

key words: evolution of language; communication; attention; tool use; reference; perception; intelligence; intentionality

Abstract.

It is argued that the capacity to focus attention is crucial for intentional communication. Intentional communication is goal-intended; directed at changing mental states and as a consequence behaviour; about a referential object common to sender and recipient; and about objects that may be context-and referent-independent. Three different kinds of attention is discerned: scanning, attention attraction, and attention-focusing. The focus of attention can, depending on the abilities of the subject, be on objects or subjects that either are contextual or stable, and it can be individual or shared. For language use, subject-subject focusing along with shared attention are necessary. This does not require Gricean metarepresentations, but basically only attention contact between the subjects and behavioral co-ordination. Language use can be compared with tool use to bring out the characteristics that distinguish informational from intentional communication. The capacities required for tool use are in several cases similar to those required for language use. A basic similarity is that both activities are used as means to an end.

ATTENTION AND TOOL-USE IN THE EVOLUTION OF LANGUAGE

1. Communication

The story of the evolution of language overlaps with the story of how communication has evolved. All animals communicate. There appears to be as many systems of communication as there are species. One way to trace the evolution of human language is by making comparisons between different communication systems. In this section, I will initially distinguish between three kinds of system, and then discuss some factors that bring out the differences between them. In section 2, I argue that the ability to focus attention has been crucial for the evolution of higher-level communication.

Intentional acts of communication are purposeful or deliberate, goal-intended, and about something else than the sender herself. That they are goal-intended means that the sender represents the goal prior to acting. In goal-directed agency, on the other hand, the goal is represented only in acting, that is, while the action is performed. The aim of intentional communication is to change the mental state of the recipient and as a consequence the behaviour. Successful intentional

communication depends on that the recipient understands and recognises the intention, or point, of the communicative act. To do so, it is not always necessary to understand the intention conceived of as a mental state of the sender. It suffices to recognise the goal of the action, and this can in many cases be done by picking up contextual clues. Behaviour is often produced in a direct connection to the goal. Reaching, to take a simplified example, is usually directed at an object in the vicinity of the subject, like another person, a glass, or a flower. Depending on the nature of the object, the number of possible interpretations of the act can be narrowed down.

I contrast intentional with mechanical and informational acts of communication. *Informational* communication is goal-directed. It aims at influencing the behaviour of the recipient to the advantage of the sender, usually by conveying information about the sender. An example of informational communication is the butterfly conveying the information that it is not edible by not being visible against the background of the flower it is sitting on. Another example is animals showing their teeth in anger. Whether the recipient benefits from communicating is irrelevant for the success of the act. Informational communication is successful when the recipient responds to the act (say, by fleeing when a warning call is emitted), or copies, or reproduces, the act (as sometimes is done with warning calls).

Mechanical communication arises when a sender influences the recipient by physical means only, as in causal interaction. To push or stand in the way are examples of mechanical communication. The act is successful when a direct response to stimuli occurs. Mechanical communication can be goal-intended as well as goal-directed.

Let me illustrate the difference between the various kinds of communication with an example. A communicative act directed at informing the recipient that a dangerous animal is nearby and that she must flee can be performed in different ways. In informational communication, which primarily aims at influencing behaviour, the goal is simply to make the recipient run. This can be done in ways that do not involve pointing out the animal to her, for instance, by emitting a warning-call. Mechanically, the goal can be achieved by throwing a stone at the recipient. Intentional communication, on the other hand, primarily aims at changing the mental state of the recipient and then have her behaviour change as a result of the change in mental state. The goal is to make the recipient direct her attention towards the animal, as a referential object common to both sender and recipient. The recipient will run as a result of having her attention directed at the animal and grasping that it is dangerous.

It can be maintained that intentional communication either encompasses and arises from mechanical and informational communication, or that it is of a different nature than the other two kinds of communication. In the following I will not

take a stand concerning this issue. In human communication, for instance, all three kinds seem to occur. Intentional communication is often connected to language-use. This certainly holds for communication between sender and recipient when they are not present in the same context, that is, are not at the same place at the same time. This is because intentional communication demands mutual attention to a common object. If sender and recipient are not present in the same context, the sender cannot direct the attention of the recipient by gaze or gestures, but needs a stand-in, say, a language, or a symbol system. In the following, I will concentrate on the relation between informational and intentional communication, since informational communication occurs within most species, while intentional communication seems to arise among humans only.

The important difference between informational and intentional communication is that the former is achieved by behavioural means, without taking the detour around the mental states of sender and recipient. It is faster and less complex than intentional communication. This makes it efficient, but also inflexible and rigid. An advantage of intentional to informational communication is that in intentional communication, the response of the recipient will be better tuned to the particular situation that the act of communication concerns, since her attention will be directed at its object.

It is quite common to spell out the difference between animal and human communication in terms of how information is conveyed. The general idea is that humans use symbols, while animals are signalling. A change in the behaviour of the recipient is caused by transmission of information, that is, signalling, either about the state of the sender or a state in the environment or both. The signal results in a manipulation of the recipient or in an interaction that will be co-operative or competitive.

