

ATTENTION AND THE EVOLUTION OF INTENTIONAL COMMUNICATION

Ingar Brinck, Department of Philosophy, Lund University
Kungshuset, Lundagård SE-222 22 Lund Sweden
ingar.brinck@fil.lu.se

ABSTRACT.

Intentional communication is perceptually based and about attentional objects. Three attention mechanisms are distinguished: scanning, attention attraction, and attention-focusing. Attention-focusing directs the subject towards attentional objects. Attention-focusing is goal-governed (controlled by stimulus) or goal-intended (under the control of the subject). Attentional objects are perceptually categorised functional entities that emerge in the interaction between subjects and environment. Joint attention allows for focusing on the same attentional object simultaneously (mutual object-focused attention), provided that the subjects have focused on each other beforehand (subject-subject attention). It results in intentional communication if the subjects attend to each other as subjects (i) capable of attending, and (ii) attending in a goal-intended way. Intentional communication is fundamentally imperative and adapted to action.

Attention and the Evolution of Intentional Communication

1. Introduction

One of the fundamental problems concerning the evolution of human language is to account for the transition from pre-intentional to intentional communication. The reason why it is important to account for this transition is that it seems to lie at the heart of the difference between human and animal communication. The clearest case of intentional communication is the use of human language, often held to depend on conceptual representations. Intentional communication can be described as aiming at indirectly influencing the behaviour of the recipient as a consequence of changing the mental state of the recipient. Pre-intentional communication, on the other hand, is aimed at directly influencing the behaviour of the recipient.

In the following my aim is to clarify the notion of intentional communication by explaining it as a form of attention-focusing that results in a particular kind of joint attention. Such an explanation is valuable for several reasons. For one thing, it allows us to understand the underlying mechanisms of intentional communication. For another, it makes it possible to judge whether other beings than humans can communicate intentionally, since the explanation dissociates the concept of intentional communication from linguistic communication.

The leap from pre-intentional to linguistic, intentional communication is often explained by the emergence of certain areas of the brain. It is very probable that these areas emerged in order to handle increasingly complex, social situations. Examples of such situations are the social co-ordination and planning pressed forward by environmental changes (the fourth glaciation); keeping up and cementing social bonds in increasingly large groups (Dunbar 1993); and cultural adaptation as a result of intraspecies competition (Donald 1991). But the question remains what made these areas of the brain evolve in the way they did, and why linguistic communication took the form it did. I suggest that ways of communicating intentionally existed before language as we now know it developed, and that they influenced the way language subsequently developed in response to the changing conditions.

I defend the view that intentional communication is characterised by the capacity, first, to focus attention in a goal-intended manner, and, moreover, to engage in joint attention towards an attentional object, one that is shared with at least one other subject. Intentional communication is triadic: sender and receiver attend to a shared object. To explain how communication can be intentional without involving human language, I will address two issues: first, in section 2, the nature of perceptual, or non-conceptual, intentionality, and, second, in section 3, how an intentional object may be determined and shared when the subjects that come to share it do not have recourse to a natural language. In section 4, an analysis of joint attention is put forward. In the final section, a view of intentional communication as fundamentally adapted to action is advocated. It is suggested that protodeclarative expressions are a special case of protoimperatives, and that the aim of both kinds of expression is to regulate behaviour.

2. Perceptual intentionality

Natural languages can be used to communicate concepts and ideas. Language makes it possible to talk about things that are not present in the context of utterance, and even about pure fantasies. These features of language-use depend on that the users have full-blown intentionality (Brentano 1874/1973). In short, intentionality is the property of a mental state to be directed, that is, be about something in the sense of having a content, and be aimed at a so-called intentional object.

The intentional object that the content represents may not exist, as, say, Santa Claus. Then we have a referent-independent representation. The representation may also be independent of any particular contextual constraints on the content. Then the representation is context-independent. For instance, the notion of a rectangle may be represented similarly in all contexts. Full-blown intentional communication can make use of both referent- and context-independent representations. Tomasello and Call (1997: 243, 262) underline that intentional communication involves a flexibility in attaining the goal. The same sign may be used in different communicative contexts, and different signs in the same context. This kind of flexibility depends on that the subject can disengage signs from contexts and consequently on context-independence. However, it does not require referent-independent signs.

In using the term full-blown when describing intentionality above, I indicate that intentional communication may come in degrees, depending on how many of the features of intentionality it displays. Perceptual intentionality is a precursor to full-blown intentionality. It consists in the capacity to focus attention - corresponding to the intentional state's aiming at an object - and perceptually categorise the attentional object - corresponding to the state's being about the object. In this section, I will give an account of the attention mechanism and explain how attention can constitute perceptual intentionality.

