COOPERATION AND THE EVOLUTION OF SYMBOLIC COMMUNICATION

Peter Gärdenfors

Lund University Cognitive Science

Kungshuset

S-222 22 Lund, Sweden

E-mail: Peter.Gardenfors@lucs.lu.se

Abstract: Humans are the only animals who can deliberately plan for future needs. Our ability to have a ”theory of mind” is also a factor that facilitates advanced forms of cooperation and communication. The main thesis of this article is that symbolic communication evolved in order to make cooperation about future goals possible.

A consequence of this position is that the meanings of symbolic expressions emerge from the cooperation between individuals. I will discuss how constraints on semantic representations evolve through processes of cooperative sharing of knowledge and plans. As a paradigmatic case, I describe pragmatic settings of referential communication and provide a model of how the roles of names, nouns and adjectives can be explained by a process of abstraction that is based on principles of cognitive economy.

1. WHAT IS THE EVOLUTIONARY ROLE OF LANGUAGE?

Homo sapiens is the only species with a symbolic language. According to evolutionary theory, there should be some selective advantage that has fostered the development of language among humans. There are many proposals for such an evolutionary force. Some of the major ideas have been that (1) language brings with it the ability to *convey information* about prey or other food or about dangers of different sorts; (2) it is a result of sexual selection; (3) language replaces the social grooming found in monkeys and apes *as an instrument for building up coalitions and other social bonds* (the so called ”gossip theory” proposed by Dunbar (1996)); or (4) language is a ”mother tongue” that evolved among kin for ”honest” communication (Fitch, this volume). However, despite all the merits these proposals may have none of them can explain why language has *not* evolved among other apes or animals.

In this paper, I will propose another advantage of symbolic language that may be evolutionarily more important than those previously suggested: (5) language makes it possible to *cooperate about future goals*. I shall prepare the ground for this thesis by first arguing that humans are the only animals that can plan for future goals. If this is correct, language would indeed be beyond the cognitive reach of other species. I will then argue that symbolic communication is necessary for advanced cooperation. Finally, as a paradigmatic example of communicating about future goals, I will analyze the cognitive and communicative prerequisites for different types of referential expressions. The evolutionary gain of being able to communicate about referents that are not yet present is that more advanced forms of long-term planning become possible. However, the basis for it all is the notion of a representation. This will be the topic of the following section.

2. CUED AND DETACHED REPRESENTATIONS

In order to understand the functions of most of the higher forms of cognition, one must rely on an analysis of how animals *represent* various things, in particular the surrounding world and what it can offer. There is an extensive debate in the literature on what should be taken to be the appropriate meaning of ”representation” in this context (see e.g. Roitblat (1982), Vauclair (1990), Gärdenfors (1996) and Grush (1997)). Here I will not go into the intricacies of the debate, but only point out that there are different *kinds* of representations. In this paper, the focus will be on the kind of representations used in cooperative communication. A key point is that in order to give an accurate analysis of many phenomena in animal and human cognition, it is necessary to distinguish between two kinds of representations: *cued* and *detached* (Gärdenfors 1996).

A *cued* representation stands for something that is present in the current external situation of the representing organism. When, for example, a particular object is categorized as food, the animal will then act differently than if the same object had been categorized as a potential mate. I am not assuming that the animal is, in any sense, *aware* of the representation, only that there is some generalizing factor that determines its behavior. In general, the represented object need not be actually present in the actual situation, but it must have been triggered by something in a recent situation. Delayed responses, in the behaviorist's sense, are also based on cued representations according to this characterization.

In contrast, *detached* representations may stand for objects or events that are neither present in the current situation nor triggered by some recent situation. A memory of something that can be evoked independently of the context where the memory was created would be an example of a detached representation. Similarly, consider a chimpanzee, who performs the following sequence of actions: walks away from a termite hill, breaks a twig, peels its leaves off to make a stick, returns to the termite hill, and uses the stick to "fish" for termites. This behavior seems impossible to explain unless it is assumed that the chimp has a detached representation of a stick and its use.

I am not claiming that it is possible to draw a sharp line between cued and detached representations. There are degrees of detachment. However, I still believe that the rough distinction between the two major kinds of representations is instrumental in that it directs our attention to key features of the representational forms.[1]

What is the main evolutionary advantage of detached representations in comparison to cued ones? In order to answer this question, I will elaborate an idea introduced by Craik (1943, p. 61):

> If the organism carries a "small-scale model" of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which are the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to react on a much fuller, safer and more competent manner to the emergencies which face it.

I will call this kind of "small-scale model" the *inner world*. The inner world is necessary for representing objects (like food and predators), places (where food or shelter can be found), actions (and their consequences), etc., even when these things are not perceptually present. The evolution of such a representational power will clearly increase the survival chances of the animal.

As a tentative definition, the inner world of an animal will in this paper be identified with *the collection of all detached representations* of the animal and their interrelations. It should be noted that I am not assuming that the animal is *aware* of its inner world, nor of the processes utilizing this construct.

