2 Sketch of an evolutionary grammar based on comparative biolinguistics

Wolfgang Wildgen

LEARNING FROM ANIMALS? CAN LINGUISTICS CONTRIBUTE TO THE QUESTION?

In any interdisciplinary endeavour that aims to link the comparative ethology of animals with linguistics, a crucial question is which theory of language is to serve as a starting point. The challenge lies in adequately specifying the grammar of human languages in an evolutionary perspective. The contribution of linguistics to such interdisciplinary research has often been inadequate for two reasons. First of all, current linguistic theories are usually based on a longstanding tradition of normative grammar and an analysis of written language—hence the relevance of grammaticality and competence in Chomsky's models. In everyday speech language use is more variable and more context-dependent and changes-in-progress are pervasive. If we compare humans to animals, the dominant informal behaviour of humans should be the starting point for comparisons and not highly formalized behaviours regulated by institutions like schools, academies, etc. Second, linguists have a historical bias towards logical (analytic) descriptions and lack dynamic or self-organizing models. Therefore classificatory devices and hierarchical knowledge trees like phrase structures are emphasized, while the underlying forces, goals, benefits, trends, and changes are neglected. As a consequence, the intrinsic relation of language to holistic action patterns or to multichannel cognition (visual imagination, musical structure) is misrepresented in the standard models. Evolutionary biologists should turn instead to cognitive linguistics (semantics) and to pragmatic and dynamic linguistics (cf. Wildgen, 1994). In an interdisciplinary cooperation between biologists, psychologists, and linguists, one must assume that new models will be necessary that are not just versions of current types of grammars. In the following sections I will sketch the features of a model suitable for cooperative research in evolutionary biology and linguistics.

A PRAGMATIST POINT OF VIEW ON THE EVOLUTION OF LANGUAGE

From its inception pragmatics has had a strong link to anthropology and evolutionary theory with its central concept of adaptation. But one must distinguish the adaptive processes found in animals, which shape instincts, from the macroprocesses of linguistic adaptation. When Whorf describes the world-view of the Hopi, he can point to a divergence in human cultural adaptation, which resulted from a separation of migration routes giving rise to the Amerindian and the European populations some 40 to 50,000 years ago. On this time scale the relevant adaptations are cultural ones, whereas changes between the *Homo erectus* (*ergaster*) and *Homo sapiens* are based on biological adaptations.

Although a gradual transition from communication in animals to language in humans is evident to the pragmatist, as it was to Darwin, several approaches in theoretical linguistics take the opposite position. Fitch et al. (2005) focus the question of language evolution on an "abstract core of computational operations, central to language and probably unique to humans" (p. 180) and they try to neglect the "myriad component mechanisms" (p. 181), which are not unique to humans. What remains is called FLN: *faculty of language in the narrow sense*. As they remark, FLN "could possibly be empty if empirical findings showed that none of the mechanisms involved are uniquely human or unique to language". For Fitch et al. the best and currently the only candidate for a component of FLN is the "computational capacities of recursion" (p. 204). All explanations that argue for an adaptive origin of this core feature are dismissed; the authors argue that only synchronic adaptation is acceptable as empirical evidence. In consequence, neo-Darwinian evolutionary explanations of human language are considered unscientific, because they are not testable here and now.

As is often the case in recent debates in linguistics, the opposing view also has proponents. Jackendoff and Pinker (2005) argue that recursion is found in the human visual system, too, and thus it is not specific to language, i.e., not part of FLN, which consequently would be empty. Moreover, they reject the "narrow syntax" view of language in the Chomskyan paradigm as scientifically unproductive. Moving beyond the controversies within Chomskyan linguistics, I will adopt a radically different approach, which highlights the *pragmatic* origin of our capacity for language, not the origin of presumably universal formal features, and draw on relevant data from paleontology, anthropology, genetics, sociolinguistics, neurolinguistics, and descriptive linguistics. Fundamental questions about the origin of humans and human language differ significantly from questions about the correct grammatical account for vowel phonemes, grammatical morphemes, or word order in dialects studied by linguists. The best one can aim for with respect to basic questions like the origin of language is to reduce the huge number of hypotheses proposed. Hence, a central goal is to overcome the myriad of ad hoc narratives that plague discussions on the origin of language.

