Major steps in the evolution of language

Dieter Wunderlich (Düsseldorf)

1. Introduction

Imagine a group of early hominids, somewhere between ape and modern homo, in the swampy shores of a big lake. Some are engaged in fish-hunting, others try to break up crabs, and a few are sitting on a rising ground, guarding their kids and gesturing about the preparation of meal, while one of them is watching around. Suddenly, he sees a crocodile. How does he communicate this news to his fellows? We don't know. Archeologists can find skulls, teeth, and bones; they can date them back and draw conclusions from them. Linguists can't find anything at all. Language is a software, which disappears when the brain dies.

Linguistic structures are quite abstract, and they can be inferred only indirectly. Even more abstract is the issue of the evolution of language. As said, there is nothing that can document early stages of language. Nevertheless, our reasoning about necessary or possible steps towards such a complicated system can converge if we take into account what we know about language in general and what we know about the evolution of modern homo from other disciplines. Many small changes occurred in various ways, but in the end, all of the evolved species died out except homo sapiens, although some of them lived longer than our own species exists. Likewise, in languages many small changes occurred in various ways, but in the end, most of the evolved communication systems disappeared. Anatomical properties, including those that we learned from fossils, can be traced back, informing us about possible and impossible stages. Likewise, properties of contemporary languages can be traced back, and one can distinguish between more and less plausible stages of language evolution.

I will put forward three claims in this paper. My first claim is that the integration of two independent evolution lines, one concerning gestures and another one concerning vocal utterances, was crucial for the evolution of human language, although this integration might have happened already in early homo one million years ago. My second claim is that the definitely last step towards human language was the invention of lexical categories such as 'noun' and 'verb', which had important structural consequences. My third claim is that syntax proper, which involves the concept of movement (or positional alternatives), only became established after separation into language families. These three claims stand against certain commonly held assumptions; in particular, that there is only one line in the evolution of human language, that this evolution culminated in the invention of syntax, and that no essential further progress has been made within single language families since that time.

2. Periods in the evolution of modern humans

As evidenced for anatomic and other biological features, evolutionary steps are minimal and produce a multiplicity of phenotypes. It is reasonable to assume that human language is the product of a similar kind of development. The many species and subspecies appearing in the hominid evolution may have differed from each other not only with respect to anatomical and physiological factors but also in terms of a variety of pre-linguistic features they had command of. Since human language is a quite complicated system involving several modules, there might have been separate evolutions for at least some of these modules. In other words, every module may have evolved in slightly different circumstances. Moreover, the integration of these modules could have led to different types of language, of which present-day human language is only one type.

Organizing the many varieties of language to be expected in the evolution of human language and its further developments, I will divide the (pre-) history of man into four major periods, based on what we know by external evidence.

Period I: *The age of hominids*. Between 7 and 2 million years ago, the now extinct hominids emerged in East Africa (as well as in the Chad), all of them distinct from the chimpanzee. They changed several anatomic as well as physiological features drastically, developed bipedalism and shifted to an albuminous nourishment. Some of the hominids including the predecessors of homo may have possessed some archaic gestural language.

Period II: *The age of homo*. Between 2 million and 100,000 years ago, several species of homo populated East Africa, and some of them also expanded to Asia and Europe, where they later coexisted with homo sapiens.

The homos lived in groups of medial size, they were able to make tools as well as other artefacts. At least some subspecies may have developed phonetic language, probably in several steps.

Period III: *The age of colonization*. Between 100,000 and 10,000 years ago, groups of modern homo sapiens colonized most parts of the world. Modern human populations settled in South Asia 70,000 years ago, in Australia 50,000 years ago, and in Europe and probably also America 40,000 years ago. By then, many of the language families known today had emerged.

Period IV: *The post-neolithic age*. Around 10,000 years ago, agriculture was introduced in several places around the world: Mesopotamia, China, Africa, and America. The size of populations increased rapidly, migrations took place, and many cultural inventions followed. As a by-product, the linguistic map of the world changed rapidly, which is still true today.

Our knowledge about the history of language only captures a fraction of period IV. Any claim about the preceding periods is necessarily speculative. Archaic language could have been simpler, but in what respects? Do the living languages show certain traces leading us to the past? And, which features must language have exhibited in the beginning?