It seems that many animal species have inner worlds. For example, the searching behavior of rats is best explained is if it is assumed that they have some form of "spatial maps" in their heads. Evidence for this, based on their abilities to find optimal paths in mazes, was collected by Tolman already in the 1930's (Tolman 1948). However, his results were swept under the carpet for many years since they were clear anomalies for the behaviorist paradigm.[2]

3. ANTICIPATORY PLANNING

One of the main evolutionary advantages of an inner world is that it frees an animal, that is seeking a solution to a problem, from dangerous trial-and-error behavior. Jeannerod (1994) says that "actions are driven by an internally represented goal rather than directly by the external world." By exploiting its inner world, the animal can *simulate* a number of different actions in order to "see" their consequences and evaluate them (also compare Grush (1997) and Barsalou (1999)). After these simulations, it can choose the most appropriate action to perform in the outer environment. Of course, the *success* of the simulations depends on how well the inner world is matched with the outer. Evolutionary selection pressures will, in the long run, result in a sufficient correspondence between the inner and the outer world. As the Norwegian poet Olav Haugen writes: "Reality is a hard shore against which the wave-borne dreamer strands."

The ability to envision various actions and their consequences is a necessary requirement for an animal to be capable of planning. Following Gulz

[1] Another caveat concerning my use of the notion of representation is that I am not making any ontological claims: I am not proposing that representations are entities with some kind of reality status. Rather, I view representations as *theoretical terms*, in the way standardly conceived of in philosophy of science (e.g., Sneed (1971)). Representations are theoretical idealizations, similar to "forces" in Newtonian mechanics, that are introduced to predict and explain empirical generalizations (cf. Lachman and Lachman 1982).

[2] Vauclair (1987) provides a more recent analysis of the notion of a "cognitive mapping."

(1991, p. 46), I will use the following criterion: An animal is planning its actions if it has a representation of a goal and a start situation, and it is capable of generating a representation of a partially ordered set of actions for itself for getting from start to goal. This criterion presupposes representations of (1) goal and start situations, (2) sequences of actions, and (3) the outcomes of actions. The representations of the actions must be detached; otherwise it is not possible for the animal to choose different actions. In brief, planning presupposes an inner world.

Ethologists, who study animal behaviour, appear to be largely in agreement that certain animal species can plan in the sense defined here (see e. g. chapters 5, 7, 8 and 9 in Ellen and Thinus-Blanc, eds., 1987, and pp. 58–61 in Gulz 1991). Yet all examples of planning among animals available in the ethological literature concern planning for *current needs*. Apes and other animals start planning because they are hungry or thirsty, tired or frightened. Their motivation comes from the present state of the body. Oakley (1961, p. 187) writes:

> Sultan, the chimpanzee observed by Kohler, was capable of improvising tools in certain situations. Tool-making occurred only in the presence of a visible reward, and never without it. In the chimpanzee the mental range seems to be limited to present situations, with little conception of past or future.

Man seems to be the only animal that can *plan for future needs*. We can *foresee* that we will be hungry tomorrow and put away some of our food; we realize that it will be cold and windy in the winter, so we build a shelter in good time. (Chimpanzees build night camps, but only for the coming night.) Gulz (1991) calls the capacity to plan for the future *anticipatory planning*.

That apes and other animals are incapable of anticipatory planning is illustrated by an experiment with chimpanzees performed by Boysen and Bernston (1995). They put peanuts in two heaps of different size placed on a table out of reach of the apes. One ape was to point at one of the heaps, and then that heap was given to the other ape, while he himself got the one that he did not point at. The result of the test was surprising. The chimpanzee repeatedly pointed at the biggest pile and was very disappointed when that pile was given to the other ape, and he himself received the small pile.

The presence of the desired food seems to make them incapable of imagining the near future, in which the other party receives the pile that they choose and they are left with the other pile. Boysen and Bernston's experiment clearly shows how difficult it is for chimpanzees to manage even the simplest form of planning for a future goal. Deacon (1997, p. 414) writes that the choice is difficult for the apes since the indirect solution (choosing the small pile) is overshadowed by the direct presence of a more attractive stimulus, namely, the big pile. They cannot suppress their perception.

If one performs the same kind of experiment with human children, they have no problem choosing the small pile – from the age of two and up. They can imagine receiving the big pile when they point at the small one. When children are younger they behave more like chimpanzees.

Why is it cognitively more difficult to plan for future needs than for current ones? The answer has to do with the different representations that are required for the two types of planning. When planning in order to satisfy current needs, one must be able to represent actions and their consequences, and to determine the value of the consequences in relation to the needs one has at that moment. But no detached representation of that need is required. To plan for future needs, on the other hand, one must also be able to represent these potential needs (and to understand that some of them will arise). The available ethological evidence so far indicates that man is the only species with the ability to imagine future wishes and to plan and act accordingly (Gulz 1991).

4. Signals and Symbols

Planning, even anticipatorily, does not presume a language. Humans, as well as animals, can simulate sequences of actions in their inner worlds (Jeannerod 1994). Such simulations are, among other things, necessary for planning. Language is, in my opinion, a latecomer on the evolutionary scene.

Language presumes the existence of an intricate inner world. In order to make this clear, I will distinguish between *signals* and *symbols*. Both signals and symbols are tools of communication. The fundamental difference between them is that *the reference of a symbol is a detached representation, while a signal refers to a cued representation*. In other words, a signal refers to something in the outer environment, while a symbol refers to the inner world. The distinction between signal and symbol is also made by Sinha (this volume). However, he has a slightly different, but compatible, view of their roles: "Whereas a communicative signal can be viewed as an instruction (perhaps coded) to *behave*, the use of symbols involves two emergent properties, *reference* and *construal*." In this article, the role of symbols in establishing references to detached objects and in construing future goal will be in focus.

Language consists of symbols – it can be used to talk about things not present in the current situation. This idea can be traced back to Hockett's (1960) notion of "displacement." Glasersfeld (1977, p. 64) expresses the point as follows:

> [W]e can talk not only about things that are spatially or temporally remote, but also about things that have no

> location in space and never happen at all [...] in order to become a symbol, the sign must be detached from input. What the sign signifies, i.e., its meaning, has to be available, regardless of the contextual situation.