SOME LEVELS OF THE EMERGENCE OF GRAMMAR AND THE HYPOTHETIC LEVEL OF A "PROTOLANGUAGE"

Despite the impression of coherence conveyed by current grammars, the language capacity does not arise from a single source or a single cause. Briefly, it is not God given, nor the result of some mythical pseudo-Darwinian mutation. Its origins are located at different points in time during the last millions of years and these represent a response to very different ecological (selective) conditions. A stratification that corresponds to this series of causing or channelling conditions and to the consecutive adaptations (in the context of the evolution of the climate, the flora, and the fauna) can be recovered from current linguistic behaviour. Features that evolved earlier have a different place in the cerebral (bodily) architecture and are functionally presupposed by more recent strata. One may postulate the following stratification:

1 *Basic level.* The most basic cognitive level contains the capacity of efficient locomotion for causal impact on the environment and action. If consciousness is added, one obtains a set of dynamic scenarios that control intentional behaviour and a grasp of causality.[1]
2 *Emergence of productive vocal articulation and auditory perception.* At this stage powerful perceptual and motor faculties for vocal communication evolve. This capacity was further elaborated with the advent of vocal communication for social comfort and control. The basic principles of phonology may have emerged in this period.
3 *A protolanguage based on a compositionally enriched lexicon.* The underlying capacity characteristic of this stage is a very systematic exploitation of the affordances of the ambient and changing ecology. This capacity was amplified by profiting from the growth of associative areas in the cortex. It probably evolved continuously in a long period between the migration of *Homo erectus* and the reign of Cro-Magnon man.
4 *The evolution of syntactically and textually complex languages.* This stage emerges with the mastering of stable valence patterns and the use of verbal art (narratives, rhetoric, song, myth). It allows a canon of myths and other text-based cultural traditions to be established. This level probably emerged with archaic *Homo sapiens* and was fully evolved and functionally exploited in Cro-Magnon populations, which created the first large cultural networks (cf. Wildgen, 2004, chap. 8).

In our search for the human capacity for language, we should respect the evolutionary stratification of the linguistic capacity of humans, which contains the levels of:

1 action/motion perception and planning (dynamic archetypes)
2 phonetic/phonological principles and routines, such as basic feature distinctions, syllable structures, rhythmic and euphonic constraints, i.e., phonetic universals and principles of phonological self-organization
3 universals of lexical fields, polysemy, metaphor, and compositionality principles for words or similar entities
4 syntactic and textual principles for the organization of larger linguistic gestalts.

Since Darwin's theory of evolution the basic idea has been that of a continuous evolution moving by infinitesimal steps. Applied to language, linguistic capacities should derive in a continuous series of steps from communicational habits and intellectual capacities of mammals and animals in general. According to Darwin, the hypothetical "protolanguage" can only be the construct of an intermediate stage which helps to fill the gap between animal communication/cognition and human communication/cognition. Recently Derek Bickerton has made several proposals for understanding *protolanguage*. He characterizes his methodology as follows:

> If there indeed exists a more primitive variety of language alongside fully developed human language, then the task of accounting for the origins of language is made much easier. No longer do we have to hypothesize some gargantuan leap from speechlessness to full language, a leap so vast and abrupt that evolutionary theory would be hard to put to account for it.
>
> (Bickerton, 1990, p. 128)

He uses data from pidgin and creole studies, data in primate and child language acquisition, and Kaspar Hauser cases. I shall instead consider the traces of semiotic activity in hominids and early man down to the emergence of writing systems as data for the reconstruction of intermediate forms of human language. This strategy has two consequences:

- Insofar as the contours of early semiotic capacities can be reconstructed from artifacts and art, one can only infer the semantics (and perhaps the pragmatics) of an earlier language capacity, not its lexicon or syntax.
- As the artifacts pertain to the cognitive level rather than to the level of linguistic expression, the reconstructed semantics must be a type of cognitive semantics.

