 0-11 |
| Show & CV | 0.10 | 0.44 | 0 | 0-3 | 0.05 | 0.21 | 0 | 0-1 |
| Show & Non-CV | 0.07 | 0.33 | 0 | 0-2 | 0.16 | 0.40 | 0 | 0-2 |
| Conventional gesture & CV | 0.05 | 0.25 | 0 | 0-2 | 0.02 | 0.13 | 0 | 0-1 |
| Conventional gesture & Non-CV | 0.04 | 0.24 | 0 | 0-2 | 0.06 | 0.24 | 0 | 0-1 |

Table 40

*Average Frequency and Proportion of Gesture-vocal Combinations (by Both Gesture and Vocalisation Type) With Gaze Coordination at 11 (n=134) and 12 months (n=64)*

| | *11 months (n=134)* | | | | | *12 months (n=64)* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| Index-finger point & CV | 0.10 | 0.45 | 0 | 0-3 | .44 | 0.20 | 0.83 | 0 | 0-6 | .69 |
| Index-finger point & Non-CV | 0.01 | 0.09 | 0 | 0-1 | .07 | 0.06 | 0.30 | 0 | 0-2 | .43 |
| Open-hand point & CV | 0.03 | 0.21 | 0 | 0-2 | .22 | 0.06 | 0.24 | 0 | 0-1 | .43 |
| Open-hand point & Non-CV | 0.04 | 0.44 | 0 | 0-5 | .33 | 0.03 | 0.17 | 0 | 0-1 | .67 |
| Give & CV | 0.07 | 0.25 | 0 | 0-1 | .60 | 0.22 | 0.55 | 0 | 0-3 | .63 |
| Give & Non-CV | 0.14 | 0.56 | 0 | 0-3 | .63 | 0.25 | 1.16 | 0 | 0-9 | .40 |
| Show & CV | 0.09 | 0.43 | 0 | 0-3 | .88 | 0.05 | 0.21 | 0 | 0-1 | 1.00 |
| Show & Non-CV | 0.07 | 0.33 | 0 | 0-2 | 1.00 | 0.16 | 0.40 | 0 | 0-2 | 1.00 |
| Conventional gesture & CV | 0.03 | 0.17 | 0 | 0-1 | .58 | 0.02 | 0.13 | 0 | 0-1 | 1.00 |
| Conventional gesture & Non-CV | 0.04 | 0.19 | 0 | 0-1 | .90 | 0.02 | 0.12 | 0 | 0-1 | .25 |

## Appendix E

Inclusive Analyses (Appendix to Chapter 5)

Descriptive statistics for infant behaviours at 11 and 12 months (and 11 & 12 months combined) are presented in Table 41. For the following analyses, we *did not* treat gesture-vocal combinations as a separate, mutually exclusive category from gesture and vocalisations, instead counting vocalisations and gestures involved in combinations in gestures and vocalisations respectively (even though they overlapped in time).

Table 41

*Average Frequency of Vocalisations and Gestures from 10 minutes Observation at 11 (n = 134) and 12 months (n = 64)*

| | *11 months (n = 134)* | | | | *12 months (n = 64)* | | | | *11 & 12 months (n = 61)* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Behaviour* | *M* | *SD* | *Median* | *Range* | *M* | *SD* | *Median* | *Range* | *M* | *SD* | *Median* | *Range* |
| **Gaze to Caregiver's Face** | 22.21 | 12.11 | 19 | 1-53 | 22.18 | 13.68 | 19 | 1-63 | 44.87 | 21.80 | 40 | 5-99 |
| **Vocalisations** | 48.89 | 29.09 | 43 | 4-173 | 41.42 | 26.24 | 37.5 | 7-128 | 89.08 | 42.98 | 83 | 17-203 |
| CV | 19.20 | 18.31 | 15 | 0-110 | 17.43 | 13.70 | 15.5 | 0-54 | 35.44 | 26.72 | 29 | 2-129 |
| Non-CV | 29.69 | 17.76 | 26.5 | 2-83 | 23.99 | 19.38 | 19 | 2-123 | 53.64 | 27.95 | 52 | 11-145 |
| **Gestures** | 2.73 | 3.36 | 2 | 0-16 | 4.57 | 6.02 | 2 | 0-30 | 7.11 | 6.80 | 5 | 0-26 |
| Index-finger point | 0.53 | 1.53 | 0 | 0-12 | 1.02 | 2.46 | 0 | 0-14 | 1.44 | 2.83 | 0 | 0-15 |
| Open-hand point | 0.39 | 1.16 | 0 | 0-10 | 0.23 | 0.55 | 0 | 0-2 | 0.72 | 1.62 | 0 | 0-10 |
| Give | 1.02 | 2.06 | 0 | 0-12 | 2.71 | 4.66 | 1 | 0-20 | 3.46 | 4.51 | 2 | 0-20 |
| Show | 0.46 | 1.13 | 0 | 0-7 | 0.41 | 0.82 | 0 | 0-4 | 0.87 | 1.49 | 0 | 0-7 |
| Conventional gesture | 0.34 | 0.88 | 0 | 0-6 | 0.20 | 0.47 | 0 | 0-2 | 0.62 | 1.17 | 0 | 0-7 |

There was one instance of significant developmental change. For the infants whose measures are reported at both measurement points, regression analyses and follow up pair-wise comparisons revealed that the frequency of giving significantly increased between 11 and 12 months (Table 42 & 43).

Table 42

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures by Gesture Type*Age at 11 and 12 months (n = 61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Index-finger Point*) | 0.52 | 0.24 | 2.20 | 300 | .028 |
| Open-hand Point | -0.03 | 0.33 | -0.10 | 240 | .920 |
| Give | 0.42 | 0.33 | 1.26 | 240 | .209 |
| Show | -0.08 | 0.33 | -0.25 | 240 | .804 |
| Conventional | -0.10 | 0.33 | -0.30 | 240 | .766 |
| 12 months | 0.39 | 0.33 | 1.19 | 300 | .235 |
| Open-hand Point * 12 months | -0.66 | 0.47 | -1.40 | 300 | .162 |
| Give * 12 months | 1.18 | 0.47 | 2.53 | 300 | .012 |
| Show * 12 months | -0.41 | 0.47 | -0.88 | 300 | .381 |
| Conventional * 12 months | -0.62 | 0.47 | -1.33 | 300 | .184 |

Table 43

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types by Age (11 and 12 months) (n = 61)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Index-finger point (11 months : 12 months) | -0.39 | 300 | -1.19 | .973 |
| Open-hand point (11 months : 12 months) | 0.26 | 300 | 0.79 | .999 |
| Give (11 months : 12 months) | -1.57 | 300 | -4.74 | < .001 |
| Show (11 months : 12 months) | 0.02 | 300 | 0.05 | 1.000 |
| Conventional (11 months : 12 months) | 0.23 | 300 | 0.69 | 1.000 |

**Do Infants' Vocalisations or Gestures Co-occur With Gaze to Caregiver's Face Above the Rate Expected by Chance?**

Figure 20 shows the expected and observed duration that vocalisations and gestures and co-occurred with gaze to caregiver's face at both 11 and 12 months (for 11 & 12 month combined, see figure 21).

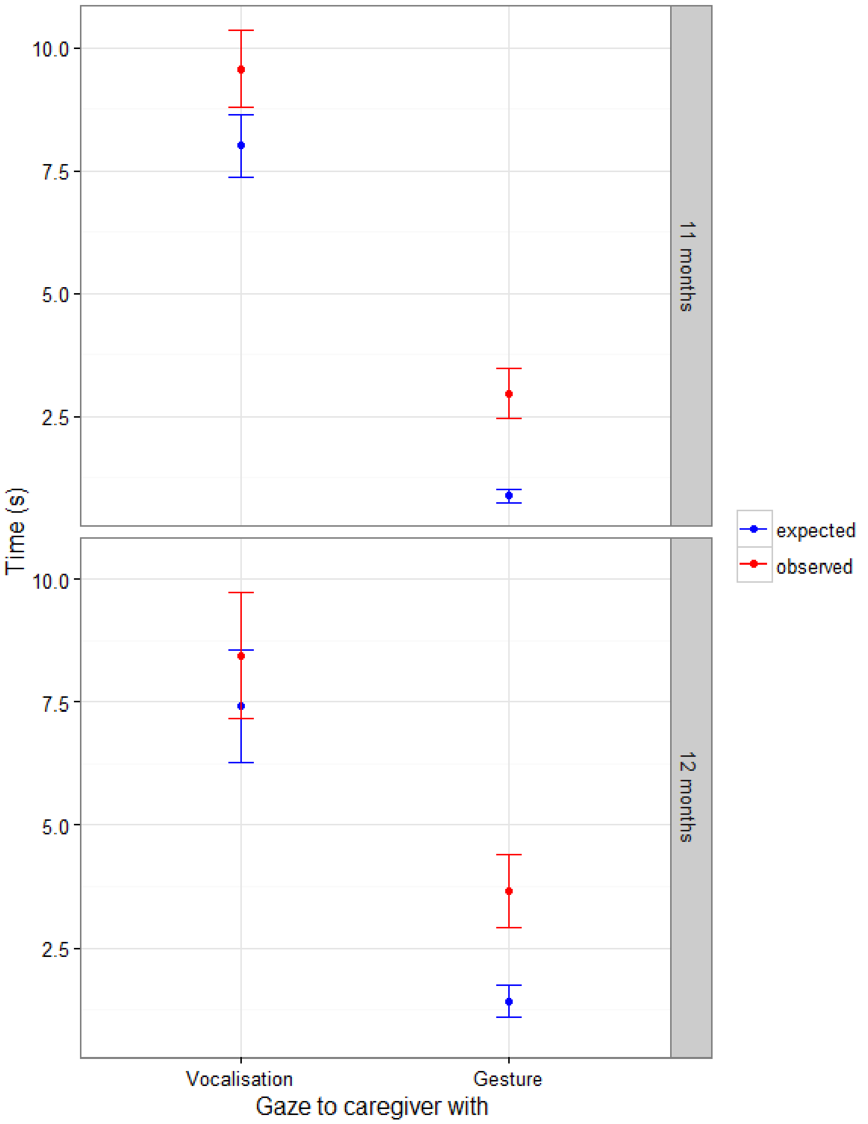


*Figure 20*. Mean expected and observed co-occurrence of vocalisations and gestures with gaze to caregiver's face at 11 (*n* = 134) and 12 months (*n* = 64). Error bars represent standard error about the mean.

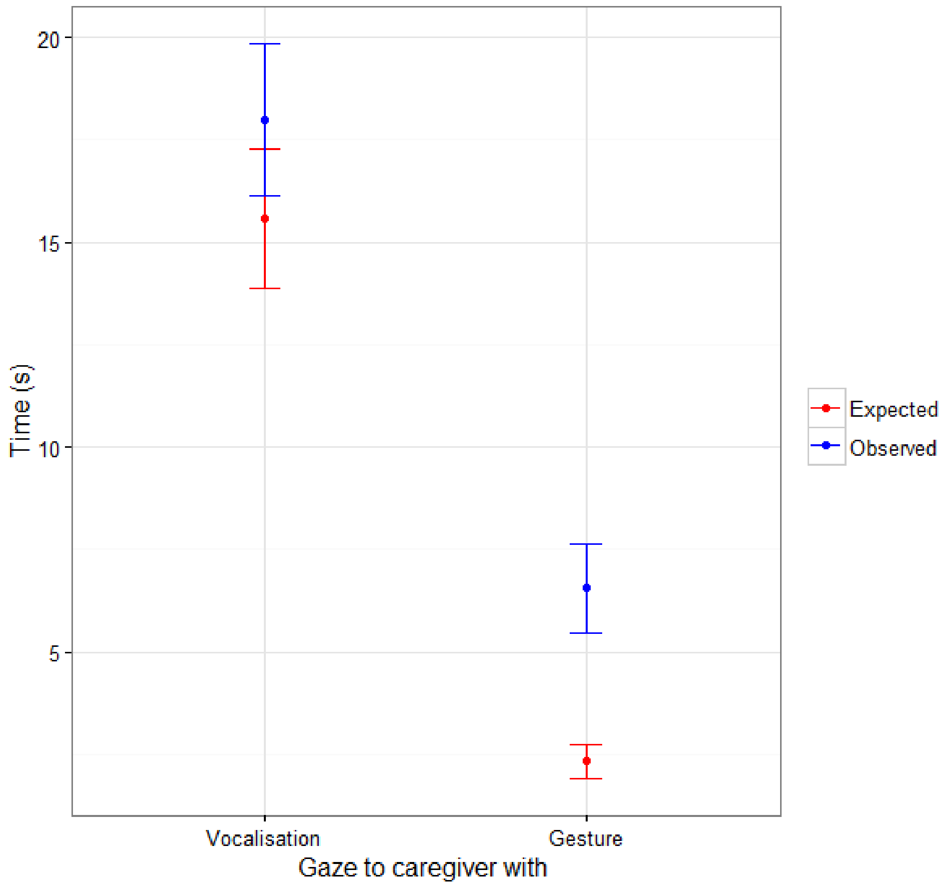


*Figure 21*. Mean expected and observed co-occurrence of vocalisations and gestures with gaze to caregiver's face at 11 & 12 months ($n$ = 61). Error bars represent standard error about the mean.

At both ages (and when collapsed across age), both vocalisations and gestures co-occurred with gaze to caregiver's face significantly above the level predicted by chance (11 months: vocalisations, $t(133) = 3.79$, $p < .001$; gestures, $t(133) = 5.30$, $p < .001$; 12 months: vocalisations, $t(63) = 2.10$ $p = .040$; gestures, $t(63) = 4.46$, $p < .001$; 11 & 12 months: vocalisations, $t(60) = 3.00$, $p = .004$; gestures, $t(60) = 5.46$, $p < .001$).

Figure 22 shows the expected and observed duration that individual vocalisation and gesture types (both produced alone and as part of gesture-vocal combinations) co-occurred with gaze to caregiver's face at both 11 and 12 months (for 11 & 12 month combined, see figure 23).

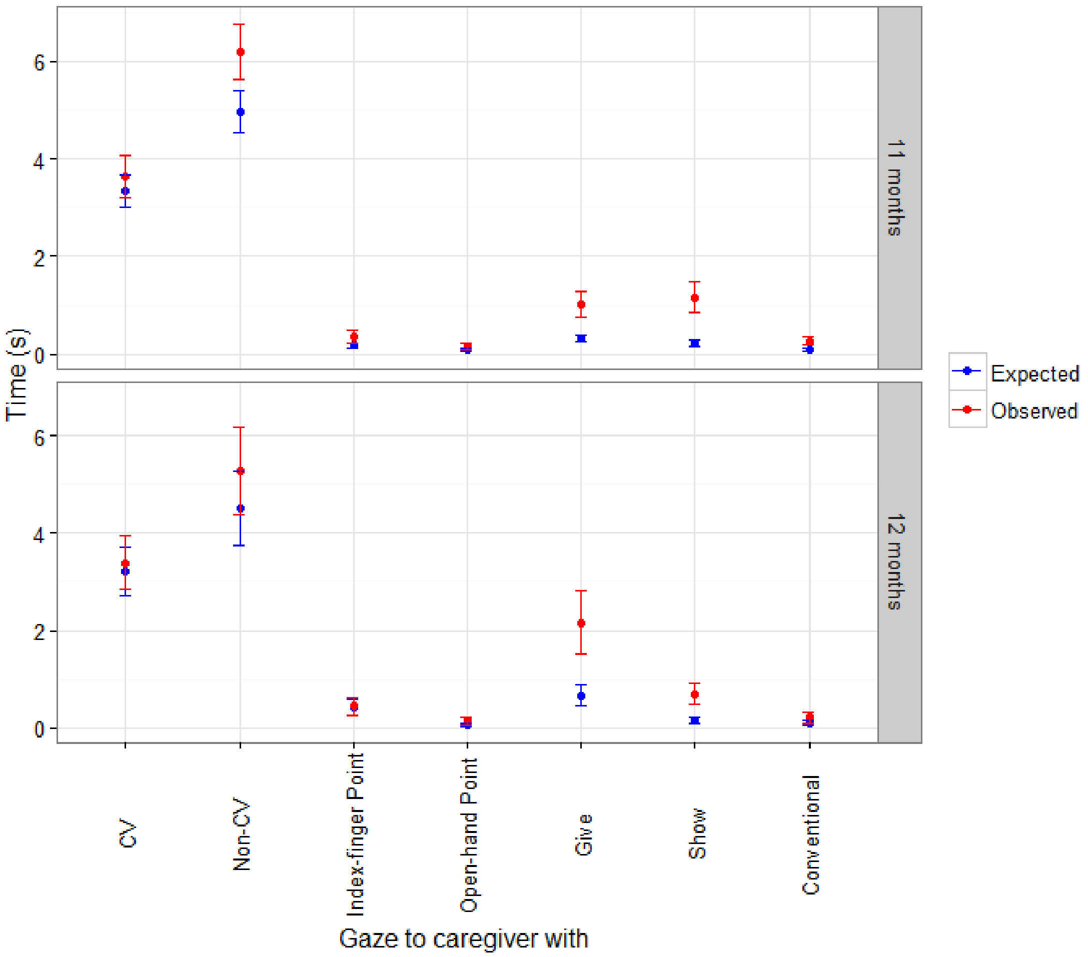


*Figure 22.* Mean expected and observed co-occurrence of individual vocalisation and gesture types co-occurred with gaze to caregiver's face at 11 ($n$ = 134) and 12 months ($n$ = 64). Error bars represent standard error about the mean.

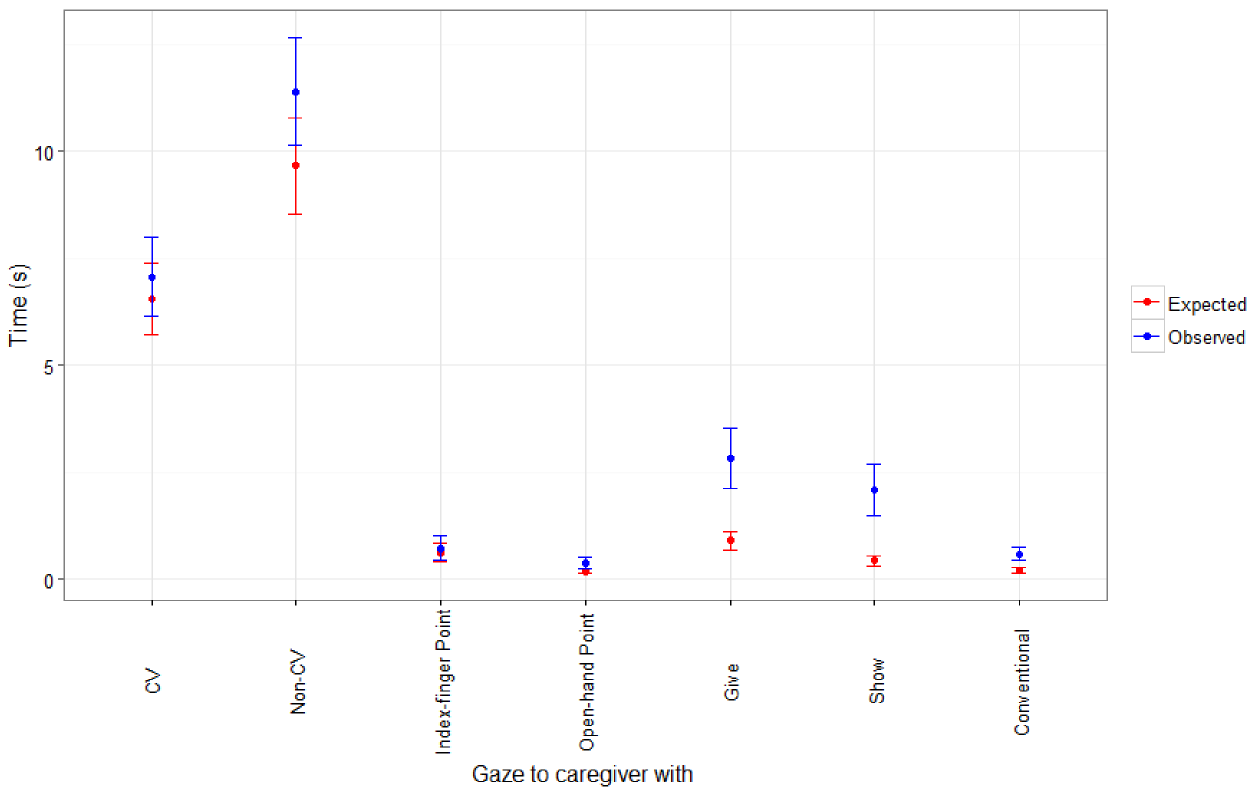


*Figure 23.* Mean expected and observed co-occurrence of individual vocalisation and gesture types co-occurred with gaze to caregiver's face at 11 & 12 months (*n* = 61). Error bars represent standard error about the mean.