With few exceptions, linguistic communication is achieved with the aid of symbols. Sjölander (1993, pp. 5–6) explains elegantly what is missing in animal communication:

> The predominant function of language is to communicate about that which is not here and not now. A dog can 'say': I am angry, I want water, I want to go out, I like you, etc. But it has no communicative means enabling it to 'say': I was angry yesterday, nor can it 'say': I will be angry if you lock me up tonight again, and I will chew up the carpet. Likewise, the dog can 'say': There is a rat here! but it cannot 'say': There is a rat in the next room.
>
> […] Clearly, if you live in the present, communicating mainly about how you feel and what you want to do in the moment, the biological signals inherent in each species are sufficient. A language is needed only to communicate your internal representation of what could be, what has been, and of those things and happenings that are not present in the vicinity.

Symbols refering to something in one person's inner world can be used to communicate as soon as the listeners have, or are prepared to add, the corresponding references in their inner worlds.[3] The actual conditions of the outer situation need not play any role for the communication to take place: two prisoners can talk fervently about life on a sunny Pacific island in the pitch dark of their cell.

Many animals have intricate systems of signals, for example, the dances of bees. However, even if their dances seem to have a kind of grammar, they still consist only of *signals*. The bees categorize, in a sophisticated way, places where nectar can be found. The crucial point is that they only use their dances in a cued manner, and thus the dances are not symbols according to my criterion.

In spite of all attempts to teach apes various forms of symbolic codes (see e.g. Savage-Rumbaugh, Shanker, and Taylor (1998)), humans seem to be the only animals that use language in a fully detached way. Even though the bonobo Kanzi's performance is quite impressive, his use of symbols is dependent on the context: they mainly express requests to "direct teacher's attention to places, things and activities" (Savage-Rumbaugh et al., (1985, p. 658)). Human children, in contrast, use language at a very early stage outside the context of request. Vauclair (1990, p. 319) notes that "the use of symbols by apes is closely tied to the achievement of immediate goals, because the referents occur in the context of behavior on their objects." This is congenial with Gulz' (1991) conclusion that only humans are anticipatory planners. My conjecture is that this capability is required for the complete detachment of language. We are still waiting for Kanzi to tell us a story by the campfire.

5. Cooperation and Communication by Symbols[4]

Human beings as well as other animals cooperate in order to reach common goals. Even seemingly simple animals like ants and bees cooperate in building complex societies. However, their cooperation is *instinctive* – they have no detached representation of the goal their collaboration is aimed at. For lack of representations, they cannot create new goals of cooperation.

Nevertheless, for many forms of cooperation among animals, it seems that representations are not needed. If the common goal is present in the actual environment, for example food to be eaten or an antagonist to fight, the collaborators need not focus on a joint representation of it before acting.

If, on the other hand, the goal is detached, i.e. distant in time or space, then a *common* representation of it must be produced before cooperative action can be taken. In other words, cooperation about detached goals requires that *the inner worlds of the individuals be coordinated.* It seems hard to explain how this can be done without evoking symbolic communication.

A problem concerning collaboration in order to reach a detached goal is that the *value* of the goal cannot be determined from the given environment, unlike a goal that is already present on the scene. The value of the future goal has to be estimated by each individual with regard to possible outcomes.

Apes seem to lack the cognitive resources that are required for imagining future goals and calculating their values. This is a decisive difference between humans and apes. In my opinion, the main reason why apes cannot cooperate about future goals is that they are not capable of anticipatory planning.

Communication by symbols is quite intricate, because the meanings of the symbols are general and defined by interrelation. As mentioned earlier, it has so far not been shown that apes can communicate in a fully symbolic way (Deacon 1997, Tomasello 1999). Rather, it seems that apes in their natural habitat mainly exploit indexicals in their signalling.

[3] For a model theoretic account of how such communication can be established, see Gärdenfors (1993). A special case of the process will be discussed in section 7.

[4] This section is based on material from Brinck and Gärdenfors (2001).

Human language is the prototype example of a symbolic communication system. Clearly, human language paves the way for long-term cooperation and for cooperation towards future goals. As Boysen and Bernston's (1995) experiment indicates, it may be hard to give up a good in possession for a future, but more precious one.

An important feature of the use of symbols in cooperation is that they can set the cooperators free from the goals that are available in the present environment. The detached goals and the means to reach them are picked out and externally shared through the linguistic medium. This kind of sharing gives humans an enormous advantage concerning cooperation in comparison to other species. I view this advantage as a strong evolutionary force behind the emergence of symbolic communication. More precisely, I believe that there has been a coevolution of cooperation about future goals and symbolic communication (cf. the "ratchet effect" discussed by Tomasello 1999, pp. 37-40).

Language is based on the use of representations as stand-ins for entities, actual or just imagined. Use of such representations replaces the use of environmental cues in communication. If I have an idea about a goal I wish to attain, I can use language to communicate my thoughts. In this way, language makes it possible for us to *share visions*.

There are many kinds of visions. Some of them are about rather concrete goals. For instance, the chief of a village can try to convince the inhabitants that they should cooperate in digging a common well that everybody will benefit from or in building a defensive wall that will increase the security of everybody. The goal requires efforts by the members of the community, but it can still have a positive net benefit for all involved.

Other visions are more abstract and distant and their potential values are hard to assess. Many religions promise a heaven after death, if you just behave according to certain norms. Such a vision is a temptation to many, even though it is impossible to know whether it can be fulfilled. An eloquent leader can depict enticing goals and convince the supporters to make radical sacrifices, even though the visionary goals are extremely uncertain.