As the evolutionary process is in principle continuous, the term "protolanguage" designates a zone between the linguistic capacities of early hominids and modern humans. In general, the existing empirical evidence is used to make an informed guess about one possible intermediate stage. In this perspective, an evolutionary grammar must account for the starting point, i.e., the symbolic capacity of the last common ancestor (LCA) some five to seven million years ago via a hypothetical protolanguage, down to the languages of modern populations currently being investigated.

THE FORMAT OF AN EVOLUTIONARY GRAMMARII[2]

The project of developing an adequate format for "evolutionary grammar" encounters problems similar to those faced by developmental grammars, grammars for linguistic change, diachronic grammars, or models of grammaticalization and grammar-genesis. Labov (1972) tried to adapt his sociolinguistic data to modified versions of a generative grammar, as did Klein and Dittmar (1979). Developmental studies adopted the generative paradigm in its Principles & Parameter version and Bickerton (1990, p. 199ff) reduces eight modules of the Chomskyan model current in 1990 to two modules: phrase structure (X-bar-theory) and verb-argument clusters. These two are then used to distinguish the protolanguage of *Homo erectus* and *neanderthalensis*, which would be structurally limited insofar as neither complex phrases nor verb argument complexes could be mastered. I shall give an alternative formulation of these two features (cf. Bickerton, 1990, pp. 189–197).

Decades of attempts to adapt the theoretical models of grammar to the concerns of interdisciplinary linguistics demonstrate that a system of rules—even one with basic categories, modules and principles—is not able to map the inherent (and not just parasitic) developmental, historical, and evolutionary character of language. The grammatical tradition of normative grammars, school grammars, and competence grammars falls short of the requirements. In general, models that deal with the dynamics of language are required (cf. Wildgen, 1994). A basic insight of interdisciplinary linguistics from socio- to neurolinguistics is that the semantic and pragmatic aspects of language are the primary focus for any explanatory enterprise in this field.

The semantics of space and time in a protolanguage

One can distinguish two subaspects: processes in space, such as spatial orientation and navigation, and temporal and rhythmical patterns. The representation of space has to do with boundaries in space and perspectives on them. A first perspective is centrifugal. It has its point of origin in the individual and bodily motions and leads to the construct of an *experienced* three-dimensional space: in front of–behind (go), above–below (climb, fall), left–right (grasp with the left hand or the right hand). This space of bodily motion involving feet and arms defines the immediate space where objects may be approached, grasped, and manipulated. The intermediate space depends on man's ecology; it can be housing (the cave, abri) or the village. The distal space contains roughly all possible itineraries, e.g., of hunting/gathering. The latter perspective is centripetal, i.e., the individual is seen as the location of effects triggered by external causes. The sky, the horizon (typical points where the sun rises or sets), the favoured direction of winds, or a ridge of mountains may be the external locus of orientation for the individual, which is at the centre of a field of force defined by these boundaries. Many myths and religions refer to this extreme locus of orientation when they interpret the fate of humans as standing under the control of distant (and often invisible) forces, e.g., God–in heaven, the devil–underground.

As soon as space is specifically organized in relation to cognition and social use, it enables a cycle of social "investment". Architecture and the spatial organization of villages (or later towns) are clear examples. The spatial organization produces a type of cyclic structure in which different partial orientation functions cooperate. Figure 2.1 depicts such a cycle.

Specific symbolic media are rooted in each subspace and coevolve with it. Thus the painted Paleolithic cave in the Franco-Cantabric culture is a specification of mythical/ritual space; it is also connected by its illusionist paintings to the ecological space of hunting. The relation is iconic, indexical (in its magical impact), and symbolic (in its abstract signs; cf. Wildgen, 2004, pp. 80–83). The dark painted cave is related to cave openings, abris, and huts where people live. The open space in front of the cave openings or huts is a public space where the production of artifacts and the distribution of shared food occur. This public space is the border to the open field of the chase and harvesting. Human action patterns occur inside a specific space or make a transition from one space to the neighbouring one. In rituals the core of these action patterns is fixed. The coding of action patterns by rituals is a social preparation/presupposition for linguistic rules/grammars.