Attention is important for any information-gathering system, since it helps the organism in handling incoming information in a diversified way. It consists in an increased awareness of something either external or internal to the subject. It can be directed at behaviours, sensations, perceptions, or conceptions, and may be involuntary as well as deliberate. It has been debated whether there is an early selection or a late selection among input in human attention. Recent research suggests that there is a selection at different stages of information-processing (Pashler 1998). Stimuli rejected at an early stage are less completely analysed than at a later stage. Semantic processing occurs late, while perceptual processing occurs early.

Attention thus proves to be complex, making use of different kinds of processing and working against different kinds of memory. Kinds of processing that may occur are sensory ones, which do not involve categorisation of input, but directly trigger motor patterns; perceptual ones, which involve categorisation according to context-sensitive and subject-dependent criteria; and conceptual ones, which rely on operations on stable categories that may enter into general, propositional inferences. Different forms of processing involve different kinds of memory, variously described as procedural memory, iconic or sensory memory, embodied memory, and semantic memory (Baddeley 1990; Glenberg 1997; Tulving, E. 1984; Tulving & Schacter 1990).

I will distinguish between three attention-mechanisms, all of which may occur within a single perceptual system: scanning, attention attraction, and attention-focusing (Brinck 1997). The distinction is intended to capture the different functions of attention in relation to increasing degrees of control and complexity of processing. The claim is that the most complex attention mechanism, attention-focusing, is necessary for achieving joint attention.

Under normal conditions animals (and humans) are immersed in a constant flow of information that provides a basic state of arousal. It functions as a background against which the subject herself and the part of her environment that she attends to stand out (Gibson 1986; Luria 1973). Scanning of the environment is continuous. It consists in a search directed at discovering possibilities to act (Gibson 1986). Registered information guides movements and triggers actions in particular contexts. Bodily movements, perception, and environmental changes are continuously attuned in an on-going, on-line co-ordination (Shanon 1993). Actions triggered during scanning are supported by pragmatic representations. These represent object attributes as affording specific motor patterns, not as cues for a given perceptual category (Jeannerod 1994). Pragmatic representations do not call for binding attributes to a single entity.

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 and events in real time (Freyd 1987). Such changes shift the direction of the scanning and the behaviour. Discrimination in scanning and attention attraction 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. It is an entirely contextual process. Thus categorisation does not occur at any of these stages of attention. Categorisation relies both on discrimination *between* instances of categories and identification of items *within* category boundaries. It arises on the level of attention-focusing.

Attention-focusing constitutes a perceptual form of intentionality directed at attentional objects. It occurs when the subject focusses on that which happened to attract her attention and categorises it perceptually (Barsalou & Prinz 1997; Mandler 1992, 1997). The subject freezes a piece of the transient reality by suppressing most of the incoming information. Simultaneously information, stored in long-term memory, is added. This is achieved by the subject's making associations and perceptual inferences based on her previous encounters with similar situations.

I maintain that attention-focusing makes it possible to focus on particular as well as types of attentional objects. Perceptual concepts representing types can be generated through the process of categorisation that is initiated by attention-focusing. Barsalou (1999) puts forward a theory of perceptual concepts that explains how this is done. He

describes perceptual concepts as simulators that simulate events or entities in their absence, analogously to the simulations that underlie mental imagery. Each simulation provides a specific way of representing or conceptualising a given kind, the kind being connected to a particular simulator.

According to Barsalou, a simulator contains the knowledge that allows a subject to represent an entity or event adequately in different situations. It is used to identify members of a category and provides categorical inferences about their structure, history and behaviour, and suggests ways of interacting with them. Perception is first parsed into schematic components. These components are integrated across instances into frames, from which simulators develop that represent types of entities. From the frames, new simulators can be created in appropriate settings.

It should be noted that perceptual concepts are not stable and general in the way that usually is required for propositional, and purely conceptual, representations and thoughts, such as logical inferences. Their instantiation is always evoked by and tuned to the context. Thus, perceptual concepts do not seem able to support fully context-independent thought in the way that general concepts are supposed to do. Nevertheless, they introduce an element of generality by providing a core that underlies the instantiations of the perceptual concept, and that keeps together different manifestations of the concept in different contexts.

Perceptual concepts have a similar function as linguistic concepts, or symbols, in the sense that they are what communication is about. But symbols are quite different from perceptual concepts. Tomasello (1998) describes symbols as social, intersubjective, and bi-directional. A competent user understands that the same symbol can be used with the same intent by different users. Deacon (1996, 1997) explains how symbols are related to each other through a complex system of relations, like opposition, substitutability, and adjacency. Symbolic reference depends on indirect reference to objects, with the help of other symbols. 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 a subject learns about an indexical relation, this knowledge must be put into a "predetermined slot" among the symbolic relations in order to qualify as a symbol. The capacity to fit the index into the slot is an instance of a logical-categorical generalisation (1996: 129).