The theory outlined here is compatible with the ”mother tongue” hypothesis presented by Fitch (this volume). Situations where common goals exist occur more frequently among kin than among non-kin. Such situations also foster ”honest” communication. Therefore it is more likely that a system of symbolic references develops in a kin group than in a group of unrelated individuals. However, a limitation of his theory seems to be that it cannot explain why only humans have a symbolic language. There is nothing in the ”mother tongue” hypothesis as presented by Fitch (this volume), that precludes that apes, for example, could have the capacity to develop a symbolic communication system. According to the theory outlined in this article, the fact that humans, but apparently no other species, can represent future goals and the inner world of others, make us uniquely prepared for symbolic communication.

6. What Aspects of the Evolution of Communication Should be Explained First?

In a sense, all communication is a sign of failure. If everybody is behaving as they should, given the circumstances, then there is no need for communication. The obvious goes without saying. When communication first appears, it is the communicative *act* in itself and the context it occurs in that is most important, not the expressive form of the act (Winter 1998, Introduction). As a consequence, the pragmatic aspects of language are the most fundamental from an evolutionary point of view. When communicative acts (later speech acts) in due time become more varied and eventually conventionalized and detached from the immediate context, one can start analysing the different *meanings* of the acts. Then semantic considerations become salient. Finally, when linguistic communication becomes even more conventionalized and combinatorially richer, certain markers, alias *syntax*, are used to disambiguate the communicative content when the context is not sufficient. Thus syntax is required only for the subtlest aspects of communication – pragmatic and semantic features are more fundamental.

This view on the evolutionary order of different linguistic functions stands in sharp contrast to mainstream contemporary linguistics. For followers of the Chomskian school, syntax is the primary study object of linguistics; semantic features are added when grammar is not enough; and pragmatics is a wastebasket for what is left over (context, deixis, etc). However, I believe that when the goal is to develop a theory of the evolution of communication the converse order – pragmatics before semantics before syntax – is more appropriate. In other words, there is much to find out about the evolution of communication, before we can understand the evolution of semantics and syntax.

In support of the position that pragmatics is evolutionarily primary, I want to point out that most human cognitive functions had been ciseled out by evolution before the advent of language. I submit that language would not be possible without all these cognitive capacities, in particular having a theory of mind and being able to represent future goals. This position is not uncontested. Some researchers argue that human thinking cannot exist in its full sense without language (e.g. Dennett 1991). Thus the emergence of language is seen as a cause of certain forms of thinking, e.g. concept formation.

However, seeing language as a cause of human thinking is like seeing money as a cause of human economics (Tomasello 1999, p. 94). Humans have been trading goods as long as they have existed. But when a monetary system does emerge, it makes economic transactions more efficient. The same applies to language: hominids have been communicating since long before they had a language, but language makes the exchange of knowledge more efficient. The analogy carries further: When money is introduced in a society, a relatively stable system of *prices* emerges. Similarly, when linguistic communication develops individuals will come to share a relatively stable system of *meanings*, i.e. components in their inner worlds that communicators can exchange between each other. In this way, language fosters a *common structure* of the inner worlds of the inidividuals in a society.

This view on the regulatory role of language gains additional support from a different direction. In a variety of computer simulations and robotic experiments (e.g. Hurford (1999), Kirby (1999), Steels (1999, this volume), Kaplan (2000)), it has been shown that a stable communicative system can emerge as a result of iterated interactions between artificial agents, even though there is nobody who determines any ”rules” of the communication. A general finding of the experiments is that the more ”speakers” and ”hearers” are involved in communication about the same outer world, the stronger is the convergence of the reference of the ”words” that are used and the faster is the convergence attained. Still, different ”dialects” in the simulated community often emerge.

7. The Evolution of Referential Expressions

Cognitive semantics places the meanings of words in the heads of people. But a general problem for such a semantic theory is: if everybody has their own inner world, how can we then talk about a representation being *the* meaning of an expression? In other words, how can individual representations, cued or detached, become social? Therefore, the question in focus of this section will be: how can language help us *share* our inner worlds?

In cooperative communication about detached goals, a particularly important case of sharing inner worlds is to *jointly refer to objects that are not present on the scene of communication.* It can be an object that is distant, such as an animal to hunt or a tree containing honey, but it can also be a not yet existing object that is to be created by the cooperation, such as a communal water well. In contrast, indexical reference, such as pointing, is sufficient for identifying referents that are present in the environment. I will take this communicative problem as paradigmatic for an analysis of what is required for symbolic communication concerning cooperation about future goals.

In the computer simulations and robotic experiments performed by Steels and others, the typical communicative situation is a ”guessing game” (Steels, this volume) where the speaker, by uttering a word, tries to make the hearer identify a particular object in the environment. It should be noted that in such guessing games (as in Wittgenstein's language games) the participants are only concerned with finding the appropriate referent among those that are present on the scene. Communicating about *non-present referents* demands more advanced representational capacities among the communicators.

In this section I want to describe three stages of abstraction in the communication about referents and the establishing of the meanings of referential expressions (partly following Winter and Gärdenfors 1998 and Gärdenfors 2000, also see Olson 1970). At each stage I shall specify the assumptions concerning the sharing of inner worlds that it requires. The fitness variables driving the abstraction process could be the strain on memory as a cost, and efficiency of communication (identifying a referent) as benefit.