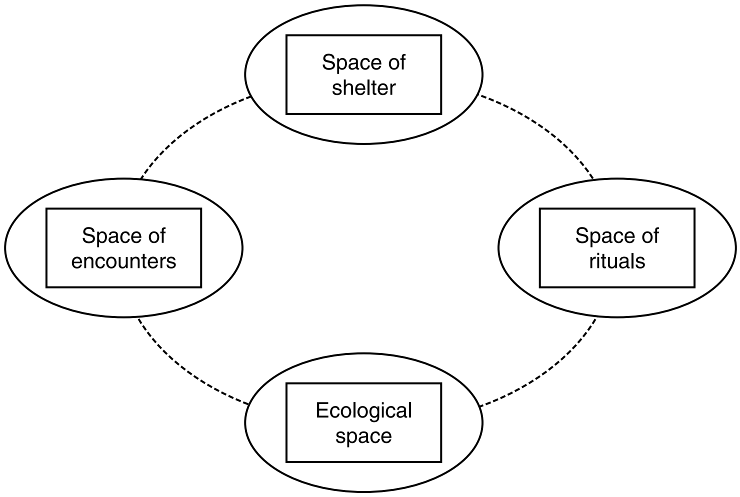

Figure 2.1 Functional subspaces and pathways that are linked to each other.

The cognizance of such schemata for orientation may only be evident in behaviour as it is in many animals; it may be gestural or it may be deictically organized in a phonic language (cf. Levinson, 2001, p. 317ff). For *Homo erectus* the cognizable space seems clear. The inner space is defined by the use of hands and instruments; the proximal space by the choice or construction of dwelling-places to which the group could return. The centripetal organization is involved in long-range excursions and migration. Since the orientation system cannot be genetically coded, it must be learned, adapted to changing contexts, and socially shared. Language is one possible solution to this problem, be it gestural (behavioural) or phonic. As humans have chosen the path of phonation, it is plausible that our ancestors began to proceed further in this direction.

The representation of *time* is rooted in the classification of multimodal sensory inputs using specific temporal rhythms (clocks). Pöppel (1994, 1997) proposed two temporal windows for multimodal integration:

- The window of 30 ms. An event becomes an object of (multimodal) perception only after a stability of 30 ms. It can then be classified, labelled, compared, i.e., further processed.
- The window of 3 s. A sequence of events can be understood as a structure. In this window the smaller units (> 30 ms) are correlated as: before–after, cause–effect, etc. This is the point where a notion of structured temporality is born.

A protolanguage must categorize events and actions (by protoverbs) and must discriminate stable entities (by protonouns). The question arises as to whether temporal, dynamic, quantitative, and qualitative *relations* between them can be mastered and if so, to what degree. This question brings us to the two basic delimitations of a protolanguage, as discussed by Bickerton (1990): first *case-frames* or *scenarios of action* (government) and second *phrase structure* (X-bar-structures). I will argue in the next sections that there are intrinsic complexity barriers which could have blocked the further elaboration of a protolanguage over a long evolutionary timespan.

Scenarios of actions and events

The event-schemata derived from catastrophe theory can be employed as a measure of complexity (Wildgen, 1982, 2005). Let us take grasping with the hand as a starting point. The action-concept GRASP involves the body (the hand) and an object and presupposes the perception of stable entities in the 30 ms window. The whole schema should fit into the 3 s window, e.g., in the sentence:

<u>The father</u> ***takes*** <u>the book</u> (from the table).
agent ***object*** ***location***

If we add two or more GRASP scenarios and integrate them spatially and temporally, more complicated scenarios can be generated:

The father ***gives***	the book	to his son	(in the living room).
agent	***object***	***patient (goal)***	***location***
The son ***hits***	the ball	with his racket	(in the yard).
agent	***patient***	***instrument***	***location***

However, these elaborations are a considerable step beyond the simple (two value) schema of grasping. They require higher cognitive controls and a social demand for such complex habitually occurring scenarios ("I" is an *intermediate* entity, which is either the object transferred or the instrument used). (See Figure 2.2.)