Several different things are intended by the term signal. For one thing, signals are not thought to have syntactic properties. This means that they do not belong to grammatical categories and are not systematic and compositional.¹The lack

¹A minimal definition of syntax says that it consists of a set of rules that allow for prediction of sequences of signals or symbols. Usually it is held that these rules are recursive. It is the rules that guarantee what is called linguistic creativity or productivity, that is, that a language-user both can understand and produce an infinite number of symbol-strings she has not encountered before. The rules confer generality on the language and account for its compositional and systematic character. That a language is compositional means that the meaning of a complex expression in part is determined by the meaning of its constituents and the way these are ordered. Systematicity on a syntactic level concerns the fact that the behaviour of the same kind of expression in different contexts is determinate, and at a semantic level that the same type expression contributes the same type meaning in different contexts - provided that it occurs in a well-formed sentence. Systematicity ensures that if you know the meaning of the expression Fa , and come to know b and c , and recognise that b and c are of the same kind as a ,

of syntax is correctly attributed to the communication systems of animals in many cases, but there are exceptions. Birdsong is one such exception. Birdsong has a complex syntax, and songs of the same species even have different dialects in different areas. Nowicki and Podos maintain that birdsong "comprises a set of minimal units arrayed in sequence with several levels of organisation" (Nowicki and Podos, 1993: 163). Notes can be ordered into phrases, and phrases into songs. The songs have minimal units of production as does natural language (phrases or sentences).

A second characteristic of signals is that they can be used inadvertently. Signals do not always convey information as a consequence of the sender's intention. This happens, for instance, when the sender cannot itself control whether to exhibit or not a certain behaviour, but behaviour patterns involuntarily are triggered by certain types of context. In such cases, signalling is completely context-determined.

The property of context-independence has often been cited as a mark of linguistic, as well as intentional, behaviour, and as distinguishing human from animal behaviour (Gärdenfors 1996; Sjölander 1995). Context-independence indicates that, to carry content about something, the sign is independent of being in the same spatiotemporal context as its object. We can further use the term referent-independence to indicate that the sign is also independent of the existence of the object. An example of a sign that is both context- and referent-independent is the word unicorn. To the contrary, signals are neither context-independent, nor referent-independent.

Signalling obviously falls under the category of informational communication. The question is what made it possible for humans to start communicate intentionally. So far I have mentioned four important characteristics of human intentional communication that is not shared by informational communication. Intentional communication is (i) goal-intended; (ii) directed at changing mental states and, as a consequence, behaviour; (iii) about a referential object common to sender and recipient; and (iv) occasionally about objects that are context-and referent-independent. In the next section, I will trace these characteristics back to a faculty that animals and humans have in common, but that has evolved further in humans: attention. Attention is important for any information-gathering system, since it helps the organism in sifting among input. There are several types of attention, some of which seem to exist only in humans.

you will then also know the meaning of *Fb* and *Fc*. Likewise, if you know *Fa*, and *G* and *H*, you will as well understand *Ga* and *Ha*.

2. Attention

Under normal conditions, subjects, be it animals or humans, are immersed in a constant flow of information. This flow provides them with an underlying, basic state of arousal. That state functions as a background against which the subject herself and the part of her environment that she attends to stand out (Luria 1973, Gibson 1986). Attention consists in an increased awareness of something either external or internal to the subject. It can be directed at sensations, perceptions, or behaviours. In case the attention is focused on one of the subject's own mental states, like sensations or experiences, then the subject will be conscious of this state.

Attention can be involuntary as well as deliberate. It may increase performance on the task at which attention is directed, since other things are not allowed to interfere with the subject's awareness. But attention can also impede performance, since the awareness of the subject is captured by the attended entity, and is not divided, or shared by other entities. The subject may as a consequence fail to pick up important information that relates to the task, but does not immediately concern the entity to which her attention is directed (Brinck, to appear).

I distinguish between scanning, attention attraction, and attention-focusing. *Scanning* of the environment is continuous. Information is registered in an active search directed at discovering discrepancies and possibilities to act. The information guides movements and triggers actions in particular contexts. There is a constant attunement between bodily movements, perception and environmental changes. Shanon (1993) describes this in terms of an equilibrium consisting in an on-going, on-line co-ordination.

Attention is *attracted* by events that are at odds with what is expected on the basis of previous experience. The perceptual system is geared to perceive changes or transitions in real time (Freyd 1987). Such changes will shift the direction of the scanning and also the line of action. In some cases, the subject will *focus* on that which happened to attract her attention (Barsalou and Prinz 1997; Brinck 1997; Mandler 1992, 1997). She is freezing a piece of the transient reality. She does so by suppressing most of the information that she is currently receiving and simultaneously adding information in making perceptual inferences based on previous encounters with similar situations.

As long as the subject keeps to scanning and occasional attention-attraction, discrimination depends on the detection of dissimilarity or similarity between items. Whether two items are discriminated or, on the contrary, perceived as similar, will depend on the context and on with what else they are juxtaposed. Attention-focusing is a kind of perceptual categorisation that makes it possible to

create general and context-independent objects of attention. It is central for the capacity to re-identify events and objects across contexts and thus for grasping identities through time and space.

It seems that the difference between attention attraction and attention-focusing constitutes a watershed that provides the basis for being able to engage in different kinds of communication. To communicate informationally, attention-focusing is not necessary.

T. Deacon (1997) has made use of Peirce's (1935: 2.274-2.308, 3.359-3.364) distinction between icons, indices, and symbols, giving it a somewhat unorthodox interpretation, to explain the difference between animal and human thinking as it surfaces in communication. As Deacon describes communication by icons and indices, it is, as far as I can see, possible without the subjects' having the capacity to focus attention. This kind of communication is prevailing among animals. Deacon argues that only humans have the capacity of symbolic thinking. Deacon holds that reference and referential competence are hierarchic, in an ascending order from iconic, over indexical, to symbolic reference.

