The Evolution of Sentential Structure

Peter Gärdenfors [†]
peter.gardenfors@lucs.lu.se

ABSTRACT

The aim of this article is to present an evolutionarily grounded explanation of why we speak in sentences. This question is seldomly addressed, neither in the Chomskian tradition nor in cognitive linguistics. I base my explanation on an analysis of different levels of communication. I identify four levels: praxis, instruction, coordination of common ground and coordination of meaning. The analysis will be focused on the evolutionary benefits of communicating about *events* as a way of coordinating actions. A cognitively grounded model of events will be outlined. My central thesis is that the communicative role of sentences is to express events.

Keywords: sentence structure, pragmatics, semantics, evolution of language, common ground, coordination of meaning, events, event construals, cooperation, indirect reciprocity.

1. Why do we speak in sentences?

In evolutionarily early forms of communication, the communicative *act* in itself and the context it occurs in were presumably more important than the expressive form of the act (Clark, 1992; Winter, 1998; Gärdenfors, 2010). As a consequence, the pragmatics of natural language is the most basic from an evolutionary point of view. When communicative acts become more varied and eventually conventionalized during hominin evolution and their contents become detached from the immediate context (Gärdenfors, 2000), one can start attending to the expanding meanings of the acts. Then semantics becomes salient. Finally, when linguistic communication becomes even more conventionalized and combinatorially richer, certain markers, a.k.a. syntax, are

used to disambiguate the contents when the context and the common ground of the interlocutors are not sufficient. According to this view, syntax is required only for the subtlest aspects of communication – pragmatic and semantic features are more fundamental.

This view of the evolutionary order of different linguistic functions stands in sharp contrast to much of 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.).

Clark (1996, p. 56) calls the Chomskian perspective the *production tradition* (focusing on the products of language) and the perspective that puts pragmatics in focus the *action tradition*. These two approaches to the evolution of language generate quite different research questions. It seems that never shall the twain meet.

There is, however, one linguistic unit that is central to both approaches: the *sentence*. In the Chomskian tradition it is taken for granted that the central goal of linguistic production is to generate sentences with a minimal structure of a noun phrase and a verb phrase. And the core linguistic data concern whether certain combinations of words are grammatical. Also in the pragmatic-semantic tradition, the sentence plays an important role. Furthermore, in analytic philosophy, sentences are central units, being the bearers of truth-values. In the tradition since Frege, a sentence expresses a proposition. But also in more cognitively oriented semantics, sentences are seen as natural units (e.g., Langacker, 1987, 2008; Talmy, 2001; Croft, 1991; Goldberg, 1995, Levin & Rapaport Hovav, 2005).

Since the sentence is so central to both research traditions, it is surprising that nobody asks *why* this unit exists. The question becomes more pressing when one compares with what is communicated by language-trained apes and other animals. Their communication never, or just by chance, exhibits sentential structures. Kanzi and his colleagues only bundle signs together without concern for whether the collection forms a sentential structure (Savage-Rumbaugh & Lewin 1994). For example, Greenfield and Savage-Rumbaugh (1990) succeed in finding a few rough patterns. Kanzi more often places the verb before the object – "hide nut" instead of "nut hide" in accordance with a language such as English. When he combines two verbs, for example "tickle bite" (which does not occur in English), he wants to do the actions in the order he mentions them. However, Kanzi's grammatical patterns

are far from consistent, and they tally poorly with the grammatical competence that Chomsky's theory of language postulates.

My aim in this article is to present an evolutionarily grounded explanation of why we speak in sentences. I will follow what Clark (1996) calls the action tradition and base my explanation on an analysis of different levels of communication. The analysis will be focused on the evolutionary benefits of communicating about *events*. My central thesis is that the communicative role of sentences is to express events.

2. Levels of communication

The obvious goes without saying. If all partners in a cooperating group perform their tasks as expected by the others, there is no need for communication. Cooperation takes place on the level of praxis. It is only when an instruction, a correction or a coordination is needed that communication plays a pragmatic role. The basic level of communication is therefore for solving problems of coordinating actions. For example, if A is carrying a heavy box, but his path is blocked by a closed door, and B does not realize the situation, A typically *instructs* or *requests* B to open the door.

