Understanding the structure of communicative interactions in infancy

Athena Vouloumanos
New York University

Kristine H. Onishi
McGill University

To appear in: H. S. Terrace & J. Metcalfe (Eds.), Social Cognition and Social Neuroscience
Abstract

Human communication is varied and complex: two (or more) agentive interlocutors who have a set of shared experiences interact using specific and intentional behaviors (verbal, written, and gestural) that transfer information. How do we come to understand communicative acts over the course of development? In this chapter, we outline a framework for understanding some important components of communicative acts, highlighting the critical roles of joint attention and agency, and propose that this framework can direct investigations into the developmental roots of understanding human communication.
Understanding the structure of communicative interactions in infancy

In a typical day, humans engage in hundreds if not thousands of communicative interactions, some verbal (“a double non-fat latte please”), some written (“stuck @ café in massive Q, cu in 20”), some gestural (a shrug at the barista faced with a temperamental espresso machine). What are the basic components and structure of human communicative acts such that we can communicate in each of these interactions? And how do we build our understanding of human communication during development? Part of the difficulty in carving a communicative act at its joints stems from its intrinsic complexity (e.g., Bruner, 1981): it requires two (or more) agentive interlocutors who have a set of shared experiences, interacting with specific and intentional behaviors that transfer information both through the literal meaning of the communicative act and through the agents’ intended meaning. Each of these aspects of communication must be resolved in order to participate in or evaluate a successful communicative interaction. We outline a framework for understanding some critical components of communicative acts, highlighting the critical roles of joint attention and agency, and propose that this framework can direct investigations into the developmental roots of understanding human communication.

Defining communication

In the biological literature, communication is commonly defined as anything that influences the behavior of others, including for example, coloring and reflex displays (Maynard Smith & Harper, 2003; Scott-Phillips, 2008). Here we restrict our discussion to communicative signals produced intentionally and flexibly to influence others (as in Tomasello, 2008; and similar to the non-natural, as opposed to natural, signs of Grice, 1957) with an assumption of cooperation between interlocutors (Clark, 1996; Grice, 1969; Levinson, 2000; Sperber & Wilson, 1995). Thus, communication will be considered a cooperative act that occurs in the context of joint attention between at least two intentional agents. Further, we focus on the information-transferring properties of communication, rather than on the question of whether communication arises primarily from adaptations for cooperation or from formal properties of language (e.g., Tomasello, 2008).

Detecting communicative acts

How do we evaluate a particular behavior as communicative? Detecting a
communicative act is not a trivial task (e.g., Scott-Phillips, Kirby & Ritchie, 2009), as entities may engage in any number of behaviors, only some of which would count as communicative. How do observers of potentially communicative behaviors distinguish situations in which communication is being attempted from those in which it is not? Since we consider communication to be a cooperative act that occurs in the context of joint attention between at least two intentional agents, in order to evaluate an act as communicative, participants (and observers) must detect whether there is an interaction involving agents and whether the interaction is intentional.

From infancy, humans are sensitive to aspects of behavior that are relevant for communicative interactions such as intention, agency, and interaction. Much of the work in this area has focused on infants’ sensitivity to the intentions of others (for reviews see Baldwin & Baird, 2001; Csibra, 2010; Meltzoff, 2007; Woodward, Sommerville & Gerson, 2009; Woodward, 2005). As early as 5-6 months, infants interpret actions such as grasping as intentional (Woodward, 1998). By 9 months, infants behave differently towards a person who is unwilling and a person who is unable to perform an action (Behne et al., 2005), reflecting their understanding that a person’s underlying intentions can change the interpretation of their actions. By 2 years, they understand the role of intentions in pretend scenarios (Rakoczy & Tomasello, 2006) and in interpreting other’s drawings (Preissler & Bloom, 2008). Thus, understanding of others’ intentions is discernable by 5-6 months of age, and develops over the next months and years.