Iconic reference is a consequence of not making distinctions between incoming information. The representational relation depends on similarity. An example is camouflage. A moth on a tree whose wings resemble the bark can avoid being eaten if the bird does not distinguish it from the tree. The bird takes the two to be similar. Indexical competence arises from relationships between icons, by bringing the latter to bear in interpretation of new stimuli. It is a kind of repeated correlation that results from associative learning, consisting in stimulus or learning-set generalisation. The correlation rests on causal links and spatiotemporal contiguity between index and icon, as when forming an association between smoke and fire.

Symbolic reference, finally, depends on indirect reference to objects, with the help of other symbols. Symbols are mapped onto symbols. The symbolic relation is an instance of a system of token-token interrelationships which are independent of their indexical function. Deacon writes that "tokens indicate one another in the sense that their presence or position in a communicative activity influences the admissibility or nonadmissibility of others." (1996: 128). When the organism learns about an indexical relation, this knowledge must be put into a "predetermined slot" among the symbolic relations. The capacity to fit the index into the slot is an instance of a logical-categorical generalisation (1996: 129).

Symbolic tokens are related to each other through a complex system of relations, like opposition, substitutability, and adjacency. Their relationship can be described with the help of an idea once put forward by Roman Jakobson (1956). One token can be substituted semantically for another along two axes: the metaphoric and

the metonymic. The first relies on similarity, the other on contiguity, both either semantically or syntactically. Let me give an example. Metaphorically, the concept *apple* can be substituted by, for instance, *ball* or *stone*, metonymically by *juice*, *fruit-tree*, *to eat*, or *red*.

Deacon (1997) maintains that humans, but not animals, can learn symbolically, and not only by association. This means that humans understand concepts as being interrelated and not primarily related to external referents. Learning symbols is an instance of symbolic recoding: a recognition of higher-order regularity among established associations. Deacon's notion of a symbol is a paradigm example of context-independent and even referent-independent representation, since the indexical relations become suppressed during learning. Symbols are also general or systematic.

I suggest that attention-focusing is a precursor of human intentionality and thus of intentional communication and of symbolic recoding. It is crucial for any activity that involves context-independent representations, such as planning, or intentional action. Brentano (1973) proposed that the capacity of (human) mental states or events to be about or be directed at objects that do not exist was the defining property of the mental. Ever since, this characteristic has been considered as a benchmark of intentionality.

In an oft-quoted passage, Brentano wrote:

Every psychic phenomenon is characterised by what the Scholastics of the Middle Ages called intentional (also indeed mental) in-existence of an object, and which we, although not with an entirely unambiguous expression, will call the relation to a content, the direction toward an object (by which here a reality is not understood), or an immanent objectivity. Every [psychic phenomenon] contains something as an object within itself, though not every one in the same way. In presentation something is presented, in judgment something acknowledged or rejected, in love loved, in hate hated, in desire desired, and so on (p. 115).

An important characteristic of intentionality is its directedness. Intentional states are about something, that is, they have a content. Directedness also captures another property of intentional states; they are aimed at an object. In the quotation above, Brentano points to two other properties of intentionality, that we already have come across under the notions of referent- and context-independence: the intentional object may not exist; and the representation of it is independent of any particular contextual constraints on the content.

The intentional relation is a peculiar one, since one of its terms does not have to exist in reality. It is not an ordinary causal relation. Brentano therefore described it as relation-like (etwas 'relativliches'). This peculiarity is both the strength and the weakness of the thesis of intentionality. It is its strength, because it brings

forth what distinguishes the mental from the purely physical. It is its weakness, because it makes it hard to explain how we ever can think about the external world if thoughts are not related to real things. This difficulty can be side-stepped in different ways, which I will not penetrate here. I will take Brentano's characterisation of the mental as given for the time being.

The state of attention-focusing has directedness in common with intentionality. One may therefore say that both perception (at least when it results in attention-focusing) and cognition have intentional features. However, directedness can in some cases be weaker than in full-blown intentionality, as I just described it. This happens if states are goal-directed instead of goal-intended. The difference is that in a goal-directed state, attention is focused as a direct consequence of attention attraction and not, as in the case of goal-intention, because the subject has an independent intention to focus her attention.

As mentioned above, attention is attracted by unexpected events. There is a predisposition to direct attention to events, that is, actions and things happening, or to changes in the environment. Events are particulars that occur at specific times and places. But the same kind of event can be instantiated at different times, although in similar contexts with respect to other things. This means that time is a variable property of event-types. Event perception is distinct from object perception in that it involves the perception of complex wholes. These wholes may involve more than one object or subject, as well as the interaction taking place between them. Information about the perceived events may be gathered from all senses, about temperature, scent, and so on, and can thus be multimodal.

By focusing her attention, the subject freezes a portion of her perceptual field. She can go on to categorise the attended information, on the basis of the entity's physical and functional properties and, if it shows any, its behaviour. The subject will then be focusing on a reified entity, one that is constituted partly by her manner of representing it, and thus by the way she interacts with it. I will call this kind of attention *object-focused* and talk about the object of attention. Attention can be focused on different entities. It seems that, in humans, ontogenetic and phylogenetic development go hand in hand concerning the kind of entity a subject is able to focus on.