However, there are situations when the communicators misunderstand each other, because of badly formulated instructions or because they have different mental models of the world. For example, if *A* commands *B* to open the door and *B* sees several possible doors, he replies "Which door". Then *A* and *B* move to the level of coordination of common ground (Clark, 1992), that is, to agree on which door *A* wants *B* to open. When this is accomplished, they return to the level of instruction and *B* can perform the desired action. Coordination of common ground can also be done as a preparation for future collaboration. As I argue below, this aspect is central from an evolutionary perspective. Everyday talk about what other people do or have done also belongs to this level.

There is a third, more severe form of misunderstanding that occurs because the addressee does not understand an expression used by the speaker, or does not understand it in the same way. For example if you say "I'll talk to the chair" and you mean the chairperson, while *chair* for me just means physical objects, I will not understand your intention. On this final level – the level of semantic

coordination – the communicators must negotiate their use of expressions until they find a sufficient agreement.

For these reasons, following Winter (1998), I want to distinguish three levels of communication, in addition to a ground level of human interaction:

Level 0: *Praxis*. On this level people interact with each other without using intentional communication.

Level 1: *Instruction*. On this level coordination of action is achieved by instruction.[1]

Level 2: *Coordination of common ground*. On this level people inform each other in order to reach a richer or better coordination. It can also be achieved via questions.

Level 3: *Coordination of meanings*. On this highest level, people negotiate the meanings of words (labels) and other communicative elements.

The four levels are used in a hierarchical manner. When one level does not function properly, a break in the communication is signalled and it moves to the next higher level. When the problem is solved the communicators signal an acknowledgment and return to the level below. One example of this is the coordination of which door to open, that was presented above. In this case the communication goes from the level of instruction to coordination of common ground and then back again. Another example, going from the second to the third level, is that if *A* is telling *B* something and uses a word that *B* does not understand, *B* can signal this and they move to the level of coordination of meaning. When this is accomplished, they return to the level of coordination of common ground. Considering the evolution of communication, it is also reasonable that the four levels emerged in the same order as well.

Clark's (1992, 1996) work on *common ground* and *uptakes* can be seen as analyses of some central forms of coordination on level 2. First, the utterances in a conversation introduce new referents or new information about referents. This, together with the participants' expectations about the other's previous knowledge, forms the common ground that the subsequent conversation can take for granted. Second, a participant often introduces a proposal for a joint project in the conversation. This proposal can be taken up by the interlocutor

---

[1] This level corresponds to the language games introduced by Wittgenstein (1953).

(or it can be rejected). The proposal and the uptake then lead to a coordination of the continuing communication.

As an illustration of coordination of knowledge about the facts of the world, consider how *definite reference* is achieved (Clark 1992, p. 107). An example is a simple communicative act such as explaining to a tourist where to find a restaurant she is looking for: it involves a complex series of further requests and information extensions, as well as corrections, nods, and interjections. Creating such a reference is a coordination problem that is rarely reduced to uttering the right word at the right time. What is required instead is a process of mutual adjustment between speaker and addressee converging on a mutual acceptance that the addressee has understood the speaker's utterance. The process is highly iterative, involving a series of reciprocal reactions and conversational moves usually concluded by assent signals. Conversational adjustments toward mutual agreement typically resort to both the discrete resources of spoken language and the continuous resources of gesture, intonation, and other bodily signals.

3. The evolutionary roles of coordination of common ground

During the evolution that lead to *Homo sapiens*, our hominin ancestors developed new forms of cooperation that made it possible to organize their societies in new ways. It is generally agreed that hominins evolved in open landscapes that favoured a long-ranging life style (Preuschoft & Witte, 1991; Hilton & Meldrum, 2004). As a part of their adaptation they changed their diet from predominantly vegetarian to more protein and fat based. The first culture along the Homo lineage is associated with the finds at Oldowan (Isaac, 1982). The Oldowan lifestyle was in a way signified by an extension in time and space. For example, there were long delays between the acquisition and the use of the tool, as well as considerable geographical distances between the sources of tool raw material sources and killing sites.