As opposed to perceptual concepts, symbols are social and general, and may be both referent- and context-independent.

Since attention-focusing makes it possible to focus on the core concept of a particular attentional object as well as on types of attentional objects, it enables several cognitive capacities. Among these we find the capacity to re-identify a particular object across time and space. This capacity depends on the organism's having access to perceptual concepts. A creature not capable of attention-focusing, and consequently not capable of forming perceptual concepts, will not perceive the world as consisting of individuals with an identity over time, but merely as a world of unrelated features, or bundles of features. When attention is continuously engaged by one item, the organism will of course be able to track the item through time and space. But as soon as there is an interruption in attention, or as soon as re-focusing is required, the item is lost. The organism will not be able to retrieve it by re-identification. In this sense, the role of perceptual concepts is to create a bridge through time, connecting disparate sense-impressions.

Attention-focusing is moreover necessary for context-independent content as used in intentional communication. Such content was characterised above as being doubly flexible: the same content can be used in different contexts of communication, and different contents can be used in similar contexts. Attention-focusing initiates categorisation which in turn introduces context-independence of content and representations. As mentioned above, categorisation also underlies the type-token distinction: tokens may appear in different contexts as instances of a certain type or category.

A further claim concerning the effects of attention-focusing is that at this level perception may be experienced in a particular way by the subject. This means that the subject may be phenomenally aware of the item (or have an experience of perceiving it), and that the item emerges as something beyond just an encoded sensory stimulation. The item can receive a value to the subject, for instance, by its being presented to the subject while evoking a particular emotion. The reason why awareness of value would arise at this level and not before is that for such awareness to exist, there must be something that can be categorised independently of particular values for the values to be perceived in themselves. Or else the value cannot be consciously experienced as an

independent property of the item, instead of as an integral part of the item. When perceptual categorisation is in place, it becomes possible for the same item to appear together with different values, and the same value can be perceived as attached to different items.

My view of attention can be compared to the one put forward by Mack and Rock (1998). They too maintain that attention is directed at an item. They explain the difference between different levels of consciousness and kinds of knowledge of the attentional object by appealing to different stages of attention. At an early level, the object is implicit, while only if attention is engaged (*cf.* my concept of attention-focusing) by the object will it become an object of conscious perception that acquires an experienced value to the subject.

Mack and Rock contend that stimuli have a value also when not consciously perceived. This explains why items, for instance, your own name, can be detected under conditions of inattention. Mack and Rock hold that we pick up high-level and complex stimuli, not simple ones from which we would construct complex objects during later stages. Nevertheless, the question whether we perceive simple or complex stimuli is moot. As mentioned above, it seems that selection among stimuli occur during the whole process of attention, and that whether we hold that selection is early or late depends on what kind of processing we focus on - sensory, perceptual, or semantic.

As a further support for the three-stage model of attention I am putting forward here, consider that to be able to benefit from contextual information in making judgements, a subject does not only need automatic discrimination (as during scanning and attention attraction) and semantic categorisation processes. Voluntarily controlled attention processes are also needed. The latter are more flexible than those that rely on semantic categorisation, in the sense that they may be tuned to the context. At the same time these processes are not fixed or driven by the environment as in scanning or attention attraction. Examples of voluntarily controlled attention processes are spatial selective attention, manipulation of mental images, controlled cross-modal representation, and spontaneously generated mental constructions (Brinck 1999, Flowers & Garbin 1989). These processes are made possible by the mechanism I call attention-focusing.

Attentional states can be focussed, that is, directed, either in a weak or in a strong sense. When directed in the weak sense, I will say that they are goal-governed. Goal-governed attention is focussed as a direct consequence of attention attraction. There is no prior intention to focus attention. In this case, attention-focusing is not induced by the subject herself, but triggered by the environment and tuned to a particular action. The goal controls attention.

When attentional states to the contrary are directed in the strong sense, they are goal-intended. Goal-intended attentional states are intentional; their directedness does not depend on being directly prompted by something external to the subject. The subject has an independent motivation to focus her attention, one that is not a direct function of attention attraction. She can operate freely on the context and re-arrange the structure of the perceptual field, for instance, by re-categorising its content according to different functions or possible actions. Vygotsky (1978) underlines the role of language for doing so, but language does not seem necessary for perceptual re-categorisation (Barsalou & Prinz 1997).

Goal-intended attention does not exclude that the subject attends to something as a function of having learnt to attend to this item in similar contexts. But for the attentional state to be goal-intended, it must be the case that, for one thing, the subject can have the same attentional state in different contexts, and, for another, that the subject can have different attentional states in the same context. If not, the case will still be one in which the external environment prompts or triggers the attentional state. Then we will not have the flexibility in attaining the goal that is mentioned by Tomasello and Call (1997).

