Recognizing Communicative Intentions in Infancy

GERGELY CSIBRA

Abstract: I make three related proposals concerning the development of receptive communication in human infants. First, I propose that the presence of communicative intentions can be recognized in others' behaviour before the content of these intentions is accessed or inferred. Second, I claim that such recognition can be achieved by decoding specialized ostensive signals. Third, I argue on empirical bases that, by decoding ostensive signals, human infants are capable of recognizing communicative intentions addressed to them. Thus, learning about actual modes of communication benefits from, and is guided by, infants' preparedness to detect infant-directed ostensive communication.

1. Ostensive Signals

According to the Gricean analysis of meaning (1989), a communicative act intends to fulfil two (or perhaps three) intentions simultaneously. By producing the act, the communicator intends (1) to generate a certain response in the audience, (2) to let the audience recognize the intention specified in (1), and (3) to make the audience fulfil the first intention on the basis of fulfilling the second one. This analysis proposes that, unlike instrumental actions, which can normally be explained by a single intention specifying the desired outcome of the action, communicative acts are generated by (at least) two intentions, one of them being a second-order intention referring to the other one. In everyday terms, Grice's insight here is that a communicator wants not only to convey a message but also that the addressee recognize his intention to do so, and would not be satisfied with the outcome of his action unless this recognition is realized.

Sperber and Wilson (1995) reformulated the same insight in different terms. Although their interpretation of how human communication works differs markedly from that of Grice, they retained the idea of the presence of a dual intention in all communicative acts. In their terms, the communicator's *informative intention* is to achieve a certain effect in the addressee by modifying her cognitive environment, while his *communicative intention* is to make his informative intention manifest for the addressee (or mutually manifest between the communicator and the addressee). Thus, ostensive communication can be achieved when the communicator manages to fulfil his informative intention *by* getting the addressee to recognize

I thank Hanoch Ben-Yami, György Gergely, Ágnes Kovács, Eugenio Parise, and Ernő Téglás for their constructive comments on earlier drafts of this paper.

Address for correspondence: Central European University, Cognitive Development Center, Hattyú u. 14, 1015 Budapest, Hungary.
Email: csibrag@ceu.hu

Mind & Language, Vol. 25, No. 2 April 2010, pp. 141–168.

this intention. In this paper, I will use Sperber and Wilson's terms 'informative intention' and 'communicative intention' to characterize the inherently dual nature of the intention that drives and explains human communicative acts.

Because a communicative intention is a second-order intention with an informative intention in its content, it presupposes the presence of an informative intention. This implies that one cannot have a communicative intention without having a corresponding informative intention. One may object to this by saying that a person can initiate a communicative interaction without having any specific informative intention in hand, and (according to G. Gergely, personal communication) young children often do this. However, it seems that these behaviours can be characterized either as not communicative but instrumental actions, intending to generate certain responses in (see Moore and D'Entremont, 2001, for an attempt to explain early pointing gestures this way), rather than making something manifest for, the addressee, or as proper communicative acts produced by the specific imperative informative intention of engaging the addressee in an enjoyable interaction.

On the other hand, one can clearly have, and act on, an informative intention without an intention to communicate, i.e. without wanting the addressee to recognize his informative intention. There are many examples of such cases of covert communication in the literature, ranging from not caring about whether the informative intention is recognized to hidden authorship of messages. Because of this asymmetry between the informative and communicative intentions, it is appropriate to ask whether the informative intention revealed by an action is accompanied by the second-order intention that would make this action a truly communicative act. For example, it is not easy to test whether young children's early pointing gestures (which are interpreted by adults as proper communicative acts) reflect their intention to make their informative intention manifest (Tomasello *et al.*, 2007), though recent studies suggest that this might be the case (Grosse, 2009).

Turning to the other side of communication, the addressee will also have to deal with the dual intention behind the communicative act. The communicator's intentions could only be fulfilled properly if, in Gricean terms, the addressee produces the intended response, and recognizes that this was the communicator's intention. In other words, the addressee should minimally attribute both a first-order informative intention and a second-order communicative intention to the communicator. (It can be argued that such attributions would not be sufficient, but a third- or fourth-order attribution, with reference to mental states of the self, is needed for proper understanding of a communicative act; see Sperber and Wilson, 1995. As this question is not relevant for the topic of this paper, I will restrict my discussion to the relation between the two intentions that correspond to the ones in Grice's first two clauses.)

Is it possible to understand an action as being driven by an informative intention without attributing the corresponding communicative intention to the actor, mirroring the case of covert communication mentioned above? Certainly, it is. One can recognize that an act has been designed to make manifest something while not making manifest the actor's intention to do so. This kind of understanding of an informative intention without the corresponding second-order communicative intention is, however, not necessarily less demanding cognitively than proper communicative understanding because it is usually accompanied by the attribution of the opposite second-order intention to the actor to hide, or at least not to make mutually manifest, the informative intention.

What about the remaining option of attributing a communicative intention without an informative intention to an action? Above I considered and dismissed the possibility of having a communicative intention without having the informative intention that it would make manifest. However, in case of receptive communication, the content of the informative intention does not have to be evident for the addressee in order to recognize its presence. In other words, one can recognize the intention to make an informative intention manifest without, or before, recovering or inferring the content of the latter intention. The paradigmatic case here is the foreign speaker whose intention to communicate something can be plainly evident to us without having a clue of what she is trying to tell us. This recognition is partial: what is recognized is *that* the communicator has an informative intention, but not *what* this intention is. One can conceptualize this process as the creation of a placeholder for the informative intention, to which the communicator's communicative intention refers, and whose content is to be specified by further cognitive effort.

The first main point of the proposal of this paper is that, from the perspective of cognitive mechanisms, the attribution of communicative and informative intentions can be temporally and procedurally (but not conceptually) separated. It is possible to recognize that one is being addressed by someone else's informative intention before being able to specify what this intention is. Several important implications follow from this proposal. First, with this separation, attributing a second-order communicative intention does not necessarily entail having to entertain the content of two embedded intentions at the same time. One can attribute a communicative intention to the communicator by recognizing the ostensive nature of his action, and then, in a successive process, can attempt to infer the content of the implied informative intention.

Second, the temporal and procedural separation between the recognition of the two intentions may potentially create a so-called binding problem. If it is not the same part of the communicative act that is interpreted as indicating the presence of an informative intention as the one giving evidence of its content, there must be ways to relate these two attributions to each other to preserve the hierarchical structure of the attributed intentions. In other words, the placeholder for the not-yet-specified (first order) informative intention should be represented as an argument of the (second order) communicative intention that is already recognized. I believe that this is not an intractable problem because spatial, temporal and conceptual factors can constrain this binding process.

Third, and most importantly for the point I am trying to make, the separability of the attribution processes that identify communicative intentions and informative intentions gives rise to the possibility of the emergence of specialized signals in production. Thus, a communicator can produce signals that are specifically designed to generate the interpretation that the communicator has a communicative intention addressed to the interpreter. If these signals precede or accompany further actions that give evidence about the content of the corresponding informative intention of the communicator, they could turn those actions into communicative actions. I call these signals, after Sperber and Wilson (1995), *ostensive signals*.

The second main point of my proposal is that the production and reception of ostensive signals could be implemented in a code-based communicative system. This is because, unlike the evidence given about an informative intention, which vastly underdetermines its content and hence requires inferential processes to (re-)construct, ostensive signals simply indicate the presence of a second-order (i.e. communicative) intention referring to an empty placeholder of the corresponding informative intention. In other words, the content of such ostensive signals does not vary across instances of communicative acts and so they do not require effortful inferential processes to be interpreted. Ostensive signals need to indicate only two things: (1) that the source is making manifest of having an informative intention and (2) who is targeted by this intention (who is the addressee). Conventionalized or pre-wired code systems could do this job perfectly.

By emphasizing that a coded (as opposed to inferential) communication system could be used to convey ostension, I do not want to imply that communicative intentions are only, or predominantly, signalled by such codes. The ostensive nature of stimuli can also be inferred from contextual factors. However, while a stimulus may be inferred to be ostensive *because* it captures one's attention (Sperber and Wilson, 1995), specialized ostensive signals command one's attention *because* they code for ostension. Ostensive signals thus create a shortcut for triggering inferential processes that would interpret accompanying actions of the same source.