3.1 Referring to absent objects

In this type of environment, it became increasingly important to jointly refer to objects that are not present on the scene.[2] If the common goal is present in the actual environment, for example food to be eaten or an antagonist to be fought, the collaborators need not communicate before acting. If, on the other hand, the goal is distant in time or space, then a common representation of it must be obtained before cooperative action can be taken. For example, building a shared dwelling requires coordinated planning of how to obtain the building material and advanced collaboration in the construction. The possibility of achieving joint attention to absent entities opens up for new forms of cooperation. This introduces selective pressures towards a communicative system that makes it possible for members of a group to share mental representations of non-present entities (Gärdenfors & Osvath, 2010; Gärdenfors, Brinck & Osvath, 2012).

Symbolic communication is based on the use of representations as stand-ins for entities, present or just imagined. This form of communication is "displaced" (Hockett, 1960) or "detached" (Gärdenfors, 2003), since it typically refers to non-present entities or events.[3] Use of such representations replaces the use of environmental cues in communication. If somebody has an idea about a goal she wishes to attain, she can use language to communicate her thoughts. In this way, language makes it possible for us to coordinate common grounds.

A wide range of communicative tasks can be performed by single words or a combination of a few words (or iconic signs). There are two main communicative situations, however, both unique to humans, where sentential structures play a crucial role: (i) cooperation for future goals and (ii) narratives, in particular gossip.

3.2 Communication for future goals

Planning for future collaboration, essentially a task of coordinating goals, requires several forms of coordination of commons ground: coordination in

---

[2] As regards cooperation among animals, there is no evidence that such communication is used.
[3] Iconic communication can also be used for displaced communication, but, unlike arbitrary symbols, the icons then exhibit some similarity with what they stand for (Gärdenfors, 2003).

space (often outside the present visual field), joint reference to absent objects, coordination of goals and coordination of actions. Such planning depends on forming joint intentions, an advanced form of intersubjectivity presumably unique to humans (Gärdenfors, 2003, 2008; Tomasello et al., 2005). A *joint plan* can be described as a combination of forming a joint intention and coordinating actions.

In previous work (Brinck & Gärdenfors, 2003; Gärdenfors, 2003, 2004; Osvath & Gärdenfors, 2005, Gärdenfors & Osvath, 2010), it has been argued that symbolic language makes it possible to efficiently cooperate about future goals. Along the same lines, Tylén et al., (2010, p. 6) write:

> Analogous to the way that manual tool use has been shown to enlarge the peripersonal space by extending the bodily action potential of arm and hand in space ..., linguistic symbols liberate human interactions from the temporal and spatial immediacy of face-to-face and bodily coordination and thus radically expand the *interaction space.*

I submit that the evolution of symbolic language generated evolutionary advantages for the individuals of a society built around cooperation toward future goals.

The transition from an animal signalling system to a symbolic language was, most likely, not made in one step. Bickerton (1990) and other researchers (e.g., Dessalles, 2007) propose that there was a stage in the evolution of language when a protolanguage, containing only the semantic components of language, was used. According to Bickerton, *Homo erectus* mastered a protolanguage and it is not until *Homo sapiens* that one finds a language with a grammatical structure. It is possible that the coordination of common ground required for forming a common plan for future actions can be achieved in a communication system that lacks syntax, that is, in a proto-language. Nevertheless, some sentential structure is necessary since a joint plan involves a series of coordinated future events. As I argue below, describing the planned events requires communication that refers to actions as well as the agent and patient of the action.