As pointed out by Shanon (1993: 303), thing-hood affords control, in this case over one's perceptual and attentional states. Objects can be manipulated in order to achieve something one desires in the external world. They can also be used to manipulate oneself, for instance, as stepping-stones in a sequence of actions or a line of reasoning. What kind of control an object affords has to do with the character of the object and its user.

We can discern different kinds of object depending on the subject's ability to focus her attention. *Contextual* objects are such that occur in a particular spatial and physical context, and only are retrieved in that context. They are entities created on the spot in interaction with the environment from performing contextbound actions, and require a particular setting. They are related to particular episodes and are remembered as such. Human contextual objects often belong to ad hoc categories (Barsalou 1983), categories that are temporary and directly introduced for a specific purpose, as when you use a small stick found on the ground to clean your shoe-soles, while you are waiting for a friend to show up. In that context, the stick will simply be a shoe-sole cleaner.

Stable objects are such that can be retrieved in different kinds of spatial and physical contexts. Many objects are stable, for instance, such that we recognise in our daily life - natural kinds like water, vegetables, dogs, cats, and so on; artefacts like tea cups and chairs; or actions like running and eating. A stable object can be put to different kinds of uses, because it can be transferred from one context to another. Stable objects thus contribute one characteristic that is important for language-use and intentional communication: generality. They pave the way for the context-independent representations that are typical for intentionality and language. They also afford long-term control, because they can be kept active in memory outside the context in which they were first encountered.

Contextual as well as stable objects can be either individual or shared. If they are *individual*, they are retrievable by, or stable, to a certain subject, but not to others. *Shared* objects, on the other hand, are retrievable, or stable, for more than one subject. Shared objects provide yet another important characteristic for intentional communication: they constitute common objects to which all participants may direct their attention.

Shared objects emerge when subjects can engage in mutual attention-focusing, during which attention is directed at the behaviour or gaze of other subjects. I call this *subject-focused attention*. In some communicative acts, subject-focused attention precedes object-focused attention, notably in those contexts that constitute precursors to intentional communication. The reason is that in intentional communication, the object has to be shared, and for it to be shared, the attention of both sender and recipient must be focused on a common object. Mutual object-focusing is achieved by co-ordination of attention. Subjects arrive at co-ordination of attention by focusing on each other and figuring out the respective objects of attention.

Communication by gestures is generally thought to be a step towards intentional communication. Gestures are shared objects that can be either contextual

or stable. Virtually all gestures that arise from so-called ontogenetic ritualisation are contextual. Ontogenetic ritualisation emerges from interaction between individuals and results in an idiosyncratic form of communication. It starts from natural ways of behaving, like a small child raising her arms to be carried when her mother reaches down for her, or the baby's moving for her mother's nipple. The next time the child makes these movements, her mother will anticipate what she wants, and subsequently the child will as well understand her own movements as means of making her mother respond in an appropriate way.

These gestures or signals emerge spontaneously from pure social, behavioural co-ordination. They depend on object- as well as subject-focusing. But for signals to acquire an intentional status, something more is required than that the communicating subjects focus on both object and subject. Their attention should furthermore be *subject-subject focused*. This means that they should focus not only on each other, but on each others attention. The reason why is that unless attention is shared, sender and recipient will treat each other as objects and not as intentional agents. That means that communication is directed directly at changing behaviour and not at influencing mental states. Ontogenetic ritualisation does not involve subject-subject attention. Tomasello (1998) remarks that infants who learn to point via ritualisation understand their pointing as a procedure for getting something done, and not as an invitation to share attention.

Gómez (1994; 1998) and Tomasello (1998) have both seized upon the role that mutual attendance between subjects plays for intentional communication. According to Gómez, a co-ordination of communicative acts with eye contact and accompanying behaviour will lead to an understanding of the communicative import of gestures. Gómez gives the example of taking somebody's hand and lead that subject to a goal (for instance, a cupboard with sweets in it), while both subjects alternatively exchange looks and direct their gazes towards the goal of the action.

A function of eye contact is to check the attentional state of the other and have attention contact. Gómez maintains that the mutual attendance to each others' focus of attention leads to attentional loops, in line with Gricean intentional loops. The difference is that Gricean loops involve propositional metarepresentations, while attentional ones do not. Gómez holds that it is possible to "understand other people's minds using first-order representations of behaviours that directly reflect mental states such as attention" (1994: p. 72).

Of course, it is a question of definition whether attention is to be thought of as a mental state. It seems that if we want to describe attention contact and mutual attendance in line with the Gricean model of communication, it would come out as something like the following. The subject intends a certain response to her focusing

her attention, and gets the response because the other subject recognises her intention behind the focusing of attention and also intends to give the desired response. Are attentional loops of this kind necessary for intentional communication?

Clearly, the procedure that gives rise to intentional communication does not have to be this advanced and complicated. Remember that intentional communication is characterised by the following four things: The referential object should be shared and possibly context- or referent-independent, and the communicative act should be goal-intended and directed at changing the behaviour of the recipient by changing her mental states. None of these things seem to require attentional loops. It is quite possible that, instead of the subjects' engaging in a complex of simultaneous and nested intentional attention loops, during which the subjects must grasp each others intentions, the behaviours of the involved subjects simply adjust over time. Except for mutual attendance, the co-ordination of behaviour depends on two factors, one long-term and one short-term. The first one is simply learning to associate certain kinds of behaviour with certain kinds of attention-focusing, directed at particular goals. This is what Deacon calls indexical reference. The second one is being sensitive to what occurs and happens in the actual context and reacting on it in real-time. If these conditions are fulfilled, intentional communication may be achieved in the end, in spite of that intentions are not involved in the procedure leading up to it.