2. Infants' Sensitivity to Ostensive Signals

The third main point of the proposal of this paper is that human infants readily interpret coded ostensive signals as indicating communicative intentions of their sources. Thus, recognizing communicative intentions (in the sense of noticing their presence rather than of accessing their content) is not the outcome but one of the sources of the development of communicative skills. In order to launch this development, some of the signals that code for ostension should be innately specified. These signals (1) must unambiguously specify that the infant is the addressee of a communicative act, (2) must be discriminable by newborns, and (3) must induce preferential orientation towards their source. There are at least three types of stimuli that satisfy these criteria: direct gaze that generates eye contact, the special intonation pattern used with infants, which is called infant-directed speech or 'motherese', and contingent reactivity to the infant's behaviour in a turn-taking manner.

2.1 Eye Contact

While prolonged eye contact that is not accompanied by further communicative signals is interpreted as intimidation and becomes quickly aversive, making a short eye contact is an ostensive signal in humans and carries a positive value (Kleinke, 1986). This is a signal that, when it is noticed, unambiguously specifies its target and it is ideal for establishing or re-establishing a communicative link between two people. Mutual looking into each other's eyes confirms that the other is 'on line' and that she is the intended addressee of a message. Deliberate avoidance of eye contact is interpreted by social partners as an attempt to escape from engaging in communication, or refusing to be the recipient of an expectable message, for example when one feels guilt, shame, or embarrassment.

Humans are sensitive to the presence of eyes from birth. Batki *et al.* (2000) have shown that human newborns, at an average age of 36 hours, prefer to look at a face with open eyes as opposed to a face with closed eyes. They interpreted this finding as evidence for an innate eye-detection mechanism that detects eyes, as a special stimulus class, in the environment, and orients babies towards them. Such a mechanism explains their result but does not demonstrate preference for eye contact. In contrast, Farroni, Csibra, Simion and Johnson (2002) found that, when they can choose between photographs of faces looking directly at them or looking away, 3-day-old newborns prefer to look at the face that appears to make eye contact with them. This is a very robust effect, unusually strong among studies with neonates, as not a single infant oriented more towards the face with averted eyes than towards the face with direct gaze. This phenomenon suggests that what newborns are sensitive to is not simply the presence of eyes but more specifically the position of the pupils/irises within the eye. The effect can also be demonstrated with schematic faces (Farroni *et al.*, 2004a) but disappears when the faces are turned upside down (Farroni *et al.*, 2006), suggesting that the preferred stimulus for newborns is not simply two eyes with the pupils in their middle, but two eyes with the pupils aligned centrally in the context of a frontal face in a canonical (i.e. upright) position. Note that this is the position that is characteristic to the prototypical mother-infant face-to-face interaction (Watson, 1972). This preference also indicates that face and gaze perception are intimately linked together from birth.

Newborns' preference for face-like patterns is usually interpreted as an adaptive orientation mechanism that ensures that infants will fixate on, and learn about, the most relevant social stimuli in their environment (Johnson and Morton, 1991). There is no doubt that early face preference fulfils this function: newborns recognize their mother's face within days after birth (Pascalis *et al.*, 1995), and infants develop more and more specialized face processing mechanisms during the first year of life. Note, however, that this function does not explain why newborns' face preference is orientation-specific. Adults and children perform much better with upright than upside-down stimuli in face recognition tasks (the so-called 'face-inversion effect'). This can plausibly be explained by the fact that they have had more experience with seeing, and acquired more expertise in recognizing, upright than inverted face orientation. Newborns, however, see faces, including their mother's face, in many different orientations (importantly, breast feeding does not normally occur in the canonical orientation). The fact that newborns do not show selective preference for non-canonical orientations suggests that the canonical orientation of a face provides some extra advantage that makes it worth being preferred. Preparedness to be a recipient of communicative acts could be the extra factor that explains this aspect of early sensitivity, because, when it comes to faces, only an upright face looking at the baby would be considered as an eye-contact stimulus, i.e. an ostensive signal.

Another aspect of newborns' preference for faces confirms further that this innate ability is based on more than a geometric face template to be matched. Gaze perception in humans is extremely sensitive to contrast polarity (Ricciardelli, Baylis and Driver, 2000). This is powerfully demonstrated by the Bogart illusion (Sinha, 2000), in which an inverted polarity ('negative') photograph of a face with eyes averted to one side appears to look towards the opposite direction. This illusion is explained by the perceptual mechanisms that attempt to read gaze direction by identifying the location of a darker spot (pupil) within a lighter area (sclera). Human eyes have a unique morphology that makes large areas of the white sclera visible (Kobayashi and Kohshima, 1997). It is possible that this unique morphology serves a human specific function, namely, to make reading gaze direction easier for our conspecifics (Tomasello *et al.*, 2007). If 'gaze' is identified by the location of a dark spot on lighter background, then a figure that does not have such spots cannot be seen as having 'gaze' and cannot be identified as a stimulus with direct gaze (i.e. eye contact). We have found that newborns' preference for 'upright' face configurations disappear if the contrast polarity of the pattern is reversed so that it would not contain dark elements on light background any more. When, however, the dark elements are provided within the white 'sclera', the preference for the face-like figure returns (Farroni *et al.*, 2005).

Eye contact triggers enhanced attention to the face in adults, children and infants alike. One consequence of this enhanced attention is superior recognition of faces with, than without, direct gaze. Hood *et al.* (2003) found that both children and adults were more accurate in memorizing faces with direct gaze than faces with averted gaze. Similar effects have also been demonstrated with 4- to 5-month-old (Farroni *et al.*, 2007) and 9- to 12-week-old (Blass and Camp, 2001) infants. One can even speculate that the fact that we all become experts of face recognition is a by-product of preferential attention to faces, and especially to the eye region, which may carry important communication signals (Gliga and Csibra, 2007).

The fact that young infants preferentially attend to faces with direct eye gaze does not necessarily entail that they 'like' to make eye contact with others. In fact, it is plausible to assume that a member of a species in which eye contact signals threat rather than ostension would also tend to orient towards such stimuli because it would be adaptive to detect an impending attack. However, plenty of data shows that infants don't just orient toward direct gaze but actually enjoy making eye contact. Three- to six-month-old infants smile less to a person after she has broken eye contact with them, even when she continues to respond to the child contingently (Hains and Muir, 1996). Five-month-olds smile at people who avert their head while maintaining eye contact with the baby (Caron *et al.*, 1997). At the same time, their smile diminishes when the adult moves her eyes only 5 degrees away, to one of their ears (Symons *et al.*, 1998). Smiling as a response to particular stimuli indicates positive affect towards those stimuli, but its evolutionary function lies in its effects on other humans (Watson, 1984). A smile on an infant face prolongs the adult's interactive behaviour that has elicited the smile, which in turn enhances the baby's chance to benefit from the maintained interaction.

2.2 Infant-directed Speech

Perhaps the most obvious communicative signal in humans is the most frequent channel of communication: speech. Although in modern Western societies we are surrounded by media that emit speech (radio, television, loudspeakers in shops, stations and trains, or computers), the human voice is still the most reliable source indicating communication. The human auditory system is specifically tuned to extract human voice and speech from noise. Newborns discriminate between speech and non-speech stimuli, and are biased towards speech (Vouloumanos and Werker, 2007), and this discrimination is subserved by dedicated cortical structures (Peña *et al.*, 2003). Human beings are born with special sensitivity to this particular medium of communication.

However, unlike eye contact, the speech signal does not specify the addressee of a communicative act directly. One cannot detect eye contact without interpreting it as an ostensive signal addressed to the receiver. In contrast, hearing speech does not unambiguously licence the conclusion of being addressed. One may disambiguate the situation from the content of the speech. If the source mentions your name, greets you in a culturally accepted manner, refers to state of affairs relevant to your actual circumstances, or to what you have said or done seconds before, this will assure you that she is talking to you.

However, if you are unable to decode the content of the speech, these latter methods are not available to you. This is the exact situation of a preverbal infant. Even if he recognizes his mother's voice and the intonation and rhythmic pattern of the language seems familiar to him, it will not tell him whether he is the one being addressed. In fact, most of the times an infant can hear speech, he is not the one who is spoken to. This is especially so in traditional societies where children are not raised in separate nursery rooms but are present during the everyday life of the extended family or community. Does that mean that an infant has no way of knowing whether a speech signal is addressed to him unless other communicative cues, like eye contact, are also available? Luckily not, as speakers can provide additional cues to indicate that they are talking to an infant.