If one abstracts from our knowledge of contemporary languages one may find some elementary functions of language and relate them to possible pieces of utterances that fulfil these functions. One may also find some elementary combinatorial operations.

Before entering this issue, some preparatory remarks are necessary.

Given the fact that there exist modern human languages with rich morphology but rudimentary syntax (such as the polysynthetic languages), and following the observation that morphology in general is less effective than syntax (see section 6), there is reason to believe that languages with poor syntax are better candidates for abstracting elementary functions of language than languages with rich syntax. Many linguistic universals assumed by syntacticians are probably not part of what we have to look for when we are searching for elementary functions of language.[1]

---

[1] A similar, more specific objection against certain syntactic universals is raised by Van Valin (to appear), who argues that languages such as Lakhota (Siouan) do not show any evidence of a cooccurence of displacement and uninterpretable features,

As an interesting fact, which is almost common-place among linguists, we have to note that none of the languages known today is a 'primitive' language. In the contrary, many languages that have few speakers but rich morphology distinguish more categories and exhibit more complexities than many of the better-known languages spoken by many speakers. What we do find is the fact that some present-day languages tend to higher generalization; in the course of their development they loose some of their categories and complexities. I am therefore convinced that contemporary languages may differ in their nature as being more or less archaic. However, they do not differ in terms of being more primitive versus less primitive, but rather in terms of being less generalized versus more generalized.

On the other hand, if we consider how human language could have looked like in the beginning, everyone is convinced that it must have been simpler than every known language. Going backwards from present languages we can find some grammaticalization patterns by which new categories were introduced, but also, we can find categories that were lost in the course of linguistic change. Therefore, in considering the historical development of language we need further ideas about unidirectional structure-building, apart from the idea that functional categories are grammaticalized unidirectionally. One important issue is generalization in terms of structural rather than semantic properties. Another one is the enrichment of the vocabulary, which gave pressure for generalization. It is also necessary to reflect upon the basic traits of language and how these may have evolved.

3. Some basic characteristics of human language

As said already, human language is a many-module system. If we reflect upon the basic traits of language, we have to pay attention to the fact that some characteristics are specific to certain modules. It is plausible to assume that each characteristic meets some selectional advantage. Moreover, the basic characteristics of human language may have different sources, and they might have emerged at different times.

---

which was claimed to be necessary in Chomsky's Minimalist Program (Chomsky 1995, 1998).

(i) *Symmetry*: Language is a speaker-hearer symmetric system. This concerns pragmatics, as well as linguistic processing.

Speaker and hearer permanently change their role; we can even speak and understand simultaneously. Moreover, everyone monitors his speaking by understanding. In general, understanding can be viewed as a kind of imitation. Language perception and language production can thus be seen as supported by a single generative system. The permanent symmetry of speaker and hearer role correlates with important features of linguistic processing.

Linguistic processing is fast: it starts right out with the beginning of an utterance. It is robust: disturbing noise, failures and gaps can be ignored. It is automatic: nobody can refrain from understanding an utterance. It is organized in a series of networks working top-down in language production, but bottom-up in language recognition. In the latter, the identification of phonemes triggers the identification of morphemes and their relations, which in turn trigger the identification of meaning elements.

Hence, symmetry has some straightforward selectional advantage. It allows for processing in a single system that adapts to all kinds of states of affairs during discourse.

(ii) *Off-line processing*: Utterances can refer to situations that are not present. This concerns semantics.

I prefer here the term 'off-line' rather than 'symbolic' because it is more general; iconic signs that imitate their meaning can more easily be performed off-line than symbols. Symbols may be the final product of off-line processing. The selectional advantage of off-line processing is again quite plausible: it allows to reflect circumstances for preventing danger, and to prepare collective activities, among other things.

(iii) *Discreteness* (or digitalization): Minimal change of features yields a different meaning. This primarily concerns segmental phonology.

Discreteness comes out by means of threshold values, and so makes independent of noise. Moreover, it allows for a simple binary combinatorics. Discreteness is inherited to CV (consonant-vowel) syllable structure, to the rhythmic alternation of strong and weak foot, and to lexical items built on these structures.[2]

(iv) *Double Articulation*: Minimal elements are meaningless, only larger elements bear meaning. This concerns the important difference between phonology and lexicon. Double articulation allows to create a larger inventory of lexical units by minimal means.