3.3 The evolutionary role of gossip and other forms of narratives

In social species, individuals often face a decision whether to cooperate or not. In the analyses of prisoners' dilemmas and similar games in standard game theory, it is taken for granted who the potential collaborators are. In practice, however, the most important question is: How do you know *who* to cooperate with? Here I agree with Dessalles (2007, p. 360): «Some of our ancestors who belonged to the first species of Homo, say, began to form sizeable coalitions. In such a 'political' context, finding good allies becomes essential». This type of information is an important special case of coordinating common ground.

Reciprocal altruism ("you scratch my back and I'll scratch yours"), is found in several animal species. *Indirect reciprocity* is a more extreme form of altruism: "I help you and somebody else will help me." The conditions for this to evolve as an evolutionary stable strategy have been modelled (e.g. Leimar & Hammerstein, 2001; Nowak & Sigmund, 2005). The key concept in Nowak and Sigmund's (2005) evolutionary model of indirect reciprocity is that of the *reputation* of an individual. An individual i's reputation is built up by members of the society observing i's behaviour towards third parties and then spreading this information to other members of the society. To wit, gossip becomes a way of achieving societal consensus about reputation (Dunbar, 1996; Slingerland et al., 2009). In this way the reputation for i being a 'selfless' helper can be known by more or less all the members of the group. The level of i's reputation can then be used by any individual when deciding whether or not to assist i in a situation of need. Reputation is not, of course, something immediately visible to others in the way of such status markers as a raised tail among wolves. Instead each individual must keep a private account of the reputation of all others with whom she interacts. Semmann et al. (2005) provide a nice experimental demonstration that building a reputation through cooperation is valuable for future social interactions, not only within but also beyond one's social group. Tirole (1996) argues that not only individual reputation, but also collective reputation plays an important role in societies: «Countries, ethnic, racial or religious groups are known to be hard-working, honest, corrupt, hospitable or belligerent» (Tirole 1996, p. 1).

In general, the communication required for functioning forms of indirect reciprocity concerns different aspects of who you can trust. The information is often conveyed in the absence of the individual discussed – and it can hence be characterized as *gossip*. Gossip normally contains expressions of the form "X

did A to Y," which involves identifying thematic roles such as agent, action and patient. Thus gossip plays a central role in the evolution of language according to the theory presented here, but it does not function as a replacement for grooming as Dunbar (1996) suggests.

3.4 Sentences are needed for the coordination of common ground

The considerations of this section provide some evolutionary reasons for why coordination of common ground is important for the forms of cooperation that seem more or less unique to humans. I have presented two forms of communication for cooperation where sentences are required: coordination of future goals and gossip that help you decide who to cooperate with. Pragmatically, they serve to coordinate the common ground of the interlocutors.

The important thing to note here is that describing planned actions as well as information about who did what to whom are special cases of describing *events*. I conjecture that the capacity to communicate about events is a watershed that distinguishes the communication of language-trained apes from that of humans. Both types of communication can be seen a special cases of *narratives*. In the following section, I will outline a cognitive theory of events that will support this position.

4. A cognitive model of events

Why then are events so central in human cognition? One central feature of events is that they are bearers of *causal relations*. An event typically contains information about an *agent* that performs an *action* related to a *patient* that leads to a *result*. Based on these components Gärdenfors and Warglien (2012) and Warglien et al. (2012) present a model of events and event categories in terms of conceptual spaces. I will here briefly outline this model.

A prototypical event is one in which the action of an agent generates a *force vector* that affects a patient causing changes in the state of the patient. The change of the properties of the patient can be described in terms of a *result vector*. As a simple example, consider the event of a person pushing a table. In this example, the force vector of the pushing is generated by an agent. The result vector is a change in the location of the patient – the table (and, perhaps,

a change in some other of its properties, e.g. it is getting warm and dusty). The result depends on the properties of the patient along with other aspects of the surrounding world: in the depicted event, e.g. frictions act as a counterforce to the force vector generated by the agent.

More technically, Warglien et al. (2012) and Gärdenfors (2014) formulate the following central criterion for event representations:

> *The two-vector condition:* A representation of an event contains at least two vectors and at least one object – a *force* vector that represents the cause of the change and a *result* vector representing a change in properties of the object.