Goal-intended attentional states make use of general perceptual concepts that represent attentional objects. The goal may be formed independently of the context. By producing an attentional object that can be shared among communicators, joint attention-focusing makes goal-intended communication possible among subjects. In the next section, I will explain how an attentional object can be shared.

3. Shared attentional objects

The recipient can know what the sender intends, and sender and recipient can share the attentional object, if they engage in joint attention. Joint attention is, as noted by Gómez (1994, 1998) and Tomasello (1998, 2001), the clue to intentional communication. Joint attention has been increasingly emphasized in research on language development and in the field of the evolution of communication (*cf.* Gómez 1994, 1998; Moore & Dunham, 1995; Tomasello 1998). The interplay between attention and intention within communicative contexts has also been studied from an experimental and technological point of view in work on human-computer interaction (Velichkovsky & Hansen, 1996; Velichkovsky, Sprenger & Unema, 1997; Vertegaal, Velichkovsky & Van der Veer, 1997). In the present and the following sections I will examine the notion of joint attention. The concepts of a shared object, to be dealt with in this section, and goal-intention (introduced in section 2) are central to the account.

Joint attention is based on the ability of two or more subjects to focus their perception simultaneously, as a consequence of attending to each other, on a shared attentional object. An attentional object that can be shared is an object, action, or event that in principle is mutually accessible, and accessible in the same way, for two or more subjects. It is such that in a particular context it is possible for two or more subjects to interact with it in a similar way on a perceptual basis. When two or more subjects do so it is shared.

Sharing does not imply that the subjects share all possible ways to interact with it. They do not have to conceive of or describe it in exactly the same way, or in fact share all the possible ways there are to interact with it. All that is required for sharing the attentional object is that there is some common way the subjects use, act on, or respond to the object, and that they exercise this way simultaneously in the same physical context. This will make them categorise the attentional object in a similar manner in that context. The knowledge that subjects use in categorising constitutes a practical rule or know-how. Anything that subjects together interact with in the manner described will be a shared object - a common prey, a tool, a toy. Shared objects do not only exist beforehand, like trees to climb in, but can be created in interaction with the environment and other subjects, like branches as (chimpanzee) nutcrackers.

Thus attentional objects are in a sense perceptual objects, based in perceptual categorisation. How such perceptually categorised objects emerge has been described by several authors. von Uexküll (1982: 74) argues that "all the properties of objects are actually nothing more than perceptual cues that are imprinted on them by the subject with which they enter into a relationship". He holds that the habitats of different kinds of animals constitute different kinds of meaningful, subjective universes, created from the interaction that each animal is capable of with its environment. According to von Uexküll, objects are functional entities, and a single item may acquire different functions in different contexts, and thus emerge as a new kind of attentional object in every distinct context. von Glasersfeld (1976) describes how sensory signals can be coordinated as a recurrent pattern in a context and in connection to a specific activity. Subsequently the pattern can be disconnected from both the presence of sensory signals and of the context in which it emerged. Then the pattern has acquired the kind of intentional properties that was described above.

How attentional objects are perceptually taken may differ between types of contexts without this posing a problem for sharing them in a particular context. On the contrary, it may be easier to share them as presented contextually, than generally. The way they are presented in a perceptual context often is connected to particular actions. Perceptual representations depend on the kind of over-all activity that the subject is engaged in. Clark (1997: 50) maintains that perception is "geared to tracking possibilities for action". Perceptual representations simultaneously describe the given contextual data and prescribe an appropriate action, by representing the data as adapted to a specific action. That the object is tuned to the circumstances of the context may facilitate discerning the object. This in turn will help different subjects to converge on the same object. Shared attentional objects thus allow for joint attention. But which is the process that leads up to joint attention?

4. Joint attention

Joint attention is the ability to engage in mutual, object-focused attention based on subject-subject attention. To put it differently: joint attention allows for focusing on the same attentional object simultaneously (mutual object-focused attention) provided that

the subjects have focused on each other beforehand (subject-subject attention). Attention can thus be focused on other subjects as well as on objects.

In mere subject-focused attention, attention is directed at the behaviour of other subjects. By, for instance, looking at each other, two subjects can detect their respective attentional objects. They do so on the basis of the direction of their respective movements in combination with a salient object that functions as a target. Co-ordination of attention based on saliency and behavioural co-ordination result in mutual object-focusing, during which the subjects can focus on the same thing simultaneously. Note that the perceptual saliency of features is connected to species-specific, as well as individual values and affordances.

But mutual object-focusing does not take us all the way to joint attention. The reason is twofold. Mutual object-focusing is behaviourally based and it is not goal-intended. It spreads automatically among subjects, as a function of attention-attraction. Consider the behaviour of a group of antelopes when a predator is approaching. The predator attracts the attention of an antelope, and the other antelopes have their attention attracted by the behaviour of the first antelope. As a consequence, they may focus their attention on the predator too. They will then share their object of attention. In case the behaviour of the predator signals that it is hungry and hunting, the situation will make the antelopes flee. Otherwise, they will keep their attention focused on the predator until it goes away. The antelopes engage in mutual object-focusing on the basis of behavioural co-ordination and contextual saliency.