7.1 Names

The starting assumption is that each object that is perceived or communicated about is represented as a point in a conceptual space (as described in Gärdenfors 2000). Conceptual spaces consist of a number of ”quality dimensions” that represent various properties of the object, such as color, size, shape, texture, sound, etc. Conceptual spaces can be seen as providing the *framework* for the knowledge that is represented in the inner worlds of individuals. Different individuals may structure their spaces differently, so there may be no immediate way of comparing them.

Properties of the objects may be changing, which means that the points representing them move around in the conceptual space as is indicated in figure 1. Furthermore, objects come into existence and disappear which means that points come and go in the representing space.

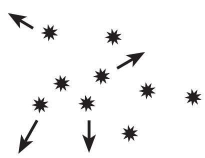

Figure 1. Points move around in the conceptual space.

Now suppose each individual in a communicative dyad has their own sets of representational points in their private conceptual space. How can we solve the paradigmatic communicative problem where the speaker wants to use symbolic language to make the hearer identify a particular object?

At the lowest level of abstraction, this communicative task is achieved by *names*. A name picks out a particular object represented as a point in the conceptual space of an individual. In figure 2, this identification is represented by encircling the representation of an object. If both participants associate the same name with the same external object, then the hearer can identify the object that the speaker intends. It should be noted that the naming mechanism puts no requirement on the alignment of the conceptual spaces of the communicating individuals, but only that their inner worlds contain an appropriate referent for the name.

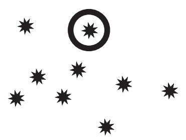

Figure 2. A name singles out a unique referent.

Even though this communicative mechanism in principle solves the task of identifying a common referent, it only works when both speakers are *acquainted with* the named object and have associated the same name with it. Furthermore, the mechanism is dependent on a stable context in the sense that entities exist in the presence of the speaker and the hearer long enough for a name to be established (by deixis or some similar pragmatic mechanism). Yet another drawback is that remembering many names is, as everybody knows, abhorrently costly for our memory.

In an evolutionary setting, there are two kinds of entities which remain relatively stable and identifiable within a community, namely, *people* and *places*. Thus one can speculate that the first stages of language contained names for people and places together with words denoting *relations* between such entities (Dunbar 1996, Worden 1996). Such a communicative system would be a protolanguage in the sense of Bickerton (1990).

7.2 Nouns

In the light of these assumptions, one should ask: how can objects that are not suitable for naming be identified? To answer this question, we must enter the second level of abstraction within the set of points in a conceptual space. This level builds on a fundamental fact about the world around us: *it is not random*. In other words, properties of objects tend to go together.

It is an interesting fact about the evolution of human thinking that, fortunately, our minds seem predisposed to detect such correlations of properties (Kornblith 1993, Holland et al 1995). A likely explanation of this capacity is that our perceptions of natural objects show correlations along several quality dimensions and, as a result of evolutionary pressures, we have developed a competence to detect these correlations.

In conceptual spaces, correlations show up as *clusters* of points. Such a cluster is marked by a circle in figure 3.

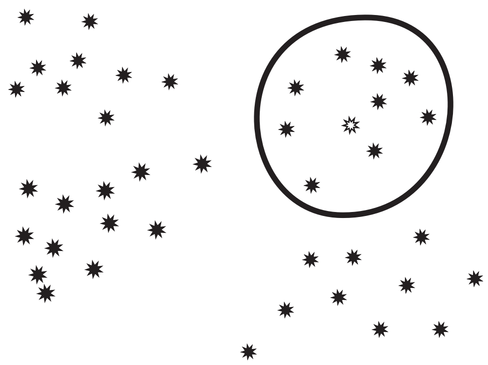

Figure 3. A noun corresponds to a cluster of correlated properties.

A paramount feature of clusters is that they, unlike points representing single objects, will remain *stable* even when objects change their properties somewhat or when new objects come into existence or old ones disappear. Thus, clusters are much more reliable as references of words than are points representing single objects. Furthermore, even if two individuals are not acquainted with the same objects represented within a cluster, their clusters may still be sufficiently similar to be matched. For this to happen, it is sufficient that we interact with the same kinds of objects and have shared socio-cultural practices. So if there is only one object from a given cluster salient in the cooperative context, it is sufficient that the communicators can identify the same cluster in their inner worlds for them to identify the object of collaboration. This level of abstraction thus puts some minimal constraints on the coordination of the conceptual spaces of the communicating individuals.

The prime linguistic tool for refering to a cluster is a *noun*. Rather than refering to the entire cluster, a noun refers to a point (representing a possible object) that functions as a stand-in for the cluster. This stand-in point, marked by a white star in figure 3, can be identified as the *prototype* of the cluster. This mechanism explains why nouns (noun phrases) have basically the same grammatical function as names.

By using a noun, the speaker indicates that she is talking about one of the elements in the cluster, by default a prototypical element, which is often sufficient for the hearer to identify the appropriate object in the context.[5]

However, a fundamental difference between objects and prototypes is that there are, in principle, an infinite number of possible objects (with different combinations of properties) while we typically work with a small number of clusters and their representing prototypes. Focussing on nouns results in a *discretization* of the space (compare Petitot 1989, p. 27).[6] Such a discretization is also necessary for a finite vocabulary.

The prototype need not represent any of the objects anybody has encountered. It is represented as a central point in the cluster associated with a noun, but no existing object need to have its representation there. Nevertheless, since different regions of the space are correlated with different properties in other domains, the possible object represented by the prototypical point will, by default, be assigned a number of properties. For example, a bird is normally small, sings, flies and builds nests in the trees. These properties form the *expectations* generated by the mentioning of a noun.