As pointed out by Gómez and others, non-intentional communicative strategies will in some cases, when the cognitive capacities are in place but not yet working, develop into intentional ones. If attentional loops were involved in behaviour that eventually leads to full-fledged intentional communication, they would, as I see it, have to be default (and not intentional) assumptions, involving tacit expectations to the effect that others have a similar capacity to attention as I do, and are able to link, for instance, visual cues, like looks or glances, to behaviour. As default assumptions at a low level, they would probably be innate.

Tomasello (1998) describes the conditions for achieving symbolic reference, by which he intends reference that is social and not merely causal. It involves shared attention to an external entity, and also shared symbols. Tomasello (1998) asserts that symbols are constituted when the recipient R understands that the sender S's attention is directed at a certain object O, and R understands that S intends R's attention to be directed at O. It is not sufficient that sender and recipient come to share the attentional object by mutual attendance and behavioural co-ordination, but the recipient should grasp that the sender intends the recipient to share the object with her. The theories of Tomasello and Gómez are similar in that they both stress the importance of intentional attention loops for intentional communication.

Furthermore, according to Tomasello, a symbol is shared when it is socially bi-directional, which means that both sender and recipient understand the effects of the symbol in the same way. That assures a role-reversal in the use of the symbol, that is, that the recipient can take the sender's role and conversely.

The conditions that Tomasello puts forward for achieving symbolic reference are rather strong. I do not see why Gricean attentional loops are necessary for social reference as such. They become necessary as soon as communication depends on being able to reason about the mental states of the other, but, by far, not all cases of intentional communication are like this.

Moreover, as far as I can see, role-reversal is not in itself necessary to share an object, although it is necessary for understanding the shared object in the same way. Only if we assume that Gricean loops are necessary for successful communication, does role-reversal become necessary. Let me describe what I mean by making an analogy. Communication can be seen as an exchange of simple-minded questions and answers.² To respond to a question, it is not necessary that you understand why the question was asked in the first place, or that you would have asked the same question, had you been in the place of the sender. You pick up the question from the context in which it is raised and do your best to answer it immediately.

Similarly, to respond appropriately to a subject's directing your attention to a certain object, you do not have to understand the object in the same way as the other subject. You respond to the object as it is given in the context common to you and the sender. You do not respond to the mental state of the sender. In the case of the question as in the case of the object, both having been put to your attention, you bring your experience and background knowledge from previous encounters with that object and with the sender to bear on the situation, and respond from your point of view of the context. Taking the role of the other is not necessary. The object, or question, is shared by being at the centre of attention and by being acted on in suitable ways that do not have to be similar or, in their turn, shared.

I do not deny that intentional communication can be very complicated, and that complexity in some cases is necessary to get the message across. But in many cases, intentional communication is less complex than we like to think. To engage in intentional communication as described above, Tomasello's conditions do not have to be fulfilled. In spite of that context-independent objects, or symbols, are used, their

²The same point can be made by exchanging the question and answer-game for a game of tennis. The story goes like this. To return a ball, it is not necessary that you understand why the ball was thrown, or that you would or even could have hit the ball in the same way, had you been in the place of the sender. You pick up the ball as it comes towards you, and do your best to return it.

import can be grasped from contextual clues in combination with general background knowledge.

But in those cases that the context is not sufficient, the recipient has to use Gricean loops and symbolic thinking. It becomes necessary to be able to reason about the mental states of others. This means that either we can make do with mutual attendance and behavioural co-ordination plus background knowledge, or we need Gricean intentional loops, and not just attentional loops. Gricean loops can, it seems, only be reached by humans.

Some animals can engage in intentional communication building on attention contact and behavioural co-ordination, but not involving exchange of higher-order intentional states. By having the capacities to focus attention on shared objects and attend mutually, they reach an intermediary level between informational and full-fledged intentional communication. Intentional communication is possible without access to something that resembles human language. The capacity of attention-focusing takes these animals all the way to intentional communication, but still, not to full-blown human communication.

3. Tool-use

It is interesting to compare tool-use and language-use from an evolutionary point of view, since they both depend on the cognitive capacity of attention-focusing. Above I have argued that attention-focusing is crucial for intentional communication. In this section, I will compare tool-use with language-use in order to show how intimately they are connected seen from an evolutionary perspective. Tool-behaviour cannot be taken as an indication of human, linguistic behaviour. Nevertheless, it seems to be an important step in that direction.

Human language has evolved from basically two sources - human psychophysical constitution and subject-subject interaction, as in co-operation, manipulation, competition, or deceit. It seems that grammar has its basis in the human way of categorising her surroundings as consisting of spatiotemporal entities characterised by their function, perceptually accessible surface properties, relations to other objects, and so on. Syntax as well as semantics appear to be motivated by the human constitution (Johnson 1987; Lakoff 1987; Lakoff and Johnson 1980). Linguistic expressions and descriptions of the world in many cases borrow their structure from the way that humans move around and function in the physical world.

The role of language, it seems, is to serve a mediating function. Linguistic expressions are used to manipulate the sender herself and her environment,

taking on and in some cases extending the role of behaviour and tools. Because of its greater complexity and almost absolute context-independence, language affords a new kind of control that easily stretches into the past and the future and over several areas.