There is also the case when attentive behaviour spreads among a group of animals, but they do not focus on the same thing. Say that something attracts the attention of one of the antelopes, and it focuses on that object. The rest of the antelopes will co-ordinate their behaviour with the one that is attending. But let us say that the object that attracted the attention of the first one is no longer present, or no longer is salient in a way that distinguishes it from the context for the other animals. Then the attentive behaviour of the herd as a whole will be co-ordinated, but without there being a common object for them to focus on. Thus there is a co-ordination of behaviour, but no mutual object-focusing.

Joint focusing of attention is, as mentioned above, based on subject-subject attention. It involves attentional co-ordination. By the term **subject-subject attention** I

mean, first, that the subjects, while attending to each other attending, actually attend to each other as subjects capable of attending. One might think that this necessarily involves having a theory of mind and being able to grasp that other subjects have mental states. This, I submit, is however not the case.

The awareness of states of attention relies on observational and behavioural evidence (*e.g.*, facial expression, gaze, body posture) that is related to a certain kind of action readiness and vigilance. Attention rests on information laid out in the environment as opposed to such attributed to mental states. Awareness of the capacity to attend in others, a capacity that is not manifested in behaviour until it actually constitutes a state of attention, consists in having an expectation of others to manifest attentional behaviour. Whiten (1993: 378) points out that attention nevertheless is not a purely behavioural state. The claim is based on a distinction between what a subject sees and what it notices: "The subject is mindreading in the sense that it is making judgements, not directly about observed behavior, but rather about what its competitors *might or might not notice.*"

The first condition for subject-subject attention (that the subjects, while attending to each other attending, actually attend to each other as subjects capable of attending) has been emphasized by J. Bruner (1998), J. C. Gómez, M. Tomasello, and others. This condition ensures that the subjects attend not only to each others behaviour, but to this behaviour as being attentional and directed. Attending to each other as subjects capable of attending may result in attention contact. But the first condition is not, I maintain, sufficient for joint attention in the absence of contextually salient objects.

Furthermore, the subjects will have to attend to each other as capable of attending in a goal-intended way, that is, in a way that is not controlled by the object of attention. The latter condition is important in order to distinguish subject-subject attention as it occurs in joint attention from mere subject-attention occurring in mutual object-focusing. Goal-intention provides for the capacity to either direct or follow the attention of the other subject in the absence of salient objects (though not in the absence of objects altogether). This means that attention-focusing can be guided by the subjects' mutual attention to each other instead of by the environment.

It should be stressed that animals capable of subject-subject attention primarily attend to each others attentional states, not to the behaviour that is a consequence of attention. However, they extract information about the attentional states of others not explicitly as being about such states, but as manifested in bodily behaviour. They can distinguish non-attentional head and body-orientation from attentional gaze, and they can engage in attention contact, during which they simultaneously check each others state of attention, for instance, by eye contact (Gómez 1994).

The antelopes in the example above were not capable of subject-subject attention. They could not make the distinction between behaving non-intentionally as a consequence of having one's attention attracted by an object and intentionally orienting oneself at a particular attentional object. They did not recognise the non-attentional bodily orientation of the first antelope as tied to a particular action, but repeated the movement without picking up the motive for doing so. Neither could they follow its gaze in order to converge on its object of attention.

Tomasello (1996) notes that reproducing behaviour intentionally, that is, by both understanding the goal of an action and the strategy used to reach the goal, requires more than sensory-matching and a perception of those changes in the environment that the behaviour to be reproduced has resulted in. He holds that intentionally reproducing behaviour requires imitative learning capacities. This is something that the antelopes in my example clearly lack. As I see it, imitative learning requires goal-intended subject-subject attention.

5. Communication and action

Directing somebody by gaze and other forms of goal-intended attentional behaviour are ways of making one's focus of attention accessible to others. It is the target of the attentional state that is shared and made available for others, not the mental states themselves or their internal content. Since attentional states are manifested behaviourally, it is not necessary to represent the mental contents of other subjects in order to make them behave as one intends. That means that intentional communication does not require a sharing of conceptual representations. Intentional communication can occur on the basis of mere perceptual intentionality and concern attentional objects.