Early humans, e.g., *Homo habilis*, already had a hand with the opposition of thumb and fingers, but some features are still linked to climbing as in gorillas and chimpanzees. The *Homo erectus* had a hand that was adapted to strong grasping (cf. Piveteau, 1991, p. 74ff). This was still true of Neanderthals. A shape of the fingers that slightly deviates from the statistical average can even be found in some human populations.

One may distinguish three types of grasping:

- the force grip (e.g., of a branch)
- the precision grip (e.g., of a small tool)[3]
- the refined grip (e.g., of a needle).

These distinctions, which have a long evolutionary history, constitute a kind of manner specification in relation to the GRASP schema. As the distinction between the types of grips shows, more elaboration appears as soon as more precise manipulations on objects and instruments are developed. The fact that a cognitive bivalent schema and a manner component were mastered does not necessarily mean that they could be transformed into phonic signals. One could even argue that the teaching of hand skills does not call for linguistic

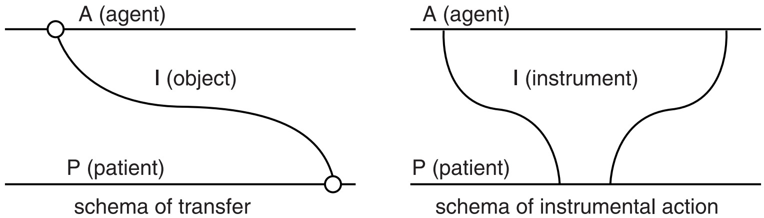

Figure 2.2 Schema of transfer and schema of instrumental action. The catastrophes contained are: A (the agent) emits I (the object) and P (the patient) catches I; in the schema of instrumental action I is the instrument that affects the patient.

instruction. If we assume frequent vocalization, inferred from the evolution of the sublaryngeal tract and a steady increase of memory due to the growth of the brain, it becomes clear that this cognitive schema and subsequent ones are preadaptations for the evolution of verbal phrases or valence patterns in sentences. Thus, in order to verbally represent important and recurrent actions in a protolanguage the cognitive schema of grasping could serve as the basis for iconic/metaphorical transfer to all kinds of manipulations on objects. As soon as instruments are employed, this schema could be iterated.

- The father (A) takes a hand-axe (B) to *move/change/kill* some object (C).
- The father (A) takes a stone/bone (B_1) to *hit/shape* the pebble (B_2) that should later *kill* the animal (C).

In this development a first barrier of complexity appears. While the GRASP schema is dynamically and topologically simple, the composition of such schemata is not simple. One needs a specific topology/geometry to restrict the degrees of freedom for such a composition. Therefore, the GRASP scenario and its bivalent (asymmetric) pattern describe the constitutive plateau of a protolanguage, from which the evolution of a full-fledged language could begin. The evolutionarily old distinction between types of grip and manners of locomotion, related to the dynamics of the legs, is a preadaptation of the manner component in spoken (or gestural) language and could well belong to the basic constituents of a protolanguage.

A set of rather abstract specifications which are often grouped together in pidgin and creole languages can be called the TMA-component (T = Time, M = Mode, A = Aspect). They are the next step that could have “evolved” in a protolanguage on the way to higher level grammars. Gestures or tonal modifications were possibly the precursors of morphemes specifically devised for that purpose. Thus, the order of emergence of grammatical features transcending the GRASP scenario could have been:

1 the elaboration of the manner component
2 the elaboration of the TMA component
3 the elaboration of valence patterns (up to valence 3 or even 4).

One has to assume simpler schemata underlying the GRASP scenario, such as the schema of stable existence. If we apply the 3 ms window, any entity that does not change within this window is a candidate. As the inputs of classification or labelling-reactions are not only spatiotemporal events but also qualities, one can assume a slow increase in quasinominal/adjectival labels as soon as memory capacities grew and corresponding social demands appeared. One could imagine that labels for animals, plants, other people, and artifacts were the first candidates for a growing lexicon. This development is also the natural continuation of classificatory capabilities in other mammals and even in birds and fish.