The structure of the event is determined by the mapping from force vector to result vector. We will call the central object of an event the *patient*. Even though prototypical event representations contain an agent, there are events without agents, for example, in events of falling, drowning, dying, growing and raining. In the limiting case when nothing happens, that is, when the result vector is the identity vector, the event is a *state*. However, identity result vectors can also be maintained by balancing forces and counterforces: for example, when a prop prevents a wall from falling. Given this representation of events, causation can then be modelled by identifying the cause with the force vector and the effect with the result vector.

The two-vector model can be seen as a form of basic schema that can be elaborated by specifying further components. To the minimal representation of an event required by the two-vector condition, a number of other entities (what linguists call thematic roles) can be added: agent, instrument, recipient, benefactive, etc.

The proposed model allows one to represent events at different levels of generality. There are subcategories of events, just as for object categories. For example, *pushing a door open* is that subcategory of *pushing a door*, where the force vector exceeds the counterforce of the door. *Pushing a door but failing to open* it is another subcategory, where the counterforce annihilates the force vector.

A limiting case of our event model, expressed linguistically by intransitive constructions such as "Victoria is walking" and "Oscar is jumping," is when the patient is identical to the agent. In these cases, the agent exerts a force on

itself: in other words, the agent modifies its own location in space or its properties.

5. Sentences express events

The traditional account within analytic philosophy is that a (declarative) sentence expresses a *proposition*. Propositions are taken to be either true or false, that is, to have truth-values. On many accounts, a proposition is identified with a set of *possible worlds*. From the point of view of semantics, as I explain in Gärdenfors (2000, Section 3.3), this is putting the cart before the horse, since possible worlds are cognitively inaccessible entities. Furthermore, most of the examples of sentences in the philosophical discussion involve states – the classical example is "Snow is white." In contrast to most philosophical theories, I do not assume that there is a semantic mapping between sentences and propositions. The reason is that the meaning of a sentence is to a large extent dependent on its *context*. For example, an ironic communicative act may drastically change the standard meaning of a sentence.

I will here not attempt to account for the drawbacks of the semantics of sentences within analytic philosophy. Instead I want to be more constructive and present the bare outlines of how sentences can be analysed on the basis of events modelled in conceptual spaces. If one takes a cognitive-communicative point of view, as presented in section 2, it is not so obvious why we express ourselves in sentences. Frege's answer that the meanings of sentences are *thoughts* is simply not sufficient, since nobody knows how a thought is identified (independently of language). So what do sentences refer to?

My basic idea is that *sentences refer to events*. Furthermore, the focus should be on *utterances* rather than on sentences. Utterances are parts of a communicative context that contributes to the meaning, while in the philosophical (and much of the linguistic) discussion, sentences are often analysed as having a meaning that is independent of the context. However, I shall here use the word sentence to take in the role of utterance as well.

Any description of an event is based on a *construal*. The *attention* of the speaker is a selection mechanism for a cognitive event representation. There are, however, other aspects of how a construal is formed (see Croft & Wood 2000, Ch. 3; Langacker, 2008, Ch. 3, for a survey). One aspect is *perspective*: For example, if I and you are located on two sides of a house, I can

say that you are behind the house, if I put myself in the centre, or I can say that you are in front of the house, if I put the house and the direction of its main side in focus. Another aspect is *categorization*: A construal must select a level of generality to describe an object, for example, *terrier*, *dog*, *mammal,* or *animal.* Yet another aspect is the relation to the *common ground* in the communicative situation. For example, when selecting whether to use a pronoun, noun, or name to refer to an individual, the speaker must consider whether the individual or the name is part of the common ground. Speakers have conversational goals in producing construals. Consequently, the construals are contextual, depending on what the conversation partner already knows or believes or will find most interesting.

What minimal elements must a construal of an event contain? A generic way of describing an event is that "something happens to something." According to the model of events outlined above, the something it happens to is the *patient* (sometimes identical with the agent). Furthermore "happen" is a placeholder for either the force or the result vector. This leads to the following thesis:

Thesis about construals: A construal of an event contains as least one vector (force or result) and one object.