(v) *Duality*: parts of utterances are memorized, thus constituting the lexicon, while complete utterances are produced combinatorially, according to morphosyntactic rules or principles. Combinatorial systems have two clear advantages in terms of economy and flexibility. They reduce the expense of storage (economy), and allow small combinatorial changes to bring about large effects (flexibility).

(vi) *Reference and predication*: Any meaningful utterance about a state of affairs expresses some predication that involves potential referents, so in addition it must be referentially anchored. This concerns semantics in its combinatorial aspects. Both a powerful system of predication and a powerful system of referencing, including reference tracking, clearly improve communication.

It is not implausible to assume that the afore-mentioned characteristics of language evolved in the given order. However, they probably have different sources. Symmetry and off-line processing most probably originate from the evolution of the hand, while discreteness and double articulation originate from the evolution of the vocal apparatus. Some kind of combinatorics may have played a role from the very beginning of archaic language. However, morphosyntax in its modern sense almost certainly is the last invention with which human language came into existence. The way in which morphosyntax is organized reflects the fundamental needs of reference and predication. I will call the first language of modern homo sapiens, from which all known language families diverged, *protolanguage*, and all preceding steps *archaic language*.

As is well-known, vocal language and gestural language are tightly connected. Usually, speakers use gestures accompanying their speech. If

---

2   In arguing along this way, Carstairs-MacCarthy (1999) assumes that the properties of CV structures passed their way towards syntax, which in my view slightly underestimates the factors that brought syntax into existence.

they are hindered to use their vocal language, they easily reorganize their gestures in more systematic ways. And deaf people, as we know, can develop a fully effective sign language. We can conclude that language is more abstract than the particular medium suggests. Furthermore, the ease of switching to sign language suggests that gestures played an important role in the evolution of language. Symmetry in the above sense is most likely to have arisen from a gestural system which is encapsulated in the hand-motoric/visual system. This is why the study of sign languages can contribute to the question of how language evolved.

4. Initial steps in the evolution of language

The best way to approach the initial conditions for the evolution of language is to consider our closest relatives among the animals. Apes use both gestures and vocal calls in ritualized ways, i.e. in fixed contexts, but their gestures are more flexible than their calls. Calls, which are addressed to the whole group, are non-symmetric, whereas face-to-face gestures have an interactive potential, hence, they are potentially symmetric. As recently has been discovered (Gallese et al. 1996, Rizzolatti et al. 1996), certain motoric brain regions of the ape are already activated if the ape watches another ape manipulating in particular ways, e.g., concerning food. This important discovery gave rise to the mirror neuron hypothesis, saying that some part of the neuronal network produces mirror images of other's behavior. The mirror neurons could have played an important role in adapting the hand to purposeful actions, and also may have enabled the imitation of actions, which led to symmetry-oriented behavior, based on the deciphering of communicative actions by resynthesis.[3] The mirror-neuron hypothesis is the first serious attempt for bridging the gap between uncoordinated individual actions and symmetric actions that can convey content (see also Rizzolatti and Arbib 1998).

The enormous development of the hand during the process of improving tool making, accompanied with corresponding developments in the brain and supported by tool making traditions, suggests that the hand could have been used for communicative purposes relatively early. The imitation of

---

[3] Liberman (1957) was the first who claimed that speech perception is based on resynthesis.

manual techniques could have been the source of off-line thinking. Highly differentiated actions with the hand can be trained. Manual actions can be performed face-to-face, they are fast, controllable, and potentially symmetric. Moreover, the hands are suitable for expressing elementary scenes more or less iconically. Deictic gestures can establish a referent, while form, position and movement of the hand can establish a predication.

The assumptions on the preconditions of human language also depend on our view of how the shift of anatomical and physiological properties in the time between 7 and 5 millions years ago was conditioned. The savanna theory (commonly accepted) claims that the predecessors of homo adapted to a savanna-like environment, while the semi-aquatic theory (which is the minority view) claims that they adapted to an environment formed by inland waters surrounded by wood, which was caused by the rising of the sea-level. In both cases, bipedalism could have evolved. However, the semi-aquatic theory offers more satisfying explanations for many of the physiological differences developed in human beings in contrast to apes – an issue that I do not want to discuss here in more detail. Suffice it to state that the semi-aquatic theory is a conceivable alternative.[4]

According to this theory, the lowering of the larynx, as well as volitional breath control, result directly from an aquatic environment rather than indirectly from conditions of bipedalism. This suggests that the vocal apparatus had its own evolution, which possibly started even earlier than tool making. In any case, there is no obvious intimate connection between the evolution of vocal articulation and the evolution of the hand. Vocal utterances may have served to synchronize actions of the group under aquatic circumstances. It is imaginable that discrete segments were invented in the phonetic system at a time where it was restricted to on-line situations.