**Vocalisation type.** At 11 months (and when collapsed across age), non-CV vocalisations co-occurred with gaze to caregiver's face significantly above the level predicted by chance (11 months: $t(133) = 3.79$, $p < .001$; 11 & 12 months: $t(60) = 3.38$, $p = .001$) while CV vocalisations did not. At 12 months, neither CV, nor non-CV vocalisations co-occurred with gaze to caregiver's face significantly above the level predicted by chance (though this was borderline for non-CV vocalisations, $t(63) = 1.93$, $p = .058$).

**Gesture type.** At 11 months, index-finger pointing, giving, showing and conventional gestures co-occurred with gaze to caregiver's face above the level predicted by chance (index-finger pointing, $t(133) = 2.00$, $p = .048$; giving, $t(133) = 3.66$, $p < .001$; showing, $t(133) = 3.65$, $p < .001$; conventional gestures, $t(133) = 2.87$, $p = .005$). At 12 months, only giving and showing co-occurred with gaze significantly above the level predicted by chance (giving, $t(63) = 3.22$, $p = .002$; showing, $t(63) = 3.36$, $p = .001$). When we collapsed across age, only giving, showing and conventional gestures co-occurred with gaze above the level predicted by chance (giving, $t(60) = 3.67$, $p = .001$; showing, $t(60) = 3.40$, $p = .001$; conventional gestures, $t(60) = 3.24$, $p = .002$). In sum, open-hand pointing did not co-occur with gaze above chance at any age, and index-finger pointing only co-occurred with gaze significantly above chance at 11 months (but not at 12 months or when collapsed across age).

**Is Intentional Communication More Likely to be Vocal or Gestural?**

Our second question concerned whether gaze coordination, taken as an index of intentional communication, was more likely to occur with vocalisations or gestures. Table 44 shows descriptive statistics including the frequency and proportion of vocalisations and gestures with gaze co-ordination.

Table 44

*Average Frequency and Proportion of Vocalisations and Gestures With Gaze Coordination at 11 months (n = 134), 12 months (n = 64), and 11 & 12 months (n = 61)*

| | *11 months (n = 134)* | | | | | *12 months (n = 64)* | | | | | *11 & 12 months (n = 61)* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Prop* | *Frequency* | | | | *Prop* | *Frequency* | | | | *Prop* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 9.15 | 7.82 | 8 | 0-37 | .19 | 8.20 | 7.75 | 6 | 0-31 | .20 | 16.93 | 11.63 | 14 | 0-51 | .19 |
| CV | 3.47 | 4.55 | 2 | 0-24 | .17 | 3.26 | 3.49 | 2 | 0-13 | .21 | 6.49 | 6.25 | 5 | 0-30 | .19 |
| Non-CV | 5.68 | 5.19 | 4 | 0-26 | .20 | 4.94 | 5.25 | 3 | 0-27 | .21 | 10.44 | 7.49 | 9 | 0-32 | .20 |
| **Gestures** | 1.60 | 2.39 | 1 | 0-12 | .56 | 2.27 | 3.53 | 1 | 0-20 | .46 | 3.80 | 4.17 | 3 | 0-20 | .54 |
| Index-finger point | 0.22 | 0.74 | 0 | 0-5 | .37 | 0.39 | 1.23 | 0 | 0-7 | .35 | 0.51 | 1.32 | 0 | 0-8 | .36 |
| Open-hand point | 0.13 | 0.64 | 0 | 0-6 | .27 | 0.09 | 0.29 | 0 | 0-1 | .36 | 0.34 | 0.99 | 0 | 0-6 | .39 |
| Give | 0.62 | 1.47 | 0 | 0-10 | .61 | 1.34 | 2.80 | 0 | 0-18 | .48 | 1.78 | 3.00 | 1 | 0-18 | .48 |
| Show | 0.46 | 1.20 | 0 | 0-8 | .92 | 0.36 | 0.76 | 0 | 0-3 | .86 | 0.80 | 1.45 | 0 | 0-7 | .89 |
| Conventional gesture | 0.19 | 0.54 | 0 | 0-3 | .57 | 0.09 | 0.34 | 0 | 0-2 | .41 | 0.36 | 0.66 | 0 | 0-3 | .63 |

**Individual behaviours.** We investigated whether single (i.e. not bouts of) vocalisations or gestures were more frequently, or proportionally more likely to be gaze coordinated.

***Frequency.*** At both ages (and when we collapsed across age), a significantly higher frequency of vocalisations were gaze coordinated than gestures, (11 months: $t(133) = 11.46$, $p < .001$; 12 months: $t(63) = 6.76$, $p < .001$; 11 & 12 months: $t(60) = 10.17$, $p < .001$).

***Proportion.*** At both ages (and when we collapsed across age) there was a significant improvement in model fit (see method) when behaviour type was added as a predictor, (11 months: $\chi^2$ (1) = 52.14, $p < .001$ (Table 45); 12 months: $\chi^2$ (1) = 27.37, $p < .001$ (Table 46); 11 & 12 months: $\chi^2$ (1) = 56.32, $p < .001$ (Table 47). At both ages (and when collapsed across age), a significantly higher proportion of infants' gestures were gaze coordinated than their vocalisations (Table 45-47).

Table 45

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* ($n$ = 6483) *and Gestures* ($n$ = 364) *With Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Behaviour Type at 11 months (n = 134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.61 | 0.08 | -21.26 | < .001 |
| Gesture | 1.89 | 0.18 | 10.39 | < .001 |

LLRI = .01, C = .73, Dxy = .47.

Table 46

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 2640) *and Gestures* (*n* = 290) *with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Behaviour Type at 12 months (n = 64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.61 | 0.14 | -11.81 | < .001 |
| Gesture | 1.51 | 0.22 | 6.73 | < .001 |

LLRI = .01, C = .77, Dxy = .53.

Table 47

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 5397) *and Gestures* (*n* = 430) *With Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Behaviour Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.58 | 0.10 | -16.10 | < .001 |
| Gesture | 1.68 | 0.16 | 10.70 | < .001 |

LLRI = .01, C = .71, Dxy = .43.

**Vocalisation type**. We investigated whether CV or non-CV vocalisations were more frequently, or proportionally more likely to be gaze coordinated.

***Frequency.*** At both ages (and when collapsed across age), a significantly higher frequency of non-CV vocalisations were gaze coordinated than CV vocalisations (11 months: $t(133) = 4.35$, $p < .001$; 12 months: $t(133) = 3.04$, $p = .003$; 11 & 12 months: $t(60) = 4.16$, $p < .001$).

***Proportion.*** Adding vocalisation type did not improve model fit at either age (or when we collapsed across age).

**Gesture type.** We investigated whether index-finger pointing, open-hand pointing, giving, showing or conventional gestures were more frequently, or proportionally more likely to be gaze coordinated.

***Frequency.*** At both ages (and when we collapsed across age) there was a significant improvement in model fit when gesture type was added as a predictor, (11 months: $\chi^2$ (4) = 24.78, $p$ < .001 (Table 48); 12 months: $\chi^2$ (4) = 32.96, $p$ < .001 (Table 49); 11 & 12 months: $\chi^2$ (4) = 31.26, $p$ < .001 (Table 50)).

Table 48

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures with Gaze Coordination by Gesture Type at 11 months (n = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.21 | 0.08 | 2.59 | .010 |
| Open-hand Pointing | -0.08 | 0.11 | -0.72 | .469 |
| Give | 0.40 | 0.11 | 3.53 | < .001 |
| Show | 0.22 | 0.11 | 1.90 | .058 |
| Conventional | -0.02 | 0.11 | -0.19 | .846 |

Table 49

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures with Gaze Coordination by Gesture Type at 12 months (n = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.39 | 0.18 | 2.18 | .030 |
| Open-hand Pointing | -0.30 | 0.25 | -1.20 | .230 |
| Give | 0.95 | 0.25 | 3.84 | < .001 |
| Show | -0.03 | 0.25 | -0.13 | .900 |
| Conventional | -0.30 | 0.25 | -1.20 | .230 |

Table 50

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures with Gaze Coordination by Gesture Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.51 | 0.22 | 2.35 | .020 |
| Open-hand Pointing | -0.17 | 0.30 | -0.56 | .576 |
| Give | 1.28 | 0.30 | 4.29 | < .001 |
| Show | 0.30 | 0.30 | 0.99 | .321 |
| Conventional | -0.15 | 0.30 | -0.50 | .621 |

Post-hoc tests pairwise tests revealed that at both ages (and when collapsed across age), a significantly higher frequency of giving was gaze coordinated than all other gestures (Table 51), with the exception of showing at 11 months only.

Table 51

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types with Gaze Coordination at 11 months (n = 134), 12 months (n = 64) and 11 & 12 months (n = 61).*

| | *11 months* | | | | *12 months* | | | | *11 and 12 months* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* |
| Index-finger point : Open-hand point | 0.08 | 532 | 0.72 | .951 | 0.30 | 252 | 1.20 | .749 | 0.17 | 240 | 0.56 | .981 |
| Index-finger point : Give | -0.41 | 532 | -3.53 | .004 | -0.95 | 252 | -3.84 | .002 | -1.28 | 240 | -4.29 | < .001 |
| Index-finger point : Show | -0.22 | 532 | -1.90 | .318 | 0.03 | 252 | 0.13 | 1.000 | -0.30 | 240 | -0.99 | .858 |
| Index-finger point : Conventional | 0.02 | 532 | 0.20 | 1.000 | 0.30 | 252 | 1.20 | .750 | 0.15 | 240 | 0.50 | .988 |
| Open-hand point : Give | -0.49 | 532 | -4.26 | < .001 | -1.24 | 252 | -5.04 | < .001 | -1.44 | 240 | -4.85 | < .001 |
| Open-hand point : Show | -0.30 | 532 | -2.63 | .067 | -0.27 | 252 | -1.08 | .818 | -0.46 | 240 | -1.55 | .529 |
| Open-hand point : Conventional | -0.06 | 532 | -0.53 | .984 | 0.00 | 252 | 0.00 | 1.000 | -0.02 | 240 | -0.06 | 1.000 |
| Give : Show | 0.19 | 532 | 1.63 | .477 | 0.98 | 252 | 3.97 | .001 | 0.98 | 240 | 3.30 | .010 |
| Give : Conventional | 0.43 | 532 | 3.73 | .002 | 1.24 | 252 | 5.04 | < .001 | 1.42 | 240 | 4.79 | < .001 |
| Show : Conventional | 0.24 | 532 | 2.10 | .223 | 0.27 | 252 | 1.08 | .818 | 0.44 | 240 | 1.49 | .570 |

***Proportion.*** At 12 months (and when collapsed across age), adding gesture type improved model fit, (12 months: $\chi^2$ (4) = 20.98, $p < .001$ (Table 52); 11 & 12 months: $\chi^2$ (4) = 24.14, $p < .001$ (Table 53). At 11 months we lacked the requisite number of gestures to run models (see method).

Table 52

*Summary of Fixed Effects from a Logistic Regression Model Fitting Index-finger Pointing (n = 65), Open-hand Pointing (n = 15), Giving (n = 171), Showing (n = 26) and Conventional Gestures (n = 13) with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Gesture Type at 12 months (n = 52)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Point*) | -0.56 | 0.37 | -1.52 | .129 |
| Open-hand Point | -0.36 | 0.38 | -0.50 | .615 |
| Give | 0.47 | 0.71 | 1.23 | .219 |
| Show | 2.78 | 0.73 | 3.65 | < .001 |
| Conventional Gesture | 0.40 | 0.76 | 0.55 | .585 |

LLRI = .06, C = .83, Dxy = .66.

Table 53

*Summary of Fixed Effects from a Logistic Regression Model Fitting Index-finger Pointing (n = 88), Open-hand Pointing (n = 43), Giving (n = 208), Showing (n = 53) and Conventional Gestures (n = 38) with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Gesture Type at 11 & 12 months (n = 52)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Point*) | -0.88 | 0.36 | -2.46 | .014 |
| Open-hand Point | 0.64 | 0.58 | 1.11 | .269 |
| Give | 0.95 | 0.39 | 2.39 | .017 |
| Show | 7.84 | 3.74 | 2.10 | .036 |
| Conventional | 1.43 | 0.58 | 2.48 | .013 |

LLRI = .05, C = .85, Dxy = .71.

Post-hoc pairwise tests revealed that at 12 months (and when collapsed across age), a significantly higher proportion of infants' show gestures were gaze coordinated than any other type of gesture (Table 54), with the exception of conventional gestures at 12 months only.

Table 54

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Proportion of Gesture Types with Gaze Coordination at 12 months (n = 52) and 11 & 12 months (n = 52)*

| | *12 months* | | | | *11 & 12 months* | | | |
|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *z* | *p* | *b* | *SE* | *z* | *p* |
| Index-finger point : Open-hand point | 0.36 | 0.71 | 0.50 | .987 | -0.64 | 0.58 | -1.10 | .804 |
| Index-finger point : Give | -0.47 | 0.38 | -1.23 | .735 | -0.94 | 0.39 | -2.39 | .117 |
| Index-finger point : Show | -2.78 | 0.76 | -3.65 | .002 | -7.84 | 3.74 | -2.09 | .222 |
| Index-finger point : Conventional | -0.40 | 0.73 | -0.55 | .983 | -1.43 | 0.58 | -2.47 | .096 |
| Open-hand point : Give | -0.83 | 0.67 | 1.24 | .729 | -0.31 | 0.50 | 0.61 | .973 |
| Open-hand point : Show | -3.14 | 0.94 | -3.33 | .008 | -7.21 | 3.73 | -1.93 | .302 |
| Open-hand point : Conventional | -0.75 | 0.91 | -0.83 | .922 | -0.79 | 0.59 | -1.33 | .673 |
| Give : Show | -2.31 | 0.73 | -3.17 | .013 | -6.90 | 3.72 | -1.86 | .342 |
| Give : Conventional | 0.07 | 0.68 | 0.11 | 1.000 | -0.48 | 0.50 | -0.96 | .872 |
| Show : Conventional | 2.39 | 0.95 | -2.53 | .085 | 6.41 | 3.75 | -1.71 | .427 |

**Developmental change.** We were interested in determining whether there were changes in the frequency, or proportion that behaviours were gaze coordinated between 11 and 12 months. There was only one instance of significant developmental change. For the infants whose measures are reported at both time points, regression analyses and follow up pair-wise comparisons revealed that the frequency of giving *with gaze coordination* significantly *increased* between 11 and 12 months ($b$ = 0.76, $t$(300) = 3.59, $p$ = .014, Table 55 for model).

Table 55

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures with Gaze Coordination by Gesture Type*Age at 11 and 12 months (n = 61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Index-finger Point*) | 0.20 | 0.15 | 1.30 | 300 | .195 |
| Open-hand Point | 0.05 | 0.21 | 0.22 | 240 | .825 |
| Give | 0.32 | 0.21 | 1.49 | 240 | .137 |
| Show | 0.23 | 0.21 | 1.08 | 240 | .279 |
| Conventional | 0.07 | 0.21 | 0.31 | 240 | .757 |
| 12 months | 0.11 | 0.21 | 0.54 | 300 | .588 |
| Open-hand Point * 12 months | -0.26 | 0.30 | -0.87 | 300 | .386 |
| Give * 12 months | 0.64 | 0.30 | 2.15 | 300 | .032 |
| Show * 12 months | -0.16 | 0.30 | -0.55 | 300 | .585 |
| Conventional * 12 months | -0.28 | 0.30 | -0.93 | 300 | .353 |

**Appendix F**

Full Regression Models (Appendix to Chapter 5)

Table 56

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gaze to Caregiver's Face, Vocalisations, Gestures* and *Combinations by Behaviour Type*Age at 11 and 12 months (n=61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Gaze to Caregiver's Face*) | 22.13 | 1.92 | 11.55 | 240 | < .001 |
| Vocalisations | 24.60 | 2.63 | 9.34 | 180 | < .001 |
| Gestures | -20.30 | 2.63 | -7.71 | 180 | < .001 |
| Combinations | -21.13 | 2.63 | -8.02 | 180 | < .001 |
| 12 months | 0.62 | 2.36 | 0.26 | 240 | .794 |
| Vocalisations * 12 months | -8.23 | 3.33 | -2.47 | 240 | .014 |
| Gestures * 12 months | 0.14 | 3.33 | 0.04 | 240 | .966 |
| Combinations * 12 months | 0.09 | 3.33 | 0.03 | 240 | .980 |

Table 57

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (All Produced Alone) by Gesture Type*Age at 11 and 12 months (n=61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Index-finger Point*) | 0.30 | 0.16 | 1.81 | 300 | .071 |
| Open-hand Point | -0.07 | 0.22 | -0.30 | 240 | .767 |
| Give | 0.38 | 0.22 | 1.73 | 240 | .086 |
| Show | -0.02 | 0.22 | -0.07 | 240 | .941 |
| Conventional | 0.05 | 0.22 | 0.22 | 240 | .825 |
| 12 months | 0.18 | 0.22 | 0.81 | 300 | .417 |
| Open-hand Point * 12 months | -0.38 | 0.31 | -1.20 | 300 | .231 |
| Give * 12 months | 0.87 | 0.31 | 2.78 | 300 | .006 |
| Show * 12 months | -0.25 | 0.31 | -0.78 | 300 | .434 |
| Conventional * 12 months | -0.39 | 0.31 | -1.25 | 300 | .211 |

Table 58

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Bouts by Bout Type*Age at 11 and 12 months (n=61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Unimodal Vocal Bouts*) | 35.69 | 1.35 | 26.37 | 180 | < .001 |
| Unimodal Gestural Bouts | -34.25 | 1.89 | -18.17 | 120 | < .001 |
| Multimodal Bouts | -34.35 | 1.89 | -18.22 | 120 | < .001 |
| 12 months | -5.75 | 1.56 | -3.69 | 180 | < .001 |
| Unimodal Gestural Bouts * 12 months | 6.40 | 2.20 | 2.90 | 180 | .004 |
| Multimodal Bouts * 12 months | 6.31 | 2.20 | 2.86 | 180 | .005 |

Table 59

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations with Gaze Coordination by Behaviour Type at 11 months (N=134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | 8.48 | 0.39 | 21.67 | < .001 |
| Gesture | -7.50 | 0.54 | -13.85 | < .001 |
| Gesture-vocal Combination | -7.87 | 0.54 | -14.53 | < .001 |

Table 60

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations with Gaze Coordination by Behaviour Type at 12 months (N=64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | 7.01 | 0.54 | 13.04 | < .001 |
| Gesture | -5.78 | 0.71 | -8.20 | < .001 |
| Gesture-vocal Combination | -5.94 | 0.71 | -8.42 | < .001 |

Table 61

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n*=6297), *Gestures* (*n*=217) *and Gesture-vocal Combinations* (*n*=147) *With Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Behaviour Type at 11 months (N=134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.68 | 0.08 | -21.57 | < .001 |
| Gesture | 2.06 | 0.26 | 7.83 | < .001 |
| Gesture-vocal Combination | 2.10 | 0.23 | 9.21 | < .001 |

LLRI = .01, C = .74, Dxy = .48.

Table 62

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n*=2504), *Gestures* (*n*=177) *and Gesture-vocal Combinations* (*n*=113) *with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Behaviour Type at 12 months (N=64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.73 | 0.14 | -12.18 | < .001 |
| Gesture | 1.99 | 0.27 | 7.26 | < .001 |
| Gesture-vocal Combination | 1.56 | 0.30 | 5.23 | < .001 |

LLRI = .01, C = .79, Dxy = .57.