Adults, and especially mothers, instinctively alter their prosody when they talk to preverbal infants. The prosody of infant-directed speech, often termed 'motherese', is characterized by higher pitch, broader pitch and amplitude variation, and lower speed than adult-directed speech. These characteristics of infant-directed speech are universal (Fernald, 1992), though culture-specific variations have also been found (e.g. Kitamura *et al.*, 2001). Several functions have been attributed to this distinctive type of speech addressed to infants: it captures infants' attention, it regulates affects, it may play a causal role in language acquisition, or it is just a by-product of the fact that infants are talked to in emotionally charged contexts. I propose that the immediate function of the infant-directed intonation pattern is much simpler: it *makes it manifest that the speech is infant-directed*. In other words, the special prosody associated with motherese indicates to the baby that he is the one to whom the given utterance is addressed, and so it serves as an ostensive signal. This signalling function turns motherese into the sibling of eye contact as being an ostensive signal, which is specifically tailored to infant recipients. If this is the case, we should see that infant-directed speech elicits the same responses as eye contact: easy and fast detection of, preferential orientation to, and positive affect towards the source of such stimuli.

Indeed, two-day-old newborns pay more attention to a source talking to them in infant-directed speech than to a source speaking in an adult-directed way (Cooper and Aslin, 1990), except when the speech is conveyed in the mother's voice, in which case the preference goes the opposite way (Hepper, Scott and Shahidullah, 1993). This seemingly paradoxical pattern of preference can be explained by two independent tendencies: a preference for the mother's voice, and a preference for motherese. Infants are familiarized to their mother's voice in the womb where it reaches them in a distorted version. The mother's body acts as a low-pass filter, which removes most of the phonetic information while preserving the prosodic and intonational aspects of speech. The maternal speech that the foetus is exposed to in the womb, however, is not infant-directed speech and so the newborn will be more familiar with the adult-directed than the infant-directed version of his mother's prosody. Thus, newborns prefer their mother's voice to a stranger's voice (Hepper *et al.*, 1993), a stranger's motherese to a stranger's adult-directed speech (Cooper and Aslin, 1990), but when these two tendencies are pitted against each other, the more familiar adult-directed prosody of the mother wins out over her motherese. This paradoxical preference is quickly rectified: by 1 month of age the two tendencies are in balance, resulting in no preference (Cooper, Abraham, Berman and Staska, 1997), and by 4 months, or possibly by 6 weeks (Mehler, Bertoncini, Barriere and Jassik-Gerschenfeld, 1978) infants prefer when their mother is talking to them in motherese.

It is very likely that sensitivity to the acoustic and/or prosodic characteristics of motherese is innate in humans, though preference for some aspects of it, like the distinctive pitch contour, may emerge later (Cooper and Aslin, 1994). The preference for motherese is present in two-day-old newborns (Cooper and Aslin, 1990), even if they are born to congenitally deaf parents who could not have trained them in special speech patterns (Masataka, 2003). Older infants prefer motherese even if the speech is delivered in a foreign language never heard before (Werker, Pegg and McLeod, 1994), and are more likely to extract motherese than adult-directed speech from acoustic noise (Colombo *et al.*, 1995). Thus, infant-directed speech is very effective in orienting infants to the speaker, and mothers use it to achieve exactly this effect. The more times they try to get their baby's attention unsuccessfully in a row, the more exaggerated motherese they produce the next time. In fact, when a mother initiates an interaction with a 4-month-old, the prosody of her first utterance is not necessarily different from her adult-directed speech. However, if her attempt to get a response from her child fails, she raises the pitch and the pitch range of her subsequent utterance (Masataka, 2003). It is interesting to note that people instinctively use the same kind of exaggerated prosody when they talk to non-infant listeners who are not expected to understand the content of the speech, such as foreigners or dogs (Hirsh-Pasek and Treiman, 1982). In other words, infants' sensitivity to motherese is complemented with adults' tendency to use this kind of prosody whenever the addressee has no means to figure out from the speech content that he or she is being addressed. It is also noteworthy that the use of motherese towards children decreases parallel with their mastering of the mother tongue.

Infants' responses to motherese, just like their responses to eye contact, also have an affective component. Werker and McLeod (1989) had 4.5- and 8-month-old infants responses to infant- and adult-directed speech rated for positive affect and found that both age groups responded more positively to infant-directed speech. They also asked naïve adults to rate the video recordings of infants in terms of friendliness, cuddliness, and likeability. When they attended to infant-directed speech, babies appeared to be more attractive to adults then when they were listening to adult-direct speech. This shows that infants' response to motherese, just like their response to eye contact, fulfils its function: it makes adults repeat their actions and prolong their infant-directed communication.

Infant-directed speech has also been suggested to play a causal role in language acquisition. Slower tempo, lengthened vowels, frequent repetitions and other characteristics of typical motherese may help the child, for example, to segment the speech input. On the other hand, high pitch and large pitch variations represent such a strong distortion of normal speech that they may make it not less but more difficult to extract information from speech (Bard and Anderson, 1983) or to discriminate between vowels (Trainor and Desjardins, 2002). I am sceptical about the claim that the primary function of the particular prosodic features of motherese is to assist language acquisition. Nevertheless, the proposal that infant-directed speech serves as an ostensive signal does imply that it may foster language acquisition as a side effect. If infants preferentially attend to the source of infant-directed speech, and it makes them receptive to any information received in these situations, motherese may have an indirect beneficial effect on acquiring language as well.

2.3 Contingent Responsivity

Imagine that you have just been jailed in a prison and put into solitary confinement. You hear knocks on the wall and recall that prisoners send each other messages via tapping on pipes or walls. Can you figure out whether what you hear is a message addressed to you or if it is just the mice rustling around? Of course, you can. Knock on the wall and see what happens. If the tapping stops while you are knocking and then returns, it seems likely that someone is trying to communicate with you. If you are still unsure, try it again for a couple of times. The more the noise appears the be related to what you are doing, the more certain you can be that there is someone connected to you via this unusual medium. You may still have to figure out the syntax and semantics of the messages embedded in the knocks (not to mention a code system for your own messages), but at least you may confidently assume that there *is* a communicative intention that the noise in the wall attempts to convey specifically to you. Contingent responsivity can be a very powerful ostensive signal.

Just like eye contact, contingent responsivity has not become an ostensive signal accidentally but rather because it is an essential structural property of normal human conversation. The basic conversational structure of a dialogue inherently entails the creation of mutual contingent responsivity, where both partners attempt to adjust the timing of their own communicative actions to the other's communicative actions. Human conversation is usually accompanied by synchronization of behaviour on many levels, of which turn-taking is probably the most fundamental phenomenon. In fact, turn-taking is a kind of interactive contingency that is qualitatively different from temporal synchrony or simultaneous mirroring reactions in that its structure is directly linked to the requirements of communication. Thus, turn-taking contingency is unique among the different levels of synchronization phenomena because it is mandatory, and because it involves complementary rather than reproductory matching of actions.

Are human newborns prepared to pick up such subtle cues from the environment? Studies that subjected newborns to operant conditioning procedures indicate that they are (e.g. Floccia *et al.*, 1997). After all, operant conditioning depends on contingent reinforcement. However, picking up the signal of contingent reactivity may not be as easy as identifying a key stimulus, like eye contact. It is not easy because detecting contingency requires producing a response, and repeating this response several times so that the contingent nature of an incoming stimulus could be verified. Human newborns are capable of producing some rudimentary actions (like head turns) voluntarily but these are rather fragile actions, and their coordinated repetition demands a lot of effort. There is, however, a simple response in which newborns are expert almost right from birth: sucking.

The function of sucking behaviour in infant mammals is to provide them nutrition from the mother's breast. However, the sucking pattern of human infants suggests that nutrition may not be the only function that this behaviour serves (Wolff, 1968). Nutritive sucking, which gets constant reinforcement from milk intake, is characterized by continuous, uninterrupted sucking, which ensures the most efficient mode of nourishment. When an infant mammal sucks a blind nipple (which is called non-nutritive sucking), its sucking pattern alternates between bursts of sucking and longer pauses. However, human infants frequently change from nutritive to non-nutritive mode of sucking even when they are on their mother's breast or receive milk from a bottle. In fact, infants start to insert breaks into their sucking pattern soon after the onset of feeding without any physiologically obvious reasons (Masataka, 2003). They can breathe normally while sucking, and are neither tired nor full of food when they pause sucking.