Both informational and intentional communication can be described from an instrumental point of view (Mead 1934; Vygotsky 1978; von Glasersfeld 1976). An instrument is used by a subject in order to do or achieve something in particular, as a kind of intermediary between subject and object. Seen from an instrumental perspective, communication aims at changing the environment in a way that is beneficiary to the sender and its kin. Language has in a broad sense had the same function as tool-use. They are both goal-directed or often even goal-intended activities, used to manipulate states either in the subject's environment or of the subject itself. These states involve in the case of language ultimately the mental states of the sender or of the recipient.

Vygotsky (1978: 54) held that the basic similarity between tool-use and language rests on their mediating function. An important difference, on the other hand, is, according to Vygotsky, that they orient human behaviour in different ways. Vygotsky claimed that tools are externally oriented, while signs are internally oriented, that is, aimed at mastering oneself (1978: 55). Mead (1934, 1938) also stressed the role of language in self-control. Intentional and meaningful behaviour seems to Mead to consist in the power to control the response of oneself and also of others. Language is a kind of co-operative process in which the sender's and the recipient's responses are adapted to each other. This view resembles the one advanced above concerning mutual attention-focusing and co-ordination of behaviour as a basis for intentional communication.

The view that control of attention and attention-focusing must already be in place before language evolves was expressed above. Language amplifies, of course, the possibility of self-control, as does more specifically the possibility to reason about one's mental states and those of others. Attention can, however, be mastered without a language. But it seems to me that without support from language, attention cannot be directed at things that the subject has never been in contact with. Moreover, the attentional field cannot be reconstructed and changed in a way that does not answer to the actual or a past manipulation of objects in the external world, but by chance, unless the subject has access to language.

As I see it, a tool is a kind of stand-in, something that we use intentionally in order to reach a goal that we cannot manage to get hold of by ourselves. Anything can function as a tool - a stick to push down fruit from a tree, a stone to crack an icy surface, the fingers, toes, and other body-parts to count on, etcetera. An object receives its identity as a tool by being used for a certain purpose

by somebody. This property tools have in common with linguistic symbols. They lose their meaning if nobody uses them. Both are arbitrary to a certain degree, although the choice of a certain object among others as a tool is motivated by the circumstances, just like the linguistic metaphors we use to get a grasp of the world are motivated by our constitution and the nature of our surroundings.

Recently, the connection between tool-use and language-use has been emphasized from various perspectives. Among other things, it has been pointed out that there seems to be a neurological overlap between the areas of the brain that control tool-use and language-use (Arbib and Rizzolatti 1996), as well as a resemblance in tool-use and language-use as to the combinatorial structure of sequences of elements in tool-behaviour and of expressions (Gibson and Ingold 1993), and finally a common ground of the two activities in that learning them relies on imitation and social behaviour (Tomasello 1996). This agrees with Vygotsky's remark that

(t)he linkage between tool use and speech affects several psychological functions, in particular perception, sensory-motor operations, and attention, each of which is part of a dynamic system of behavior (p. 31).

Two important facts about tool-use and language-use are that they both take place out in the open and that they are social skills. This has several consequences as regards learning and the generality of thought and language.

When attention is directed at the behaviour of others, as in subject-attention, and it is the behaviour that is reified as a shared object (or activity), behaviour can be not only observed, but also learned by those who attend to it. Tomasello (1996) brings up three kinds of learning: mimicry, imitation, and emulation. He describes learning by mimicry as reproduction of behaviour on a sensory-motor level that does not involve any cognitive processes. Imitative learning reproduces not only the behaviour as such, but the behaviour as fit to a certain goal-directed or intentional context. If, for example, hand-waving is imitated, it is reproduced together with a certain goal, to say hello to passers-by, to stop a car, or whatever. If it is mimicked, only the motor-activity is reproduced.

In emulation learning, focus is on the shared object of the behaviour and not on the behaviour itself. It constitutes an understanding of the cause and effects involved in achieving a change in the world. But an understanding of the relation between the behaviour and the strategy to achieve the goal does not occur. Examples of emulation are when chimpanzees learn how to open a nut and find food inside or how to get rid of sand from food by holding the food under water. The observer understands how a change in the world can be accomplished by a certain kind of

behaviour, but when she tries to reach the same result, she does not reproduce the exact behaviour.

The implication is that in emulation-based behaviour, the subject has to develop a skill all by herself and cannot rely on the technical achievements of other agents. On the contrary, behaviour based on imitative learning and involving shared objects does not have to be renegotiated by each generation, but may be passed on between them. Imitation allows for accumulated knowledge of groups or communities. Knowledge of tools and of language can be passed on between subjects through imitation.

Tomasello (1996: 324) asserts that imitation involves an understanding of the intention behind a certain behaviour and the changes the behaviour causes in the environment. Imitative behaviour is distinguished from mimicking by this sort of understanding, which explains how the imitator can reproduce the behaviour on purpose and in a suitable new setting. It seems to me that imitation mixes the object-focused and the subject-subject-focused strategies in attention-focusing, by putting an emphasis both on the goal of the behaviour and the behaviour as performed by a certain subject. To bring the two kinds of focus together is, as we have seen, also an important step towards learning how to communicate intentionally.