Cognitively, the manufacturing of stone tools and a fortiori of tools shaped with the help of stone tools goes beyond the GRASP scenario. One hand or one foot must fix the pebble, the other hand grasps the stone or bone that hits the pebble. Finally, the planned chipping off subtracts material from the chosen stone and after several strokes the desired sharp edge of the pebble is produced. This scenario involves two objects, two hands, and a change in the shape of the pebble (the separation of parts from it). This schema contains four symmetric "grasping/emitting" subschemata (simple instrumental action) and one further "emitting" schema. The integration of the shaping by the tool is on a higher level of complexity (it has two force dimensions) and has structural stability only under very specific conditions. In fact, a linguistic description of the action normally requires more than one basic sentence pattern in actual languages. In this sense Paleolithic industries drove the evolution of symbolic tools like language.[4]

The complexity of (nominal) phrases

Bickerton suggested that in order to organize a descriptive (nominal) language one needs three structural layers:

> (a) a generic class, X; (b) the properties peculiar to particular members of that class (large, with a dark red cover, of Mary's) and (c) the specification of the complete individual in terms of abstract relations such as quantity, proximity, familiarity, and so on (a, this, there)
>
> (Bickerton, 1990, p. 195)

His proposal again reflects a position typical of generative grammar (in 1990). In Chomsky (1995) many of the specific features of the Principles & Parameters model are abandoned, because they "appear to be computationally irrelevant" (p. 389). I think that there is no need to follow the traditional X-bar-schema in an "evolutionary grammar". The primary reason is that the function of determiners (a, this, there) differs from other specifiers (attributes), e.g., deixis or the anchoring of an utterance in context. Their evolution probably followed the evolution of manual and/or visual deixis and later incorporated these devices into the phonic code making them more reliable and less context dependent. This function may be called indexical and exhibits another evolutionary path linked to traces and to effective binding between language and nonlinguistic action.

The head (e.g., the noun) and its attribute (e.g., the adjective) share the basic feature (+ nominal) and the problem to be solved is the risk of improperly blending two or more semantic spaces. A semantic space may be conceived of as defined by a set of independent features, which are either polar oppositions or graded scales. Ideally, a semantic space should be homologous to an imagined space, i.e., the hearer/speaker should be able to "flatten" it until it has no more than three dimensions. The imaginative content of meaning is only accessible under this restriction. A very general notion of mental spaces (without topological restrictions) is employed by Fauconnier (1997). The notion of conceptual integration put forward by Fauconnier and Turner (2002) elaborates this approach, taking up insights from the cognitive theory of metaphor. I suggest that these proposals could be applied to the phenomena discussed in this section.

If every noun or adjective is associated with a place in a semantic space, then the mapping of one place in space A to another in space B is a problem insofar as the spaces are different and may not be easily transformed into a conjunct space A×B. This is possible if A is *father* and B is *old*; in this case *old* + *father* has a new, well-defined place, insofar as *age* is an implicit feature of humans. It is not the case, if one tries to combine *father* with adjectives like: *narrow, deep*, or *fluid* and *quadratic*. Moreover, if A has *n* dimensions and B has *m*, then A×B may have *n*+*m* dimensions (if space A is not contained in space B or vice versa) and any increase in dimensionality creates new instabilities; this is a general result of topological dynamics. Another problem is that the mapping of a space A to space B under deformation (insecurity, vagueness, variation) easily produces chaotic results as experiments with video feedback have shown (Wildgen, 1998). Thus Bickerton (1990) dramatically underestimated the problem posed by (iterated) attributes of nouns. On the basis of these considerations, one can specify an *evolutionary barrier* concerning the blending operation necessary for groups with two or more constituents. Hence, the proper hypothesis is that such products of semantic composition were not (or not stably and reliably) accessible to speakers of a protolanguage.