Young infants distinguish between agents (entities with the capacity for intentional, goal-oriented behavior) and non-agents, and do so using a variety of cues such as self-propelled or variable motion (Csibra, 2008; Gergely, Nádasdy, Csibra & Bíró, 1995; Luo & Baillargeon, 2005), eyes and eye gaze (Johnson, Slaughter & Carey, 1998), having an effect on other objects (Bíró & Leslie, 2007; Király, Jovanovic, Prinz, Aschersleben & Gergely, 2003), (Bíró & Leslie, 2007) and contingent behavior (Johnson et al., 1998; Johnson, Alpha Shimizu & Ok, 2007; Premack & Premack, 1997). Infants also differentiate between the interactions of agents and non-agents. By 6 months of age, infants expect people, but not objects, to respond contingently (Legerstee, 1997), and 15- to 18-month-olds have different expectations for human-human and human-chair interactions (Golinkoff, 1975; Golinkoff & Kerr, 1978). Unlike their expectations for the movements of inert objects, infants do not necessarily expect the actions of agents to be governed by spatio-temporal properties such as continuity (Kuhlmeier, Bloom & Wynn, 2004), instead accepting that agents can act contingently and at a distance from other entities (Spelke, Phillips & Woodward, 1995). For infants, the actions of agents are
self-governed and can be contingent on the behaviors of other entities.

Although infants can evaluate an entity’s intentions, differentiate between potential agents and non-agents, and hold different expectations about how people interact with people versus objects, their understanding of interlocutors’ potential for interaction and agency as key structural components of communicative acts remains to be examined.

**Functions of communicative acts**

Even with the simplifying assumptions we have adopted in this chapter, communicative acts are diverse and challenging to characterize, due in part to the complexity introduced by the social context of language. Communication serves many functions including behavior regulation (e.g., making requests, protesting), generating a state of joint attention (e.g., commenting on the world), and encouraging social interaction (e.g., attention seeking, social games; Bruner, 1981). (For more ways of classifying communicative functions, see Searle 1976; Tomasello, 2008.)

Children seem to understand that communicative acts can serve a range of functions. Infants produce different communicative acts that elicit specific reactions; an infant’s cry quickly gets a parent’s attention, while reaching up towards the ceiling often leads to getting picked up. Infants in ambiguous situations (for example, when faced with a borderline slope that they have a 50% probability of successfully descending) actively seek emotional signals from trusted adults and use them to modulate their behavior (e.g., descending the borderline slope more often following their mother’s positive than negative emotion; Tamis-LeMonda et al., 2008; Hoehl, this volume). Older children appear to be sensitive to more sophisticated aspects of communicative acts, for example, differentiating the propositional content (an utterance’s compositional meaning or meaning that can be derived from the individual words and their combination: “My homework is due tomorrow”, meaning that a particular document must be delivered shortly) from the illocutionary force (the speaker’s intent: “My homework is due tomorrow”, meaning that the speaker can't go to the movies now; Searle, 1976). For example, 3-year-olds interpret assertive and imperative speech acts differently, correcting the speaker when she failed to describe actions correctly but correcting the receiver when she did not follow the speaker’s instructions (Rakoczy & Tomasello, 2009). Children expect speech to be truthful, i.e., to reflect the actual state of the world, and detect violations in truthfulness (consistent with the Gricean maxim of Quality, be truthful, Grice, 1989; Koenig & Echols, 2003). Children are also sensitive to the different communicative functions of gestures, correctly inferring, for example, that an adult’s gaze direction and pointing gesture indicate either the location of a hidden toy...
This sensitivity suggests that children are not just following gaze, but also infer the intention before the gaze and have some understanding that both partners know that they are attending to the same location in a shared psychological state (Carpenter & Call, this volume; Tomasello, 2008). In other words, children understand shared communicative acts with others as states of joint attention and use their knowledge of shared experience to determine the function of a given communicative act. Function, however, is but one aspect of communicative acts. Understanding communicative acts also requires understanding of their internal structure and constituent components.

**Components of a communicative act**

We identify several critical components of communicative acts based on the adult literature; these components underlie the success of communicative acts in transferring information. First, *shared intentionality*: communication is a joint activity requiring coordinated action between interlocutors (the sender and the receiver; Clark, 2002), which in turn requires shared psychological states between participants, or joint attention (Tomasello & Carpenter, 2007). Second, *directional signaling*: communication requires transfer of information from the sender (the source of the signal who has the information) to the receiver. Third, *mutually-readable formats*: the communicative signal must be in a format that is understood by both sender and receiver, used by the sender to encode information and the receiver to decode information. Fourth, *referential specificity*: the communicative signal must use specific forms titrated to the relevant level of meaning (i.e., using the appropriate wording; Clark, 1979).