Table 63

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Unimodal Vocal, Unimodal Gestural and Multimodal Bouts with Gaze Coordination by Bout Type at 11 months (N=134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Unimodal vocal bouts*) | 6.75 | 0.29 | 23.22 | < .001 |
| Gestural bouts | -5.95 | 0.40 | -14.79 | < .001 |
| Multimodal bouts | -5.94 | 0.40 | -14.77 | < .001 |

Table 64

*Regression Model fitting Frequency of Unimodal Vocal, Unimodal Gestural and Multimodal Bouts with Gaze Coordination by Bout Type at 12 months (N=64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Unimodal vocal bouts*) | 5.83 | 0.46 | 12.76 | < .001 |
| Gestural bouts | -4.88 | 0.61 | -7.98 | < .001 |
| Multimodal bouts | -4.72 | 0.61 | -7.23 | < .001 |

Table 65

*Summary of Fixed Effects from a Logistic Regression Model Fitting Unimodal Vocal* (*n*=4761), *Unimodal Gestural* (*n*=176) *and Multimodal* (*n*=183) *Bouts with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Bout Type at 11 months (N=134)*

| | *B* | *SE* | *Z* | *p* |
|---|---|---|---|---|
| Intercept (*Unimodal vocal bouts*) | -1.60 | 0.07 | -21.83 | < .001 |
| Gestural bouts | 2.11 | 0.25 | 8.37 | < .001 |
| Multimodal bouts | 1.99 | 0.25 | 7.80 | < .001 |

LLRI = .01, C = .74, Dxy = .48.

Table 66

*Summary of Fixed Effects from a Logistic Regression Model Fitting Unimodal Vocal* (*n*=1912), *Unimodal Gestural* (*n*=145) *and Multimodal* (*n*=127) *Bouts with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Bout Type at 12 months (N=64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Unimodal vocal bouts*) | -1.63 | 0.14 | -11.78 | < .001 |
| Gestural bouts | 1.27 | 0.32 | 3.95 | < .001 |
| Multimodal bouts | 1.88 | 0.26 | 7.21 | < .001 |

LLRI = .02, C = .78, Dxy = .57.

Table 67

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n*=1040) *and Non-CV Vocalisations* (*n*=1464) *with Gaze Coordination (1= Gaze Coordinated, 0 = Not) by Vocalisation Type at 12 months (N=64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*CV Vocalisation*) | -1.88 | 0.16 | -11.85 | < .001 |
| Non-CV Vocalisations | 0.26 | 0.12 | 2.18 | .029 |

LLRI = .002, C = .76, Dxy = .51.

Table 68

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) with Gaze Coordination by Gesture Type at 11 months (N=134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.10 | 0.06 | 1.78 | .076 |
| Open-hand Pointing | -0.04 | 0.08 | -0.48 | .630 |
| Give | 0.31 | 0.08 | 3.87 | < .001 |
| Show | 0.17 | 0.08 | 2.18 | .030 |
| Conventional | 0.02 | 0.08 | 0.28 | .780 |

Table 69

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) with Gaze Coordination by Gesture Type at 12 months (N=64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.14 | 0.11 | 1.29 | .199 |
| Open-hand Pointing | -0.14 | 0.15 | -0.93 | .352 |
| Give | 0.72 | 0.15 | 4.79 | < .001 |
| Show | 0.02 | 0.15 | 0.10 | .918 |
| Conventional | -0.08 | 0.15 | -0.52 | .605 |

Table 70

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gesture-vocal Combinations with Gaze Coordination by Gesture Type at 12 months (N=64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.27 | 0.10 | 2.69 | .008 |
| Open-hand Pointing | -0.17 | 0.13 | -1.29 | .199 |
| Give | 0.21 | 0.13 | 1.56 | .120 |
| Show | -0.06 | 0.13 | -0.47 | .641 |
| Conventional | -0.23 | 0.13 | -1.76 | .080 |

Table 71

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (All Produced Alone) with Gaze Coordination by Gesture Type*Age at 11 and 12 months (n=61)*

| | *B* | *SE* | *t* | *df* | *p* |
|---|---|---|---|---|---|
| Intercept (*11 months*: *Index-finger Point*) | 0.13 | 0.94 | 1.39 | 300 | .165 |
| Open-hand Point | -0.02 | 0.13 | -0.14 | 240 | .886 |
| Give | 0.22 | 0.13 | 1.66 | 240 | .098 |
| Show | 0.13 | 0.13 | 1.00 | 240 | .317 |
| Conventional | 0.07 | 0.13 | 0.50 | 240 | .617 |
| 12 months | -0.05 | 0.13 | -0.38 | 300 | .707 |
| Open-hand Point * 12 months | -0.06 | 0.18 | -0.34 | 300 | .733 |
| Give * 12 months | 0.49 | 0.18 | 2.66 | 300 | .008 |
| Show * 12 months | -0.05 | 0.18 | -0.27 | 300 | .791 |
| Conventional * 12 months | -0.08 | 0.18 | -0.44 | 300 | .658 |

**Appendix G**

Gaze Coordination at 11 & 12 months (Appendix to Chapter 5 and 6)

All infants gazed to their caregiver's face and all produced both CV and non-CV vocalizations. Fewer infants produced gestures (either alone or in a gesture-vocal combination), with 85% (*n* = 52) producing one or more gestures. Most commonly, infants produced give gestures (produced by 64% of infants, *n* = 39), but a number also produced show gestures (39%, *n* = 24), index finger pointing (39%, *n* = 24), open hand pointing (30%, *n* = 18) and conventional gestures (34%, *n* = 21). Sixty-two percent (*n* = 38) produced combinations.

Descriptive statistics for infant behaviours at 11 & 12 months are presented in Table 72.

Table 72

*Average Frequency of Vocalisations, Gestures and Gesture-vocal combinations from 20 minutes Observation at 11 & 12 months (n = 61)*

| *Behaviour* | *M* | *SD* | *Median* | *Range* |
|---|---|---|---|---|
| **Gaze to Caregiver's Face** | 44.87 | 21.80 | 40 | 5-99 |
| **Vocalisations** | 85.84 | 41.41 | 82 | 17-197 |
| CV | 33.82 | 25.66 | 29 | 2-125 |
| Non-CV | 52.02 | 27.05 | 48 | 11-145 |
| **Gestures** | 4.41 | 5.11 | 2 | 0-22 |
| Index-finger point | 0.77 | 1.79 | 0 | 0-9 |
| Open-hand point | 0.26 | 0.74 | 0 | 0-4 |
| Give | 2.41 | 3.23 | 1 | 0-13 |
| Show | 0.49 | 0.92 | 0 | 0-4 |
| Conventional gesture | 0.48 | 1.13 | 0 | 0-7 |
| **Gesture-vocal combinations** | 2.70 | 3.48 | 1 | 0-14 |
| CV | 1.43 | 2.31 | 0 | 0-12 |
| Non-CV | 1.27 | 2.22 | 0 | 0-13 |
| Index-finger point | 0.67 | 1.94 | 0 | 0-12 |
| Open-hand point | 0.46 | 1.43 | 0 | 0-10 |
| Give | 1.05 | 2.03 | 0 | 0-12 |
| Show | 0.38 | 0.78 | 0 | 0-3 |
| Conventional gesture | 0.15 | 0.40 | 0 | 0-2 |
| Index-finger point & CV | 0.51 | 1.76 | 0 | 0-12 |
| Index-finger point & Non-CV | 0.16 | 0.52 | 0 | 0-3 |
| Open-hand point & CV | 0.25 | 0.60 | 0 | 0-2 |
| Open-hand point & Non-CV | 0.21 | 1.07 | 0 | 0-8 |
| Give & CV | 0.43 | 0.83 | 0 | 0-4 |
| Give & Non-CV | 0.61 | 1.62 | 0 | 0-11 |
| Show & CV | 0.16 | 0.52 | 0 | 0-3 |

| | | | | |
|---|---|---|---|---|
| Show & Non-CV | 0.21 | 0.52 | 0 | 0-2 |
| Conventional gesture & CV | 0.08 | 0.28 | 0 | 0-1 |
| Conventional gesture & Non-CV | 0.07 | 0.25 | 0 | 0-1 |

**Do Infants' Vocalisations, Gestures and Combinations Co-occur with Gaze to Caregiver's Face Above the Rate Expected by Chance?**

Figure 24 shows the expected and observed duration that vocalisations, gestures and combinations co-occurred with gaze to caregiver's face at 11 & 12 months.

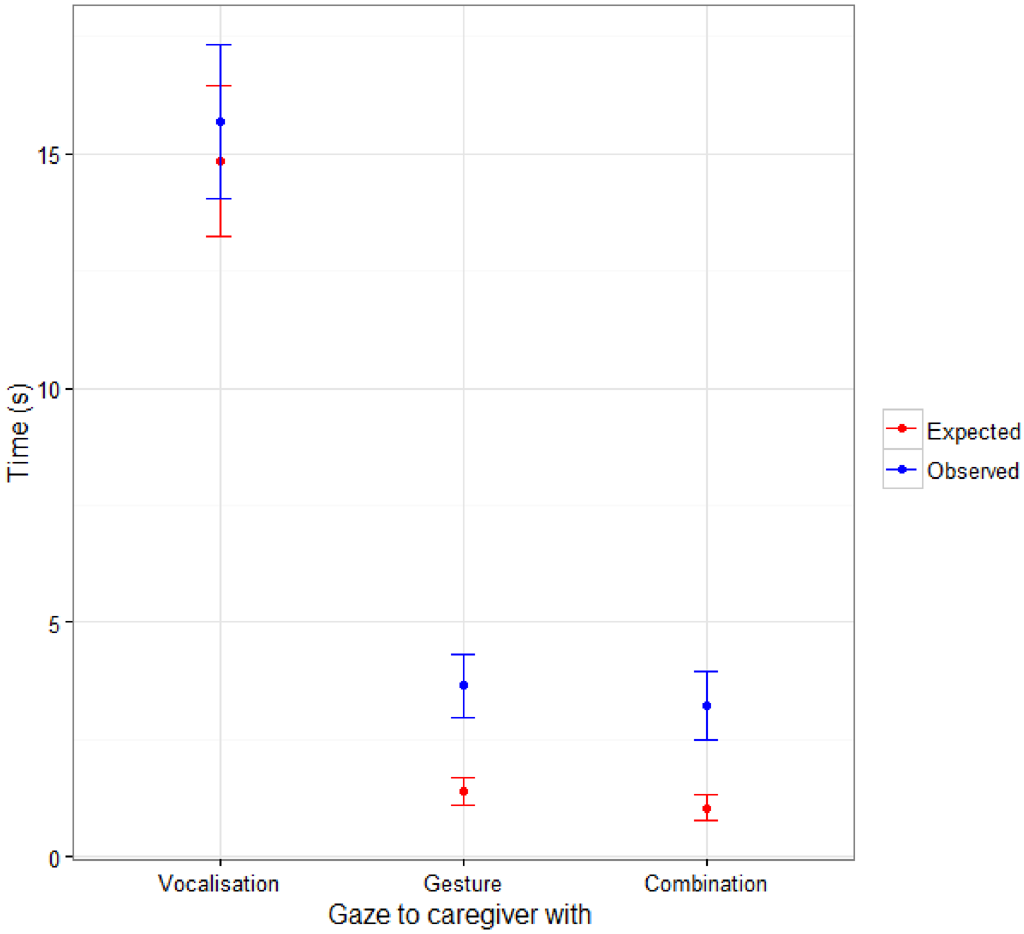


*Figure 24*. Mean expected and observed co-occurrence of vocalisations, gestures and gesture-vocal combinations with gaze to caregiver's face at 11 & 12 months ($n$ = 61). Error bars represent standard error about the mean.

Dependent t-tests revealed that gestures and combinations co-occurred with gaze significantly above the level predicted by chance (gestures, $t(60) = 4.87$, $p < .001$, combinations, $t(60) = 4.18$, $p < .001$), however vocalisations did not.

Figure 25 shows the expected and observed duration that individual vocalisation and gesture types (both produced alone and as part of combinations) co-occurred with gaze to caregiver's face at 11 & 12 months.

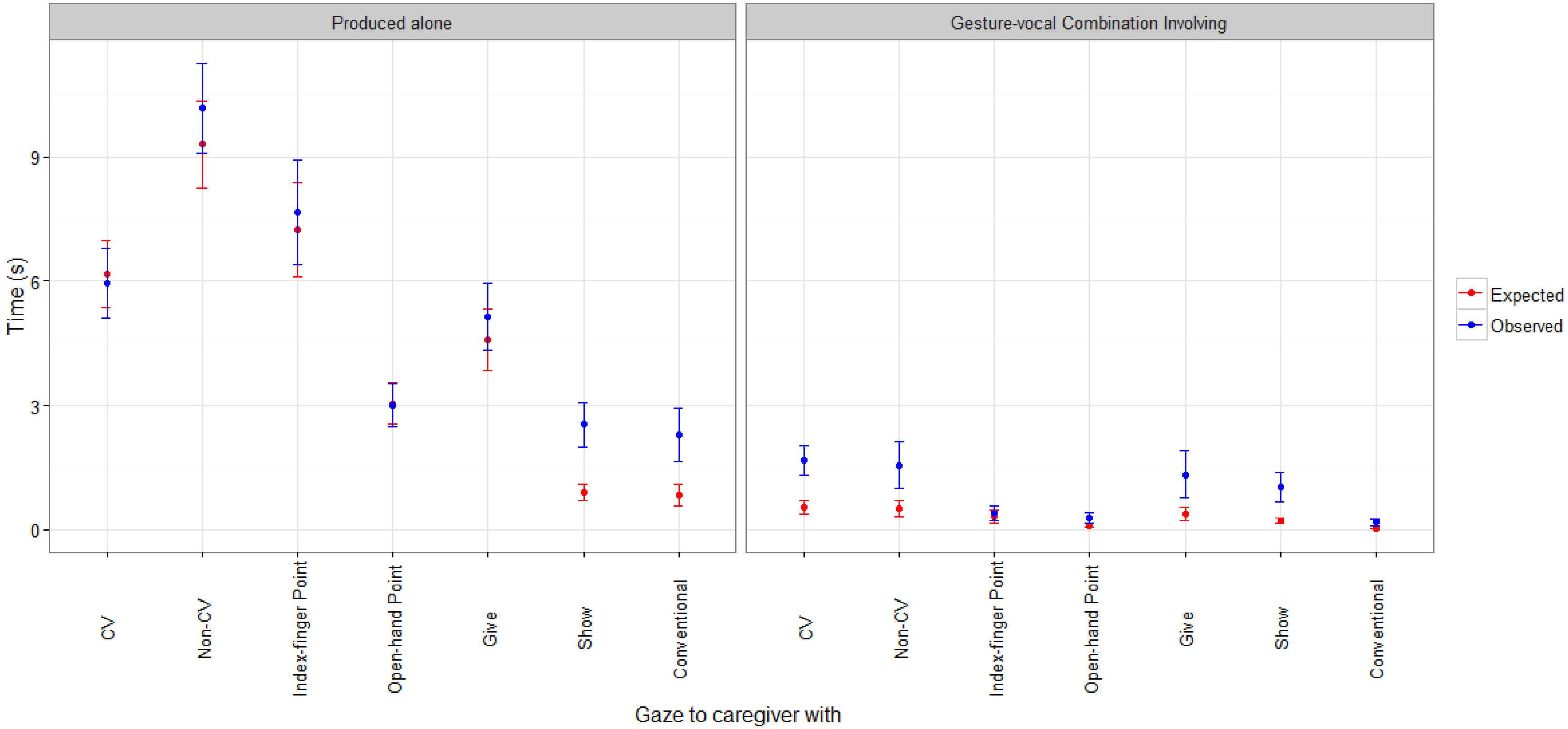


*Figure 25.* Mean expected and observed co-occurrence of individual vocalisation and gesture types (both produced alone and as part of gesture-vocal combinations) co-occurred with gaze to caregiver's face at 11 & 12 months ($n$ = 61). Error bars represent standard error about the mean.

**Vocalisation type.** Neither CV nor non-CV vocalisations co-occurred with gaze significantly above the level predicted by chance. However, combinations involving CV and non-CV vocalisations co-occurred with gaze above the level predicted by chance (combinations involving CV vocalisation(s), $t(60) = 3.80$, $p < .001$; involving only non-CV vocalisations, $t(60) = 2.72$, $p = .008$).

**Gesture type.** Only showing and conventional gestures co-occurred with gaze significantly above the level predicted by chance (showing, $t(60) = 4.30$, $p < .001$; conventional gestures, $t(60) = 3.35$, $p = .001$). Index-finger pointing, open-hand pointing and giving did not co-occur with gaze above the level predicted by chance.

A different pattern of results occurred for the type of gestures included in combinations. Combinations involving all types of gestures co-occurred with gaze above the level predicted by chance (combinations involving index-finger pointing, $t(60) = 5.96$, $p < .001$; open-hand pointing, $t(60) = 5.48$, $p < .001$; giving, $t(60) = 5.99$, $p < .001$; showing, $t(60) = 4.72$, $p < .001$; conventional gestures, $t(60) = 3.56$, $p = .001$).

**Do Infants Appear to Intentionally Co-ordinate their Vocalisations and Gestures as Combinations or Does this Occur by Chance?**

Dependent t-tests revealed that vocalisations and gestures co-occurred significantly above the level predicted by chance, $t(60) = 4.87$, $p < .001$.

**Vocalisation type.** Both types of vocalisations co-occurred with gestures significantly above the level predicted by chance (CV, $t(60) = 4.25$, $p < .001$; non-CV, $t(60) = 3.28$, $p = .002$).

**Gesture type.** All gestures (with the exception of conventional gestures) co-occurred with vocalisations significantly above the level predicted by chance (index finger pointing, $t(60) = 2.41$, $p = .019$; open hand pointing, $t(60) = 2.02$, $p = .048$; giving, $t(60) = 2.61$, $p = .011$; showing, $t(60) = 2.65$, $p = .010$).

**Vocalisation and gesture type.** CV vocalisations with index-finger pointing, open hand pointing, and giving, and non-CV vocalisations with and showing with giving co-occurred significantly above the level predicted by chance (index-finger pointing and CV, $t(60) = 2.23$, $p = .030$; open hand pointing and CV, $t(60) = 2.23$, $p = .030$; giving and CV, $t(60) = 2.52$, $p = .015$; giving and non-CV, $t(60) = 2.06$, $p = .044$), all other combinations did not co-occur above the level predicted by chance.

## Is Intentional Communication More Likely to be Vocal, Gestural or Part of a Combination?

Our second question concerned whether gaze coordination, taken as an index of intentional communication, was more likely to occur with vocalisations, gestures and combinations. Table 73 shows descriptive statistics including the frequency and proportion of vocalisations, gestures and combinations with gaze coordination.

Table 73

*Average Frequency and Proportion of Vocalisations, Gestures and Gesture-vocal Combinations With Gaze Coordination at 11 &12 months (n = 61)*

| | *Frequency* | | | | *Proportion* |
|---|---|---|---|---|---|
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 15.10 | 10.56 | 13 | 0-51 | .18 |
| CV | 5.56 | 5.79 | 4 | 0-30 | .17 |
| Non-CV | 9.54 | 6.69 | 9 | 0-26 | .19 |
| **Gestures** | 2.15 | 2.71 | 1 | 0-12 | .51 |
| Index-finger point | 0.21 | 0.76 | 0 | 0-5 | .28 |
| Open-hand point | 0.11 | 0.47 | 0 | 0-3 | .39 |
| Give | 1.14 | 1.92 | 0 | 0-8 | .45 |
| Show | 0.43 | 0.87 | 0 | 0-4 | .85 |
| Conventional gesture | 0.26 | 0.60 | 0 | 0-3 | .62 |
| **Gesture-vocal combinations** | 1.68 | 2.46 | 1 | 0-12 | .57 |
| CV | 0.89 | 1.36 | 0 | 0-6 | .66 |
| Non-CV | 0.79 | 1.79 | 0 | 0-11 | .51 |
| Index-finger point | 0.31 | 0.96 | 0 | 0-6 | .43 |
| Open-hand point | 0.25 | 0.91 | 0 | 0-6 | .42 |
| Give | 0.64 | 1.56 | 0 | 0-10 | .59 |
| Show | 0.38 | 0.78 | 0 | 0-3 | 1.00 |
| Conventional gesture | 0.10 | 0.30 | 0 | 0-1 | .69 |
| Index-finger point & CV | 0.25 | 0.89 | 0 | 0-6 | .47 |
| Index-finger point & Non-CV | 0.07 | 0.31 | 0 | 0-2 | .43 |
| Open-hand point & CV | 0.11 | 0.37 | 0 | 0-2 | .40 |
| Open-hand point & Non-CV | 0.13 | 0.67 | 0 | 0-5 | .73 |
| Give & CV | 0.30 | 0.62 | 0 | 0-3 | .74 |
| Give & Non-CV | 0.34 | 1.25 | 0 | 0-9 | .45 |
| Show & CV | 0.16 | 0.52 | 0 | 0-3 | 1.00 |

| | | | | | |
|---|---|---|---|---|---|
| Show & Non-CV | 0.21 | 0.52 | 0 | 0-2 | 1.00 |
| Conventional gesture & CV | 0.07 | 0.25 | 0 | 0-1 | .80 |
| Conventional gesture & Non-CV | 0.03 | 0.18 | 0 | 0-1 | .50 |

As predicted, vocalisations were more frequently gaze-coordinated than gestures and combinations. However, a far higher *proportion* of gestures and combinations were gaze-coordinated, compared to vocalisations. The following analyses focus on whether these predicted differences are significant, and also explores whether different vocalisation or gesture types are more frequently, or proportionally more likely to be gaze-coordinated.