There is, however, a consistent effect on their cessation of sucking. Mothers tend to respond to such pauses by jiggling the child or the bottle, which they never do while the baby is sucking (Kaye, 1977). When asked, mothers say that they do this to encourage the infant to suck more, but the amount of milk intake is not affected by this maternal behaviour. In fact, sucking returns only after the mother stops jiggling, and if she jiggles the child during sucking, it interrupts sucking (Kaye and Wells, 1980). This pattern of interaction between mother and child resembles the turn-taking pattern of normal human conversation: sucking and jiggling inhibit each other, while finishing either of these behaviours induces a response from the partner. One can even call this type of synchronization of behaviour between mother and infant a 'proto-conversation' (Bateson, 1979). However, there is no indication that information exchange would take place during this interaction.

What is then the function of this synchronization? If this phenomenon were simply a by-product of sucking and nursing behaviours, it would also occur in other species but there is no evidence for this. Additionally, human infants seem to expect a response whenever they pause sucking. Masataka (1993; cited in Masataka, 2003) experimentally compared the sucking pattern of two-week-old infants whose pauses were contingently responded to by jiggling with groups of infants who received no jiggling during sucking or were jiggled randomly, i.e. non-contingent with their pauses. He found that, although the total duration of sucking was not different across groups, pauses occurred more frequently when they were *not* reciprocated by a response. It seems as if infants may have intended to test whether contingent responsivity characterized their mother's responses, and the only way to make sure that it did was to emit a signal (in this case, a pause) and wait for a response. Interestingly, the first non-cry vocalizations (cooing) at 8 weeks of age also occurred during pauses during sucking, especially if there was no response from the mother, as if babies demanded a response to their signal, or tested whether the maternal response may in fact be controlled by another type of response (apart from pausing) that they were now capable of producing. Contingent responsivity, therefore, is not just an emergent pattern of the earliest mother-infant interactions but a situation that is actively sought by the infant.

As infants' repertoire of reliably executed voluntary movements, and especially facial expressions, extends, they have more and more chance to detect and test contingent responsivity from the environment. A large body of literature describes a specific aspect of early mother-infant interaction, which is most characteristic between 2 and 6 months of age. Mothers and babies at this age are frequently observed as they are engaged in face to face play situations. They vocalize, smile at, gesture to and touch each other, and these acts are carefully coordinated in more than one dimension. The temporal pattern of this type of coordination resembles conversational turn taking, and the phenomenon was again dubbed as proto-conversation (Bateson, 1979).

Several lines of evidence suggest that infants are not just mere reactive participants in these turn taking events but actively seek out and try to prolong these situations. Murray and Trevarthen (1985) developed a technique to test whether infants were sensitive to the inherent contingency involved in these interactions. They established a video communication link between mother and infant, who were in separate rooms, whereby both parties could see and hear the other via a television set. After an initial phase of communication through this medium, the researchers switched off the live link from the mother to the infant, and replaced it with a replay of the mother's behaviour during the earlier phase. Thus, in this second phase infants saw and heard their mother communicating with them, but her behaviour was literally de-coupled from what the babies were doing. Murray and Trevarthen reported that 6 to 12-week-old infants changed their behaviour dramatically in this second phase. They looked less toward the television set showing their mother, smiled less at her and frowned much more than they did in the first phase during live connection. This finding indicates that young infants are indeed sensitive to the contingent structure of these interactions and prefer them to random stimulation (Bigelow, 2001; Nadel *et al.*, 1999; Stormark and Braarud, 2004; see also Bigelow and Rochat, 2006).

These types of early interactions received a lot of attention and were proposed to provide evidence for the innate sociability of human infants. Beyond proto-conversation, the phenomenon has been termed 'dyadic interaction' (Stern, 1977) or 'primary intersubjectivity' (Trevarthen, 1979), and has been attributed various functions, including sharing mental states (Trevarthen and Aitken, 2001), 'affect attunement' (Stern, 1985), 'mutual affect regulation' (Gianino and Tronick, 1992), or promoting identification and attachment. I propose that these early dyadic interactions primarily serve a simpler function: identifying and responding to an ostensive signal, and practising this process.

The fact that young infants enjoy contingent interactions even in the absence of another human being also suggests that the sensitivity to contingent responsivity does not necessarily imply a sharing of emotional states or identification with the source of contingency. John Watson (1972) observed that, after a training phase, 2-month-old infants started to smile and coo at a mobile that they could control by their leg kick. He argued that these positive reactions to an inanimate object were not simply signs of pleasure of being in control but should be interpreted in terms of the effect they may exert on the environment. Since the immediate effect of this response is a prolongation of the action that elicited it, we may assume that this is the primary function of the smiling response. Infants seek out, and try to prolong, contingent interactions because they serve as ostensive signals.

2.4 Learning About Ostensive Signals

As this short review suggests, human infants are well prepared to pick up ostensive signals. The sensitivity to these signals is innate and can be demonstrated within days after birth. Eye contact is visual, motherese is auditory, and contingency is an amodal signal of communicative intent, *yet all* evoke the same response: orientation towards the source of the cue. The fact that these signals can come from multiple modalities ensures that if any of them is blocked because of environmental circumstances (occluders, noise, etc.), or disabilities of either the infant or the parents (e.g. blindness), further channels may still be available for identifying ostension.

It is noteworthy that all kinds of ostensive signals tend to elicit positive affective responses, primarily smiling. Infant researchers usually emphasize the interpersonal aspect of these responses and their mutuality. Because ostensive signals act as rewards for infants, stimuli associated with them also acquire a reinforcement value, and become secondary reinforcers themselves. This is how the infant's own name becomes an ostensive signal itself: it occurs in synchrony with eye contact and contingent play, and usually uttered in infant-directed prosody. It is worth noting that this is the earliest word infants can recognize around 4.5 months of age (Mandel, Jusczyk and Pisoni, 1996; Newman, 2005), but it is unlikely that it functions as a lexical item referring to the self at this age. Instead, this word must have acquired a special status via strong association with the other ostensive signals like eye contact or motherese, and its 'meaning' for an infant is entirely defined by pragmatic rather than semantic factors. From about 6 months, infants spontaneously turn their head when their name is called, showing that they interpret this word as a vocative.

In fact, the four ostensive signals discussed above tend to occur together: the mother looks into her child's eyes, starts a contingent play with him, and calls his name in motherese. Flooding infants with all these stimuli concurrently not just helps them to recognize the presence of a communicative intention but also serves as a calibration process. As we have seen, the innate specification of ostensive signals is rather rudimentary: two darker blobs at the upper parts of objects, high pitched and slow speech, and a wide range of levels of contingency. During the first couple of months of life, the wide range of stimuli that meet these specifications is gradually narrowed down to the actual range of variation that the infant encounters. At two months of age a schematic face is no longer as attractive as a real one (Johnson and Morton, 1991), though the face of a monkey is still within the preferred range until 6 months (Pascalis, de Haan and Nelson, 2002). A similar 'perceptual narrowing' can also be observed in the preference for contingent stimuli (Bigelow, 1998), and the change of the optimal parameters of infant-directed speech during the first year of life could also be explained by a calibration process that tunes infants' sensitivity to the actual cultural variants present in their environment. The co-occurrence of ostensive signals during infant-directed communication helps babies verify that an ambiguous stimulus (for example, low pitched but slow speech) is likely to be an ostensive signal specific to the particular source. A similar learning process could explain how new stimuli, like the infant's name, acquire the status of ostensive signals or how new parameters are added to the ones pre-specified. For example, while newborns prefer eye contact only in canonical frontal faces, by 4 months they show the same responses to direct gaze on a laterally oriented face (Farroni *et al.*, 2004b; Grossmann *et al.*, 2008), and to accompanying facial signals, such as an eyebrow-raise (Grossmann *et al.*, 2008). During the learning process that establish these new stimuli as ostensive signals, simultaneous signals from other modalities can help the infants figure out the relevant invariances and narrow the range of preference down to those stimuli that tend to co-occur with other ostensive signals.

3. Expectations Generated by Ostensive Signals

Infants' initial responses to signals of communicative intention can be interpreted as preparedness to create optimal conditions for receiving the source's manifestations about the content of the corresponding informative intention. They orient towards the source of these signals: they inspect the face making eye contact with them, turn towards the perceived auditory origin of infant-directed speech and/or the calling of their name, and look for the source that responds to them contingently. They also produce signals, like smiling and cooing, that act as acknowledgements, letting the source know that they have decoded the ostensive signals and that they are ready to receive further evidence about the informative intention of the source. However, in order to argue that these stimuli serve not just as attentional cues but also as ostensive signals, it is critical to show that they make an effort to fill the empty placeholder of an informative intention that ostensive signals are supposed to set up.