As I see it, intentional understanding of actions also requires that one can divide the behaviour into a set of separate components and categorise the elements involved in the action. The reason is that to be able to reproduce the behaviour in the right circumstances and transfer it from one particular setting to another, one must grasp how the different components are related in order to reach the goal. It is not sufficient just to see that they follow on each other mechanically or causally.

If a subject conceives of a certain behaviour as a sequence of separable components, it will moreover be possible for the subject to reuse some of the components in new settings and perhaps to change the order of them. This feature of re-combination is something that tool-use and language-use have in common. It gives rise to generality and generativity, although the structure or syntax that lies behind language-use is immensely more complicated than the one behind tool-use.

Vygotsky pointed out that attention-focusing is necessary for tool-use. But he also claimed that there is a crucial difference in the practical intelligence of animals and humans (children, in this case), that depends on the capacity of children to break loose from the given structure of the field of attention. Children can reconstruct their perception and master their attention, and one way to do this is to use

external representations, for instance, indexical expressions.³ According to Vygotsky, symbolic activity has a specific organizing function that "penetrates the process of tool-use and produces fundamentally new forms of behavior." (1978: 24). This means that tool-use in animals and humans are radically different, and that the difference comes with the human ability to engage in symbolic activity and entertain context- and referent-independent thoughts. Language becomes a second-order tool.

Wynn (1993) maintains that there is another difference between tool-use and language-use, except for context-independence. It is that tool behaviour appears to be more idiosyncratic than language-use. It is social in the sense that the learning of it depends on knowledge that is transmitted from one member of the society to another and on tradition. But the way the individual executes the behaviour may be attuned to his or her own conditions to such an extent that the segments of a behaviour that leads to a certain goal may be quite different in two individuals.

Wynn's observation suggests that learning of tool behaviour is based on emulation. This would mean that tool-use and language-use rely on different kinds of learning, something that apparently points to an important difference between the two as far as their evolution goes. Does this speak against the analogy between tool-use and language-use that I have been making? Not really, since different kinds of tool-use rely on different kinds of learning. Furthermore, language-use is also to a certain extent idiosyncratic. Rather, Wynn's remark brings forth that tool-use and language-use alike may depend on emulation as well as imitation.

As activities, then, the use of both tools and language relies on a capacity to use external representations in order to reach the desired goal. Added to the perceptual categories arising from attention-focusing, that were brought up in section 2, is the *social space*, also with an origin in attention-focusing. The social space prepares for advanced forms of learning. It does not only depend on the existence of perception and logical categories, but may, in turn, shape them.

The emergence of an open, social space is crucial both for advanced forms of intentionality and language-use. Unfortunately, theories about thought and language have sometimes stressed the role of an inner environment containing referent-independent representations of strategy and goal to such an extent that the link to the execution of the intended action and the problems that such an execution may present to the subject have been forgotten. An inner environment is not by itself sufficient for advanced forms of intentionality, like long-term or co-operative planning (Brinck and Gärdenfors, to appear; Ingold 1993a). Plans evolve and are tested in real-life situations in which other subjects happen to intervene and

³Cf. the work of Ballard on deictic representation, attention-focusing, and reconstruction of the perceptual field: Ballard (1991); Ballard, Hayhoe, Pook and Rao (1997).

sometimes even are necessary for the execution of the plan and in which the circumstances for various reasons may be unforeseen. The capacity to perform an action on purpose demands sensitivity to the circumstances in the external world.

Likewise, it is not clear how a subject can understand the point of an action, that is, exhibit the kind of understanding required to imitate and learn a purposive behaviour, unless she can carry it out in real life. Understanding is reached by performing the action in a social setting. It is not primarily an intellectual feat, but a practical one. Cognitive faculties are designed as activities in the world (Shanon 1993: 343). As pointed out in section 2, understanding does not always involve intentional loops, but can result from attention-contact and behavioural co-ordination, which issues from indexical reference and sensitivity to information that is given in the actual context.

Intentionality and understanding thus arise when actions actually are performed. Taking the step from the inner to the outer sphere does not only consist in the execution of a rehearsed role, but the execution still is part of the strategy to reach the desired goal. The execution may itself cause changes in the plan (Brinck, to appear). Both tool-use and linguistic behaviour are in this sense activities and not intelligences.⁴ They do not primarily consist in theoretical knowledge and cannot be learned, or planned, separately from the goings on in real life.

So, what is it about tool-use that makes it relevant for the story of how language-use has evolved? Tool-behaviour manifests two characteristics that seem to be necessary not only for language-use but also for full-blown intentionality: first, the kind of understanding that lies behind imitation, characterized by generality, purposive or intentional goal-direction, and a combination of theory and praxis, and second, the capacity to externalize one's representations and start using them as tools in the same way as humans use language as a tool. Humans use language in a general and systematic way for getting messages across and to reach desired goals, that is, they use language as a means to an end.

Ingar Brinck, Department of Philosophy, Lund University, Sweden
ingar.brinck@fil.lu.se

⁴Tim Ingold (1993b) makes a similar point. He underlines that tool-use and speech play an important part in the evolution of human beings "as animals endowed with specific capabilities of action" (p. 444). P. C. Reynolds (1993) draws attention to the social or co-operative character of tool-use.