The self-organization of a grammatical system

Pinker and Bloom (1990) think that the central human feature of language is syntax and that an evolutionary theory of language should therefore explain the selection criteria for the syntactic abilities of modern humans. In contrast, Kirby (2000) argues that compositionality (and thus syntax) may emerge if the size of the lexicon, i.e., the meanings associated with linguistic expressions, increases and if individuals learn from utterances. In a computer simulation he shows that after a stage of random invention (and noise) a sudden change occurs, after which: “The number of meanings covered increases dramatically, as does the size of the grammar” (pp. 313–314). A further stable state emerges when the number of meanings increases and the size of the grammar decreases. The resulting grammar is not only compositional, but also “groups all the objects . . . under one syntactic category . . . and all the actions . . . under a second category” (p. 314). Given a certain size of the lexicon and a process of language learning, a syntax can evolve with no Darwinian selection required. In search of further answers one may turn to the mechanics of sign production, i.e., to phonic evolution. MacNeilage (1998) proposed a mechanism in his “frame/content theory” of the evolution of speech production, which leads to higher order organization in language. Studdert-Kennedy (2000, p. 171) states in his comment on it: "Both short- and long-term phonetic memory were also essential pre-adaptations for syntax. . . . Without a pre-adapted system for storing phonetic structure independently from its meaning, syntax could not have begun to evolve."

The purely syntactic problem of chaining elements of an existent vocabulary therefore does not require a specific endowment or evolutionary processes enabling it. The deeper problem is that of semantic compositionality, because the mapping/blending of spaces with different topologies and accounting for the dynamics inherent in verbs is crucial for sentential units. This is the central problem that early humans had to resolve in order to allow for stable and reliable communication via phrases and sentences. In order to arrive at a level of syntactic behaviour, early humans had to accommodate two major factors:

- They needed the cognitive capacity for a stable solution to semantic compositionality. This is the *cost* of higher order language capacity.
- There had to be a social demand for a compositional level of referentiality in communities of humans. If this demand is satisfied, the *gain* of higher order language capacity becomes apparent.

Rewarding situations probably often arose by chance and the evolving species spontaneously used the "dormant" capacity. With the increase in population density and in networks of supraregional communication and exchange among modern humans such a system very likely became necessary. As soon as it was developed, it spurned long traditions of language use down to modern times. Since language is deeply rooted in human biological endowment, the turning point in the use of cognition for language must lie before the rise of modern man, i.e., before at least 100 ky BP (thousands of years before the present) and probably even before 200 ky BP. Thus, the central question is not how syntax came about, but what made it rewarding to use the available cognitive potential for syntax. The pay-off could be social or individual, which in turn could lead to higher social competence and thus to social gain.

A cognitive and functional explanation should take the role of language in (silent) thinking as its point of departure. If thinking is a kind if silent speaking learned in early childhood, the complexity of an internalized language could provide advantages for silent deliberation and planning, even if it is not always uttered in social communication. The intellectual advance would indirectly lead to advantages for the community, e.g., having (and listening to) individuals able to act strategically and to plan effectively or to solve other difficult problems. These communities could be selected for their superiority in competitive situations. The role of specialists (technicians, artists, and scientists) is highlighted by the first large civilizations like that of the Egyptians. The artists that created Paleolithic paintings and sculptures possibly belonged to groups marked by their intellectual and practical superiority. This line of thought, which seems promising, is sketched in Foley (1997, p. 73).

Major levels of an evolutionary (biological) grammar and the transition towards a culturally based grammar

From the stratified capacities and barriers discussed here, we can derive a first sketch of the grammar of a protolanguage. It specifies three hierarchical levels:

1. *stable* entities: no change in the perceptual and classificatory time window and recurrence as pattern (statistical relevance)
2. *dynamic aspects* of *entities* in change and motion ("force-dynamics")
3. the *bivalent* GRASP schema (capture or emission).

This allows for the accumulation of a lexicon of protonouns/-adjectives and protoverbs. The combinatorial possibilities depend on context. The grammar of the protolanguage is based on these protoclasses and their implicit dynamical binding forces (valences) and inherent causal links (force dynamics). The complexity barriers may explain why further conditions of control on the combinatorial/mapping/blending semantics had to be satisfied in order to arrive at a more complex and less context-dependent grammar. I have mentioned three basic restrictions, which apply to protomanner adverbials, to a compact TMA component, and to recursive constructions of specifier phrases with a protonominal head.