On the basis of the notion of a construal of an event, I can now formulate a fundamental connection between the semantics of sentences (utterances) and events:

Thesis about sentences: A (declarative) sentence typically expresses a construal of an event.

From a communicative perspective, one can ask why sentences have such a fundamental status, in comparison to other compositions of words. I do not believe there is a unique answer to the question, but I will base my analysis on the levels of communication discussed in section 2. On level 1, instruction, sentences are often not required. When sitting at a dinner table "Salt!" may function as request, albeit not a polite one (polite requests are often concealed as questions). Or if standing in front of a door "Open!" may be an efficient speech act, since the addressee (the agent) and the object are contextually given. When it comes to level 2, coordination of common ground, the situation is different. Here the communication typically concerns agents, patients,

actions, and results that are not present in the context of the utterance. Then the thesis about sentences implies that at least an agent or a patient (expressed by a noun phrase) and a force vector or a result vector (parts of a construal of an event) are elements in what is expressed. Thus the two main components noun phrase and verb phrase have to be present in a linguistic description of an event. This generates a *semantic* explanation of why a combination of a noun phrase and a verb phrase is so fundamental.

The upshot is that on the coordination level of communication, sentences are indeed central units. On level 3, when coordinating their meanings of words, the partners may rely on *definitional generics* as a special tool (Lawler, 1973). For example, "Whales are mammals" and "A wrench is a tool for fastening nuts" are used to express elements of the meaning of *whale* and *wrench*. Grammatically, generics are sentences, but semantically they are atypical since they describe generic information about concepts rather than about events.

For this reason, sentences are natural units of a semantic theory, albeit not as central as philosophers and some linguists want it. In brief, the model of events and the thesis about construals explain the necessity of the central components of a sentence. They thus provide a motivation for a sentence being a *cognitive unit* of communication.

An analogy to a construal is perhaps visual perception leading to judgments of the form "category X is at location Y." Searle (1983, p. 40) writes: «The content of the visual perception ... is always equivalent to a whole proposition. Visual experience is never simply *of* an object but rather it must always be *that* such and such is the case». He says that his experience of a station wagon must also be an experience of, for example, a station wagon in front of me. Given the thesis about construals, however, the analogy does not capture all kinds of event construals, in particular not the dynamic aspects. Talmy (1988, p. 61) summarizes the position succinctly:

> All of the interrelated factors in any force-dynamic pattern are necessarily copresent wherever that pattern is involved. But a sentence expressing that pattern can pick out different subsets of the factors for explicit reference–leaving the remainder unmentioned–and to these factors it can assign different syntactic roles within alternative construc-tions.

For example, the sentences "Victoria hits Oscar" and "Oscar is hit by Victoria" describe the same event with the aid of two different construals, where Victoria and Oscar, respectively, are put in focus.[4]

Consequently, no simple mapping exists between the role taken in an event and the designation of subject, object or oblique. A sentence expresses a construal representing a particular focus on an event. In English (and many other languages), the most focused role is designated subject and the secondary focus is designated object. Givón (2001) calls these *primary* and *secondary topics*. He writes that topicality «is fundamentally a cognitive dimension, having to do with the focus on one or two important event-or-state participants during the processing of multi-participant clauses» (Givón, 2001, p. 198). As Croft (2012, pp. 252–253) notes, this phenomenon creates problems for all argument realization rules that are based on thematic roles.[5] In agreement with Givón (2001, p. 198), I see topicality not as directly part of event representation, but rather as a central element of the construal process. This setup avoids the problems that arise when event representation and construal are conflated.[6]

Conclusion: From event thinking to sentence structure

This article has been written from what Clark (1996) calls the action tradition. In support of the position that pragmatics is evolutionarily primary, it is clear that most human cognitive functions had been chiselled out by evolution before the advent of language. Language would not be possible without these cognitive capacities, in particular having rich intersubjectivity, having a memory system that includes episodic memory, and being able to represent future goals (see Gärdenfors, 2003, 2007; Gärdenfors et al., 2012).