**Individual behaviours.** We investigated whether vocalisations, gestures or combinations were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** There was a significant improvement in model fit when behaviour type was added as a predictor, $\chi^2$ (2) = 125.56, $p$ < .001 (Table 74). Planned contrasts revealed that a significantly higher frequency of vocalisations were gaze-coordinated than both gestures and combinations, $b$ = 6.59, $t(120)$ = 14.15, $p$ < .001.

Table 74

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations with Gaze Coordination by Behaviour Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | 15.10 | 0.83 | 18.28 | .000 |
| Gesture | -12.95 | 1.08 | -12.03 | .000 |
| Gesture-vocal Combination | -13.42 | 1.08 | -12.47 | .000 |

***Proportion.*** There was a significant improvement in model fit when behaviour type was added as a predictor, $\chi^2$ (2) = 61.31, $p$ < .001 (Table 75).

Table 75

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 5200), *Gestures* (*n* = 266) *and Gesture-vocal Combinations* (*n* = 164) *With Gaze Coordination (1= Gaze-coordinated, 0 = Not) by Behaviour Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Vocalisation*) | -1.67 | 0.10 | -16.53 | .000 |
| Gesture | 2.14 | 0.22 | 9.91 | .000 |
| Gesture-vocal Combination | 1.65 | 0.19 | 8.65 | .000 |

LLRI = .01, C = .73, Dxy = .46.

Planned contrasts revealed that a significantly lower proportion of infants' vocalisations were gaze-coordinated than their gestures and combinations, *b* = -0.95, *z* = -0.08, *p* < .001.

**Vocalisation type**. We investigated whether CV or non-CV vocalisations (whether produced alone or part of a combination) were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** Regarding vocalisations produced alone, dependent t-tests revealed that a significantly higher frequency of non-CV vocalisations were gaze-coordinated than CV vocalisations, *t*(60) = 4.63, *p* < .001. Gaze coordination was not significantly more or less frequent with combinations involving either vocalisation type.

***Proportion.*** When a model was fitted to vocalisations (produced alone), adding vocalisation type did not improve model fit. When similar models were fitted to datasets including combinations, adding vocalisation type again did not improve model fit.

**Gesture type.** We investigated whether index-finger pointing, open-hand pointing, giving, showing or conventional gestures (whether produced alone or part of a combination) were more frequently, or proportionally more likely to be gaze-coordinated.

***Frequency.*** There was a significant improvement in model fit when gesture type was added as a predictor, $\chi^2$ (4) = 37.87, *p* < .001 (Table 76).

Table 76

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) with Gaze Coordination by Gesture Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.21 | 0.14 | 1.58 | .116 |
| Open-hand Pointing | -0.10 | 0.18 | -0.55 | .584 |
| Give | 0.93 | 0.18 | 5.04 | .000 |
| Show | 0.21 | 0.18 | 1.16 | .247 |
| Conventional | 0.05 | 0.18 | 0.27 | .789 |

Post-hoc pairwise tests revealed that a significantly higher frequency of giving was gaze-coordinated than all other gestures (Table 77).

Table 77

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types (Produced Alone) with Gaze Coordination at 11 & 12 months (n = 61).*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Index-finger point : Open-hand point | 0.10 | 240 | 0.55 | .982 |
| Index-finger point : Give | -0.93 | 240 | -5.04 | .000 |
| Index-finger point : Show | -0.21 | 240 | -1.16 | .774 |
| Index-finger point : Conventional | -0.05 | 240 | -0.27 | .999 |
| Open-hand point : Give | -1.03 | 240 | -5.59 | .000 |
| Open-hand point : Show | -0.31 | 240 | -1.71 | .430 |
| Open-hand point : Conventional | -0.15 | 240 | -0.82 | .925 |
| Give : Show | 0.71 | 240 | 3.88 | .001 |
| Give : Conventional | 0.88 | 240 | 4.78 | .000 |
| Show : Conventional | 0.16 | 240 | 0.89 | .899 |

We constructed a similar model to determine if there were significant differences in the frequency that combinations with different types of gestures were gaze-coordinated. There was a significant improvement in model fit when gesture type was added as a predictor, $\chi^2$ (4) = 10.72, *p* = .030 (Table 78).

Table 78

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gesture-vocal Combinations with Gaze Coordination by Gesture Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger pointing*) | 0.31 | 0.13 | 2.46 | .015 |
| Open-hand Pointing | -0.07 | 0.17 | -0.38 | .706 |
| Give | 0.33 | 0.17 | 1.92 | .056 |
| Show | 0.07 | 0.17 | 0.38 | .705 |
| Conventional | -0.21 | 0.17 | -1.23 | .221 |

Post-hoc tests revealed that, a significantly higher frequency of combinations involving giving were gaze-coordinated than those involving conventional gestures (Table 79).

Table 79

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gesture Types (Produced as part of Gesture-vocal Combinations) with Gaze Coordination at 11 & 12 months (n = 61).*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Index-finger point : Open-hand point | 0.07 | 240 | 0.38 | .996 |
| Index-finger point : Give | -0.33 | 240 | -1.92 | .309 |
| Index-finger point : Show | -0.07 | 240 | -0.38 | .996 |
| Index-finger point : Conventional | 0.21 | 240 | 1.23 | .735 |
| Open-hand point : Give | -0.40 | 240 | -2.30 | .149 |
| Open-hand point : Show | -0.13 | 240 | -0.76 | .942 |
| Open-hand point : Conventional | 0.15 | 240 | 0.85 | .915 |
| Give : Show | 0.27 | 240 | 1.54 | .536 |
| Give : Conventional | 0.55 | 240 | 3.15 | .016 |
| Show : Conventional | 0.28 | 240 | 1.61 | .494 |

***Proportion.*** Unlike in the main analyses, when we collapsed across age, there were enough gestures (though not enough gesture-vocal combinations) to conduct an analyses to determine if higher proportions of any type of gesture were gaze-coordinated. There was a significant improvement in model fit when gesture type was added as a predictor, $\chi^2$ (4) = 22.07, $p < .001$ (Table 80). Post-hoc pairwise tests revealed that no gestures were proportionally more or less gaze-coordinated than others.

Table 80

*Summary of Fixed Effects from a Logistic Regression Model Fitting Index-finger Pointing (n = 47), Open-hand Pointing (n = 15), Giving (n = 145), Showing (n = 30) and Conventional Gestures (n = 29) (All Produced Alone) with Gaze Coordination (1= Gaze-coordinated, 0 = Not) by Gesture Type at 11 & 12 months (n = 51)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Point*) | -1.54 | 0.91 | -1.70 | .089 |
| Open-hand Point | -0.47 | 5.38 | -0.09 | .931 |
| Give | 1.36 | 0.95 | 1.43 | .152 |
| Show | 7.12 | 5.01 | 1.42 | .155 |
| Conventional | 2.08 | 1.08 | 1.92 | .054 |

LLRI = .07, C = .92, Dxy = .84.

## Appendix H

Descriptive Statistics (Appendix to Chapter 6)

The following tables show the average frequency with which vocalisations, gestures and combinations were produced by infants, and the frequency and proportion with which these were gaze-coordinated at 11 & 12 months for infants whom we had 15-month (table 84) and 24-month (Table 85) expressive vocabulary measures.

Table 81

*Average Frequency of Vocalisations, Gestures and Gesture-vocal Combinations Produced and the Frequency and Proportion Produced With Gaze Coordination at 11 & 12 months with Language Outcomes at 15 months (n = 53)*

| | *Produced* | | | | | | *Produced With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *By Infants* | | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *N* | *%* | *M* | *SD* | *Med* | *Range* | *M* |
| **Gaze to Caregiver's Face** | 44.38 | 23.09 | 39 | 5-99 | 53 | 100% | - | - | - | - | - |
| **Vocalisations** | 88.31 | 39.19 | 84 | 17-197 | 53 | 100% | 15.82 | 10.99 | 14 | 0-51 | .18 |
| CV | 35.66 | 26.33 | 30 | 2-125 | 53 | 100% | 5.98 | 6.05 | 4 | 0-30 | .17 |
| Non-CV | 52.64 | 22.80 | 52 | 12-120 | 53 | 100% | 9.84 | 6.93 | 9 | 0-26 | .19 |
| **Gestures** | 4.75 | 5.35 | 2 | 0-22 | 44 | 83% | 2.33 | 2.84 | 2 | 0-12 | .53 |
| Index-finger point† | 0.89 | 1.90 | 0 | 0-9 | 18 | 34% | 0.25 | 0.81 | 0 | 0-5 | .28 |
| Open-hand point** | 0.30 | 0.79 | 0 | 0-4 | 9 | 17% | 0.13 | 0.50 | 0 | 0-3 | .39 |
| Give | 2.55 | 3.39 | 1 | 0-13 | 31 | 58% | 1.24 | 2.02 | 0 | 0-8 | .49 |
| Show† | 0.51 | 0.97 | 0 | 0-4 | 16 | 30% | 0.45 | 0.91 | 0 | 0-4 | .89 |
| Conventional gesture† | 0.51 | 1.20 | 0 | 0-7 | 13 | 25% | 0.26 | 0.62 | 0 | 0-3 | .56 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Gesture-vocal combinations** | 2.60 | 3.36 | 1 | 0-14 | 33 | 62% | 1.61 | 2.37 | 1 | 0-12 | .57 |
| CV† | 1.38 | 2.27 | 0 | 0-12 | 26 | 49% | 0.86 | 1.28 | 0 | 0-6 | .67 |
| Non-CV† | 1.22 | 2.22 | 0 | 0-13 | 25 | 47% | 0.76 | 1.80 | 0 | 0-11 | .51 |
| Index-finger point† | 0.58 | 1.80 | 0 | 0-12 | 12 | 23% | 0.30 | 0.99 | 0 | 0-6 | .44 |
| Open-hand point** | 0.47 | 1.51 | 0 | 0-10 | 9 | 17% | 0.25 | 0.94 | 0 | 0-6 | .40 |
| Give† | 1.02 | 1.98 | 0 | 0-12 | 22 | 42% | 0.59 | 1.49 | 0 | 0-10 | .56 |
| Show† | 0.38 | 0.79 | 0 | 0-3 | 12 | 23% | 0.38 | 0.79 | 0 | 0-3 | 1.00 |
| Conventional gesture** | 0.15 | 0.41 | 0 | 0-2 | 7 | 13% | 0.09 | 0.30 | 0 | 0-1 | .64 |
| Index-finger point & CV** | 0.47 | 1.77 | 0 | 0-12 | 8 | 15% | 0.25 | 0.92 | 0 | 0-6 | .53 |
| Index-finger point & Non-CV* | 0.11 | 0.38 | 0 | 0-2 | 5 | 9% | 0.06 | 0.30 | 0 | 0-2 | .40 |
| Open-hand point & CV** | 0.25 | 0.62 | 0 | 0-2 | 8 | 15% | 0.11 | 0.38 | 0 | 0-2 | .38 |
| Open-hand point & Non-CV* | 0.23 | 1.14 | 0 | 0-8 | 4 | 8% | 0.13 | 0.71 | 0 | 0-5 | .66 |
| Give & CV† | 0.42 | 0.80 | 0 | 0-4 | 15 | 28% | 0.27 | 0.53 | 0 | 0-2 | .70 |
| Give & Non-CV† | 0.59 | 1.65 | 0 | 0-11 | 14 | 26% | 0.32 | 1.28 | 0 | 0-9 | .43 |
| Show & CV** | 0.17 | 0.55 | 0 | 0-3 | 6 | 11% | 0.17 | 0.55 | 0 | 0-3 | 1.00 |
| Show & Non-CV** | 0.21 | 0.49 | 0 | 0-2 | 9 | 17% | 0.21 | 0.49 | 0 | 0-2 | 1.00 |
| Conventional gesture & CV* | 0.08 | 0.27 | 0 | 0-1 | 4 | 8% | 0.06 | 0.23 | 0 | 0-1 | .75 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Conventional gesture & Non-CV* | 0.08 | 0.27 | 0 | 0-1 | 4 | 8% | 0.04 | 0.19 | 0 | 0-1 | .50 |

* denotes behaviour not included as a predictor in any models (fewer than 10% of infants producing)

** denotes behaviour not included as a predictor in proportional models (fewer than 10 infants producing)

† denotes behaviour not included as a predictor in proportional models (fewer than 50% producing the behaviour more than twice)

Table 82

*Average Frequency of Vocalisations, Gestures and Gesture-vocal Combinations Produced and the Frequency and Proportion Produced With Gaze Coordination at 11 & 12 months with Language Outcomes at 24 months (n = 49)*

| | *Produced* | | | | | | *Produced With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *By Infants* | | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *N* | *%* | *M* | *SD* | *Med* | *Range* | *M* |
| **Gaze to Caregiver's Face** | 44.41 | 22.38 | 40 | 5-99 | 49 | 100% | - | - | - | - | - |
| **Vocalisations** | 87.80 | 39.81 | 83 | 17-197 | 49 | 100% | 14.54 | 9.32 | 13 | 0-38 | .17 |
| CV | 35.96 | 26.63 | 31 | 2-125 | 49 | 100% | 5.37 | 5.03 | 4 | 0-24 | .15 |
| Non-CV | 51.84 | 25.29 | 48 | 12-145 | 49 | 100% | 9.18 | 6.30 | 9 | 0-24 | .18 |
| **Gestures** | 4.96 | 5.37 | 3 | 0-22 | 43 | 88% | 2.33 | 2.80 | 2 | 0-12 | .49 |
| Index-finger point† | 0.88 | 1.94 | 0 | 0-9 | 16 | 33% | 0.27 | 0.84 | 0 | 0-5 | .32 |
| Open-hand point** | 0.33 | 0.82 | 0 | 0-4 | 9 | 18% | 0.14 | 0.52 | 0 | 0-3 | .39 |
| Give | 2.69 | 3.37 | 1 | 0-13 | 32 | 65% | 1.17 | 1.84 | 0 | 0-8 | .42 |
| Show† | 0.51 | 0.98 | 0 | 0-4 | 15 | 31% | 0.45 | 0.91 | 0 | 0-4 | .88 |
| Conventional gesture† | 0.55 | 1.23 | 0 | 0-7 | 14 | 29% | 0.31 | 0.65 | 0 | 0-3 | .63 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Gesture-vocal combinations** | 2.87 | 3.34 | 2 | 0-13 | 34 | 69% | 1.68 | 2.19 | 1 | 0-9 | .54 |
| CV† | 1.66 | 2.48 | 1 | 0-12 | 27 | 55% | 1.01 | 1.43 | 0 | 0-6 | .64 |
| Non-CV† | 1.21 | 1.71 | 0 | 0-8 | 24 | 49% | 0.67 | 1.27 | 0 | 0-5 | .48 |
| Index-finger point† | 0.80 | 2.13 | 0 | 0-12 | 14 | 29% | 0.35 | 1.03 | 0 | 0-6 | .39 |
| Open-hand point† | 0.57 | 1.58 | 0 | 0-10 | 11 | 22% | 0.31 | 1.00 | 0 | 0-6 | .42 |
| Give† | 1.00 | 1.57 | 0 | 0-7 | 21 | 43% | 0.56 | 1.07 | 0 | 0-6 | .56 |
| Show† | 0.37 | 0.73 | 0 | 0-3 | 12 | 24% | 0.37 | 0.73 | 0 | 0-3 | 1.00 |
| Conventional gesture** | 0.14 | 0.35 | 0 | 0-1 | 7 | 14% | 0.10 | 0.31 | 0 | 0-1 | .71 |
| Index-finger point & CV† | 0.63 | 1.94 | 0 | 0-12 | 10 | 20% | 0.31 | 0.98 | 0 | 0-6 | .47 |
| Index-finger point & Non-CV** | 0.16 | 0.51 | 0 | 0-3 | 6 | 12% | 0.04 | 0.20 | 0 | 0-1 | .33 |
| Open-hand point & CV† | 0.31 | 0.65 | 0 | 0-2 | 10 | 20% | 0.14 | 0.41 | 0 | 0-2 | .40 |
| Open-hand point & Non-CV** | 0.27 | 1.19 | 0 | 0-8 | 5 | 10% | 0.16 | 0.75 | 0 | 0-5 | .73 |
| Give & CV† | 0.50 | 0.90 | 0 | 0-4 | 15 | 31% | 0.33 | 0.67 | 0 | 0-3 | .70 |
| Give & Non-CV† | 0.50 | 0.98 | 0 | 0-4 | 13 | 27% | 0.22 | 0.59 | 0 | 0-3 | .41 |
| Show & CV** | 0.14 | 0.41 | 0 | 0-2 | 6 | 12% | 0.14 | 0.41 | 0 | 0-2 | 1.00 |
| Show & Non-CV** | 0.22 | 0.51 | 0 | 0-2 | 9 | 18% | 0.22 | 0.51 | 0 | 0-2 | 1.00 |
| Conventional gesture & CV* | 0.08 | 0.28 | 0 | 0-1 | 4 | 8% | 0.08 | 0.28 | 0 | 0-1 | 1.00 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Conventional gesture & Non-CV* | 0.06 | 0.24 | 0 | 0-1 | 3 | 6% | 0.02 | 0.14 | 0 | 0-1 | .33 |

* denotes behaviour not included as a predictor in any models (fewer than 10% of infants producing)

** denotes behaviour not included as a predictor in proportional models (fewer than 10 infants producing)

†denotes behaviour not included as a predictor in proportional models (fewer than 50% producing the behaviour more than twice)

## Appendix I

Inclusive Analyses (Appendix to Chapter 6)

We focus first on whether the *raw frequency* of behaviours (e.g. vocalisations, gestures, and individual vocalisation and gesture types), and the frequency of these behaviours *with gaze coordination* at 11 & 12 months predict expressive vocabulary at 15 months. We then focus on whether the *proportion* of behaviours with gaze coordination at 11 & 12 months predict expressive vocabulary at 15 months. Finally, all analyses are then repeated to with expressive vocabulary at 24 months as the outcome. For the following analyses, we did not treat gesture-vocal combinations as a separate, mutually exclusive category from gesture and vocalisations.

### Predicting Expressive Vocabulary at 15 Months

Table 83 shows the average frequency of infant behaviours at 11 & 12 months for the infants for whom we had language outcomes at 15 months.

Table 83

*Average Frequency of Vocalisations and Gestures Produced and the Frequency and Proportion Produced With Gaze Coordination at 11 & 12 months with Language Outcomes at 15 months (n = 53)*

| | *Produced* | | | | | | *Produced With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *By Infants* | | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *N* | *%* | *M* | *SD* | *Med* | *Range* | *M* |
| **Gaze to Caregiver's Face** | 44.38 | 23.09 | 39 | 5-99 | 53 | 100% | - | - | - | - | - |
| **Vocalisations** | 91.36 | 40.54 | 86 | 17-203 | 53 | 100% | 17.55 | 11.87 | 16 | 0-51 | 0.19 |
| CV | 37.23 | 27.37 | 30 | 2-129 | 53 | 100% | 6.89 | 6.48 | 5 | 0-30 | 0.18 |
| Non-CV | 54.13 | 23.78 | 53 | 12-135 | 53 | 100% | 10.66 | 7.69 | 9 | 0-32 | 0.20 |
| **Gestures** | 7.35 | 7.08 | 5 | 0-26 | 45 | 85% | 3.90 | 4.33 | 3 | 0-20 | 0.53 |
| Index-finger point | 1.47 | 2.86 | 0 | 0-15 | 21 | 40% | 0.53 | 1.40 | 0 | 0-8 | 0.35 |
| Open-hand point† | 0.77 | 1.71 | 0 | 0-10 | 16 | 30% | 0.36 | 1.03 | 0 | 0-6 | 0.38 |
| Give | 3.56 | 4.72 | 1 | 0-20 | 34 | 64% | 1.83 | 3.13 | 0 | 0-18 | 0.48 |
| Show† | 0.89 | 1.58 | 0 | 0-7 | 19 | 36% | 0.83 | 1.53 | 0 | 0-7 | 0.92 |
| Conventional gesture† | 0.66 | 1.22 | 0 | 0-7 | 19 | 36% | 0.36 | 0.65 | 0 | 0-3 | 0.60 |

† denotes behaviour not included as a predictor in proportional models (fewer than 50% producing more than 2 behaviours)

**Raw frequency.** We first explored whether raw frequencies of behaviours at 11 & 12 months (Table 83) predicted expressive vocabulary at 15 months.