Interpreting human communication is an inferential process that takes into account many factors, including the history of the conversation, and common grounds in episodic events and in semantic background, when reconstructing the informative intention of the communicator. Much of the information that adults use in inferring the meaning of communicative signals is not available to young infants because of their thin background knowledge, lack of linguistic skills and insufficient cognitive resources, such as memory. Yet, there must be constraints that guide infants' inferences about the content of communication addressed to them.

I propose that infants apply several assumptions in order to interpret ostensive communication directed to them. These assumptions will constrain the search space that infants survey in the effort to attribute an informative intention to the communicator, or, in other words, they implement biases that will influence the interpretation of inherently ambiguous communicative acts. I discuss two such assumptions here (there must be many more of them): referential expectation and the assumption of genericity. Both of these assumptions deserve a longer treatment, which would go beyond the scope of this paper. Here I only outline the evidence that suggests that these assumptions are deployed when infants find themselves addressed by ostensive signals.

3.1 Referential Expectation

A widespread view among researchers of infant communication is that referential communication does not emerge before the second half of the first year of life. Young infants are said to be restricted to dyadic and affective communication, characterized by 'primary intersubjectivity', which does not involve objects external to the communicative partners (e.g. Adamson and McArthur, 1995; Butterworth, 2004; Masataka, 2003; Tomasello, 1999; Trevarthen and Aitken, 2001). While it is certainly true that the frequency of overtly triadic communication (which includes the infant, an adult, and an object in their shared attention) increases enormously during the first year of life, it is not entirely absent in the early months, as evidenced by recent studies (Parise *et al.*, 2008; Reid and Striano, 2005; Striano and Stahl, 2005; Tremblay and Rovira, 2007).

Gaze following is one of the behaviours that establish a triadic situation between a person, an object or event she is looking at, and another person who follows her gaze to the object/event. Infants spontaneously follow others' gaze, especially from 9 months of age. The 'lean' interpretation of this behaviour holds that this is an operant response performed to get a reward (Corkum and Moore, 1998), while 'rich' interpretations construe the phenomenon as evidence of infants' understanding of perception and attention (Baron-Cohen, 1991). An alternative explanation of this behaviour is that it reflects the assumption that gaze shifts are potential deictic referential signals of communicators (Csibra, unpublished manuscript). If this is the case, gaze following should be strongly tied to situations in which the source of the gaze is a communicator, that is, when ostensive signals indicating an infant-directed communicative intention are present.

Although gaze following is most evident in spontaneous behaviour towards the end of the first year of life, more sensitive measures reveal a tendency for both overt and covert gaze following in very young infants. (Covert gaze following, or 'gaze cueing', is the phenomenon when an observed eye movement causes a shift of the observer's locus of attention to the same direction.) Even newborns follow the gaze shift on a face (Farroni *et al.*, 2004a), and so do 3- to 6-month-old infants (D'Entremont *et al.*, 1997; Hood *et al.*, 1998). Crucially, covert gaze following occurs in 4-month-olds only when the gaze shift is preceded by eye contact (Farroni *et al.*, 2003). In fact, a clear head turn, which has been shown to produce overt gaze-following in infants as young as 6 months (Gredebäck *et al.*, 2008), elicits this behaviour only if it is preceded by an ostensive signal such as eye contact, or infant-directed greeting (Senju and Csibra, 2008), or contingent responsivity (Deligianni *et al.*, in preparation). Ostensive signals have been shown to induce gaze following in older infants as well (e.g. Johnson *et al.*, 1998), and such a response is more likely to occur if the gaze shift is accompanied by further communicative signals, such as verbalization and pointing (Flom *et al.*, 2004). In fact, infants seem only to detect the correspondence between gaze shift direction and potential target objects when a period of direct gaze (i.e., eye contact) precedes the gaze shift (Senju, Csibra and Johnson, 2008).

Gaze following is not the only behaviour that indicates the strong tie between the presence of ostensive signals and referential interpretation of actions in infants. In ostensive contexts, infants do not just follow others' gaze but expect to find a referent (Csibra and Volein, 2008), interpret verbal labels as co-referring to the deictically highlighted object (Gliga and Csibra, 2009), and find hidden objects signalled by referential looking or pointing in a search game (Behne *et al.*, 2005). All these findings point to the conclusion that ostensive communicative signals specifically provoke a search for potential referents, which, for pre-verbal infants, are best specified by non-verbal deictic signals. Note that most of the above studies were carefully designed to exclude the possibility that the directional signals carried a communicative meaning by themselves, as they were physically the same in ostensive and non-ostensive contexts. Thus, they acquired a communicative ('non-natural': Grice, 1957) meaning by the preceding or accompanying ostensive signals, illustrating infants' expectation for referential content.

3.2 Genericity Assumption

Children learn a lot by simply observing adults' actions, preferences, object use etc. However, their learning pattern changes qualitatively when they have the opportunity to acquire the same information by communication directed at them. This is evidenced by what actions they imitate (Southgate *et al.*, 2009), how they evaluate sample composition (Rhodes *et al.*, in press), and how much they explore novel objects (Bonawitz *et al.*, 2009). One way to characterize the stance children take when they are specifically targeted by adult communication is that they are biased to assume generalizability of the content of the informative intention signified by the ostensive signals (Csibra and Gergely, 2009).

For example, 18-month-old infants learn about people's personal preferences by observing their object-directed emotional expressions (Repacholi and Gopnik, 1997) whether or not these displays are preceded by ostensive signals (Egyed *et al.*, 2007). However, they seem to learn different things in the two situations. When they observe the emotional expressions in a non-communicative context, they do not think that the actor's personal preference extends to other people. In contrast, when the emotional expressions are presented *for them* ostensively, they apparently think that the same preference applies to others as well. Thus, they generalize the object valence information acquired from a single person to everyone, or, in other words, they treat this information as the property of the object and/or its kind rather than of the actor. This suggests that, for infants, the presence of ostensive signals licenses the inductive leap from a single observation to a more general conclusion that extends to other people and/or other objects of the same kind (Egyed *et al.*, 2007).

Similar effects of inductive generalization or kind-based individuation are reported in studies that taught infants verbal labels for objects (Xu, 2002). It seems, however, that while verbal labelling may not be, the presence of ostensive signals is necessary for making infants treat novel information about an object as a potentially generic property of its kind. When the hidden functional properties of novel artefacts are ostensively demonstrated for 10-month-olds, they tend to represent those objects in terms of their functional kind and infer the numerosity of objects at a hidden location accordingly. In the absence of ostensive signals, however, they fail to use the same causal information for object individuation (Futó *et al.*, in preparation). A bias to receive generic information about objects in ostensive contexts can even be detected in younger babies. When 9-month-olds perceive a novel object as the referent of ostensive communication, they remember better its permanent, kind-relevant properties, such as its shape, than its transient properties, such as its location. In contrast, when infants' attention is drawn to the same object by non-communicative means, they tend to focus on its location more than on its visual features (Yoon *et al.*, 2008).

Just like referential expectation, the assumption of genericity that characterizes infants' effort to make sense of communication directed to them represents a processing bias rather than an irrefutable presumption. Children can and do acquire non-generalizable information by communication that is restricted to the 'here and now'. But unless they can reconstruct from some parts of the adult's manifestation a local, episodic content for the informative intention, they assume, by default, that they are supposed to learn something generic. This assumption may lead to misinterpretation of communication and to the characteristic errors that infants commit in search tasks before the age when they understand the episodic nature of the game (Topál *et al.*, 2008).

4. Potential Challenges

In this paper, I propose that infants' sensitivity and responses to ostensive signals demonstrate that they recognize others' communicative intentions and make them search for the potential content of the corresponding informative intentions. This perspective requires the reinterpretation of communicative development of human infants, and I anticipate several challenges to this account. Here I outline my response to three of the potential objections that (1) claim non-human-specificity of sensitivity to ostensive signals, (2) offer alternative explanations for infants' responses, or (3) express doubts about recognition of communicative intentions in infancy.

4.1 Human Specificity

The stimuli that are interpreted above as innately pre-wired ostensive signals will also generate preferential attention in many species. Eye contact carries important social information in non-human mammals because it signals threat and/or dominance. Similarly, all species that can be subjected to operant learning procedures must be sensitive to R-S contingencies. A potential challenge here is that infants' sensitivity to, and preference for, these signals may simply represent a species-universal tendency to pick up naturally meaningful signals from the environment.