Research on understanding and development of communication has focused mainly on children’s understanding of shared intentionality (1) and of referential specificity (4), but less is known about their understanding of the role of directional signaling (2) and mutually-readable formats (3) for successful communication.

**1. Shared intentionality**

Shared intentionality is a mental construct so can neither be observed nor can its understanding be tested directly. However, observers have access to behavioral indices of shared intentionality, including proximity of interlocutors, coordinated (contingent) actions, and co-localized eye gaze. Shared intentionality provides a basis for making inferences about the information states of interlocutors, and more specifically for establishing common ground.
between interlocutors, where common ground includes the history of the participants’
interactions and the current information state of the sender and receiver given the particular
situation [this notion has also been discussed as “forms of life” (Wittgenstein, 1953), “joint
attentional formats” (Bruner, 1981), and “common conceptual ground” (Clark, 1996; Tomasello,
2008)].

To understand the content of communicative acts, observers assume that others will
communicate using principles of relevance (Sperber & Wilson, 1995) and informativeness
(Grice, 1989), both of which rely on establishing common ground. In adult communication,
information unknown to the receiver is identified as potentially relevant, and thus likely to be the
content of the current communication (Sperber & Wilson, 1995), and the response of the
recipient is expected to be appropriate to the communication (Albright, Cohen, Malloy, Christ &
Bromgard, 2004). Children appear to use relevance principles before informativeness principles
to evaluate others’ communication. By 3 to 5 years of age, children are sensitive to violations of
Grice’s maxim of Relation (be relevant), preferring the puppet who uses this relevance maxim,
over puppets who use only informativeness maxims: Quantity (say enough but not more than
necessary) or Quality (be truthful; Eskritt, Whalen & Lee, 2008). In some circumstances, 6-year-
olds can appropriately use informativeness maxims. For example, consistent with the maxim of
Quantity, when presented with several cards with clowns and asked to “give the clown,” children
selected a card with a clown alone over one with a clown with flowers or two clowns (since, they
presumably infer that if the speaker had wanted the cards with flowers or with two clowns, they
should have qualified their description; Surian & Job, 1987).

The ability to use joint attention contexts to infer shared intentionality has its roots in
infancy. Even 18-month-olds behave in a manner consistent with the principle of relevance
(Carpenter & Call, this volume; Moll, Koring, Carpenter & Tomasello, 2006; Southgate,
Chevallier & Csibra, 2009). For example, 18-month-olds use others’ discourse (e.g., "Oh, great,
look!") to infer the focus of attention (i.e., the new object or object part) and therefore what is
relevant (Moll et al., 2006). And when infants of this age were shown a toy moving into a house,
if the goal location was mutually known to the infants and the experimenter, infants preferentially
imitated the new information (about the manner of motion), suggesting that they considered this
new information to be the relevant focus of the communication (Southgate et al., 2009). Thus,
infants use behavioral indices of shared intentionality to interpret communicative acts.

More direct evidence for infants’ understanding of shared intentionality comes from
studies showing they respond differentially to information presented during interactions with and
without ostensive communicative cues (direct eye gaze, child-directed speech; Gliga & Csibra, 2009; Senju & Csibra, 2008; Yoon, Johnson & Csibra, 2008). These behavioral markers of joint attention signal to infants the presence of a knowledgeable speaker who is providing generalizable information.

Thus, although shared intentionality is not directly observable, children are sensitive to behavioral cues of shared intentionality that signal contexts of joint attention. Infants and young children thus seem to have a basic understanding of the role of shared intentionality in communicative acts.

2. Directional signaling

Infants are sensitive to cause and effect, and causal direction within events (Bullock & Gelman, 1979; Golinkoff, 1975; Leslie & Keeble, 1987). However, there is no direct evidence that children watching a third-party interaction understand that a communicative signal must travel in a particular direction, from an informed sender to an uninformed receiver, in order for information to be transferred. A related line of work suggests that older children understand something about the importance of the source of a communicative signal. For example, 3- and 4-year-olds preferentially learn a new word from a previously reliable rather than unreliable speaker (Birch, Vauthier & Bloom, 2008; Pasquini, Corriveau, Koenig & Harris, 2007; Vanderborght & Jaswal, 2009), demonstrating knowledge that the specific source of the signal matters. Although children infer the causality of physical events and are selective in choosing the source of communicative signals, their understanding that transferring information requires directionality has yet to be established.