Maybe this position at first glance seems to run counter to the view about intentional communication put forward at the outset. There it was claimed that intentional communication ultimately aims at changing behaviour by taking the detour over mental states. I have described the process behind non-linguistic intentional communication as one that does not involve explicitly taking into account the mental states either of oneself or others. But as indicated by, for instance, Whiten (1993), attention is not a purely behavioural state. It involves cognitive capacities, although this is not necessarily recognised by those communicating. **Bodily states of attention have a mental counterpart. The point is that attention lies between pure behaviour and reflected intention or theories of mind. That is why communication based on perceptual intentionality and joint attention can provide a bridge between the transmission of information and the transmission of linguistic messages.**

The imperative character of intentional communication reveals the close connection between perception and agency (Proust 1999). Perceptual demonstratives are primarily used to initiate action, not to inform. They do not state something about the referent, they show or display it, inviting the audience to respond. Intentional communication is a tool for making the audience do something, say, get something for the subject, share the subject's attitudes or reactions to the object in a certain context, and so on. Gómez, Sarriá, and Tamarit (1993) and Gómez (1994) describe the role of so-called protoimperatives to be to regulate the behaviour of others, and as a means by which the subject indirectly can make desirable changes in the environment by exploiting the behaviour of others. An important function of intentional communication is to regulate behaviour, that is, to initiate co-operative or competitive actions.

Sometimes the difference between protoimperatives and protodeclaratives is emphasized. It is suggested that there is a clear-cut and important difference between the two, from a developmental as well as an evolutionary point of view. I rather see protodeclaratives as a special case of protoimperatives. That means that they evolve directly from protoimperatives and do not constitute a distinct way of communicating. A similar view is, it seems, advanced by Gómez, Sarriá, and Tamarit (1993).

I agree with Gómez, Sarriá, and Tamarit (1993) that both protoimperatives and protodeclaratives are based in a first-order understanding of the external manifestations of attention and emotion. They assert, however, that while the function of

protoimperatives is to regulate action and behaviour, the sender's goal in using protodeclaratives is to monitor the other subjects' reaction to shared objects. But I do not think that it is the reactions *per se* that are interesting, but what they may lead to. According to Gómez *et al*, protodeclaratives do not seem to be involved in the regulation of behaviour. But as I see it, the two functions blend.

The function of protodeclaratives, I maintain, is to make it possible for the sender to evaluate the item that communication converges on. Monitoring the reactions of other subjects is a way to make an evaluation of the shared object. The evaluation will have a common, or in some cases even social, basis. It will not primarily be based on the sender's own reaction to the object, but on the observations that the sender makes of the reactions of others. Evaluation, in turn, serves the goal of action. It underlies taking proper action, the goal of which may be, as in the case of protoimperatives, to regulate or exploit the behaviour of others.

The aim of attentional and emotional exchange surrounding a shared item is consequently to close in upon an evaluation of the item. The evaluation will be based on a mutual recognition and exchange of attitudes towards the item. Thus, in contrast to in the case of protoimperatives, the attentional object is not only taken in together with the actions and values that it directly affords. It is perceived together with values that emerge in interaction with other subjects. The interaction builds on monitoring the evaluative attitudes others display towards the object in a particular context. Still, the ultimate goal is the same in the cases of protoimperatives and of protodeclaratives: to regulate behaviour.

The view that the pragmatic, action-related function of communication is basic receives much support from work within speech act theory and pragmatic linguistics (Grice 1989; Searle 1969; Sperber & Wilson 1986). For instance, Givon (1989) maintains that the difference between declaratives, imperatives, and interrogatives is graded. These three so-called prototype peaks of speech acts are cross-linguistic. They span a multi-dimensional mental space of several non-discrete socio-psychological dimensions. Within this space, the speech acts shade into each other.

What conclusion can be drawn from the scarcity of declaratives in nonhuman primates? Tomasello and Call (1997) suggest that it "may evidence of a lack of understanding that others intentionally perceive and attend to the world in ways that

may be influenced and shared.” The correctness of this remark depends on what is meant by understanding. If understanding is thought to be a conceptual, higher-level kind of reflection it is probably right. If understanding, on the other hand, means to be able to tacitly predict actions of others, to share an activity and continue a series of shared actions in an accurate way, then it is not so. Then the understanding does not have to be explicit neither to observers, nor to users. It is something that the subject exercises and displays to others in praxis. And it does seem that nonhuman primates have a practical grasp of the fact that perception of and attention to the world can be influenced and shared.

The latter view of understanding was advanced by Wittgenstein (1953). Understanding an activity is to be able to follow the rules that govern the activity in a correct way. Whether the rule is regulative, that is, such that governs an activity that existed before the rule appeared (as social rules may define and refine bonds between relatives), or constitutive, that is, such that constitutes an activity (as the rules of chess do), understanding the rule consists in being able to follow the procedure it establishes. One argument behind this view of understanding is that higher-level thought simply is not necessary to know how to go on to follow those rules that govern shared activities, such as language use. The rule-following capacity is not fundamentally theoretical, or even conceptual, but a practical skill that is formed in actual social interaction.