My response to this challenge is simply that the same signals can carry different meanings for different species. As far as it can be decided on the scarce data available, the development of sensitivity to these signals differs across species. A study by Myowa-Yamakoshi *et al.* (2003), for example, found preference for human direct gaze stimuli in three infant chimpanzees who were very familiar with humans. These chimpanzees selectively chose to look at a human face with eyes open rather than eyes closed, and a face with direct gaze rather than averted gaze, but only if the facial features were arranged normally (not scrambled across the face). This is reminiscent of the human newborn's preference for eye contact, but some aspects of these findings suggests that the underlying mechanisms might not be the same. For example, the preference for direct gaze is present at birth in human infants (Farroni *et al.*, 2002, 2004a, 2005), while the chimpanzees in this study were 10 to 32 weeks old. In addition, the chimps equally preferred direct gaze in laterally oriented (3/4 view) faces as well, a feat that human infants do not achieve before they are 4 months old (Farroni *et al.*, 2004b). Recently it has also been reported that rhesus monkey mother-infant pairs are involved in face-to-face communication, including exchange of mutual gaze and lipsmacking, which resembles human mother-infant interactions (Ferrari *et al.*, 2009). It is important to add, however, that these interactions dramatically decrease after the first month of life, which suggests that they have a role in establishing mother-infant affiliation rather than serving to establish a stable communication system.

Third, non-human animals' responses to ostensive signals do not suggest that they interpret these stimuli as indicating that further information is to be expected that would enable to infer the source's informative intention. Many non-human species, like dolphins (Pack and Herman, 2004), goats (Kaminski, Riedel, Call and Tomasello, 2005), ravens (Bugnyar, Ströwe and Heinrich, 2004), monkeys and apes (Emery, 2000; Itakura, 2004) follow the gaze of others. However, they hardly do so in infancy (Ferrari *et al.*, 2000; Tomasello *et al.*, 2001) and do not do so preferentially in ostensive contexts. In other words, non-human animals orient towards stimuli that attract their attention and readily interpret others' meaningful (for example, goal-directed) behaviour, but do not expect that certain attention-getting stimuli indicate that the source's behaviour would carry meaning beyond its natural interpretation. In other words, they do not recognize those stimuli as ostensive signals.

There is, however, one notable exception to this conclusion. Domesticated dogs are well known of their readiness to respond to human communicative signals, including ostensive signals that indicate that they are being instructed by a human, and referential signals, like pointing and gazing (Kaminski, 2009). Such sensitivity to human non-verbal communication may even lead them to commit similar search errors to human infants, though it does not seem to be coupled with an assumption of genericity (Topál *et al.*, 2009). Dogs' sensitivity to human ostensive signals can be well explained as an evolutionary adaptation to the human environment in which dogs have been bred for thousands of years (Miklósi, 2008).

4.2 Alternative Explanations

A frequent criticism against the kind of proposals that I advance in this paper is that the phenomena in hand can also be explained by simpler (even 'lower level') psychological mechanisms. For example, one can say that infants' responses to ostensive signals simply represent their increased attention, and it is unnecessary to invoke concepts such as recognition of communicative intentions to account for them. My response to this challenge is that the suggestion that processes of selective attention are involved in the responses to ostensive signals is not an alternative explanation but a fundamental constituent of the proposal I defend here. Ostensive signals, if recognized as such, must command the attention of the addressee to the source in order to perceive the further manifestations that may reveal its informative intention. However, increased attention does not explain all aspects of infants' responses to ostensive signals. For example, it does not specify why infants smile at the source of these stimuli when they do not do so at other attention-grabbing events, such a light flash.

A more refined challenge could then claim that the stimuli that I identified as ostensive signals carry reward value and hence they attract infants' attention and produce a positive emotional response. While this explanation accounts for infants' emotional response, it does not tell us why these stimuli are rewarding. Eye contact, motherese or contingent responses do not feed or protect the baby, and their attractiveness cannot be explained by secondary reinforcement, as they are preferred already at birth. However, if these stimuli are recognized as ostensive signals, they should be attractive because they signal the potential acquisition of a commodity that is valuable itself: information. In addition, this interpretation explains why ostensive signals elicit the *expression*, rather than just the feeling, of positive emotion in the form of smiles: it provides feedback to the communicator.

Perhaps then ostensive signals play a role in affective attachment and parent-infant bonding, which would account for the reciprocal feedback that infants produce in terms of smiles. While emotionally charged interactions between parents and babies will certainly contribute to their bonding, such a function would not explain why infants' responses are promiscuous in terms of the source of these signals. Even 4-day-old newborns show some recognition of their mother (Pascalis *et al.*, 1995), but infants readily respond to ostensive signals even when they originate from unfamiliar people or even from non-human sources (Watson, 1972).

A further potential proposal could be that these signals simply indicate the presence of conspecifics, which, in a social species like humans, is important information, and could trigger attention and affiliative responses of the kind that infants produce. The problem with this explanation is that ostensive signals are not the only, and perhaps not even the best, indicators of social partners. Upside-down faces, people looking some other direction, or any kind of speech could, and do, signal the presence of conspecifics for infants but do not produce the same responses as ostensive signals.

In addition, none of the above alternative accounts offers an explanation for why infants tend to interpret further stimuli originating from the source of the ostensive signals differently from those experienced in a non-ostensive context. The expectations of referentiality and genericity suggest that infants' primary reactions to ostensive signals, which are implemented in the general cognitive mechanisms of attention and stimulus evaluation, serve a functional role in optimizing the input conditions for receiving further stimuli from the same source. In other words, infants act as if they have created an empty placeholder to be filled.

4.3 Does Decoding Ostensive Signals Amount to Recognition of Communicative Intentions?

I have argued above that infants, by decoding ostensive signals, recognize the communicative intentions of communicators as evidenced by their responses to these signals and to the following stimuli. But am I entitled to call this process 'recognition'? Even if you accept that infants' responses are tied to signals that reliably indicate infant-directed communicative acts and that their reactions and expectations help them to infer the likely content of the informative intention, you may say that this does not amount to recognition that others have communicative intentions.

What is missing here? Some would say that recognition implies awareness or phenomenal experience of identifying an ostensive signal as such. I do not see why this would be a requirement for using this term. After all, even adults are not always aware of the meaning of stimuli, including ostensive signals, to which they respond appropriately. Another form of this objection is that the term 'recognition' implies having a concept of COMMUNICATIVE INTENTION, which infants are unlikely to possess. I think that the evidence I have reviewed in this paper suggests that they do possess such a concept, though the identification of the instances of this concept is initially restricted to cases when the infants themselves, and not others, are addressed. Thus, infants' concept of COMMUNICATIVE INTENTION may not be as rich as that of adults and may not be accessible to metacognitive processes, but it is sufficient for recognizing certain instances of such a concept in the world and for invoking the appropriate inferential processes about them. In other words, they possess such a concept in the same sense as they possess the concept of OBJECT (Carey, 2009).

There is, however, a question that one can legitimately ask about infants' recognition of communicative intentions. Attributing a communicative intention is attributing a second-order intention, even when it refers to an empty placeholder (because the content of the first-order informative intention, embedded in the communicative intention, has not yet been identified). Is there any evidence that infants recognize ostensive signals as indicating a second-order intention referring to another specific intention rather than taking them as reliable signs for anticipating further informative stimuli? How do we know that infants bind an ostensive signal conceptually, rather than serially and spatially, to the subsequent stimuli that they get to anticipate? As far as I can see, no empirical study has addressed this question yet. Until such data are available, it remains a hypothesis, though a plausible one, that infants' responses to ostensive signals provide evidence for recognizing communicative intentions.

5. Conclusions

I have argued above that (1) communicative intentions can be manifested by specialized ostensive signals, and (2) human infants recognize some of these signals as such from birth. The recognition of communicative intentions can help them to comprehend infant-directed communication by triggering inferential processes that identify and interpret manifestations of the informative intentions whose presence is implied by the ostensive signals. These inferential processes might be implemented in specialized comprehension modules (Sperber and Wilson, 2002) or in more general mechanisms that are constrained by specific biases (such as the assumptions of referentiality and genericity).