---

[4] The aquatic hypothesis was first raised by Hardy (1960) and subsequently favored by Elaine Morgan in many writings (among others, Morgan 1990, 1997), which were heavily disputed as 'pseudoscience' (e.g., Krywaniuk 1996). Only a few authors (e.g., Verhaegen 1988) considered this hypothesis seriously. However, independent evidence from geological, hydrological and biochemical research suggests that some of the basic assumptions are justified. Particular interesting are the findings by Crawford et al. (1999) that the cerebral expansion could only be supported by marine food.

One of the biggest puzzles of language evolution in my view is the fact that at some time the two independent developments, that of the hand and that of the vocal apparatus, met - maybe in situations where individuals, when they were engaged in tool-making, had to communicate with each other. The latest stage in which this integration has taken place must be early in the age of homo.

The processing of temporally adjacent phonetic segments allows some combinatorics of CV structures. The off-line processing of gestures allows the emergence of concepts such as 'temporally subsequent' and 'causally determined' by means of iconicity, hence a combinatorics of meaningful units. As soon as vocal gestures were combined with manual gestures, some even more effective linguistic system emerged. One possible factor is that each system contributed important characteristics, symmetry on the side of manual gestures, and discreteness on the side of vocalic gestures. In the end, the vocal system took preference, possibly because it was advantageous for two reasons: vocal utterances are not bound to face-to-face situations, and they can be performed even when the hands are used for other purposes.

A more concrete scenario could have been as follows. In the gestural system, manipulation scenes were imitated and became stepwise more complex along the following lines. (i) *Off-line processing of possible scenes*: A specialized lexicon for scenes ('propositions') and scenic factors emerged. Participants such as I, you, and others were introduced indexically, such that a first distinction between reference and predication became possible. (ii) *Increasing variation*: The gestural articulation of a complex scene involved juxtaposition, modification, and the expression of causal chains, to a large extent iconically. (iii) *Start of modularity*: lexicon vs. composition. Elements of the scenic lexicon were decomposed, and the resulting units again combined in various ways. Thus, expressivity and precision could be improved.

Once vocal utterances were associated with manual gestures, the system could become even more expressive and flexible.[5] Besides their function as attention and arrangement signals, vocalic gestures accompanying

---

[5] Given the organization of brain areas it is conceivable that first facial gestures accompanied the manual ones, and only later vocalic gestures were associated.

the manual ones could support reference to absent participants and modify gestural predication. Finally, vocalic gestures could be detached from the gestures they were associated with, which opened a new potential.

Vocalic utterances may first have accompanied gestures that largely functioned iconically, but as soon they were detached from the gestures, they could take over the task symbolically. While they still represented the same concept as the gestures, the relationship between the vocalic utterance and the concept became arbitrary.[6] Lexical elements represented vocally could be used in situations other than face-to-face. Modifications by manual gestures could be replaced by much more precise vocal gestures. Even indexical gestures could be accompanied by corresponding vocal gestures, thus getting a place within symbolic organization. Types of vocal gestures were distinguished and combined. Complex lexical entries were divided into such parts that contribute reference and others that contribute predication. Probably at that stage, several subtypes of lexical entries could be distinguished on semantic grounds.

Given this background we can try to characterize in what respect the emergence of protolanguage, identifying the linguistic potential of modern homo sapiens at the time 100,000 years ago, marked an important progress in the evolution of language.

5. The start of human language: protolanguage

The final step towards human language, so I would like to argue, was the invention of the noun-verb distinction in the lexicon.[7] Prototypically, the verb is the bearer of a predication in time, and the noun is an attribute for possible referents in space. Such a partition into categories can be seen only as the product of a structural generalization based on already existing

---

[6] This can answer the question raised by Hurford (2003), namely how the central link between meanings and sounds was established. The sounds replaced a gesture for which this link was not arbitrary.