If the scarcity of declaratives in nonhuman primates does not depend on a failure of understanding that attention to the world can be influenced and shared, what will explain it? One possibility might be that humans often theorise and make hypotheses about the nonpresent or the possible. To do so, they need to assume certain conditions, from which other conditions can be inferred. These assumptions are made with the help of declaratives. Hence, declaratives may have developed because they are needed in counterfactual reasoning about states that do not exist. Nonhuman primates do not seem to have a need for such reasoning in their natural habitat. They communicate about items that are present in their environment. What made the capacity for theorising and hypothesizing develop in humans may be changing social conditions as a result of population growth and increasingly large groups trying to keep together. What may have caused such a population growth we do not know for sure.

The keeping together of and providing for large interrelated groups of individuals require planning and being able to think about individuals that are not continuously present in the group. Social bonds too may be complex and difficult to keep track of. Planning will involve many indeterminate factors, as to times and places, participants in activities, and the nature of the activity to be performed. These factors can in turn be combined in different ways. It is not difficult to see why a flexible way of thinking, and thoughts with a content that is not constrained by what is present to the thinker in space and time, would be advantageous.

An additional reason for the scarcity of declaratives in nonhuman primates is related to the function that I claim protodeclaratives have. It might be that the increasingly large groups made humans develop skills in converging on evaluations of shared objects. Being able to converge on an evaluation, I maintain, is based on monitoring the external manifestations of attention and emotion of others. This is the function of protodeclaratives. Such convergence may be valuable in large groups as well as in smaller groups where members come and go, perhaps as a result of being part of several smaller groups. To nonhuman primates, on the other hand, living in more stable or homogenous groups, an interest in the value that an attentional object directly affords may suffice.

References

- Baddeley, A. 1990. *Human Memory. Theory and Practice*. Hove and London: Lawrence Erlbaum Ass.
- Barsalou, L.W. 1999. "Perceptual symbol systems" *Behavioral and Brain Sciences* 22(4): 577-660
- Barsalou, L.W. and Prinz, J.J. 1997. "Mundane creativity in perceptual symbol systems". In T.B. Ward, S.M. Smith, and J. Vaid (eds), *Creative Thought: An Investigation of Conceptual structures and Processes*. Washington, DC: American Psychological Association, 267-307.
- Brentano, F. 1874/1973. *Psychology from an Empirical Standpoint*. London: Routledge & Kegan Paul.
- Brinck, I. 1997. *The Indexical 'I'*. Dordrecht: Kluwer Academic Publishers.
- Brinck, I. 1999. "Non-conceptual content and the distinction between implicit and explicit knowledge". *Behavioral and Brain Sciences* 22(5): 760-761.
- Bruner, J. 1998. "Routes to reference". *Pragmatics & Cognition* 6 (1/2): 209-227.
- Clark, A. 1997. *Being There*. Cambridge, Mass.: MIT Press.
- Deacon, T. 1996. "Prefrontal cortex and symbol learning". In B.M. Velichkovsky and D.M. Rumbaugh (eds), *Communicating Meaning: The Evolution and Development of Language*. Mahwah, NJ: Lawrence Erlbaum, 103-138.
- Deacon, T. 1997. *The Symbolic Species: The Coevolution of Language and Brain*. Cambridge, MA: MIT Press.
- Donald, R.I.M. 1991. *Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.
- Dunbar, R.I.M. 1993. "Co-evolution of neocortex size, group size and language in humans". *Behavioral and Brain Sciences* 16(4): 681-735.
- Flowers, J.H. and Garbin, C. 1989. "Creativity and perception". In J.A. Glover, R.R. Ronning, and C.R. Reynolds (eds), *Handbook of Creativity*. New York and London: Plenum Press, 147-162.
- Freyd, J.J. 1987. "Dynamic mental representation". *Psychological Review* 94: 427-438.
- Gibson, J.J. 1986. *The Ecological Approach To Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Ass.
- Givon, T. 1989. *Mind, Code, and Context*. Hillsdale, NJ: Lawrence Erlbaum Ass.
- Glenberg, A. 1997. "What memory is for". *Behavioral and Brain Sciences* 20(1): 1-19.
- Gómez, J.C. 1994. "Mutual awareness in primate communication: a Gricean approach". In S.T. Parker, R.W. Mitchell, and M.L. Boccia (eds), *Self-Awareness in Animals and Humans*. Cambridge: Cambridge University Press, 61-80.
- Gómez, J.C. 1998. "Some thoughts about the evolution of LADS, with special reference to TOM and SAM". In P. Carruthers and J. Boucher (eds), *Language and Thought*. Cambridge: Cambridge University Press, 76-93.
- Gómez, J.C., Sarriá, E., and Tamarit, J. 1993. "The comparative study of early communication and theories of mind: ontogeny, phylogeny, and pathology". In S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (eds), *Understanding Other Minds. Perspectives from Autism*. Oxford: Oxford University Press, 397-426.
- Grice, P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