My proposal is not intended to solve dilemmas about communication and mental state attribution in infancy (see e.g. Breheny, 2006). Instead, it offers a different way to think about the development of early comprehension. Decoding an ostensive signal entails attributing an intention (in fact, a second-order intention) to communicators from the outset. In this perspective, infants' comprehension skills develop slowly not because they need much experience to understand what communicative intentions, or generally mental states, are, but because they need time to learn about conventional modes of communication, to refine inferential processes for the interpretation of communicative acts, and to figure out when mental states attributed on other, non-communicative bases should be considered in the reconstruction of informative intentions (Southgate, Chevallier and Csibra, in press). This development is driven by infants' capability to recognize communicative intentions by decoding ostensive signals.

Central European University
Budapest

References

Adamson, L. and McArthur, D. 1995: Joint attention, affect, and culture. In C. Moore and P. Dunham (eds), *Joint Attention: Its Origins and role in Development*. Hillsdale, NJ: Lawrence Erlbaum, 205–21.

Bard, E. G. and Anderson, A. H. 1983: The unintelligibility of speech to children. *Journal of Child Language*, 10, 265–92.

Baron-Cohen, S. 1991: Precursors to a theory of mind: understanding attention in others. In A. Whiten (ed.), *Natural Theories of Mind. Evolution, Development, and Simulation of Everyday Mindreading*. Oxford: Blackwell.

Bateson, M. C. 1979: The epigenesis of conversational interaction: a personal account of research and development. In M. Bullowa (ed.), *Before Speech: The Beginnings of Human Communication*. Cambridge: Cambridge University Press, 63–77.

Batki, A., Baron-Cohen, S., Wheelwright, S., Connelan, J. and Ahluwalia, J. 2000: Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, 23, 223–29.

Behne, T., Carpenter, M. and Tomasello, M. 2005: One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8, 492–99.

Bigelow, A. E. 1998: Infants' sensitivity to familiar imperfect contingencies in social interaction. *Infant Behavior and Development*, 21, 149–62.

Bigelow, A. E. 2001: Discovering self through other: infants' preference for social contingency. *Bulletin of the Menninger Clinic*, 65, 335–46.

Bigelow, A. E. and Rochat, P. 2006: Two-month-old infants' sensitivity to social contingency in mother-infant and stranger-infant interaction. *Infancy*, 9, 313–25.

Blass, E. M. and Camp, C. A. 2001: The ontogeny of face recognition: eye contact and sweet taste induce face preference in 9- and 12-week-old human infants. *Developmental Psychology*, 37, 762–74.

Bonawitz, E., Shafto, P., Gweon, H,, Chang, I., Katz, S. and Schulz, L. 2009: The double-edged sword of pedagogy: modeling the effect of pedagogical contexts on preschoolers' exploratory play. *Proceedings of the Thirty-first Cognitive Science Society*.

Breheny, R. 2006: Communication and folk psychology. *Mind & Language*, 21, 74–107.

Bugnyar, T., Stöwe, M. and Heinrich, B. 2004: Ravens, *Corvus corax*, follow gaze direction of humans around obstacles. *Proceeding of the Royal Society B: Biological Sciences, B* 271, 1331–36.

Butterworth, G. 2004: Joint visual attention in infancy. In G. Bremner and A. Slater (eds), *Theories of Infant Development*. Oxford: Blackwell, 317–54.

Caron, A. J., Caron, R., Roberts, J. and Brooks, R. 1997: Infant sensitivity to deviations in dynamic facial-vocal displays: the role of eye regard. *Developmental Psychology*, 33, 802–13.

Carey, S. 2009: *The Origin of Concepts*. Oxford: Oxford University Press.

Colombo, J., Frick, J. E., Ryther, J. S., Coldren, J. T. and Mitchell, D. W. 1995: Infants' detection of analogs of 'motherese' in noise. *Merrill-Palmer Quarterly*, 41, 104–13.

Cooper, R. P., Abraham, J., Berman, S. and Staska, M. 1997: The development of infant's preference for motherese. *Infant Behavior and Development*, 20, 477–88.

Cooper, R. P. and Aslin, R. N. 1990: Preference for infant-directed speech in the first month after birth. *Child Development*, 61, 1584–95.

Cooper, R. P. and Aslin, R. N. 1994: Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development*, 65, 1663–77.

Corkum, V. and Moore, C. 1998: The origins of joint visual attention in infants. *Developmental Psychology*, 34, 28–38.

Csibra, G. and Gergely, G. 2009: Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–53.

Csibra, G. and Volein, A. 2008: Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology*, 26, 1–11.

Deligianni, F., Senju, A., Gergely, G. and Csibra, G. in preparation: Gaze-contingent objects elicit the illusion of communication in preverbal infants.

D'Entermont, B., Hains, S. M. J. and Muir, D. W. 1997: A demonstration of gaze following in 3-to 6- months-olds. *Infant Behavior and Development*, 20, 569–72.

Egyed, K., Király, I., Krekó, K., Kupán, K. and Gergely, G. 2007: Understanding object-referential attitude expressions in 18-month-olds: the interpretation switching function of ostensive-communicative cues. Poster presented at the Biennial Meeting of the SRCD, Boston, August 2007.

Emery, N. J. 2000: The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24, 581–604.

Farroni, T., Csibra, G., Simion, F. and Johnson, M. H. 2002: Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the USA*, 99, 9602–05.

Farroni, T., Johnson, M. H. and Csibra, G., 2004b: Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, 16, 1320–26.

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D. and Csibra, G. 2005: Newborns' preference for face-relevant stimuli: effects of contrast polarity. *Proceedings of the National Academy of Sciences of the USA*, 102, 17245–50.

Farroni, T., Mansfield, E. M., Lai, C. and Johnson, M. H. 2003: Infants perceiving and acting on the eyes: tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, 85, 199–212.

Farroni, T., Massaccesi, S., Menon, E. and Johnson, M.H. 2007: Direct gaze modulates face recognition in young infants. *Cognition*, 102, 396–404.

Farroni, T., Massaccesi, S., Pivitori, D., Simion, F. and Johnson, M. H. 2004a: Gaze following in newborns. *Infancy*, 5, 39–60.

Farroni, T., Menon, E. and Johnson, M. H. 2006: Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology*, 95, 298–308.

Fernald A. 1992: Human maternal vocalizations to infants as biologically relevant signals: an evolutionary perspective. In J. H. Barkow, L. Cosmides and J. Tooby (eds), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press, 391–428.

Ferrari, P. F., Kohler, E., Fogassi, L. and Gallese, V. 2000: The ability to follow eye gaze and its emergence during development in macaque monkeys. *The Proceedings of the National Academy of Sciences*, 97, 13997–14002.

Ferrari, P. F., Paukner, A., Ionica, C. and Suomi, S. J. 2009: Reciprocal face-to-face communication between rhesus macaque mothers and their newborn infants. *Current Biology*, 19, 1768–72.

Floccia, C., Christophe, A. and Bertoncini, J. 1997: High-amplitude sucking and newborns: the quest for underlying mechanisms. *Journal of Experimental Child Psychology*, 64, 175–98.

Flom, R., Deák, G. O., Phill, C. G. and Pick, A. D. 2004: Nine-month-olds' shared visual attention as a function of gesture and object location. *Infant Behavior and Development*, 27, 181–94.

Futó, J., Téglás, E., Csibra, G. and Gergely, G. in preparation: Communicative function demonstration induces kind-based artifact representation in preverbal infants.

Gianino, A. and Tronick, E. Z. 1992: The mutual regulation model: the infant's self and interactive regulation a coping and defensive capacities. In: T. M. Field, P. M. McCabe and N. Schneiderman (eds), *Stress and Coping across Development*. Hillsdale, NJ: Lawrence Erlbaum, 47–68.

Gliga, T. and Csibra, G. 2007: Seeing the face through the eyes: a developmental perspective on face expertise. *Progress in Brain Research*, 164, 323–39.

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

Gredebäck, G., Theuring, C., Hauf, P. and Kenward, B. 2008: The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy*, 13, 533–43.

Grice, H. P. 1957: Meaning. *The Philosophical Review*, 64, 377–88.

Grice, P. 1989: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grosse, G. 2009: Communicative intentions in infancy: some empirical evidence. Paper presented at the 17th Annual Conference of the European Society for Philosophy and Psychology, August 2009, Budapest, Hungary.

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C. and Csibra, G. 2008: Early cortical specialization for face-to-face communication in human infants. *Proceedings of the Royal Society, London, B*, 275, 2803–11.

Hains, S. M. J. and Muir, D. W. 1996: Infant sensitivity to adult eye direction information. *Child Development*, 67, 1940–51.

Hepper, P. G., Scott, D. and Shahidullah., S. 1993: Newborn and fetal response to maternal voice. *Journal of Reproductive and Infant Psychology*, 11, 147–53.