[7] A similar position is held by Heine and Kuteva (2002). From a different background, they argue that at the earliest conceivable stage, human language might have possessed only two types of linguistic entities, nouns and verbs.

semantic classifications of various kinds.[8] It might have taken many steps to reach such a result.

The noun-verb distinction has enormous structural impacts on possible grammars.

(i) It partitions the lexicon of nonpropositional units into two subsets, thereby assigning the members of the subsets certain combinatorial structural properties. It thus transforms semantic properties into combinatorial ones, and it opens the way of further subpartitions in the lexicon in terms of combinatorial possibilities.

(ii) It introduces clause-internal structure. Besides the possibility of concatenating simple 'propositions', the propositional expressions can now be decomposed, and a clause-internal structure can be established by the combination of nouns and verbs. Both nouns and verbs have referential *and* predicative properties, so that the former distinction between indexical (referential) and iconic (predicative) signs can be made more sophisticated.

(iii) Being a structural generalization, the noun-verb distinction becomes subject to formal operations that convert nouns into verbs, and verbs into nouns. This opens the way for clause-internal recursion. Two categories that on principle can be converted into each other form a recursive system: a nominalized verb can saturate the argument requirement of a verb, while at the same time it requires arguments of its own. Such a

---

[8] One piece of evidence that existing categories have a structural impact on the establishment of grammatical rules comes from a very stringent experiment by Indefrey (2002). In a cross-sectional study he investigated the acquisition of a minor inflectional rule of German which requires that all masculine nouns ending in Schwa follow the so-called weak declension, ending with *–n* in all singular forms (N.sg *der Hase* – G/D/A.sg *des/dem/den Hasen* 'hare'). Indefrey found that the older the children were the more likely they were to reanalyze the set of forms acquired by root learning to a productive rule with which they then could also master new words, but even some adults did not perform this step. The relevant feature of the rule was gender, a categorical feature, rather than animacy, a semantic feature, contrary to what has been claimed in the literature. In a subsequent simulation study Indefrey found a substantial difference between the human learner and a connectionist network based on frequency: the latter was unable to simulate the rule because it was not able to override the statistical effects of the feature 'animacy'. I believe that structural sensitivity of this kind is the major factor that creates grammars, and that its introduction marks the last crucial step towards human language.

situation also allows for the creation of attitude verbs, i.e., verbs that specify mental attitudes or speech acts with respect to elementary scenes ('propositions').

The combinatorial system of protolanguage should not be confused with the grammar of any language known today. Proto-morphosyntax was probably a much simpler system than that exemplified by contemporary languages, and in many respects rather vague and context-dependent.

From what we know from today's languages, it can reasonably be deduced that arguments could often remain implicit, and only the most salient ones were realized. Arguments could be realized by pronominal markers relating to person, number or semantic class, being attached to the verb or noun, and full nouns could be added as adjuncts. A complex predication might have been expressed by the juxtaposition of two verbs ('verbal compounding'), and possibly also by a juxtaposition interrupted by a common argument ('serial verb construction').

On this basis it is conceivable how later refinements might have been developed: One of the verbs was categorized as a relational adverb (or a pre-/postposition), or it was categorized as a voice category (such as causative, or applicative), or it was augmented by temporal-aspectual information, and thus shifted to a finite verb.

Protolanguage must have included also some reference tracking device for clauses that are juxtaposed or modified. Here, again, several scenarios are conceivable: One of the arguments was categorized as the designated argument ('grammatical subject'), which allowed, among others, to establish control verbs and relative clauses, or a marker was introduced for the option of a shared argument ('same subject'), for the suppression of an argument ('passive, antipassive'), or for the addition of a further argument ('4th person, obviative person, oblique').

Such a concept of 'protolanguage' seems to be rich enough to constitute the 'mother' of all languages that developed from that state within specific language traditions or families.

As is well-known, some contemporary languages belong to the 'rich morphology' type, especially the polysynthetic languages of America and New Guinea. On the other hand, there exist isolating languages of a purely

syntactic type. The majority of languages shows some mixture of morphology (regulating word-internal structure) and syntax.