3. Mutually readable formats

Do children understand that certain formats (speech, pointing) allow for efficient communication of information while others (coughing, nose-scratching) generally do not, and that these formats should be shared across interlocutors? Children seem to know that the communicative signal must use conventional means to accomplish a specific action or goal (Clark, 1979). For example, conventional requests ("What time do you close tonight?") are responded to with information ("at six"), while less conventionally worded requests ("Do you close before seven tonight?") receive a response to the question ("yes") as well as the information ("at six"). And even 13- and 19-month-old infants understand that speech is used conventionally across individuals, such that different people are expected to share labels for
objects (Buressh & Woodward, 2007; Graham, Stock & Henderson, 2006), though they are not expected to share, for example, preferences.

Although there is no evidence that younger infants understand the conventionality of communication formats, infants as young as 6 months show understanding that some formats are more efficient vehicles for information transfer. For example, based on hearing speech, infants can make some reasonable inferences: (1) that a human is present (Vouloumanos, Druhen, Hauser & Huizink, 2009), (2) that a person shares the same language as other people (Kinzler, Dupoux & Spelke, 2007), (3) that the entity spoken to is more likely to be another person than an inanimate object (Legerstee, Barna & DiAdamo, 2000), and (4) that speech labels correspond to the number of objects (Xu, 2002) and to category membership (Fulkerson & Waxman, 2007). While speech is a primary symbolic communicative signal for humans, humans may also communicate using other formats which infants also understand as transmitting information. For example, infants can use a pointing gesture produced for a third party as a tool for finding a hidden toy (Gräfenhain, Behne, Carpenter & Tomasello, 2009). Infants thus understand that some formats (speech, pointing) can transmit information, and that these are conventionally shared across individuals. Recent studies in our labs show that by the end of their first year, infants recognize that a sender’s speech or pointing gesture can communicate to a receiver which of two objects is the target (Krehm, Onishi & Vouloumanos, 2010; Martin, Onishi & Vouloumanos, 2010). Thus, while infants can evaluate some acts, but not others, as transferring information, there is as yet no direct evidence that children understand that the format of a communicative act must be mutually readable, that is, for a communicative act to be successful, that the sender must use a format that the receiver has the capability to perceive and knowledge to decode.

4. Referential specificity

The communicative signal must use specific forms with a level of specificity appropriate for the interlocutors’ common ground and for the situation (Clark, 1979). However, not all formats can titrate degrees of referential specificity equally. Speech is uniquely flexible in its capacity to express degrees of specificity; it allows a sender to merely draw attention to what is salient ("wow!") or to unambiguously pick out a specific entity in the world (“your grandmother’s Eames chair”). Not every instance of speech is specific. If I say “chair” in a room full of chairs, I have not availed myself of speech’s potential for specificity but if I say the same utterance in a room full of tables and lamps, I likely have succeeded in being highly referentially specific.
Throat-clearing, coughing, expressions of delight, disgust, or surprise do not unambiguously pick out a chair in a room of tables and lamps without some common ground—knowledge that the sender likes, dislikes, or is not expecting a chair, or some other concurrent cue like eye gaze. Nonspeech vocalizations are more dependent on deictic factors (temporal contiguity between the chair and the vocalization, simultaneous eye gaze or gestures toward the chair) whereas speech can be referentially specific, communicating meaning more independently of the context.

The appropriateness of a particular speech form for communicative success is modulated by the common ground between interlocutors. Common ground includes a range of levels of potential shared background knowledge including geographical and cultural communities as well as specific individually shared knowledge (Brennan & Clark, 1996; Isaacs & Clark, 1987). For example, if you are a New Yorker (of a certain era) then you know about CBGBs; if you live in the US then you know about the Grand Canyon; and if you are human you know about fingers. Correct determination of group membership affects the appropriate referential form for an utterance (“fingers” vs. "the appendages at the tips of the upper limbs of the organism") and can even affect the message's content (for a child, a spleen is something found inside your body; for adults, an organ found in vertebrates for removing red blood cells).