We initially tested whether the frequency of gaze to caregiver's face, vocalisations or gestures individually predicted 15-month expressive vocabulary. Only the frequency of gaze to caregiver's face improved on the null model, $F(1,51) = 6.90$, $p = .011$, $R^2 = .12$.

We then explored whether the frequency of specific types of vocalisation or gestures predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on a null model with no predictors.

***Vocalisation type.*** For models with the frequency of CV or non-CV vocalisations as predictors, only the frequency of CV vocalisations improved on the null model, $F(1,51) = 9.91$, $p = .003$, $R^2 = .16$.

***Gesture type.*** For models with the frequency of specific gesture types as predictors, only the frequency of index-finger pointing improved on the null model, $F(1,51) = 5.63$, $p = .021$, $R^2 = .10$.

***Final model.*** We next tested whether the three predictors (frequency of gaze to caregiver's face, CV vocalisations and index-finger pointing) explained separate variance. We systematically tested for improvement in fit, combining the three predictors. The final model is presented in Table 84.

Table 84

*Regression Model fitting Frequency of CV Vocalisations and Index-finger Pointing at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 1.54 | 5.39 | 0.29 | .777 |
| Frequency of CV Vocalisations | 0.39 | 0.11 | 3.53 | .001 |
| Frequency of Index-finger Pointing | 3.01 | 1.06 | 2.84 | .006 |

$R^2 = .28$, $F(2,50) = 9.68$, $p < .001$.

**Frequency of behaviours with gaze coordination.** We next investigated whether the frequency of behaviours *with gaze coordination* at 11 & 12 months predicted expressive vocabulary at 15 months.

We initially tested whether the frequency of vocalisations or gestures with gaze coordination individually predicted 15-month expressive vocabulary. Only the frequency of vocalisations with gaze coordination improved on the null model, $F(1,51) = 10.81$, $p = .002$, $R^2 = .18$.

We then explored whether the frequency of specific types of vocalisation or gesture with gaze coordination predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** For models with the frequency of CV or non-CV vocalisations with gaze coordination as predictors, only the frequency of CV vocalisations with gaze coordination improved on the null model, $F(1,51) = 23.45$, $p < .001$, $R^2 = .32$.

***Gesture type.*** For models with the frequency of specific gesture types with gaze coordination as predictors, only the frequency of index-finger pointing with gaze coordination improved on the null model, $F(1,51) = 4.11$, $p = .048$, $R^2 = .07$.

***Final model.*** We next tested whether the two compatible predictors (frequency of CV vocalisations and index-finger pointing with gaze coordination) explained separate variance. The final model is presented in Table 85.

Table 85

*Regression Model fitting Frequency of CV Vocalisations and Index-finger Pointing with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 5.24 | 4.24 | 1.24 | .222 |
| Frequency of CV Vocalisations with Gaze Coordination | 2.04 | 0.47 | 4.36 | .000 |
| Frequency of Index-finger Pointing with Gaze Coordination | 2.24 | 2.18 | 1.03 | .308 |

$R^2 = .33$, $F(2,50) = 12.27$, $p < .001$.

**Proportion of behaviours with gaze coordination.** We next investigated whether the proportion of behaviours that were gaze-coordinated at 11 & 12 months (Table 83) predicted expressive language at 15 months.

We initially tested whether the proportion of vocalisations or gestures with gaze coordination individually predicted 15-month expressive vocabulary (on subsets where all participants had produced each type of behaviour– see Method). Only the proportion of vocalisations with gaze coordination improved on the null model $F(1,51) = 6.39$, $p = .015$, $R^2 = .11$.

We then explored whether the proportion of specific types of vocalisation or gestures with gaze coordination predicted 15-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** For models with the proportion of CV or non-CV vocalisations with gaze coordination as predictors, only the proportion of non-CV vocalisations with gaze coordination improved on the null model, $F(1,51) = 4.28$, $p = .044$, $R^2 = .08$.

***Gesture type.*** When we tested the proportion of index-finger pointing or giving with gaze coordination as a predictor (as these were the only gesture types it was possible to test – see Method), it did not improve on the null model.

***Final model.*** The final model consists only of the proportion of vocalisations with gaze coordination, and is presented in Table 86.

Table 86

*Regression Model fitting Proportion of Vocalisations with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 6.09 | 6.58 | 0.93 | .359 |
| Proportion of Vocalisations with Gaze Coordination | 76.06 | 30.08 | 2.53 | .015 |

$R^2 = .11$, $F(1,51) = 6.39$, $p = .015$.

**Does Gaze-coordinated Behaviour Better Predict Expressive Vocabulary at 15 Months than the Raw Frequency of Behaviours?**

We wanted to determine whether gaze-coordinated behaviours were better predictors of later expressive vocabulary than the frequency with which they were produced regardless of gaze coordination, as this would follow if they were early instances of intentional communication. A number of gaze-coordinated behaviours predicted expressive vocabulary at 15 months (as explored above), and so we consider this question for these behaviours. Table 87 shows the variance explained by predictors relating to gestures and 15-month expressive vocabulary, as demonstrated by both $R^2$ and AIC.

Table 87

*Summary of Regression Models ($R^2$ and AIC) fitting Predictors relating to Vocalisations and CV Vocalisations at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| *Predictor(s)* | $R^2$ | *AIC* |
|---|---|---|
| ***Predictors relating to Vocalisations*** | | |
| Frequency of Vocalisations (*non-significant predictor*) | .03 | 495.39 |
| Frequency of Vocalisations with Gaze Coordination | .17 | 487.01 |
| Frequency of Vocalisations + Frequency of Gaze | .13 | 491.92 |
| ***Predictors relating to CV Vocalisations*** | | |
| Frequency of CV Vocalisations | .16 | 487.78 |
| Frequency of CV Vocalisations with Gaze Coordination | .32 | 477.14 |
| Frequency of CV Vocalisations + Frequency of Gaze | .24 | 484.57 |

Table 87 shows that the frequency of CV vocalisations with gaze coordination was a better predictor of expressive vocabulary at 15 months than the frequency of CV vocalisations (as denoted by $R^2$) and also was a better predictor than a model with both the frequency of CV vocalisations and the frequency of gaze to caregiver’s face as predictors (as denoted by AIC). The raw frequency of CV vocalisations is correlated with the frequency of CV vocalisations with gaze coordination (*r* = .74). The same pattern was also true for vocalisations (regardless of sub-type), but the variance explained was far less.

The frequency of index-finger pointing with gaze coordination ($R^2$ = .07) was not a better predictor than the frequency of index-finger pointing ($R^2$ = .10).

**Predicting Expressive Vocabulary at 24 Months**

Table 88 shows the average frequency of infant behaviours at 11 & 12 months for the infants for whom we had language outcomes at 15 months.

Table 88

*Average Frequency of Vocalisations and Gestures Produced and the Frequency and Proportion Produced With Gaze Coordination at 11 & 12 months with Language Outcomes at 24 months (n = 49)*

| | *Produced* | | | | | | *Produced With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *By Infants* | | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *N* | *%* | *M* | *SD* | *Med* | *Range* | *M* |
| **Gaze to Caregiver's Face** | 44.41 | 22.38 | 40 | 5-99 | 49 | 100% | - | - | - | - | - |
| **Vocalisations** | 91.25 | 40.99 | 85 | 17-203 | 49 | 100% | 16.37 | 10.33 | 15 | 0-39 | 0.18 |
| CV | 37.87 | 27.73 | 31 | 2-129 | 49 | 100% | 6.41 | 5.65 | 5 | 0-25 | 0.17 |
| Non-CV | 53.38 | 25.54 | 52 | 12-145 | 49 | 100% | 9.95 | 6.68 | 9 | 0-25 | 0.19 |
| **Gestures** | 7.83 | 6.59 | 6 | 0-26 | 44 | 90% | 3.97 | 3.78 | 3 | 0-15 | 0.51 |
| Index-finger point | 1.67 | 3.04 | 0 | 0-15 | 22 | 45% | 0.59 | 1.44 | 0 | 0-8 | 0.37 |
| Open-hand point† | 0.90 | 1.77 | 0 | 0-10 | 18 | 37% | 0.43 | 1.09 | 0 | 0-6 | 0.39 |
| Give | 3.69 | 4.22 | 2 | 0-17 | 34 | 69% | 1.73 | 2.28 | 1 | 0-9 | 0.45 |
| Show† | 0.88 | 1.56 | 0 | 0-7 | 19 | 39% | 0.82 | 1.51 | 0 | 0-7 | 0.92 |
| Conventional gesture† | 0.69 | 1.25 | 0 | 0-7 | 19 | 39% | 0.41 | 0.70 | 0 | 0-3 | 0.65 |

† denotes behaviour not included as a predictor in proportional models (fewer than 10 infants producing the behaviour, and fewer than 50% producing the behaviour more than twice)

**Raw frequency.** We tested whether raw frequencies of behaviours at 11 & 12 months (Table 83) predicted expressive language at 24 months.

We initially tested whether the frequency of gaze to caregiver's face, vocalisations or gestures individually predicted 24-month expressive vocabulary. However, no predictors improved on the null model.

We then explored whether the frequency of specific types of vocalisation or gestures predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on a null model with no predictors.

***Vocalisation type.*** Neither predictor improved on the null model when we tested the frequency of CV or non-CV vocalisations as predictors.

***Gesture type.*** For models with the frequency of specific gesture types as predictors, the frequency of index-finger pointing improved on the null model, $F(1,47) = 4.87$, $p = .032$, $R^2 = .09$, as did the frequency of showing, $F(1,47) = 6.47$, $p = .014$, $R^2 = .12$.

***Final model.*** We next tested whether the two compatible predictors (frequency of index-finger pointing and showing) explained separate variance. We systematically tested for improvement in fit, combining the two predictors. The final model is presented in Table 89.

Table 89

*Regression Model fitting Frequency of Index-Finger Pointing and Showing at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 314.70 | 28.23 | 11.15 | .000 |
| Frequency of Index-finger Pointing | 15.30 | 7.55 | 2.03 | .049 |
| Frequency of Showing | 34.88 | 14.67 | 2.38 | .022 |

$R^2 = .19$, $F(2,46) = 5.50$, $p = .007$.

**Frequency of gaze coordination.** We next investigated whether the frequency of behaviours *with gaze coordination* at 11 & 12 months predicted expressive vocabulary at 24 months.

We initially tested whether the frequency of vocalisations or gestures with gaze coordination individually predicted 24-month expressive vocabulary. Only the frequency of gestures with gaze coordination improved on the null model, $F(1,47) = 5.72$, $p = .021$, $R^2 = .11$.

We then explored whether the frequency of specific types of vocalisation or gesture with gaze coordination predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model.

***Vocalisation type.*** Neither predictor improved on the null model when we tested the frequency of CV or non-CV vocalisations as predictors with gaze coordination.

***Gesture type.*** For models with the frequency of specific gesture types with gaze coordination as predictors, only the frequency of showing with gaze coordination improved on the null model, $F(1,47) = 5.75$, $p = .021$, $R^2 = .11$.

***Final model.*** The final model consists only of the frequency of showing with gaze coordination, and is presented in table 90.

Table 90

*Regression Model fitting Frequency of Gestures with Gaze Coordination at 11 & 12 months to Expressive Vocabulary at 15 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 340.20 | 26.72 | 12.73 | .000 |
| Frequency of Showing with Gaze Coordination | 37.62 | 15.69 | 2.40 | .021 |

$R^2 = .11$, $F(1,47) = 5.75$, $p = .021$.

**Proportion of gaze coordination.** We next investigated whether the proportion of behaviours that were gaze-coordinated at 11 & 12 months (Table 83) predicted expressive language at 24 months.

We initially tested whether the proportion of vocalisations or gestures with gaze coordination individually predicted 24-month expressive vocabulary (on subsets where all participants had produced each type of behaviour– see Method). However, no predictors improved on the null model. We then explored whether the proportion of specific types of vocalisation or gestures with gaze coordination predicted 24-month expressive vocabulary. We tested each predictor individually, to determine if they improved on the null model, but again, no predictors improved on the null model.

**Does Gaze-coordinated Behaviour Better Predict Expressive Vocabulary at 24 Months than the Raw Frequency of Behaviours?**

We wanted to determine whether gaze-coordinated behaviours were better predictors of later expressive vocabulary than the frequency with which they were produced regardless of gaze coordination, as this would follow if they were early instances of intentional communication. A number of gaze-coordinated behaviours predicted expressive vocabulary at 24 months (as explored above), and so we consider this question for these behaviours. Table 91 shows the variance explained by predictors relating to gestures and 24-month expressive vocabulary, as demonstrated by both $R^2$ and AIC.

Table 91

*Summary of Regression Models ($R^2$ and AIC) fitting Predictors relating to Gestures at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| *Predictor(s)* | $R^2$ | *AIC* |
|---|---|---|
| Frequency of Gestures (*non-significant predictor*) | .07 | 644.85 |
| Frequency of Gestures with Gaze Coordination | .11 | 642.88 |
| Frequency of Gestures + Frequency of Gaze | .08 | 646.59 |

Table 91 shows that the frequency of gestures with gaze coordination was a better predictor of expressive vocabulary at 24 months than the frequency of gestures (as denoted by $R^2$) and also was a better predictor than a model with both the frequency of gestures and the frequency of gaze to caregiver's face as predictors (as denoted by AIC). The raw frequency of gestures is correlated with the frequency of gestures with gaze coordination ($r$ = .80).

For showing gestures we were unable to establish whether gaze coordination better predicted later expressive vocabulary. The frequency of showing, was predictive of later expressive language outcomes, as was the frequency of showing with gaze coordination. However, we are unable to determine whether gaze coordination is a better predictor than the raw frequency, as showing is almost always produced with gaze coordination (see table 83). As such the predictors are highly correlated ($r$ = .99).

The frequency of index-finger pointing with gaze coordination (non-significant predictor $F(1,47) = 1.14$, $p > .05$, $R^2 = .02$) was not a better predictor than the frequency of index-finger pointing ($R^2 = .09$). Furthermore, there were no further gaze-coordinated behaviours (other than discussed above) that were significant predictors of 24-month expressive vocabulary.

**Appendix J**

Additional Regression Models (Appendix to Chapter 6)

Table 92

*Regression Model fitting Frequency of Gaze to Caregiver's Face and CV Vocalisations at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | -5.20 | 7.53 | -0.69 | .493 |
| Frequency of Gaze to Caregiver's Face | 0.32 | 0.14 | 2.34 | .023 |
| Frequency of CV Vocalisations | 0.32 | 0.12 | 2.69 | .010 |

$R^2$ = .23, $F$(2,50) = 7.50, $p$ = .001.

Table 93

*Regression Model fitting Frequency of Gaze to Caregiver's Face and Vocalisations at 11 & 12 months to Expressive Vocabulary at 15 months (n=53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | -0.31 | 9.42 | -0.03 | .973 |
| Frequency of Gaze to Caregiver's Face | 0.35 | 0.15 | 2.37 | .022 |
| Frequency of Vocalisations | 0.06 | 0.09 | 0.68 | .503 |

$R^2$ = .13, $F$(2,50) = 3.64, $p$ = .033.

Table 94

*Regression Model fitting Frequency of Gaze to Caregiver's Face and Gestures at 11 & 12 months to Expressive Vocabulary at 24 months (n=49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 302.64 | 54.48 | 5.55 | .000 |
| Frequency of Gaze to Caregiver's Face | 0.93 | 1.23 | 0.76 | .454 |
| Frequency of Gestures | 5.46 | 5.12 | 1.07 | .292 |

$R^2$ = .06, $F$(2,46) = 1.52, $p$ = .230.

## Appendix K

Inclusive Analyses (Appendix to Chapter 7)

Here we consider the questions posed in the introduction of Chapter 7, but do not treat gesture-vocal combinations as a mutually exclusive category from gestures or vocalisations.

### Are Certain Types of Infant Behaviour, or Gaze-coordinated Behaviours Frequently, or Proportionally More Likely to be Met with a Response?

Descriptive statistics for caregiver responses at 11 and 12 months separately as well as 11 & 12 months combined are presented in Tables 95-97. These show both the frequency and proportion with which infants produced behaviours that were met with a caregiver response regardless of gaze coordination. Additionally, they show both the frequency and proportion with which infants produced specifically gaze-coordinated behaviours that were met with a caregiver response.

Table 95

*Average Frequency and Proportion of Infant Behaviours and Infant Behaviours with Gaze Coordination with Caregiver Responses at 11 months (n = 134)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 11.47 | 9.46 | 9 | 0-53 | .23 | 3.05 | 3.50 | 2 | 0-19 | .29 |
| CV | 4.94 | 6.29 | 3 | 0-44 | .24 | 1.25 | 2.17 | 0 | 0-15 | .30 |
| Non-CV | 6.53 | 5.35 | 5 | 0-30 | .22 | 1.80 | 2.37 | 1 | 0-14 | .28 |
| **Gestures** | 2.01 | 2.59 | 1 | 0-12 | .71 | 1.23 | 1.96 | 0 | 0-11 | .75 |
| Index-finger point | 0.39 | 1.24 | 0 | 0-9 | .69 | 0.16 | 0.58 | 0 | 0-4 | .71 |
| Open-hand point | 0.22 | 0.75 | 0 | 0-7 | .59 | 0.09 | 0.42 | 0 | 0-3 | .72 |
| Give | 0.88 | 1.74 | 0 | 0-11 | .90 | 0.58 | 1.39 | 0 | 0-10 | .93 |
| Show | 0.31 | 0.82 | 0 | 0-5 | .71 | 0.30 | 0.82 | 0 | 0-5 | .70 |
| Conventional gesture | 0.20 | 0.61 | 0 | 0-5 | .64 | 0.11 | 0.34 | 0 | 0-2 | .64 |

Table 96

*Average Frequency and Proportion of Infant Behaviours and Infant Behaviours with Gaze Coordination with Caregiver Responses at 12 months (n = 64)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 9.60 | 8.75 | 7.5 | 0-37 | .23 | 2.50 | 2.95 | 2 | 0-13 | .31 |
| CV | 4.02 | 4.88 | 2 | 0-23 | .21 | 1.00 | 1.35 | 1 | 0-5 | .35 |
| Non-CV | 5.58 | 5.57 | 4.5 | 0-33 | .24 | 1.50 | 2.08 | 1 | 0-10 | .31 |
| **Gestures** | 2.90 | 4.33 | 1 | 0-20 | .62 | 1.46 | 2.38 | 1 | 0-13 | .70 |
| Index-finger point | 0.56 | 1.64 | 0 | 0-8 | .46 | 0.17 | 0.72 | 0 | 0-4 | .32 |
| Open-hand point | 0.09 | 0.34 | 0 | 0-2 | .36 | 0.05 | 0.21 | 0 | 0-1 | .50 |
| Give | 1.85 | 3.46 | 0 | 0-16 | .68 | 0.93 | 1.92 | 0 | 0-12 | .80 |
| Show | 0.27 | 0.62 | 0 | 0-3 | .68 | 0.22 | 0.54 | 0 | 0-2 | .63 |
| Conventional gesture | 0.13 | 0.38 | 0 | 0-2 | .59 | 0.09 | 0.34 | 0 | 0-2 | 1.00 |

Table 97

*Average Frequency and Proportion of Infant Behaviours and Infant Behaviours with Gaze Coordination with Caregiver Responses at 11& 12 months (n = 61)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| **Vocalisations** | 20.48 | 15.12 | 16 | 0-68 | .22 | 5.37 | 4.46 | 5 | 0-19 | .31 |
| CV | 8.84 | 10.25 | 4 | 0-53 | .23 | 2.21 | 2.61 | 1 | 0-13 | .36 |
| Non-CV | 11.64 | 8.25 | 11 | 0-44 | .22 | 3.16 | 2.95 | 3 | 0-11 | .27 |
| **Gestures** | 4.88 | 4.97 | 3 | 0-20 | .68 | 2.63 | 2.94 | 2 | 0-13 | .71 |
| Index-finger point | 0.92 | 2.05 | 0 | 0-9 | .56 | 0.25 | 0.85 | 0 | 0-5 | .39 |
| Open-hand point | 0.39 | 1.09 | 0 | 0-7 | .48 | 0.19 | 0.62 | 0 | 0-3 | .57 |
| Give | 2.57 | 3.55 | 1 | 0-16 | .76 | 1.37 | 2.19 | 1 | 0-12 | .85 |
| Show | 0.64 | 1.11 | 0 | 0-5 | .78 | 0.57 | 1.07 | 0 | 0-5 | .75 |
| Conventional gesture | 0.36 | 0.82 | 0 | 0-5 | .60 | 0.25 | 0.51 | 0 | 0-2 | .73 |

We first investigated whether the type of behaviour (either vocalisations or gestures) and whether or not such behaviour was gaze-coordinated meant that they were more frequently, or proportionally more likely to be met with a response.