References

- Arbib, M. A. and G. Rizzolatti (1996) "Neural Expectations: A Possible Evolutionary Path from Manual Skills to Language" in *Communication and Cognition* Vol. 29, Nr. 3/4, 393-424
- Ballard, D. (1991) "Animate Vision" *Artificial Intelligence* 48, 57-86
- Ballard, D., M. Hayhoe, P. Pook, R. Rao (1997) "Deictic Codes for the Embodiment of Cognition" *Behavioral and Brain Sciences* 20 (4), 723-767
- Barsalou, L., "Ad hoc categories" *Memory & Cognition* 11 (1983) pp. 211-227
- Barsalou, L. & Prinz, J., "Mundane Creativity in Perceptual Symbol Systems" in T. B. Ward, S. M. Smith & J. Vaid (eds.) *Creative Thought: An Investigation of Conceptual Structures and Processes* American Psychological Association: Washington, DC (1997)
- Brentano, F. (1973) *Psychology from an Empirical Standpoint*, London: Routledge & Kegan Paul
- Brinck, I. (1997) *The Indexical 'I'*, Dordrecht: Kluwer Academic Publishers
- Brinck, I. (to appear) "Procedures and Strategies: Context-Dependence in Creativity" *Proceedings from the International Conference on Discovery and Creativity 1998* Tilburg: Tilburg University Press
- Brinck, I. and P. Gärdenfors (to appear) "Representation and Self-Awareness in Intentional Agents" *Synthese*
- Deacon, T. (1997) *The Symbolic Species: The Coevolution of Language and Brain*, Cambridge, MA: MIT Press
- Deacon, T. (1996) "Prefrontal Cortex and Symbol Learning" in *Communicating Meaning: The Evolution and Development of Language*, eds. B. M. Velichkovsky and D. M. Rumbaugh Mahwah, NJ: Lawrence Erlbaum
- Freyd, J. J. (1987) "Dynamic Mental Representation" *Psychological Review*, vol. 94, 427-438
- Gärdenfors, P. (1996) "Cued and Detached Representations in Animal Cognition" *Behavioural Processes* 36, 263-273
- Gibson, J. J. (1986) *The Ecological Approach To Visual Perception*, Hillsdale, NJ: Lawrence Erlbaum Ass.
- Gibson, K. and T. Ingold, eds. (1993) *Tools, Language and Cognition in Human Evolution*, Cambridge: Cambridge University Press
- Glaserfeld, E. von (1976) "The Development of Language as Purposive Behavior" in eds. S. R. Harnad, H. D. Steklis, and J. Lancaster *Annals of the New York Academy of Science*, 280: 212-226
- Gómez, J. C. (1994) "Mutual Awareness in Primate Communication: A Gricean Approach" in *Self-Awareness in Animals and Humans*, eds. S. T. Parker, R. Mitchell, and M. Boccia Cambridge: Cambridge University Press
- Gómez, J. C. (1998) "Some Thoughts about the Evolution of LADS, with Special Reference to TOM and SAM" in *Language and Thought* eds. P. Carruthers and J. Boucher Cambridge: Cambridge University Press
- Ingold, T. (1993a) Introduction to Part V, in Gibson, K. and T. Ingold
- Ingold, T. (1993b) "Tool-use, Sociality and Intelligence" in Gibson, K. and T. Ingold
- Jakobson, R. and M. Halle (1956) *Fundamentals of Language*, Gravenhage
- Johnson, M. (1987) *The Body In the Mind*, Chicago: University of Chicago Press
- Lakoff, G. and M. Johnson (1980) *Metaphors We Live By*, Chicago: University of Chicago Press
- Lakoff, G. (1987) *Women, Fire, and Dangerous Things*, Chicago: University of Chicago Press
- Luria, A. R. (1973) *The Working Brain*, New York: Basic Books
- Mandler, J. (1992) "How to Build a Baby: II. Conceptual Primitives" *Psychological Review* Vol. 99 No. 4, 587-604

- Mandler, J. (1997) "Preverbal Representation and Language" in *Language and Space* eds. P. Bloom, M. Peterson, L. Nadel, and M. Garrett Cambridge, Mass.: MIT Press
- Mead, G. H. (1934) *Mind, Self, and Society*. Chicago: University of Chicago Press
- Mead, G. H. (1938) *The Philosophy of the Act* Chicago: University of Chicago Press
- Nowicki, S., and J. Podos (1993) "Complexity, Coupling, and Contingency in the Production of Birdsong" in *Perspectives in Ethology, Vol. 10: Behaviour and Evolution*, eds. Bateson *et al.* New York: Plenum Press
- Peirce, C. S. (1935) *Collected Papers*, Cambridge, Mass.: Harvard University Press
- Reynolds, P. C. (1993) "The Complementation Theory of Language and Tool-Use" in Gibson, K. and T. Ingold
- Shanon, B. (1993) *The Representational and the Presentational*, New York: Harvester Wheatsheaf
- Sjölander, S. (1995) "Some Cognitive Breakthroughs in the Evolution of Cognition and Consciousness" *Evolution and Cognition* 1, 3-11
- Tomasello, M. (1996) "Do Apes Ape?" in *Social Learning in Animals: The Roots of Culture*, eds. C. M. Heyes and B. G. Galef New York: Academic Press, Inc.
- Tomasello, M. (1998) "Reference: Intending that Others Jointly Attend" *Pragmatics & Cognition* Vol. 6 (1/2): 229-243
- Vygotsky, L. S. (1978) *Mind in Society. The Development of Higher Psychological Processes*, Cambridge, Mass.: Harvard University Press
- Wynn, T. (1993) "Layers of Thinking in Tool Behaviour" in Gibson, K. and T. Ingold