The separation of common ground from privileged ground (information known only to oneself) may be challenging for children, and even adults (Keysar, Barr & Horton, 1998). Children sometimes fail to use common ground in discourse, producing utterances that are ambiguous for others, and confusing ambiguous and informative utterances directed to them (Ackerman, 1981; Glucksberg, Krauss & Weisberg, 1966; Rosenberg & Cohen, 1964). This failure in communication has been ascribed either to children being egocentric, in the Piagetian tradition, or to computational difficulties that arise when trying to map the communicative message to the set of possible referents (see Surian & Job, 1987).

Interestingly, it may be children’s understanding of an agent’s intent behind communication—that a speaker intends to communicate something to a receiver—that leads them to erroneously infer that communication has been successful even when the message lacks the appropriate level of referential specificity (by using illocutionary performative force instead of the locutionary content of the utterance, as in Austin, 1962). Children are thus fooled by ambiguous communicative attempts because they assume the receiver is cooperative and responding to the sender appropriately, even when the referent of the utterance is underspecified (Grice, 1989). Although children are biased towards interpreting utterances consistently with the
speaker’s intent (Shatz, 1978), they can be induced to evaluate the referential specificity of ambiguous communication appropriately when they cannot trust the intent of the speaker or when the importance of the content is emphasized (Ackerman, 1981). Children’s understanding of a critical component of communicative acts, intentionality, may thus interfere with their successful use of other components.

Children build their understanding of different communicative acts over time. By the age of 2 years, they can discuss specific topics and negotiate action; after 2, children can moderate politeness, deixis, and some forms of indirect speech; by 4, children can adjust utterances depending on the knowledge of the listener (Ninio & Snow, 1996). More direct evidence that common ground can modulate referential specificity comes from the finding that 5- and 6-year-old children adjust their speech to a partner’s perspective and to information that is in common ground (Nadig & Sedivy, 2002). Recent work shows that rudimentary understanding of common ground may emerge earlier, as 4-year-olds can sometimes detect ambiguity in utterances (Nilsen, Graham, Smith & Chambers, 2008), and even 2-year-olds take into account what others know when communicating, for example, gesturing more to the location of a hidden object when the receiver was absent than when she was present during a hiding event (Liszkowski, Carpenter & Tomasello, 2007; O’Neill, 1996). Children can thus use and make judgments about the referential appropriateness of particular forms for transferring information between interlocutors, taking into account the interlocutors’ common ground.

Conclusions

Human communication is complex. In order to understand communicative acts, observers must detect potentially communicative situations by (1) identifying agents, (2) detecting whether the agents are interacting, and (3) evaluating whether their interaction is intentional. Observers must further understand four critical components of communicative acts, (1) that the sender and receiver have shared intentionality established through joint attention, (2) the sender is directing a signal to a receiver, (3) that the signal format must be mutually readable, and (4) that the referential specificity of the signal is appropriate. Although infants can detect agents and intentions, whether they use this understanding in the service of understanding communicative acts is not fully known. More is known about children’s understanding of shared intentionality, and how they produce and interpret utterances while taking into account common ground. However, children’s understanding of the importance of the
direction of the communicative signal and about the role of a mutually readable format for successful communicative interactions remains to be examined. Our proposed framework can guide investigations of how we build an understanding of human communication during development. A better understanding of the structure and components of communicative interactions elucidates the developmental roots of communicative competence, which provides a foundation for our social and cultural life as humans.
Author Note

Athena Vouloumanos, Department of Psychology, New York University, and Kristine H. Onishi, Department of Psychology, McGill University.

Correspondence concerning this article should be directed to Athena Vouloumanos, Department of Psychology, New York University, 6 Washington Place, New York, NY, USA 10003. E-mail: athena.vouloumanos@nyu.edu, or Kristine H. Onishi, Department of Psychology, McGill University, 1205 Dr Penfield Avenue, Montreal, QC, Canada H3A 1B1. E-mail: kris.onishi@mcgill.ca.
References