In summary, my thesis is that we communicate using sentential structures because human cooperation benefitted evolutionarily (and still benefits) from

---

[4] Croft (2012, p. 256) describes the passive voice as a *deprofiling* of the causal chain from the agent to the patient. This can be expressed in my terminology by saying that the patient is made the focus (or topic) of the event.

[5] Also Jackendoff (1987, p. 380) accords: «*Subject* is a syntactic relation, not a conceptual one, and syntactic subjects can hold a variety of [thematic] roles».

[6] Mapping subject to primary topic and object to secondary is not the only way of mapping the elements of an event onto language: In *ergative* languages, the agent is mapped onto the subject and the patient to object. Thus, in such a language, the sentence "Oscar is hit" is expressed as object plus verb.

communication about events. I have argued that the basic semantic referent of a sentence is an event (or a state as a special case). Other animals do not communicate about events since they neither have the mental capacities to cooperate about future goals, nor to cooperate via indirect reciprocity, and therefore they have no need for such communication.

My hypothesis that humans have more advanced cognitive representations of events than other species tallies well with the hypothesis that human are, more or less, the only species that has *episodic memory* that allows us to remember individual events and the order in which they have occurred (Tulving, 1985). This is the memory we use when we think of previous episodes and experiences we have encountered, or when we elicit from our memory events we have learned from conversations with others. Without episodic memory we cannot relate or recount anything. Tulving claims that only human beings have episodic memory, but this thesis has been contested by researchers in animal cognition (Osvath, 2010).

Despite all efforts, the apes' linguistic communication is very limited. In the best cases, they reach the level of a two-year-old human child. They seldom create combinations of more than two words. Most of what they communicate is about something they want. A typical example is when the chimpanzee Nim Chimpsky signs "Nim milk, give milk, Laura give Nim milk, more milk." The apes never *tell* anything, but already two-and-a-half-year-old human infant can tell rudimentary stories. Narration presumes event representations. Therefore, a speculative explanation of why language-trained apes do not tell stories is that they do not mentally represent event or do not represent them in a way that corresponds to the structure of sentences.

The narrative ability is central for human communication. But there seem to be no examples showing that Kanzi masters narration. We are still waiting for him to tell a story by the campfire. So, even if Kanzi understands many spoken utterances, they are coupled to a limited repertoire of communicative functions. They mainly consist of his obeying requests and expressing wishes. He remains on level 1 of communication and never engages in coordinating a common ground (level 2), let alone explains the meaning of words (level 3).

ACKNOWLEDGEMENTS

REFERENCES

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

Brinck, I. & Gärdenfors, P. (2003). Cooperation and communication in apes and humans. *Mind and Language, 18*, 484-501.

Clark, H. (1992). *Arenas of Language Use*. Chicago: University of Chicago Press.

Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Croft, W. (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

Croft, W. (2012). *Verbs: Aspect and Argument Structure*. Oxford: Oxford University Press.

Croft, W. & Wood, C. (2000). Construal operations in linguistics and artificial intelligence. In L. Albertazzi (Ed.), *Meaning and Cognition: A Multidisciplinary Approach*. Amsterdam: John Benjamins, 51-78.

Dessalles, J.-L. (2007). *Why We Talk*. Oxford: Oxford University Press.

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

Gärdenfors, P. (2003). *How Homo Became Sapiens: On the Evolution of Thinking*. Oxford: Oxford University Press.

Gärdenfors, P. (2008). Evolutionary and developmental aspect of intersubjectivity. In H. Liljenström, and P. Århem (Eds.), *Consciousness Transitions – Phylogenetic, Ontogenetic and Physiological Aspects*. Amsterdam: Elsevier, 281–385.

Gärdenfors, P. (2010). Evolution and semantics. In P. C. Hogan (Ed.), *Cambridge*

Encyclopedia of the Language Sciences. Cambridge: Cambridge University Press, 748-750.