***Frequency.*** At both ages (and when we collapsed across age), there was a significant improvement in model fit when gaze coordination (11 months: $\chi^2$ (1) = 48.69, $p$ < .001; 12 months: $\chi^2$ (1) = 18.48, $p$ < .001; 11 & 12 months: $\chi^2$ (1) = 27.78, $p$ < .001), and behaviour type (11 months: $\chi^2$ (1) = 105.32, $p$ < .001; 12 months: $\chi^2$ (1) = 39.23, $p$ < .001; 11 & 12 months: $\chi^2$ (1) = 52.51, $p$ < .001) were added as predictors. There were also interaction effects between behaviour type and gaze coordination (11 months: $\chi^2$ (1) = 88.44, $p$ < .001 (Table 98); 12 months: $\chi^2$ (1) = 23.61, $p$ < .001 (Table 99); 11 & 12 months: $\chi^2$ (1) = 43.51, $p$ < .001 (Table 100)). Post-hoc tests (Table 101) revealed that at 11 months, caregivers responded significantly more frequently to vocalisations than gestures (both gaze-coordinated and not gaze-coordinated). At 12 months (and when we collapsed across age), this was only the case for infant behaviours that were not gaze-coordinated (i.e. gaze-coordinated vocalisations were not responded to significantly more frequently than gaze-coordinated gestures). At both ages, caregivers responded significantly more frequently to vocalisations that were *not* gaze-coordinated than those that were gaze-coordinated.

Table 98

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations and Gestures that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 months (N = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *Vocalisation*) | 3.05 | 0.36 | 8.40 | < .001 |
| No Gaze Coordination | 5.38 | 0.40 | 13.28 | < .001 |
| Gesture | -1.81 | 0.49 | -3.73 | < .001 |
| No Gaze Coordination: Gesture | -5.84 | 0.57 | -10.20 | < .001 |

Table 99

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations and Gestures that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 months (N = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *Vocalisation*) | 2.50 | 0.55 | 4.57 | < .001 |
| No Gaze Coordination | 4.59 | 0.66 | 6.99 | < .001 |
| Gesture | -1.04 | 0.66 | -1.59 | .117 |
| No Gaze Coordination: Gesture | -4.62 | 0.93 | -4.97 | < .001 |

Table 100

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations and Gestures that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *Vocalisation*) | 5.37 | 0.88 | 6.13 | < .001 |
| No Gaze Coordination | 9.74 | 1.00 | 9.75 | < .001 |
| Gesture | -2.74 | 1.13 | -2.42 | .018 |
| No Gaze Coordination: Gesture | -10.13 | 1.41 | -7.17 | < .001 |

Table 101

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Vocalisations and Gestures that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 (n = 134) and 12 months (n = 64).*

| | *11 months* | | | | *12 months* | | | | *11 & 12 months* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* |
| ***Gaze-coordinated: Not*** | | | | | | | | | | | | |
| Vocalisations | -5.38 | 266 | -13.28 | < .001 | -4.59 | 126 | -6.99 | < .001 | -9.74 | 120 | -9.75 | < .001 |
| Gestures | 0.46 | 266 | 1.14 | .665 | 0.03 | 126 | 0.04 | 1.000 | 0.39 | 120 | 0.39 | .980 |
| ***Gaze-coordinated*** | | | | | | | | | | | | |
| Vocalisations: Gestures | 1.81 | 133 | 3.73 | .002 | 1.04 | 63 | 1.59 | .392 | 2.74 | 60 | 2.42 | .084 |
| ***Not Gaze-coordinated*** | | | | | | | | | | | | |
| Vocalisations: Gestures | 7.65 | 133 | 15.74 | < .001 | 5.66 | 63 | 8.62 | .001 | 12.87 | 60 | 11.39 | < .001 |

***Proportion.*** At both ages (and when we collapsed across age) there was a significant improvement in model fit when gaze coordination (11 months: $\chi^2$ (1) = 19.67, $p$ < .001; 12 months: $\chi^2$ (1) = 8.19, $p$ = .004; 11 & 12 months: $\chi^2$ (1) = 17.43, $p$ < .001) and behaviour type (11 months: $\chi^2$ (1) = 92.99, $p$ < .001 (Table 102); 12 months: $\chi^2$ (1) = 55.16, $p$ < .001 (Table 103); 11 & 12 months: $\chi^2$ (1) = 80.38, $p$ < .001 (Table 104)) were added as predictors. There were no interaction effects. At both ages (and when we collapsed across age) caregivers were significantly more likely to respond to gestures than to vocalisations. Caregivers were also significantly more likely to respond to behaviours that were gaze-coordinated than those that were not gaze-coordinated.

Table 102

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 6483) *and Gestures* (*n* = 364) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 11 months (n = 134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.41 | 0.07 | -21.60 | < .001 |
| Gaze Coordination | 0.60 | 0.09 | 6.57 | < .001 |
| Gesture | 2.02 | 0.15 | 13.79 | < .001 |

LLRI = .02, C = .73, Dxy = .45.

Table 103

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* ($n$ = 2640) *and Gestures* ($n$ = 290) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 12 months (n = 64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.47 | 0.12 | -12.78 | < .001 |
| Gaze Coordination | 0.52 | 0.15 | 3.59 | < .001 |
| Gesture | 1.52 | 0.16 | 9.52 | < .001 |

LLRI = .02, C = .76, Dxy = .52.

Table 104

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* ($n$ = 5397) *and Gestures* ($n$ = 430) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.46 | 0.10 | -15.30 | < .001 |
| Gaze Coordination | 0.58 | 0.10 | 6.01 | < .001 |
| Gesture | 1.85 | 0.13 | 14.25 | < .001 |

LLRI = .01, C = .71, Dxy = .43.

**Type of vocalisation**. We investigated whether CV or non-CV vocalisations were more frequently, or proportionally more likely to be responded to.

***Frequency.*** We conducted the same analyses, for responses that were semantically contingent. At both ages (and when we collapsed across age) there was a significant improvement in model fit when gaze coordination, (11 months: $\chi^2$ (1) = 97.38, $p$ < .001; 12 months: $\chi^2$ (1) = 34.48, $p$ < .001; 11 & 12 months: $\chi^2$ (1) = 55.27, $p$ < .001), and vocalisation type (11 months: $\chi^2$ (1) = 7.10, $p$ = .008 (Table 105); 12 months: $\chi^2$ (1) = 4.44, $p$ = .035 (Table 106); 11 & 12 months: $\chi^2$ (1) = 4.00, $p$ = .045 (Table 107)) were added as predictors. There were no interaction effects. At both ages (and when we collapsed across age) caregivers responded significantly more frequently to non-CV than CV vocalisations. Additionally, caregivers responded significantly more frequently to both types of vocalisations that were *not* gaze-coordinated than those that were gaze-coordinated.

Table 105

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations that were Met with a Caregiver Response by Gaze Coordination and Vocalisation Type at 11 months (N = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *CV Vocalisation*) | 1.12 | 0.28 | 3.99 | < .001 |
| No Gaze Coordination | 2.69 | 0.25 | 10.81 | < .001 |
| Non-CV Vocalisation | 0.80 | 0.30 | 2.69 | .008 |

Table 106

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations that were Met with a Caregiver Response by Gaze Coordination at 12 months (N = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *CV Vocalisation*) | 0.86 | 0.38 | 2.27 | .025 |
| No Gaze Coordination | 2.29 | 0.37 | 6.18 | < .001 |
| Non-CV Vocalisation | 0.78 | 0.37 | 2.11 | .039 |

Table 107

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations that were Met with a Caregiver Response by Gaze Coordination and Vocalisation Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *CV Vocalisation*) | 1.99 | 0.66 | 3.00 | .003 |
| No Gaze Coordination | 4.87 | 0.59 | 8.30 | < .001 |
| Non-CV Vocalisation | 1.40 | 0.69 | 2.02 | .048 |

***Proportion.*** At both ages, and when we collapsed across age there was a significant improvement in model fit when gaze coordination was added as a predictor (11 months: $\chi^2$ (1) = 26.23, $p < .001$ (Table 108); 12 months: $\chi^2$ (1) = 21.67, $p < .001$ (Table 109); 11 & 12 months: $\chi^2$ (1) = 28.03, $p < .001$ (Table 110)). There was no effect of vocalisation type, or interaction effects. At both ages (and when we collapsed across age) caregivers were significantly more likely to respond to gaze-coordinated vocalisations than those that were not gaze-coordinated.

Table 108

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 2553) *and Non-CV Vocalisations* (*n* = 3930) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination at 11 months (n = 134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.44 | 0.07 | -21.53 | < .001 |
| Gaze Coordination | 0.59 | 0.10 | 5.90 | < .001 |

LLRI = .004, C = .71, Dxy = .42.

Table 109

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 1112) *and Non-CV Vocalisations* (*n* = 1528) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination at 12 months (n = 64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.48 | 0.11 | -13.19 | < .001 |
| Gaze-coordinated | 0.57 | 0.12 | 4.73 | < .001 |

LLRI = .008, C = .71, Dxy = .43.

Table 110

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 2154) *and Non-CV Vocalisations* (*n* = 3243) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination at 11 & 12 months (N = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.50 | 0.09 | -16.55 | < .001 |
| Gaze-coordinated | 0.62 | 0.10 | 6.24 | < .001 |

LLRI = .01, C = .69, Dxy = .38.

**Type of gesture.** We investigated whether index-finger pointing, open-hand pointing, giving, showing or conventional gestures were more frequently, or proportionally more likely to be responded to.

***Frequency.*** At 11 months, there was a significant improvement in model fit when gaze coordination was added as a predictor, $\chi^2$ (1) = 7.30, *p* = .007. At both ages (and when collapsed across age) there was a significant improvement in model fit when gesture type was added as a predictor (11: months $\chi^2$ (4) = 33.76, *p* < .001; 12 months: $\chi^2$ (4) = 43.74, *p* < .001 (Table 112); 11 & 12 months: $\chi^2$ (4) = 50.26, *p* < .001) were added as predictors. At 11 months (and when we collapsed across age) there were interaction effects between behaviour type and gaze coordination (11 months: $\chi^2$ (4) = 20.96, $p$ < .001 (Table 111); 11 & 12 months: $\chi^2$ (4) = 12.04, $p$ = .017 (Table 113)). Post-hoc tests (Table 114) revealed that at 11 months (and when we collapsed across age), caregivers responded significantly more frequently to gaze-coordinated giving all other gaze-coordinated gestures. At 11 months, caregivers responded significantly more frequently to giving that was not gaze-coordinated than showing that was not gaze-coordinated, however when we collapsed across age, caregivers responded more frequently to giving that was not gaze-coordinated than open-hand pointing, showing and conventional gestures that were not gaze-coordinated. At 11 months, caregivers responded significantly more frequently to giving and showing that was not gaze-coordinated than these gestures when they were gaze-coordinated. Additionally, caregivers responded significantly more frequently to gaze-coordinated showing and giving than when these were not gaze-coordinated. Post hoc tests revealed that at 12 months (Table 115), caregivers responded significantly more frequently to giving than all other gestures.

Table 111

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 months (N = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze-coordinated*: *Index-finger Pointing*) | 0.16 | 0.06 | 2.56 | .011 |
| No Gaze Coordination | 0.08 | 0.08 | 1.03 | .301 |
| Open-hand Pointing | -0.07 | 0.09 | -0.79 | .428 |
| Giving | 0.42 | 0.09 | 4.89 | < .001 |
| Showing | 0.14 | 0.09 | 1.66 | .097 |
| Conventional | -0.04 | 0.09 | -0.52 | .603 |
| No Gaze Coordination: Open-hand Pointing | -0.04 | 0.11 | -0.37 | .711 |
| No Gaze Coordination: Giving | -0.35 | 0.11 | -3.27 | .001 |
| No Gaze Coordination: Showing | -0.36 | 0.11 | -3.41 | .001 |
| No Gaze Coordination: Conventional | -0.10 | 0.11 | -0.94 | .347 |

Table 112

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 12 months (N = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Pointing*) | 0.28 | 0.11 | 2.47 | .014 |
| No Gaze Coordination | -0.01 | 0.07 | -0.08 | .937 |
| Open-hand Pointing | -0.23 | 0.15 | -1.55 | .122 |
| Giving | 0.64 | 0.15 | 4.26 | < .001 |
| Showing | -0.15 | 0.15 | -0.98 | .326 |
| Conventional | -0.22 | 0.15 | -1.45 | .149 |

Table 113

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 & 12 months (N = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordination: Index-finger Pointing*) | 0.25 | 0.16 | 1.53 | .127 |
| No Gaze Coordination | 0.43 | 0.19 | 2.19 | .030 |
| Open-hand Pointing | -0.05 | 0.22 | -0.23 | .817 |
| Giving | 1.13 | 0.22 | 5.06 | < .001 |
| Showing | 0.33 | 0.22 | 1.47 | .143 |
| Conventional | 0.00 | 0.22 | 0.00 | .999 |
| No Gaze Coordination: Open-hand Pointing | -0.42 | 0.28 | -1.54 | .125 |
| No Gaze Coordination: Give | -0.60 | 0.28 | -2.19 | .029 |
| No Gaze Coordination: Show | -0.93 | 0.28 | -3.39 | .001 |
| No Gaze Coordination: Conventional | -0.56 | 0.28 | -2.02 | .044 |

Table 114

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gestures that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 (n = 134) and 11 & 12 months (n = 61)*

| | *11 months* | | | | *11 & 12 months* | | | |
|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* |
| ***Gaze-coordinated: Not*** | | | | | | | | |
| Giving | 0.27 | 665 | 3.60 | .013 | 0.18 | 300 | 0.91 | .996 |
| Showing | 0.28 | 665 | 3.78 | .007 | 0.51 | 300 | 2.61 | .219 |
| ***Gaze-coordinated*** | | | | | | | | |
| Index-finger Pointing: Giving | -0.42 | 532 | -4.89 | < .001 | -1.13 | 240 | -5.06 | < .001 |
| Open-hand Pointing: Giving | -0.49 | 532 | -5.69 | < .001 | -1.18 | 240 | -5.29 | < .001 |
| Giving: Showing | 0.28 | 532 | 3.23 | .043 | 0.80 | 240 | 3.59 | .015 |
| Giving: Conventional | 0.47 | 532 | 5.42 | < .001 | 1.13 | 240 | 5.06 | < .001 |
| ***Not Gaze-coordinated*** | | | | | | | | |
| Open-hand Pointing: Giving | -0.18 | 532 | -2.09 | .533 | -1.00 | 240 | -4.48 | .001 |
| Giving: Showing | 0.29 | 532 | 3.39 | .026 | 1.13 | 240 | 5.07 | < .001 |
| Giving: Conventional | 0.22 | 532 | 2.53 | .257 | 1.08 | 240 | 4.85 | < .001 |

Table 115

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gestures that were Met with a Caregiver Response by Gesture Type at 12 months (n = 64)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Index-finger point : Open-hand point | 0.23 | 252 | 1.55 | .529 |
| Index-finger point : Give | -0.64 | 252 | -4.26 | < .001 |
| Index-finger point : Show | 0.15 | 252 | 0.98 | .863 |
| Index-finger point : Conventional | 0.22 | 252 | 1.45 | .597 |
| Open-hand point : Give | -0.88 | 252 | -5.82 | < .001 |
| Open-hand point : Show | -0.09 | 252 | -0.57 | .979 |
| Open-hand point : Conventional | -0.02 | 252 | -0.10 | 1.000 |
| Give : Show | 0.79 | 252 | 5.25 | < .001 |
| Give : Conventional | 0.86 | 252 | 5.71 | < .001 |
| Show : Conventional | 0.07 | 252 | 0.47 | .990 |

***Proportion.*** We lacked the data to properly investigate proportional differences for either gestures or combinations.

**Do Responses to Infant Behaviours at 11 Months Cause them to be Produced More Frequently at 12 Months?**

If infants have an active role in their vocabulary learning, in where they are motivated to more frequently produce specific vocalisations or gestures to elicit specific responses, we could expect social shaping of preverbal behaviours, whereby infants increase the amount that they produce specific behaviours (i.e., vocalisations or gestures) between 11 and 12 months because they get a response at 11 months. We first investigated whether the frequency or proportion of infant behaviours that were met with a response at 11 months predicted the frequency that they were produced at 12 months, controlling for the frequency that they were produced at 11 months. Then, as we were particularly interested in whether there was social shaping of gaze-coordinated (i.e. intentionally communicative) behaviours, we conducted the same analyses, focusing on the frequency of gaze-coordinated infant behaviours. However, we only found evidence of social shaping in one case that is was possible to test.

When we tested whether the proportion that infants' gestures that were met with a response at 11 months shaped the frequency that they produced gestures at 12 months, we found a significant effect, $F(1,35) = 7.46$, $p = .010$ (Table 116). This suggested that the proportion of gestures that were met with a response negatively predicted the frequency of 12 month gestures. The two predictors were not highly correlated, $r = -.03$. This result is likely due to a low frequency of gestures observed at either 11 or 12 months.

Table 116

*Regression Model fitting the Proportion that Gestures that were Met with a Caregiver Response at 11 months to the Frequency that Infants Produced Gestures at 12 months (Controlling for the Frequency that Infants Produced Gestures at 11 months) (n = 34)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 12.85 | 2.95 | 4.35 | < .001 |
| Frequency of Gestures at 11 months | -0.07 | 0.27 | -0.26 | .800 |
| Proportion of Gestures that were met with a response at 11 months | -8.74 | 3.20 | -2.73 | .010 |

$R^2 = .18$, $F(2,35) = 3.75$, $p = .034$.

**Predicting expressive language**

We posed a number of questions (questions 3 - 6, see *introduction*) that relate to later expressive vocabulary outcomes. For analyses relating to these questions, we opted to collapse the 11- and 12-month data (as in Chapter 6), in order to provide enough variance in measures regarding infants' gestures (see Table 97 for descriptives).

**Does the Frequency or Proportion of Infant Behaviours that are Met with a Response Predict 15-month Expressive Vocabulary?**

We first explored whether the frequency with which infant behaviours were produced and met with a response predicted 15-month expressive vocabulary.

After determining which predictors improve on the null model, we also consider whether predictive relationships between behaviours held after controlling for semantically contingent speech (question 5). As a first step we checked whether the overall rate of semantically contingent IDS at 11 & 12 months predicted expressive vocabulary at 15 months regardless of whether it was produced in response to an infant behaviour. However neither the frequency of semantically contingent speech nor the proportion of IDS that was semantically contingent was predictive of 15-month expressive vocabulary. We return to this point, after determining whether the frequency with which vocalisations, gestures and combinations are met with a caregiver response predict later expressive vocabulary (question 4).

The frequency with which a number of behaviours were produced by infants and were met with a response predicted 15-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,51) = 10.12$, $p = .002$, $R^2 = .17$), and specifically CV vocalisations ($F(1,51) = 19.71$, $p < .001$, $R^2 = .28$) and index-finger pointing ($F(1,51) = 6.06$, $p = .017$, $R^2 = .11$). To check the extent to which these behaviours explained separate variance, we systematically tested for improvements in fit when combining appropriate predictors (see method). The final model is presented in table 117.

.

Table 117

*Regression Model fitting Frequency of CV Vocalisations and Index-finger Pointing that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 5.63 | 3.87 | 1.46 | .152 |
| Frequency of CV Vocalisations that were met with a response | 1.22 | 0.26 | 4.68 | < .001 |
| Frequency of Index-finger Pointing that was met with a response | 3.95 | 1.39 | 2.84 | .007 |

$R^2$ = .38, *F*(2,50) = 15.24, *p* < .001.