Among the objects represented in the conceptual space of an individual, there may be several layers of clusters, depending on how finely one wants to partition the space. However, there tends to be a privileged way of clustering the objects that will generate the *basic categories* in the sense of prototype theory (see for example Rosch 1978). This is the set of clusters that provides the most "economic" way of partitioning the world. What is "economic" depends, among other things, on the *practices* of the members of the community. Economy goes hand in hand with learnability: the basic categories are also those that are first learned by children.

7.3 Adjectives

Basic level nouns partition the conceptual space only in a rather coarse way. Using nouns presumes that the communicators have *represenations the same clusters*, which is a much less severe assumption than that they are acquainted with the same individuals. However, in some communicative contexts even this presumption delimits the communicative capacities. One example of such a context is when the speaker and hearer face a class of objects that all fall under the same noun and the speaker needs to identify one of the objects in the class, but has no name for it.

There are two solutions to this referential problem. The first is to introduce a finer level of granularity when identifying clusters. This strategy leads to the introduction of *subordinate* nouns (ostrich instead of bird, Volvo instead of car, etc). The drawback from the viewpoint of cognitive economy is, as in the case of names, that learning a large number of subordinate nouns demands a rich memory. However, if a finer categorization helps you solve new problems, the cost of remembering many nouns may be worth the benefits. (As a matter of fact, being an expert in an area involves having a large number of subordinate concepts, i.e. having a finely partitioned set of clusters.)

The second solution is to introduce a third level of abstraction. A fundamental strategy to distinguish points *within* a cluster that has been determined by correlated properties is to identify a feature that does *not* covary with other properties of the cluster. This is the basic mechanism for generating the *dimensions* of communication. For example, the color of an object often does not covary with other properties. In figure 4, the color dimensions are indicated (in one dimension only) by different shades of grey.

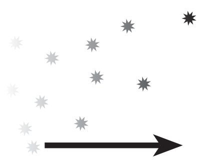

Figure 4. Adjectives single out dimensions.

Domains that are singled out by this process will be expressed by *adjectives* in natural language (see also Givón 1984). For example, to identify a particular car in a parking lot, one can say "the red car" (color domain) or "the big car" (size domain). The most useful adjectives are those that can be used with a large class of nouns, such as color or size words.

In principle, adjectives can be used to refer without a noun. For example, you may use an expression such as "the red one" to identify an object that is present in the communicative context (where the noun phrase "one" serves as a placeholder for a noun). However, in most cases, an adjective is used to give further information about a specific object. (Adjectives can hardly be used in isolation to identify an object that is not present.)

---

[5]Some further aspects of referential communication, in particular the relevance of *contrast classes*, are treated in Winter and Gärdenfors (1998). Barsalou and Prinz (1997, p. 297) emphasize the role of embodiment: "Variable embodiment ensures that different individuals can match their perceptual symbols optimally to their perceptions. […] Thus, if one individual represents color categories in a somewhat idosyncratic manner, his or her perceptual symbols will reflect this structure, such that they will be optimally tuned to match subsequent perceptions of color."

[6] This process is related to the phenomenon of categorical perception.

The combination of an adjective plus a noun allows you to identify a referent with less load on memory than subcategories of nouns. In elementary communication-economic terms: if you have a vocabulary with m nouns and n adjectives, you can use these m + n words to express m x n combinations. This multiplicativity of referential power does not apply to subcategories of nouns. Another aspect of communicative economy is that when you are faced with a situation where a noun covers several potential referents, you should select an adjective that picks out a *maximally informative* dimension within the cluster that represents the noun. Speakers are in general excellent at intuitively selecting the right dimension in a given communicative context. These considerations show that adjectives contribute substantially to the cognitive economy of communication. The cost is that the use of adjectives presupposes that communicators share dimensions. This presupposition demands a rather strict alignment of the conceptual spaces of the communicators, which is why adjectives involve a higher level of abstraction and coordination than names and nouns.

The thesis that adjectives are more abstract tools for communication than are names and nouns is supported by data from child language, as is witnessed by the following quotation from Smith (1989, p. 159):[7]

> [T]here is a *dimensionalization* of the knowledge system. […] Children's early word acquisitions suggest such a trend. Among the first words acquired by children are the names for basic categories – categories such as *dog* and *chair*, which seem well organized by overall similarities. Words that refer to superordinate categories (e.g., *animal*) are not well organized by overall similarity, and the words that refer to dimensional relations themselves (e.g., *red* or *tall*) appear to be understood relatively late […].

Social interactions will generate a need for representations where the dimensional structure is represented by a small number of values on each dimension. As a matter of fact, dimensional adjectives generally come in polarity pairs: heavy – light, tall – short, etc.

Freyd (1983) argues that knowledge about the world, by the fact that it is *shared* in a language community, imposes constraints on individual representations. She argues that the structural properties of individuals' knowledge domains have evolved because "they provide for the most efficient sharing of concepts," and proposes that a dimensional structure with a small number of values on each dimension will be especially "shareable."

This shareability process is continually ongoing: the interplay between individual and social structures is in eternal co-evolution. The effects are magnified when communication takes place between many individuals (cf. the simulations by Steels and others). Freyd hypothesizes that the mechanism will over time create a *grid* of fairly stabilized and discrete values on a few dimensions. When communicating about objects, the grid, with its corresponding combinations of adjectives, will generate a class of *communicable references.* Referential meanings outside this class cannot be easily shared in communication since they are not directly codable.