Gärdenfors, P. (2014). *Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, MA: MIT Press

Gärdenfors, P., Brinck, I. & Osvath, M. (2012). Coevolution of cooperation, cognition and communication. In F. Stjernfelt, T. Deacon & T. Schilhab (Eds.), *New Perspectives of the Symbolic Species*. Berlin: Springer, 193-222.

Gärdenfors, P. & Osvath, M. (2010). Prospection as a cognitive precursor to symbolic communication. In R. Larson, V. Déprez, & H. Yamakido (Eds.), *Evolution of Language: Biolinguistic Approaches*. Cambridge: Cambridge University Press, 103-114.

Gärdenfors, P. & Warglien, M. (2012). Using conceptual spaces to model actions and events. *Journal of Semantics*, *29*, 487-519.

Givón, T. (2001). *Syntax, Vol I*. Philadelphia, PA: John Benjamins.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press

Greenfield, P. M. & Savage-Rumbaugh, E. S. (1990). Grammatical combination in Pan paniscus. Processes af learning and invention. In S. T Parker & K. H. Gibson (Eds.) *Language" and intelligence in monkeys and apes: Comparative developmental perspectives*, Cambridge: Cambridge University Press, 540-578.

Hilton C. E. & Meldrum D. J. (2004). Striders, runners and transporters. In D. J. Meldrum, & C. E. Hilton (Eds.), *From Biped to Strider: The Emergence of Modern Human Walking, Running and Resource Transport*. New York: Kluwer Academic, 1-8.

Isaac, G. (1982). The earliest archaeological traces. In *Cambridge History of Africa, Vol. 1: From the Earliest Times to c. 500 BC*. Cambridge: Cambridge University Press.

Jackendoff, R. (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar, Vol. 1*. Stanford, CA: Stanford University Press,

Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

Lawler, J. M. (1973). Studies in English generics. *University of Michigan Papers in Linguistics.*

Leimar, O. & Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocation, *Proceedings of the Royal Society London B, 268,* 745–53

Levin, B. & Rappaport Hovav, M. (2005). *Argument Realization.* Cambridge: Cambridge University Press.

Nowak, M. A. & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature, 437,* 1291–8

Osvath, M. (2010). *Planning Primates: A search for Episodic Foresight.* Lund: Lund University Cognitive Studies 148.

Osvath, M. & Gärdenfors, P. (2005). Oldowan culture and the evolution of anticipatory cognition. Lund: *Lund University Cognitive Studies 122.*

Preuschoft H. & Witte H. (1991). Biomechanical reasons for the evolution of hominid body shape. In Coppens Y. & Senut B. (Eds.) *Origine de la Bipedie Chez les Hominides.* Paris: Cahiers de Paleoanthropologie Editions du CNRS, 55-79.

Savage-Rumbaugh, E. S. & Lewin, R. (1994). *Kanzi: The Ape at the Brink of the Human Mind.* New York, NY: Wiley.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press.

Semmann, D., Krambeck, H.-J. & Milinski, M. (2005). Reputation is valuable within and outside one's own social group. *Behavioral Ecology and Sociobiology 57,* 611–616.

Slingerland, I., Mulder, M., van der Vaart, E. & Verbrugge, R. (2009). A multi-agent systems approach to gossip and the evolution of language. In Taatgen, N.A. & van Rijn, H. (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society.* Austin: Cognitive Science Society, 1609–1614.

Talmy, L (2000). *Toward a Cognitive Semantics, Vol 1: Concept Structuring Systems.* Cambridge, MA: MIT Press.

Tirole, J. (1996). A theory of collective reputations (with application to the persistence of corruption and to firm quality), *The Review of Economic Studies, 64,* 1-22.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28, 675-735.

Tulving, E. (1985). How many memory systems are there? *American Psychologist 40*, 385-398.

Warglien, M, Gärdenfors, P. & Westera, M (2012). Event structure, conceptual spaces and the semantics of verbs. *Theoretical Linguistics*, 38, 159-193.

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

Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.