COMPARATIVE BIOLINGUISTICS

If we want to understand human language, we cannot simply study animal behaviour, cognition, and communication. Since Jacob von Uexküll's proposals in his "*Bedeutungslehre*", far more radical comparisons have been made. These led to the field of "biosemiotics", which is characterized as follows:[5]

- The life sphere is permeated by sign processes and signification. Whatever an organism senses also means something to it.
- All organisms are born into a *semiosphere*, i.e., into a world of meaning and communication.
- The *semiosphere* places constraints or boundary conditions on the environment (Uexküll's "*Umwelt*").[6]
- It is plausible that the semiotic affordances and demands of populations were a decisive challenge to their success or even to their survival.

In this perspective, we can learn from very primitive animals and finally relate the basic laws of sign behaviour to the origins of life (Wildgen, 2006). This perspective may elicit enthusiasm or provoke rejection. In the case of mammals and primates, the parallels between human and animal behaviour are rather obvious, so that a systematic comparison seems promising. However, the Darwinian principle of continuity forces us to go beyond our immediate cousins and house pets to ask: What are the deeper principles in the realm of life that make communication and language possible or even necessary? It seems that beyond physical laws and chemical reactions, the facts of life refer to a new type of meaning, function, and purpose, which is characteristic of complex living systems. An important feature is that they operate simultaneously on different levels of magnitude. Thus one must be aware that the choice of a specific scale (from micro to macro) leads to very different insights concerning the living organism. What is extremely complicated (even chaotic) at a lower level (physical, chemical) may show a very regular (almost simple) organization at a higher level. René Thom (1974) suggested that the morphology of language is much simpler and in its organization more basic than the underlying neural or physiological morphologies. Even the physics of a process rendered by verbs like *run, catch, give, send, tie*, etc., are many orders of complexity beyond that of the linguistic patterns.

Contrary to reductionist views, the facts concerning the human mind and language are not beyond human grasp. They are the first things we can understand because we embody them. We are able to ask and answer questions concerning nature beyond our bodies only through symbolic competence. This creates an epistemological dilemma. Humans use their linguistic competence to do research in the natural sciences, but this competence cannot be assessed in the same way as questions in the natural sciences. This dilemma motivates a kind of garden path strategy in the humanities. First one starts in the direction of central concern: How can we understand human beings (ourselves)? This path ends after several steps and scientific progress becomes more and more difficult; the results are vague, insecure. Then using human symbolic competence, one returns to the analysis of physical, chemical, and biological aspects. With this apparatus one returns to the former garden path and hopes to move beyond the previous point of failure. One probably gets one step further, but must again undertake new expeditions into nature in order to take a third step down the garden path. Thus, learning from animals and from plants, biomolecules, chemical dynamics, etc. is a necessary strategy, but one should return to humans, their minds and language, even if the risk of failing along this path is daunting.

NOTES

1 The instrumental behaviour of chimpanzees living today shows elements of such a capacity. However, it remains, an open question whether the common ancestor of humans and chimpanzees had the same capacity (probably not).

2 This section is a partial, revised version of a paper presented at the Max Planck Institute for Evolutionary Anthropology in Leipzig on 29 May 2002; see http://www.fb10.uni-bremen.de/homepages/wildgen/pdf/evolutionarygrammar.pdf

3 In the evolution of pongids the origin of the precision grip seems to be a critical transition which allowed "grasping predation of certain species of insects at the terminal ends of bushes and shrubs" and this "opened a niche for primate evolution" (Quiatt & Reynolds, 1993, p. 123). It had as consequence the "conversion of active behaviour to crepuscular and diurnal phases of activity" (p. 123).

4 In the sense of Cassirer's "symbolic forms", technology already produces symbolic forms. Therefore, the impact on language would be a transfer between different types of symbolic forms (cf. Wildgen, 2004, p. 9).

5 Cf. Hoffmeyer (1996) and the definition of "biosemiotics" on his homepage: http://www.imbf.ku.dk/MolBioPages/abk/PersonalPages/Jesper/Hoffmeyer.html

6 Jakob Johann von Uexküll (1864–1944) worked at the Institute of Physiology, University of Heidelberg, and at the Zoological Station in Naples. In his book *Umwelt und Innenwelt der Tiere* (1909) he introduced the term "*Umwelt*" to denote the experienced world of an organism.