It should also be noted that representational availability of a domain normally precedes explicit *awareness* of the domain. In other words, even if a domain is exploited in linguistic communication, the communicators are often not able to *refer* to the domain itself. Such a capacity would presume an even higher level of abstraction than the three levels discussed in this section. In support of this position, it can be noted that children learn to use color words before they can engage in abstract talk of color in general. A related phenomenon from child language is that adjectives that denote contrasts within one domain are often used also for other domains. Thus, three- and four-year-olds confuse "high" with "tall," "big" with "bright," etc (Carey, 1985).

There is potentially an unlimited number of dimensions in conceptual spaces that are grounded in perception. This could be an insurmountable problem when coordinating the spaces of several individuals. However, even though the class of adjectives is open ended, linguistic space has a limited number of dimensions. Furthermore, cooperative communication highlights which dimensions are relevant (in a particular society). Which they are is to a large extent dependent on the practices of the society. Success in communicative tasks lead to a *stabilization* of the perceptual dimensions of the individuals and make them shared in a community. Following an earlier analogy, it can be said that, like money, language is a *social good.*

In this section I have modeled an abstraction process concerning communication about referents. The arguments suggest that common dimensional structures are likely to emerge as a consequence of the requirement that cooperation about future goals be highly dependent on shared knowledge.

This stance on symbolic communication leads to a chicken and egg problem: are conceptual spaces prerequisites for successful communication or are they emergent results of successful communication? The answer, it seems to me, is "both." As is argued in Gärdenfors (2000), the dimensions in conceptual spaces have several origins. This section has added yet another: *communication* is a catalyst for geometrically structured meanings. The analysis also

[7] Also see Smith and Sera (1992, p. 132).

indicates the *semantic* functions of different word classes (in contrast to traditional linguistic theory that defines word classes in terms of syntactic features).

8. Conclusion: Cooperation Begat Language

Recent literature on animal cognition has, to a large extent, focused on the deceptive capacities of different species (e.g. Whiten and Byrne 1988, Byrne 1995) often in terms of so called Machiavellian intelligence (Whiten and Byrne, 1997). This tendency has spilled over to the debate on the evolution of human cognition. However, a general conclusion to be drawn from this article is that, as regards the human species, the development of advanced forms of cooperation is more important when explaining the evolution of language.

Advanced cooperation demands access to detached representations, and the capacity to communicate about such representations. Therefore, the efficiency of communication about a detached goal will be a bottleneck in changing the strategic situation of the group. The core argument of this article is that without the aid of symbolic communication, we would not be able to share visions about the future. We need it in order to convince each other that a future goal is worth striving for.

The key question for cooperation on the basis of symbolic communication is thus: How do we communicate the detached representations of our inner worlds? In my opinion, the emergence of sharable conceptual spaces provides the first steps of an answer. I believe that the benefits of advanced cooperation are so extensive that they are the major evolutionary forces behind the emergence of symbolic language. In this sense, cooperation begets language.

The theory presented in this article also explains why only humans have language. Being able to cooperate about future goals requires detached representations of goals as well as a theory of mind. As far as we know, both these cognitive capacities are uniquely human.

Acknowledgement

I want to thank the participants of the Altenberg conference on the Evolution of Communication Systems as well as Ingar Brinck and David de Léon for their helpful comments.

References

Barsalou, L. W. (1999): ”Perceptual symbol systems,” *Behavioral and Brain Sciences 22*, 577-609.

Barsalou, L. W. and Prinz, J. J. (1997), “Mundane creativity in perceptual symbol systems,” in Ward, T. B., Smith, S. M. and Vaid, J., eds., *Creative Thought: An Investigation of Conceptual Structure and Processes*, American Psychological Association, Washington, DC, 267-307.

Bickerton, D. (1990), *Language and Species*, University of Chicago Press, Chicago.

Boysen, S. and Bernston, G. (1995), ”Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes),” *Journal of Experimental Psychology and Animal Behavior Processes 21*, 82-86.

Brinck, I. and Gärdenfors, P. (2001), ”Co-operation and communication in apes and humans,” *Lund University Cognitive Studies 88, Lund.*

Byrne, R. (1995), *The Thinking Ape: Evolutionary Origins of Intelligence*, Oxford University Press, Oxford.

Carey, S. (1985), *Conceptual Change in Childhood*, MIT Press, Cambridge, MA.

Craik, K. (1943), *The Nature of Explanation*, Cambridge University Press, Cambridge.

Deacon, T. W. (1997), *The Symbolic Species: The Co-evolution of Language and the Brain*, Norton, New York, NY.

Dennett, D. (1991), *Consciousness Explained*, Little, Brown and Company, Boston, MA.

Donald, M. (1991), *Origins of the Modern Mind*, Harvard University Press, Cambridge, MA.

Dunbar, R. (1996), *Grooming, Gossip and the Evolution of Language*, Harvard University Press, Cambridge, MA.

Ellen, P. and Thinus-Blanc, C., eds. (1987), *Cognitive Processes and Spatial Orientation in Animal and Man: Volume I Experimental Animal Psychology and Ethology,* Martinus Nijhoff Publishers, Dordrecht.

Fitch, W. T. (to appear), ”Evolving honest communication systems: Kin selection and ’mother tongues’.”

Freyd, J. (1983), “Shareability: the social psychology of epistemology,” *Cognitive Science 7*, 191-210.

Gärdenfors, P. (1993), “The emergence of meaning,” *Linguistics and Philosophy 16*, 285-309.

Gärdenfors, P. (1996), ”Cued and detached representations in animal cognition,” *Behavioural Processes 36*, 263-273.