Hirsh-Pasek, K. and Treiman, R. 1982: Doggerel: motherese in a new context. *Journal of Child Language*, 9, 229–37.

Hood, B. M., Macrae, C. N., Cole-Davies, V. and Dias, M. 2003: Eye remember you! The effects of gaze direction on face recognition in children and adults. *Developmental Science*, 6, 69–73.

Hood, B. M., Willen, J. D. and Driver, J. 1998: Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, 9, 131–4.

Itakura, S. 2004: Gaze-following and joint visual attention in nonhuman animals. *Japanese Psychological Research*, 46(3), 216–26.

Johnson, M. H. and Morton, J. 1991: *Biology and Cognitive Development: The Case of Face Recognition*. Oxford: Blackwell.

Johnson, S., Slaughter, V. and Carey, S. 1998: Whose gaze will infant follow? The elicitation of gaze following in 12-months-olds. *Developmental Psychology*, 1, 233–38.

Kaminski, J. 2009: Dogs *Canis familiaris*) are adapted to receive human communication. In A. Berthoz and Y. Christen (eds), *Neurobiology of 'Umwelt': How Living Beings Perceive the World*. Berlin: Springer, 103–7.

Kaminski, J., Riedel, J., Call, J. and Tomasello, M. 2005: Domestic goats Capra hircus) follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, 69, 11–18.

Kaye, K. 1977: Towards the origin of dialogue. In H. R. Schaffer (ed.), *Studies in Mother-Infant Interaction*. New York: Academic Press, 89–117.

Kaye, K. and Wells, A. J. 1980: Mothers' jiggling and the burst-pause pattern in neonatal feeding. *Infant Behavior and Development*, 3, 29–46.

Kitamura, C., Thanavishuth, C., Burnham, D. and Luksaneeyanawin S. 2001: Universality and specificity in infant-directed speech: pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, 24, 372–92.

Kleinke, C. L. 1986: Gaze and eye contact: a research review. *Psychological Bulletin*, 100, 78–100.

Kobayashi, H. and Kohshima, S. 1997: Unique morphology of the human eye. *Nature*, 387, 767–68.

Mandel, D. R., Jusczyk, P. W. and Pisoni, D. B. 1995: Infants' recognition of the sound patterns of their own names. *Psychological Science*, 6, 314–17.

Masataka, N. 2003: *The Onset of Language*. Cambridge: Cambridge University Press.

Mehler, J., Bertoncini, J., Barriere, M. and Jassik-Gerschenfeld, D. 1978: Infant recognition of mother's voice, *Perception*, 7, 491–97.

Miklósi, Á. 2008: *Dog Behaviour, Evolution, and Cognition*. Oxford: Oxford University Press.

Moore, C. and D'Entremont, B. 2001: Developmental changes in pointing as a function of parent's attentional focus. *Journal of Cognition and Development*, 2, 109–29.

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

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M. and Matsuzawa, T. 2003: Preference for human direct gaze in infant chimpanzees *Pan troglodytes: Cognition*, 89, B53–B64.

Nadel, J., Carchob, I., Kervella, C., Marcelli, D. and Réserbat-Plantey, D. 1999: Expectancies for social contingency in 2-month-olds. *Developmental Science*, 2, 164–73.

Newman, R. S. 2005: The cocktail party effect in infants revisited: listening to one's name in noise. *Developmental Psychology*, 41, 352–362.

Pack, A. A. and Herman, L. M. 2004: Bottlenose dolphins *Tursiops truncatus*) comprehend the referent of both static and dynamic human gazing on pointing in an object choice task. *Journal of Comparative Psychology*, 118, 160–71.

Parise, E., Reid, V. M., Stets, M. and Striano, T. 2008: Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience*, 3, 141–50.

Pascalis, O., de Haan, M. and Nelson, C. A. 2002: Is face processing species-specific during the first year of life? *Science*, 296, 1321–23.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C. and Febre-Genet, M. 1995: Mother's face recognition by neonates: a replication and an extension. *Infant Behavior and Development*, 18, 79–85.

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F. and Mehler, J. 2003: Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the USA*, 100, 11702–5.

Reid, V. M. and Striano, T. 2005: Adult gaze influences infant attention and object processing: implications for cognitive neuroscience. *European Journal of Neuroscience*, 21, 1763–6.

Repacholi, B. M. and Gopnik, A. 1997: Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental Psychology*, 33, 12–21.

Rhodes, M., Gelman, S. A. and Brickman, D. in press: Children's attention to sample composition in learning, teaching, and discovery. *Developmental Science.*

Ricciardelli, P., Baylis, G. and Driver, J. 2000: The positive and negative of human expertise in gaze perception *Cognition*, 77, B1–B14.

Senju, A. and Csibra, G. 2008: Gaze following in human infants depends on communicative signals. *Current Biology*, 18, 668–71.

Senju, A., Csibra, G. and Johnson, M. H. 2008: Understanding the referential nature of looking: infants' preference for object-directed gaze. *Cognition*, 108, 303–19.

Sinha, P. 2000: Here's looking at you, kid. *Perception*, 29, 1005–8.

Southgate, V., Chevallier, C. and Csibra, G. 2009: Sensitivity to communicative relevance tells young children what to imitate. *Developmental Science*, 12, 1013–19.

Southgate, V., Chevallier, C. and Csibra, G. in press: Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science.*

Sperber, D. and Wilson, D. 1995: *Relevance: Communication and Cognition.* Second Edition. Oxford: Blackwell.

Sperber, D. and Wilson, D. 2002: Pragmatics, modularity and mind-reading. *Mind & Language*, 17, 3–23.

Stern, D. 1977: *The First Relationship: Infant and Mother*. Cambridge, MA: Harvard University Press.

Stern, D. N. 1985: Affect attunement. In J. D. Call, E. Galenson and R. L. Tyson (eds), *Frontiers of Infant Psychiatry, Vol. 2*. New York: Basic Books, 3–14.

Stormark, K. J. and Braarud, H. C. 2004: Infants' sensitivity to social contingency: a 'double video' study of face-to-face communication between 2- and 4-month-olds and their mothers. *Infant Behavior and Development*, 27, 195–203.

Striano, T. and Stahl, D. 2005: Sensitivity to triadic attention in early infancy. *Developmental Science*, 8, 333–43.

Symons, L. A., Hains, S. M. J. and Muir, D. W. 1998: Look at me: five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development*, 21, 531–6.

Tomasello, M. 1999: *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M., Carpenter, M. and Liszkowski, U. 2007: A new look at infant pointing. *Child Development*, 78, 705–22.

Tomasello, M., Hare, B. and Fogleman, T. 2001: The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour*, 61, 335–43.

Tomasello, M., Hare, B., Lehmann, H. and Call, J. 2007: Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of Human Evolution*, 52, 314–20.

Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á. and Csibra, G. 2008: Infant perseverative errors are induced by pragmatic misinterpretation. *Science*, 321, 1831–4.

Trainor, J. and Desjardins, R. N. 2002: Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, 9, 335–40.

Tremblay, H. and Rovira, K. 2007: Joint visual attention and social triangular engagement at 3 and 6 months. *Infant Behavior and Development*, 30, 366–79.

Trevarthen, C. 1979: Communication and cooperation in early infancy: a description of primary intersubjectivity. In M. Bullowa (ed.), *Before Speech: The Beginning of Human Communication*. Cambridge: Cambridge University Press, 321–47.

Trevarthen, C. and Aitken, K. J. 2001: Infant intersubjectivity: research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42, 3–48.

Vouloumanos, A. and Werker, J. F. 2007: Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10, 159–64.

Watson, J. S. 1972: Smiling, cooing, and 'The Game'. *Merrill-Palmer Quarterly*, 18, 323–39.

Watson, J. S. 1984: Memory and learning: analysis of three momentary reactions of infants. In R. Kail and N. Spear (eds), *Comparative Perspectives on the Development of Memory*. Hillsdale, NJ: Lawrence Erlbaum, 159–79.

Werker, J. F. and McLeod P. J. 1989: Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, 43, 230–46.

Werker, J. F., Pegg, J. E. and McLeod, P. J. 1994: A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17, 321–31.

Wolff, P. H. 1968: The serial organization of sucking in the young infant. *Pediatrics*, 42, 943–56.

Xu, F. 2002: The role of language in acquiring object kind~concepts in infancy. *Cognition*, 85, 223–250.

Yoon, J. M. D., Johnson, M. H. and Csibra, G. 2008: Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 13690–5.