**Proportion.** We next explored whether the proportion that infant behaviours were met with a response were predictors of 15-month expressive vocabulary. We used the same method as for frequency analyses, regarding granularity of predictors (i.e. testing main types, followed by sub-types) on subsets where all participants had produced each type of behaviour (see Method).

Only the proportion of vocalisations (*F*(1,51) = 5.32, *p* = .025, $R^2$ = .09) and specifically CV vocalisations (*F*(1,51) = 6.05, *p* = .017, $R^2$ = .11 (Table 118)) that were met with a response improved on the null model.

*Table 118*

*Regression Model fitting Proportion of CV Vocalisations that were met with a Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 8.95 | 5.73 | 1.56 | .125 |
| Proportion of CV Vocalisations that were met with a Response | 52.09 | 21.18 | 2.46 | .017 |

$R^2$ = .11, *F*(1,51) = 6.05, *p* = .017.

**Does the Frequency or Proportion of *Gaze-coordinated* Infant Behaviours that are Met With a Response Predict 15-month Expressive Vocabulary?**

We were particularly interested in whether responses to gaze-coordinated (i.e. intentionally communicative) behaviours were especially predictive of later expressive vocabulary. We conducted the same analyses as above but considered whether the frequency or proportion of infant behaviours that were *gaze-coordinated* and met with a response were predictors of 15-month expressive vocabulary.

**Frequency**. The frequency with which a number of behaviours were produced with gaze coordination by infants and were met with a response predicted 15-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,51) = 19.51$, $p < .001$, $R^2 = .28$), specifically CV vocalisations ($F(1,51) = 34.29$, $p < .001$, $R^2 = .40$) and index-finger pointing ($F(1,51) = 4.57$, $p = .037$, $R^2 = .08$). The final model was similar to when we considered responsiveness to these behaviours regardless of gaze coordination, except that it explained much more variance (Table 119).

Table 119

*Regression Model fitting Frequency of Gaze-coordinated CV Vocalisations and Index-finger Pointing that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 6.54 | 3.47 | 1.89 | .065 |
| Frequency of *Gaze-coordinated* CV Vocalisations that were met with a response | 5.57 | 0.99 | 5.62 | < .001 |
| Frequency of *Gaze-coordinated* Index-finger Pointing that was met with a response | 5.29 | 3.00 | 1.77 | .083 |

$R^2 = .44$, $F(2,50) = 19.42$, $p < .001$.

***Proportion.*** We conducted the same analyses but considered whether the proportion of specific infant behaviours that were *gaze-coordinated* and met with a response were predictors of 15-month expressive vocabulary.

Only the proportion of vocalisations ($F$(1,51) = 10.83, $p$ = .002, $R^2$ = .18 (Table 120)) and specifically CV vocalisations ($F$(1,51) = 8.06, $p$ = .006, $R^2$ = .14) that were *gaze-coordinated* and met with a response improved on the null model.

Table 120

*Regression Model fitting Proportion of Gaze-coordinated Vocalisations that were Met with a Caregiver Response at 11 & 12 months to Expressive Vocabulary at 15 months (n = 53)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 5.96 | 5.43 | 1.10 | .278 |
| Proportion of Vocalisations that were *Gaze-coordinated* and Met with a Response | 253.64 | 77.06 | 3.29 | .002 |

$R^2$ = .18, $F$(1,51) = 10.83, $p$ = .002.

**Do Relationships Between Infant Behaviours Met with a Response and 15-month Expressive Vocabulary Hold When the Amount of Caregiver's Semantically Contingent IDS is Controlled For?**

We wanted to check whether the relationship between caregivers' responsiveness to specific infant behaviours and 15-month expressive vocabulary (as investigated above) was explained instead by rates of semantically contingent IDS by caregivers. However, when we added the frequency of semantically contingent IDS or the proportion of IDS that was semantically contingent as a predictor to the models in Tables 117-120, they were non-significant predictors. This suggests that the relationship between caregivers' responsiveness to specific infant behaviours and 15-month expressive vocabulary (as investigated above) is not due to these behaviours being a proxy for rates of semantically contingent IDS.

**Does Caregiver Responsiveness to Infant Behaviours Mediate the Relationship Between Infants' Production of these Behaviours and 15-month Expressive Vocabulary?**

In Chapter 6, we demonstrated that 15-month expressive vocabulary was predicted by the frequency with which infants produced a number of behaviours, and in some cases, the frequency or proportion with which these behaviours were gaze-coordinated. In the current chapter, we have demonstrated that responses to some of these behaviours also predicts expressive vocabulary at 15 months. The question arises, then, as to whether some infant behaviours predict outcomes *because* they are responded to. To recap results so far, Table 121 summarises the predictive relationships we have demonstrated throughout this chapter and Chapter 6 and expressive vocabulary at 15 months.

Table 121

*Predictors of 15-month expressive vocabulary (with $R^2$)*

| | *Vocalisations* | | | |
|---|---|---|---|---|
| *Predictors of 15-month expressive vocabulary* | *All* | *CV* | *Non-CV* | *Index-finger pointing* |
| *Frequency* | | | | |
| Frequency produced* | | .16 | | .10 |
| Gaze-coordinated* | .18 | .32 | | .07 |
| Met with a response | .17 | .28 | | .11 |
| Gaze-coordinated and met with a response | .28 | .40 | | .08 |
| *Proportion* | | | | |
| Gaze-coordinated* | .11 | | .08 | |
| Met with a response | .09 | .11 | | |
| Gaze-coordinated and met with a response | .18 | .14 | | |

** relationship established in Chapter 6*

For all the behaviours in Table 121 where both the infant behaviour and the caregiver response predicted outcomes, we conducted mediation analyses to determine whether caregiver responsiveness was the mechanism by which these behaviours predicted 15-month vocabulary. Caregiver responsiveness did not mediate the relationship between the frequency of index-finger pointing (or gaze-coordinated index-finger pointing) and 15-month expressive vocabulary. For other behaviours, we found mediating effects of caregiver responsiveness, discussed in detail below. Note, that we did not investigate the proportional measures for these, as the frequency measures explained more variance and thus were of greater interest.

***Vocalisations***. Figure 26 shows that the relationship between the frequency of gaze-coordinated vocalisations and 15-month expressive vocabulary was mediated by the frequency with which these vocalisations were met with a response, $b = 0.41$, $p = .01$.

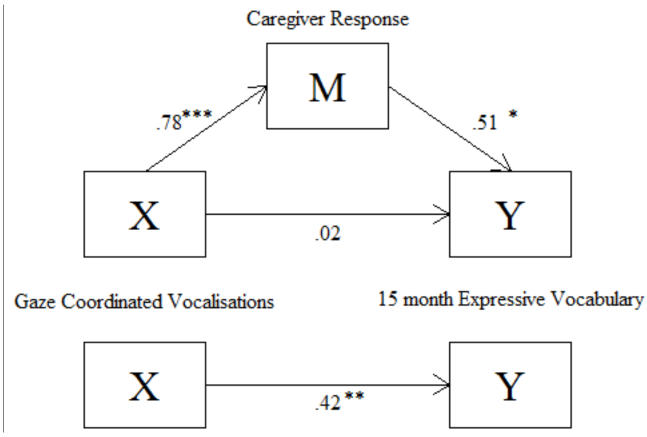


*Figure 26.* Standardized regression coefficients for the relationship between frequency of gaze-coordinated vocalisations and 15-month expressive vocabulary, mediated by the frequency with which these were met with a response. * p < .05, ** p < .01, *** p < .001

***CV Vocalisations.*** Figure 27 shows that the relationship between the frequency of CV vocalisations (and *gaze-coordinated* CV vocalisations) and 15-month expressive vocabulary was mediated by the frequency with which these vocalisations were met with a response, (regardless of gaze coordination: $b = 0.59$, $p < .01$; gaze-coordinated: $b = 0.40$, $p < .01$).

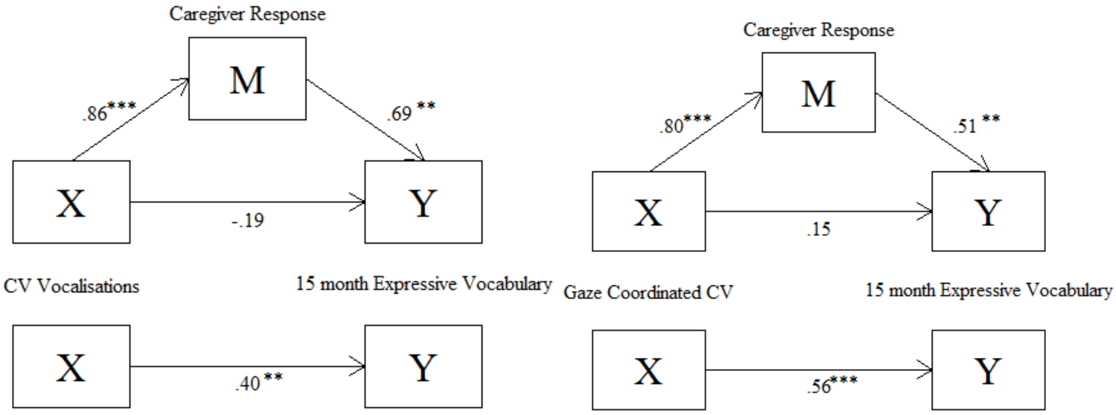


*Figure 27.* Standardized regression coefficients for the relationship between frequency of CV vocalisations (left) or gaze-coordinated CV vocalisations (right) and 15-month expressive vocabulary, mediated by the frequency with which these were met with a response. ** p < .01, *** p < .001

**Does the Frequency or Proportion of Infant Behaviours that are Met with a Response Predict 24-month Expressive Vocabulary?**

Having investigated how caregiver responses to infant behaviours predict 15-month outcomes, we then turned our attention to 24-month outcomes. Again, we focus on questions 3-6 posed in the introduction.

We initially tested whether responses to specific types of infant behaviour (vocalisations and gestures, and the sub-types) at 11 & 12 months predicted 24-month expressive vocabulary. As before, we first considered the frequency with which infants produced behaviours that were responded to, and then proportional measures.

As for the 15-month analyses, as a first step, we checked whether the overall rate of semantically contingent IDS at 11 & 12 months predicted expressive vocabulary at 24 months regardless of whether it was produced in response to an infant behaviour. The frequency of IDS that was semantically contingent at 11 & 12 months was a significant predictor of 24-month expressive vocabulary, $F(1,47) = 8.40$, $p = .006$, $R^2 = .15$, however the proportion of IDS that was semantically contingent did not. Unlike for 15-month expressive vocabulary, controlling for the frequency of semantically contingent IDS also affected some final models. We have highlighted where this is the case.

**Frequency.** We explored whether the frequency of specific infant behaviours that were met with a response were predictors of 24-month expressive vocabulary. We used the same method as for predicting 15-month expressive vocabulary (regarding granularity of predictors to be tested).

The frequency with which a number of behaviours were produced by infants and were responded to predicted 24-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,47) = 5.24$, $p = .027$, $R^2 = .10$), specifically CV vocalisations ($F(1,47) = 5.14$, $p = .028$, $R^2 = .10$), gestures ($F(1,47) = 4.64$, $p = .036$, $R^2 = .09$), specifically index-finger pointing ($F(1,47) = 5.61$, $p = .022$, $R^2 = .11$), open-hand pointing ($F(1,47) = 5.76$, $p = .020$, $R^2 = .11$) and showing ($F(1,47) = 7.09$, $p = .011$, $R^2 = .13$).

We systematically tested for improvements in fit when combining appropriate predictors (see Method). The final model is presented in Table 122, Model 1. However, additionally, when we added the frequency of semantically contingent IDS as a predictor to this model, though it was a non-significant predictor it reduced the role of vocalisations that were met with a response to non-significant (Table 122, Model 2). This suggests that the frequency of vocalisations that were met with a response at 11 & 12 months did not predict expressive vocabulary at 24 months when we controlled for the frequency of semantically contingent IDS, however the same was not true for open-hand pointing or showing.

Table 122

*Regression Model fitting Frequency of Vocalisations, Index-finger Pointing, Open-hand Pointing and Showing that were Met with a Caregiver Response (model 1) and Frequency of Semantically Contingent IDS (model 2) at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| **Model 1** | | | | |
| Intercept | 261.29 | 38.12 | 6.85 | < .001 |
| Frequency of Vocalisations that were met with a response | 3.97 | 1.37 | 2.91 | .006 |
| Frequency of Index-finger Pointing that was met with a response | 18.45 | 9.19 | 2.01 | .051 |
| Frequency of Open-hand Pointing that was met with a response | -46.41 | 17.27 | -2.69 | .010 |
| Frequency of Showing that was met with a response | 39.27 | 18.06 | 2.17 | .035 |
| $R^2 = .39, F(4,44) = 7.05, p < .001.$ | | | | |
| **Model 2** | | | | |
| Intercept | 178.76 | 64.18 | 2.79 | .008 |
| Frequency of Vocalisations that were met with a response | 2.36 | 1.69 | 1.40 | .170 |
| Frequency of Index-finger Pointing that was met with a response | 14.44 | 9.38 | 1.54 | .131 |
| Frequency of Open-hand Pointing that was met with a response | -44.78 | 17.01 | -2.63 | .012 |
| Frequency of Showing that was met with a response | 42.16 | 17.85 | 2.36 | .023 |
| Frequency of Semantically Contingent IDS | 0.65 | 0.41 | 1.58 | .121 |
| $R^2 = .42, F(5,43) = 6.34, p < .001.$ | | | | |

**Proportion.** We next explored whether the proportion with which specific infant behaviours were met with a response were predictors of 24-month expressive vocabulary. We used the same method as for frequency analyses, regarding granularity of predictors (i.e. testing main types, followed by sub-types) on subsets where all participants had produced each type of behaviour (see Method).

The proportion of vocalisations ($F(1,47) = 9.99$, $p = .003$, $R^2 = .18$), specifically non-CV vocalisations ($F(1,47) = 10.28$, $p = .002$, $R^2 = .18$) and giving ($F(1,32) = 4.98$, $p = .033$, $R^2 = .13$) that were met with a response improved on null models. Combining the two compatible predictors (proportion of non-CV vocalisations and giving) did not result in an improvement of fit, suggesting that they explained similar variance. Additionally, when we controlled for the frequency of semantically contingent IDS by caregivers, neither significantly predicted expressive language at 24 months.

**Does the Frequency or Proportion of Gaze-coordinated Infant Behaviours that are Met with a Response Predict 24-month Expressive Vocabulary?**

As we were particularly interested in whether responses to gaze-coordinated (i.e. intentionally communicative) behaviours were especially predictive of later expressive vocabulary. We conducted the same analyses as above but considered whether the frequency or proportion of infant behaviours that were *gaze-coordinated* and met with a response were predictors of 24-month expressive vocabulary.

**Frequency.** The frequency with which a number of behaviours were produced with gaze coordination by infants and were met with a response predicted 24-month expressive vocabulary (as they improved on the null model). These were vocalisations ($F(1,47) = 7.54$, $p = .009$, $R^2 = .14$), specifically CV vocalisations ($F(1,47) = 4.66$, $p = .036$, $R^2 = .09$), gestures ($F(1,47) = 5.89$, $p = .019$, $R^2 = .11$), specifically giving ($F(1,47) = 4.30$, $p = .044$, $R^2 = .08$) and showing ($F(1,47) = 6.01$, $p = .018$, $R^2 = .11$).

The final model included the frequency of gaze-coordinated vocalisations, and showing as significant predictors of 24-month expressive vocabulary (Table 123, Model 1). However, when we added the frequency of semantically contingent IDS as a predictor, it was a significant predictor, and furthermore reduced the role of the frequency of gaze-coordinated vocalisations that were met with a response to non-significant (Table 123, Model 2). This suggests that the frequency of gaze-coordinated vocalisations that were met with a response at 11 & 12 months did not predict expressive vocabulary at 24 months when we controlled for the frequency of semantically contingent IDS, however the same was not true for gaze-coordinated showing.

Table 123

*Regression Model fitting Frequency of Gaze-coordinated and Semantically Contingently Responded To Vocalisations and Showing at 11 & 12 months to Expressive Vocabulary at 24 months (n = 49)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| **Model 1** | | | | |
| Intercept | 276.28 | 36.74 | 7.52 | < .001 |
| Frequency of *Gaze-coordinated* Vocalisations that were met with a response | 13.27 | 5.56 | 2.39 | .021 |
| Frequency of *Gaze-coordinated* Showing that was met with a response | 43.15 | 20.91 | 2.06 | .045 |
| *$R^2$ = .21, F(2,46) = 6.16, p = .004* | | | | |
| **Model 2** | | | | |
| Intercept | 129.08 | 67.86 | 1.90 | .064 |
| Frequency of *Gaze-coordinated* Vocalisations *that were met with a response* | 7.21 | 5.78 | 1.25 | .218 |
| Frequency of *Gaze-coordinated* Showing that was met with a response | 51.39 | 20.06 | 2.56 | .014 |
| Frequency of Semantically Contingent IDS | 0.94 | 0.37 | 2.53 | .015 |
| *$R^2$ = .31, F(3,45) = 6.71, p = .001* | | | | |

**Proportion.** Finally, we conducted the same analyses but considered whether the proportion of specific infant behaviours that were *gaze-coordinated* and met with a response were predictors of 24-month expressive vocabulary.

Only the proportion of vocalisations ($F(1,47) = 8.03$, $p = .007$, $R^2 = .15$) and specifically non-CV vocalisations ($F(1,47) = 7.05$, $p = .011$, $R^2 = .13$) improved on the null model.

However, when we controlled for the frequency of semantically contingent IDS by caregivers, the proportion of vocalisations did not predict expressive language at 24 months.

**Does Caregiver Responsiveness to Infant Behaviours Mediate the Relationship Between Infants' Production of these Behaviours and 24-month Expressive Vocabulary?**

Through this chapter, and Chapter 6, we have demonstrated that frequency and proportional measures of infant behaviours, their gaze coordination and the extent to which they are responded to by caregivers predict expressive vocabulary at 24 months. Table 124 demonstrates what predictive relationships we have demonstrated throughout this chapter and Chapter 6 between and expressive vocabulary at 24 months.

Table 124

*Predictors of 24-month expressive vocabulary (with $R^2$)*

| *Predictors of 24-month expressive vocabulary* | *Vocalisations* | | | *Gestures* | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *All* | *CV* | *Non-CV* | *All* | *Pointing* | *Index-finger Pointing*** | *Open-hand* | *Giving* | *Showing* |
| *Frequency* | | | | | | | | | |
| Frequency produced* | | | | | | .09 | | | .12 |
| Gaze-coordinated* | | | | .11 | | | | | .11 |
| Met with a response | .10 | .10 | | .09 | | .11 | .11 | | .13 |
| Gaze-coordinated and met with a response | .14 | .09 | | .11 | | | | .08 | .11 |
| *Proportion* | | | | | | | | | |
| Gaze-coordinated* | | | | | | | | | |
| Met with a response | .18 | | .18 | | | | | .13 | |
| Gaze-coordinated and met with a response | .15 | | .13 | | | | | | |

** relationship established in Chapter 6*

*** negative predictor of expressive language*

For all of the behaviours in Table 124 where both the infant behaviour and the caregiver response predicted outcomes, we conducted mediation analyses to determine whether caregiver responsiveness was the mechanism by which these behaviours predicted 24-month vocabulary. However, caregiver responsiveness did not mediate the relationship between any of these behaviours and 24-month expressive vocabulary.

# Appendix L

Extra Gesture-vocal Combination Descriptives (Appendix to Chapter 7)

The following tables show the average frequency with which specific types of gesture-vocal combinations were produced by infants and met with a caregiver response at 11 (Table 125), 12 (Table 126) and 11 & 12 months combined (Table 127).