Gärdenfors, P. (2000), *Conceptual Spaces: The Geometry of Thought*, MIT Press, Cambridge, MA.

Givón, T. (1984), *Syntax – a Functional-Typological Introduction, Vol. 1*, John Benjamins Amsterdam.

Grush, R. (1997), ”The architecture of representation,” *Philosophical Psychology 10*, 5-23.

Gulz, A. (1991), *The Planning of Action as a Cognitive and Biological Phenomenon*, Lund University Cognitive Studies 2, Lund.

Hockett, C. F. (1960), “The origin of speech,” *Scientific American 203(3)*, 88-96.

Holland, J. H., Holyoak, K. J., Nisbett, R. E. and Thagard, P. R. (1995), *Induction: Processes of Inference, Learning, and Discovery*, MIT Press, Cambridge, MA.

Hurford, J. (1999), ”The evolution of language and languages,” in Dunbar, R., Knight, C. and Power, C., eds., *The Evolution of Culture*, Edinburgh University Press, Edinburgh, 173-193.

Jeannerod, M. (1994), “The representing brain, neural correlates of motor intention and imagery,” *Behavioral and Brain Sciences*, 187-202.

Kaplan, F. (2000), *L'émergence d'un lexique dans une population d'agents autonomes*, Ph. D. thesis, Laboratoire d'Informatique de Paris 6, Paris.

Kirby, S. (1999), *Function, Selection and Innateness: The Emergence of Language Universals*, Oxford University Press, Oxford.

Kornblith, H. (1993), *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*, MIT Press, Cambridge, MA.

Lachman, R. and Lachman, J. L. (1982), “Memory representations in animals: Some metatheoretical issues,” *Behavioral and Brain Sciences 5*, 380-381.

Oakley, K. P. (1961): ”On man's use of fire, with comments on tool-making and hunting”, in Washburn, S. L., eds., *Social Life of Early Man*, Aldine Publishing Company, Chicago, ss. 176-193.

Olson, D. R. (1970), “Language and thought – aspects of a cognitive theory of semantics,” *Psychological Review*, **77**, 257–273.

Petitot, J. (1989), “Morphodynamics and the categorical perception of phonological units,” *Theoretical Linguistics 15*, 25-71.

Roitblat, H. L. (1982), “The meaning of representation in animal memory,” *Behavioral and Brain Sciences 5*, 353-372.

Rosch, E. (1978), “Prototype classification and logical classification: the two systems,” in Scholnik, E., ed., *New Trends in Cognitive Representation: Challenges to Piaget’s Theory*, Lawrence Erlbaum Associates, Hillsdale, NJ, 73–86.

Savage-Rumbaugh, E. S., Rumbaugh, D. M, and McDonald, K. (1985), “Language learning in two species of apes,” *Neuroscience and Biobehavioral Review 9*, 653-665.

Savage-Rumbaugh, E. S., Shanker, S. G. and Taylor, T. J. (1998), *Apes, Language and the Human Mind*, Oxford University Press, Oxford.

Sinha, C. (to appear), ”The epigenesis of symbolization.”

Sjölander, S. (1993), “Some cognitive breakthroughs in the evolution of cognition and consciousness, and their impact on the biology of language,” *Evolution and Cognition 3*, 1-10.

Smith, L. B. (1989), “From global similarities to kinds of similarities – the construction of dimensions in development,” in Vosniadou, S. and Ortony, A., eds. *Similarity and Analogical Reasoning*, Cambridge University Press, Cambridge.

Smith, L. B. and Sera, M. D. (1992), “A developmental analysis of the polar structure of dimensions,” *Cognitive Psychology 24*, 99-142.

Sneed, J. (1971), *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.

Steels, L. (1999) *The Talking Heads Experiment*, Laboratorium, Antwerpen.

Steels, L. (to appear), ”Social and cultural learning in the evolution of human communication.”

Tolman, E. C. (1948), “Cognitive maps in rats and men,” *Psychological Review 55*, 189-208.

Tomasello, M. (1999), *The Cultural Origins of Human Cognition*, Harvard Unversity Press, Cambridge, MA.

Vauclair, J. (1987), “A comparative approach to cognitive mapping,” in P. Ellen and C. Thinus-Blanc, eds., *Cognitive Processes and Spatial Orientation in Animal and Man: Volume I, Experimental Animal Psychology and Ethology,* Martinus Nijhoff Publishers, Dordrecht, 89-96.

Vauclair, J. (1990), “Primate cognition: From representation to language,” in S. T. Parker and K. R. Gibson, eds., *“Language” and Intelligence in Monkeys and Apes*, Cambridge University Press, Cambridge, 312-329.

Von Glasersfeld, E. (1977), “Linguistic communication: theory and definition,” in D. M. Rumbaugh, ed., *Language Learning by a Chimpanzee: The LANA Project*, Academic Press, New York, 55-71.

Whiten, A. and Byrne, R. W. (1988), “Tactical deception in primates,” *Behavioral and Brain Sciences 11*, 233-273.

Whiten, A. and Byrne, R. W., eds. (1997), *Machiavellian Intelligence II: Evaluations and Extensions*. Cambridge University Press, Cambridge.

Winter, S. (1998), *Expectations and Linguistic Meaning*, Lund University Cognitive Studies 71, Lund.

Winter, S. and Gärdenfors, P. (1998), “Evolving social constraints on individual conceptual representations,” *Lund University Cognitive Studies 69*, Lund.

Worden, R. P. (1996), “Primate social intelligence,” *Cognitive Science 20*, 579-616.