- Jeannerod, M. 1994. "The representing brain: neural correlates of motor intentions and imagery". *Behavioral and Brain Sciences* 17(2): 187-202.
- Luria, A.R. 1973. *The Working Brain*. New York: Basic Books.
- Mack, A. and Rock, I. 1998. *Inattention blindness*. Cambridge, Mass.: MIT Press.
- Mandler, J.M. 1992. "How to build a baby: II. Conceptual primitives". *Psychological Review* 99(4): 587-604.
- Mandler, J.M. 1997. "Preverbal representation and language". In P. Bloom, M.A. Peterson, L. Nadel, and M.F. Garrett (eds), *Language and Space*. Cambridge, Mass.: MIT Press, 365-384.
- Moore, C. and Dunham, P.J. (eds) 1995. *Joint Attention. Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum Ass.
- Pashler, H. (ed) 1998. *Attention*. Hove: Psychology Press.
- Proust, J. 1999. "Indexes for action". *Revue Internationale de Philosophie* 3(209): 321-345.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
- Shanon, B. 1993. *The Representational and the Presentational*. New York: Harvester Wheatsheaf.
- Sperber, D. and Wilson, D. 1986. *Relevance: Communication and Cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. 1996. "Do apes ape?". In J. Galef, and C. Heyes (eds), *Social Learning in Animals. The Roots of Culture*. New York: Academic Press, 319-345.
- Tomasello, M. 1998. "Reference: intending that others jointly attend". *Pragmatics & Cognition* 6(1/2): 229-243.
- Tomasello, M. 2001. "Perceiving intentions and learning words in the second year of life". In M. Bowerman and S. Levinson (eds), *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 132-158.
- Tomasello, M. and Call, J. 1997. *Primate Cognition*. New York: Oxford University Press.
- Tulving, E. 1984. "How many memory systems are there?". *American Psychologist* APA Award Addresses 40(4): 385-398.
- Tulving, E. and Schacter, D.L. 1990. "Priming and human memory systems". *Science* 247: 301-306.
- Velichkovsky, B.M. and Hansen, J.P. 1996. "New technological windows into mind: There is more in eyes and brains for human-computer interaction". In *Proceedings of ACM CHI96 Conference of Human Factors in Computing Systems*. Vancouver, Canada: ACM, 496-503.
- Velichkovsky, B.M., Sprenger, A. and Unema, P. 1997. "Towards gaze-mediated interaction: Collecting solutions of the "Midas touch problem". In S. Howard, J. Hammond & G. Lindgaard (eds.), *Human-Computer Interaction: INTERACT'97* (Sydney, July 14-18th), London: Chapman & Hall.
- Vertegaal, R., Velichkovsky, B.M. & Van der Veer, G. 1997. "Catching the eye: Management of joint attention in teleconferencing and cooperative work". *ACM SIGCHI Bulletin*, 29(4), 87-99.
- von Glasersfeld, E. 1976. "The development of language as purposive behavior". In S.R. Harnad, H.D. Steklis, and J. Lancaster (eds), *Origins and Evolution of Language and Speech*. *Annals of the New York Academy of Science* 280: 212-226.
- von Uexküll, J. 1982. "The theory of meaning". *Semiotica* 42(1): 25-82.

- Vygotsky, L.S. 1978. *Mind in Society*. Cambridge, Mass.: Harvard University Press.
- Whiten, A. 1993. "Evolving a theory of mind: the nature of non-verbal mentalism in other primates". In S. Baron-Cohen, H. Tager-Flusberg, and D.J. Cohen (eds), *Understanding Other Minds. Perspectives from Autism*. Oxford: Oxford University Press, 367-395.
- Wittgenstein, L. 1953. *Philosophical Investigations*. New York: MacMillan.

ACKNOWLEDGEMENTS.

I would like to thank Peter Gärdenfors, Lund University Cognitive Science, and Joëlle Proust, CNRS Institut Jean Nicod in Paris, for valuable comments. Thanks also to the audience at the Conference of the Evolution of Language, Paris, April 2000, where an earlier draft of this paper was presented.

ABOUT THE AUTHOR.

Ingar Brinck graduated in Philosophy at Lund University, with a PhD on the indexical 'I'. She is specialising in cognitive science and philosophy of mind and language. She has been teaching in philosophy and cognitive science since 1988, and currently holds a research position in philosophy at Lund University. She has published papers on creativity, self-consciousness, and representation, as well as a book, *The Indexical 'I'. The First Person in Thought and Language*. (Dordrecht: Kluwer, 1997). Now she is working on the evolution of language and on perceptual demonstratives. She is a member of the committee of the European Society of Philosophy and Psychology.