Table 125

*Average Frequency and Proportion of Gesture-vocal Combinations (by Vocalisation and Gesture Type) and Gesture-vocal Combinations with Gaze Coordination with Caregiver Responses at 11 months (n = 134)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| Index-finger point & CV | 0.20 | 0.85 | 0 | 0-8 | .86 | 0.08 | 0.37 | 0 | 0-3 | .80 |
| Index-finger point & Non-CV | 0.04 | 0.31 | 0 | 0-3 | .33 | 0.01 | 0.09 | 0 | 0-1 | 1.00 |
| Open-hand point & CV | 0.04 | 0.21 | 0 | 0-1 | .53 | 0.01 | 0.12 | 0 | 0-1 | .50 |
| Open-hand point & Non-CV | 0.06 | 0.53 | 0 | 0-6 | .55 | 0.03 | 0.27 | 0 | 0-3 | .80 |
| Give & CV | 0.12 | 0.41 | 0 | 0-3 | .89 | 0.07 | 0.25 | 0 | 0-1 | 1.00 |
| Give & Non-CV | 0.17 | 0.63 | 0 | 0-4 | .82 | 0.13 | 0.53 | 0 | 0-3 | .97 |
| Show & CV | 0.08 | 0.39 | 0 | 0-3 | .83 | 0.07 | 0.38 | 0 | 0-3 | .81 |
| Show & Non-CV | 0.03 | 0.17 | 0 | 0-1 | .58 | 0.03 | 0.17 | 0 | 0-1 | .58 |
| Conventional gesture & CV | 0.04 | 0.24 | 0 | 0-2 | .83 | 0.03 | 0.17 | 0 | 0-1 | 1.00 |
| Conventional gesture & Non-CV | 0.02 | 0.19 | 0 | 0-2 | .40 | 0.01 | 0.12 | 0 | 0-1 | .40 |

Table 126

*Average Frequency and Proportion of Gesture-vocal Combinations (by Vocalisation and Gesture Type) with Contingent Caregiver Responses at 12 months (n = 64)*

| | *Regardless of Gaze Coordination* | | | | | *With Gaze Coordination* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| Index-finger point & CV | 0.20 | 0.95 | 0 | 0-7 | .57 | 0.11 | 0.56 | 0 | 0-4 | .44 |
| Index-finger point & Non-CV | 0.06 | 0.24 | 0 | 0-1 | .50 | 0.02 | 0.12 | 0 | 0-1 | .17 |
| Open-hand point & CV | 0.06 | 0.30 | 0 | 0-2 | .36 | 0.05 | 0.21 | 0 | 0-1 | .75 |
| Open-hand point & Non-CV | 0.00 | 0.00 | 0 | 0-0 | .00 | 0.00 | 0.00 | 0 | 0-0 | .00 |
| Give & CV | 0.30 | 0.69 | 0 | 0-3 | .83 | 0.19 | 0.54 | 0 | 0-3 | .82 |
| Give & Non-CV | 0.27 | 0.89 | 0 | 0-6 | .58 | 0.16 | 0.69 | 0 | 0-5 | .65 |
| Show & CV | 0.03 | 0.17 | 0 | 0-1 | .67 | 0.03 | 0.17 | 0 | 0-1 | .67 |
| Show & Non-CV | 0.13 | 0.38 | 0 | 0-2 | .78 | 0.13 | 0.38 | 0 | 0-2 | .78 |
| Conventional gesture & CV | 0.02 | 0.13 | 0 | 0-1 | 1.00 | 0.02 | 0.13 | 0 | 0-1 | 1.00 |
| Conventional gesture & Non-CV | 0.05 | 0.21 | 0 | 0-1 | .75 | 0.02 | 0.12 | 0 | 0-1 | 1.00 |

Table 127

*Average Frequency and Proportion of Gesture-vocal Combinations (by Vocalisation and Gesture Type) and Gesture-vocal Combinations with Gaze Coordination with Caregiver Responses at 11 & 12 months (n = 61)*

| | *Regardless of Gaze Coordination* | | | | | *With Semantically Contingent Caregiver Response* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Frequency* | | | | *Proportion* | *Frequency* | | | | *Proportion* |
| *Behaviour* | *M* | *SD* | *Med* | *Range* | *M* | *M* | *SD* | *Med* | *Range* | *M* |
| Index-finger point & CV | 0.31 | 1.12 | 0 | 0-7 | .61 | 0.11 | 0.58 | 0 | 0-4 | .31 |
| Index-finger point & Non-CV | 0.11 | 0.45 | 0 | 0-3 | .64 | 0.02 | 0.13 | 0 | 0-1 | .17 |
| Open-hand point & CV | 0.13 | 0.43 | 0 | 0-2 | .45 | 0.07 | 0.25 | 0 | 0-1 | .58 |
| Open-hand point & Non-CV | 0.13 | 0.78 | 0 | 0-6 | .45 | 0.07 | 0.40 | 0 | 0-3 | .40 |
| Give & CV | 0.38 | 0.79 | 0 | 0-4 | .91 | 0.27 | 0.58 | 0 | 0-3 | .89 |
| Give & Non-CV | 0.40 | 1.06 | 0 | 0-6 | .60 | 0.25 | 0.81 | 0 | 0-5 | .76 |
| Show & CV | 0.15 | 0.51 | 0 | 0-3 | .86 | 0.15 | 0.51 | 0 | 0-3 | .86 |
| Show & Non-CV | 0.15 | 0.40 | 0 | 0-2 | .75 | 0.15 | 0.40 | 0 | 0-2 | .75 |
| Conventional gesture & CV | 0.07 | 0.25 | 0 | 0-1 | .80 | 0.07 | 0.25 | 0 | 0-1 | 1.00 |
| Conventional gesture & Non-CV | 0.03 | 0.18 | 0 | 0-1 | .50 | 0.02 | 0.13 | 0 | 0-1 | .50 |

**Appendix M**

Full Regression Models (Appendix to Chapter 5)

Table 128

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 months (n = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Vocalisation*) | 2.64 | 0.28 | 9.36 | < .001 |
| No Gaze Coordination | 5.40 | 0.31 | 17.19 | < .001 |
| Gesture | -1.86 | 0.39 | -4.76 | < .001 |
| Gesture-vocal Combination | -2.15 | 0.39 | -5.51 | < .001 |
| No Gaze Coordination: Gesture | -5.74 | 0.44 | -12.91 | < .001 |
| No Gaze Coordination: Gesture-Vocal Combination | -5.56 | 0.44 | 12.51 | < .001 |

Table 129

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 12 months (n = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Vocalisation*) | 1.91 | 0.41 | 4.67 | < .001 |
| No Gaze Coordination | 4.73 | 0.53 | 8.93 | < .001 |
| Gesture | -1.09 | 0.53 | -2.06 | .041 |
| Gesture-vocal Combination | -1.20 | 0.53 | -2.26 | .025 |
| No Gaze Coordination: Gesture | -4.51 | 0.75 | -6.02 | < .001 |
| No Gaze Coordination: Gesture-Vocal Combination | -5.03 | 0.75 | -6.71 | < .001 |

Table 130

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations, Gestures and Gesture-vocal Combinations that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Vocalisation*) | 4.45 | 0.68 | 6.57 | < .001 |
| No Gaze Coordination | 9.89 | 0.80 | 12.39 | < .001 |
| Gesture | -2.90 | 0.92 | -3.16 | .002 |
| Gesture-vocal Combination | -3.29 | 0.92 | -3.59 | < .001 |
| No Gaze Coordination: Gesture | -9.88 | 1.13 | -8.76 | < .001 |
| No Gaze Coordination: Gesture-Vocal Combination | -10.33 | 1.13 | -9.15 | < .001 |

Table 131

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Vocalisations, Gestures and Combinations that were Met with a Caregiver Response by Gaze Coordination and Behaviour Type at 11 (n = 134) and 12 months (n = 64).*

| | *11 months* | | | | *12 months* | | | | *11 & 12 months* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *T* | *p* | *b* | *df* | *T* | *p* |
| ***Gaze Coordinated: Not*** | | | | | | | | | | | | |
| Vocalisations | -5.40 | 399 | -17.19 | < .001 | -4.73 | 189 | -8.93 | < .001 | -9.89 | 180 | -12.39 | < .001 |
| Gestures | 0.33 | 399 | 1.07 | .895 | -0.22 | 189 | -0.42 | .998 | -0.01 | 180 | -0.01 | 1.000 |
| Combinations | 0.16 | 399 | 0.50 | .996 | 0.30 | 189 | 0.56 | .993 | 0.44 | 180 | 0.56 | .994 |
| ***Gaze Coordinated*** | | | | | | | | | | | | |
| Vocalisations: Gestures | 1.86 | 266 | 4.76 | < .001 | 1.09 | 126 | 2.06 | .315 | 2.90 | 120 | 3.16 | .024 |
| Vocalisations: Combinations | 2.15 | 266 | 5.51 | < .001 | 1.20 | 126 | 2.26 | .217 | 3.29 | 120 | 3.59 | .006 |
| Gestures: Combinations | 0.29 | 266 | 0.75 | .975 | 0.11 | 126 | 0.20 | 1.000 | 0.39 | 120 | 0.43 | .998 |
| ***Not Gaze Coordinated*** | | | | | | | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Vocalisations: Gestures | 7.59 | 266 | 19.45 | < .001 | 5.60 | 126 | 10.57 | < .001 | 12.78 | 120 | 13.93 | < .001 |
| Vocalisations: Combinations | 7.71 | 266 | 19.75 | < .001 | 6.23 | 126 | 11.76 | < .001 | 13.63 | 120 | 14.85 | < .001 |
| Gestures: Combinations | 0.12 | 266 | 0.30 | 1.000 | 0.63 | 126 | 1.19 | .842 | 0.84 | 120 | 0.92 | .941 |

Table 132

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 6297), *Gestures* (*n* = 217) *and Combinations (n* = 147) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 11 months (n = 134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.44 | 0.06 | -22.80 | < .001 |
| Gaze Coordination | 0.54 | 0.09 | 5.93 | < .001 |
| Gesture | 2.28 | 0.20 | 11.31 | < .001 |
| Combination | 2.13 | 0.21 | 10.15 | < .001 |

LLRI = .02, C = .72, Dxy = .45.

Table 133

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 2504), *Gestures* (*n* = 177) *and Combinations (n* = 113) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 12 months (n = 64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.53 | 0.12 | -13.03 | < .001 |
| Gaze Coordination | 0.44 | 0.14 | 3.24 | .001 |
| Gesture | 1.83 | 0.21 | 8.89 | < .001 |
| Combination | 1.67 | 0.23 | 7.18 | < .001 |

LLRI = .02, C = .76, Dxy = .52.

Table 134

*Summary of Fixed Effects from a Logistic Regression Model Fitting Vocalisations* (*n* = 5200), *Gestures* (*n* = 266) *and Combinations (n* = 164) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination and Behaviour Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination: Vocalisation*) | -1.50 | 0.09 | -16.09 | < .001 |
| Gaze Coordination | 0.51 | 0.10 | 5.21 | < .001 |
| Gesture | 2.08 | 0.17 | 12.07 | < .001 |
| Combination | 1.99 | 0.21 | 9.31 | < .001 |

LLRI = .02, C = .73, Dxy = .46.

Table 135

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Vocalisation Type at 11 months (n = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept ( *Gaze Coordinated*: *CV Vocalisation*) | 0.89 | 0.27 | 3.33 | .001 |
| No Gaze Coordination | 2.70 | 0.24 | 11.16 | < .001 |
| Non-CV Vocalisation | 0.85 | 0.28 | 3.05 | .003 |

Table 136

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination at 12 months (n = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept ( *Gaze Coordinated*: *CV Vocalisation*) | 0.54 | 0.35 | 1.54 | .126 |
| No Gaze Coordination | 2.36 | 0.36 | 6.56 | < .001 |
| Non-CV Vocalisation | 0.82 | 0.36 | 2.28 | .026 |

Table 137

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Vocalisations (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Vocalisation Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *CV Vocalisation*) | 1.49 | 0.63 | 2.38 | .019 |
| No Gaze Coordination | 4.94 | 0.57 | 8.60 | < .001 |
| Non-CV Vocalisation | 1.47 | 0.65 | 2.27 | .027 |

Table 138

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Vocalisation Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*) | 0.58 | 0.09 | 6.27 | < .001 |
| No Gaze Coordination | -0.22 | 0.10 | -2.26 | .026 |

Table 139

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 2453) *and Non-CV Vocalisations* (*n* = 3844) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination at 11 months (n = 134)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.46 | 0.06 | -22.54 | < .001 |
| Gaze Coordination | 0.51 | 0.10 | 5.08 | .001 |

LLRI = .003, C = .71, Dxy = .41.

Table 140

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 1040) *and Non-CV Vocalisations* (*n* = 1464) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Gaze Coordination at 12 months (n = 64)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.54 | 0.12 | -12.82 | < .001 |
| Gaze Coordinated | 0.47 | 0.16 | 3.00 | < .001 |

LLRI = .002, C = .72, Dxy = .44.

Table 141

*Summary of Fixed Effects from a Logistic Regression Model Fitting CV* (*n* = 2055) *and Non-CV Vocalisations* (*n* = 3145) *that were Met with a Caregiver Response (1= Response, 0 = No response) by Vocalisation Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Intercept (*No Gaze Coordination*) | -1.52 | 0.09 | -17.23 | < .001 |
| Gaze Coordinated | 0.54 | 0.10 | 5.41 | < .001 |

LLRI = .004, C = .69, Dxy = .38.

Table 142

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 months (n = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Index-finger Pointing*) | 0.09 | 0.04 | 2.08 | .038 |
| No Gaze Coordination | -0.00 | 0.06 | -0.08 | .940 |
| Open-hand Pointing | -0.05 | 0.06 | -0.76 | .446 |
| Giving | 0.29 | 0.06 | 4.89 | < .001 |
| Showing | 0.11 | 0.06 | 1.76 | .079 |
| Conventional | -0.02 | 0.06 | -0.37 | .710 |
| No Gaze Coordination: Open-hand Pointing | 0.03 | 0.08 | 0.35 | .725 |
| No Gaze Coordination: Giving | -0.16 | 0.08 | -2.06 | .040 |
| No Gaze Coordination: Showing | -0.18 | 0.08 | -2.33 | .020 |
| No Gaze Coordination: Conventional | -0.00 | 0.08 | -0.05 | .958 |

Table 143

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 12 months (n = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Pointing*) | 0.13 | 0.09 | 1.57 | .118 |
| No Gaze Coordination | 0.04 | 0.06 | 0.78 | .435 |
| Open-hand Pointing | -0.14 | 0.11 | -1.26 | .210 |
| Giving | 0.51 | 0.11 | 4.52 | < .001 |
| Showing | -0.09 | 0.11 | -0.84 | .403 |
| Conventional | -0.13 | 0.11 | -1.12 | .265 |

Table 144

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Gestures (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordination: Index-finger Pointing*) | 0.25 | 0.10 | 2.54 | .012 |
| No Gaze Coordination | 0.00 | 0.07 | 0.02 | .987 |
| Open-hand Pointing | -0.19 | 0.13 | -1.51 | .133 |
| Giving | 0.66 | 0.13 | 5.29 | < .001 |
| Showing | -0.07 | 0.13 | -0.59 | .558 |
| Conventional | -0.12 | 0.13 | -0.98 | .329 |

Table 145

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gestures (Produced Alone) that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 months (n = 134)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| ***Gaze Coordinated: Not*** | | | | |
| Showing | 0.19 | 665 | 3.38 | .027 |
| ***Gaze Coordinated*** | | | | |
| Index-finger Pointing: Giving | -0.29 | 532 | -4.89 | < .001 |
| Open-hand Pointing: Giving | -0.34 | 532 | -5.66 | < .001 |
| Giving: Showing | 0.19 | 532 | 3.14 | .057 |
| Giving: Conventional | 0.32 | 532 | 5.27 | .001 |
| ***Not Gaze Coordinated*** | | | | |
| Giving: Showing | 0.21 | 532 | 3.49 | .019 |

Table 146

*Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Gestures (Produced Alone) that were Met with a Caregiver Response by Gesture Type at 12 months (n = 64) and 11 & 12 months (n = 61)*

| | *12 months* | | | | *11 & 12 months* | | | |
|---|---|---|---|---|---|---|---|---|
| | *b* | *df* | *t* | *p* | *b* | *df* | *t* | *p* |
| Index-finger point : Open-hand point | 0.14 | 252 | 1.26 | .718 | 0.19 | 240 | 1.51 | .558 |
| Index-finger point : Give | -0.51 | 252 | -4.52 | < .001 | -0.66 | 240 | -5.29 | < .001 |
| Index-finger point : Show | 0.09 | 252 | 0.84 | .918 | 0.07 | 240 | 0.59 | .977 |
| Index-finger point : Conventional | 0.13 | 252 | 1.12 | .797 | 0.12 | 240 | 0.98 | .865 |
| Open-hand point : Give | -0.65 | 252 | -5.78 | < .001 | -0.85 | 240 | -6.80 | < .001 |
| Open-hand point : Show | -0.05 | 252 | -0.42 | .994 | -0.12 | 240 | -0.92 | .888 |
| Open-hand point : Conventional | -0.02 | 252 | -0.14 | 1.000 | -0.07 | 240 | -0.53 | .984 |
| Give : Show | 0.60 | 252 | 5.36 | < .001 | 0.74 | 240 | 5.87 | < .001 |
| Give : Conventional | 0.63 | 252 | 5.64 | < .001 | 0.79 | 240 | 6.27 | < .001 |
| Show : Conventional | 0.03 | 252 | 0.28 | .999 | 0.05 | 240 | 0.39 | .995 |

Table 147

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 months (n = 134)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Index-finger Pointing*) | 0.09 | 0.04 | 2.41 | .016 |
| No Gaze Coordination | 0.06 | 0.04 | 1.46 | .145 |
| Open-hand Pointing | -0.04 | 0.05 | -0.86 | .392 |
| Giving | 0.11 | 0.05 | 2.14 | .033 |
| Showing | 0.01 | 0.05 | 0.29 | .775 |
| Conventional | -0.04 | 0.05 | -0.86 | .392 |
| No Gaze Coordination: Open-hand Pointing | -0.04 | 0.06 | -0.77 | .440 |
| No Gaze Coordination: Giving | -0.17 | 0.06 | -2.96 | .003 |
| No Gaze Coordination: Showing | -0.16 | 0.06 | -2.71 | .007 |
| No Gaze Coordination: Conventional | -0.08 | 0.06 | -1.42 | .157 |

Table 148

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 12 months (n = 64)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Index-finger Pointing*) | 0.16 | 0.05 | 3.25 | .001 |
| No Gaze Coordination | -0.06 | 0.03 | -2.04 | .042 |
| Open-hand Pointing | -0.10 | 0.07 | -1.55 | .123 |
| Giving | 0.15 | 0.07 | 2.33 | .020 |
| Showing | -0.05 | 0.07 | -0.83 | .405 |
| Conventional | -0.10 | 0.07 | -1.55 | .123 |

Table 149

*Summary of Fixed Effects from a Linear Regression Model Fitting Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 & 12 months (n = 61)*

| | *B* | *SE* | *t* | *p* |
|---|---|---|---|---|
| Intercept (*Gaze Coordinated*: *Index-finger Pointing*) | 0.13 | 0.08 | 1.60 | .111 |
| No Gaze Coordination | 0.16 | 0.09 | 1.77 | .077 |
| Open-hand Pointing | 0.00 | 0.12 | 0.00 | 1.000 |
| Giving | 0.38 | 0.12 | 3.31 | .001 |
| Showing | 0.16 | 0.12 | 1.42 | .156 |
| Conventional | -0.05 | 0.12 | -0.42 | .671 |
| No Gaze Coordination: Open-hand Pointing | -0.16 | 0.13 | -1.25 | .211 |
| No Gaze Coordination: Giving | -0.41 | 0.13 | -3.14 | .002 |
| No Gaze Coordination: Showing | -0.46 | 0.13 | -3.51 | .001 |
| No Gaze Coordination: Conventional | -0.23 | 0.13 | -1.76 | .080 |

Table 150

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 12 months (n = 64)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| Gaze Coordinated: Not | 0.06 | 319 | 2.04 | .042 |
| Open-hand point : Give | -0.25 | 252 | -3.88 | .001 |
| Give : Show | 0.21 | 252 | 3.17 | .015 |
| Give : Conventional | 0.25 | 252 | 3.88 | .001 |

Table 151

*Abridged Summary of Post-hoc Pairwise Comparisons (using Tukey's method of correction) between Frequency of Combinations that were Met with a Caregiver Response by Gaze Coordination and Gesture Type at 11 & 12 months (n = 61)*

| | *b* | *df* | *t* | *p* |
|---|---|---|---|---|
| ***Gaze Coordinated: Not*** | | | | |
| Showing | 0.30 | 300 | 3.20 | .049 |
| | | | | |
| ***Gaze Coordinated*** | | | | |
| Index-finger Pointing: Giving | -0.38 | 240 | -3.31 | .035 |
| Open-hand Pointing: Giving | -0.38 | 240 | -3.31 | .035 |
| Giving: Conventional | 0.43 | 240 | 3.74 | .009 |