There is reason to believe that the morphology-syntax distinction itself was not part of protolanguage. No language that adapts a particular effective procedure in the course of its development will loose it later, unless it develops structural generalizations that serve the same function. If we can show that syntax proper is more effective and more flexible than morphology, there would be no motivation for a language to loose this kind of syntax and to redevelop to a polysynthetic language.

Therefore, I assume that syntax proper, in contrast to morphology, was not part of protolanguage, but evolved in the age of colonization (period III above) as a historical innovation in some human populations. Such an evolution was facilitated by universal factors not necessarily belonging to a specific linguistic capacity, such as 'Express topic first', 'Express scope by position', 'Move only along the shortest way', 'Preserve a given structure by symmetry'. In general, I assume that syntax proper has clear advantages for a group with increasing size or an increasing number of external contacts.

6. The development of the morphology-syntax distinction

If we looks at what is possible in morphology as opposed to syntax, we can see that morphology is less efficient than syntax, both in terms of flexibility and expressivity.

(i)   Morphological structures usually exhibit a strict ordering of morphemes with respect to the lexical core, while syntactic structures often allow positional alternatives, and the first or last constituent can be designated.

(ii) There is no internal agreement or binding between constituents in a morphological structure, while both are possible between constituents of a syntactic structure.[9]

(iii) Morphological structures usually do not exhibit scope ambiguities, while syntactic structures often do.

---

[9]   Morphological reflexives reduce the number of arguments by semantic identification; they do not need a morphological antecedent.

(iv) Morphological structures usually do not include topic and focus marking, while syntactic structures often do.

(v) Word forms are often cross-classified in a paradigm structure, which is subject to memorization and can suffer from infelicities such as syncretism, gaps and other irregularities, while syntactic structures show hierarchical, recursive patterns.

(vi) Morphological structures usually exhibit more lexicalized irregularities than syntactic structures.

The differences outlined in (i) to (iv) can be captured by one single factor, i.e., the possibility of movement (or positional alternatives). Movement is only found in syntax.

From the view of flexible, and at the same time effective computation, syntactic structures clearly outrank morphological structures, and the question arises why there is morphology at all. One advantage of morphological structures is that they are more readily processible if they are lexicalized.

In a small population (of, say, around 300 people) it is an advantage to use stored items that are commonly known because they are the faster processible the more frequent they are, while in a larger population or a population with many contacts there is a need for marking topic and focus, and the use of stored items is less felicitous, since people do not share the same set of items.

As pointed out, many advantages of syntax result from its positional options. One factor that triggers syntactic movement is grammatical topicalization, which I think was not present in protolanguage. What was marked there rather was the argument role in focus. In my typological studies I found various examples where argument linking devices were reanalyzed as topic constructions, but never the reverse.

A few words of caution are necessary. The reader should not get the impression that I consider protolanguage to be a simple morphological system in modern terms, a system to which the structure-dependent constraints usually associated with Universal Grammar (UG) are added later.

First, I do not consider protolanguage to be morphological in a contemporary sense (as opposed to syntax). There was simply no difference between word structure and clausal structure. Paradigms, considered to be

one of the main domains of morphology, certainly played no important role in protolanguage; most of the inflectional categories did not exist. Paradigm structure is easily produced by cross-classification.

Second, some languages that appear to be mainly syntactic because of their isolating nature (Khoisan languages, Chinese), nevertheless display typical morphological phenomena in the sense I just outlined, such as verbal compounding and incorporation. These languages appear syntactic because they did not develop morpho-phonological alternations within their structural complexes.

Third, even if the concept of movement (or positional alternatives), which is central to syntax, emerged later than protolanguage, this does not mean that the conditions on movement are not innate to human beings. In the contrary, they must be innate because syntax has been developed, as I believe, independently in various languages. One suggestion is that these conditions, mainly reflecting locality and parallelism, were already present in the geometrical system being part of the visual system.

As already indicated, I am convinced that language evolution did not come to its end with the establishment of protolanguage, although only features of protolanguage are innate. Languages further developed in important ways by tradition.

In the linguistic literature, the concept of UG, claimed to be innate to all human beings, is loaded with many specific structural conditions found in contemporary human languages. In this respect, one should be sceptical. It suffices to assume that the search for structural rather than semantic generalizations is the most selective factor. UG is available only until maturation, it enables children to learn every human language in predictable steps, and it provides them with the faculty to judge structures not learned.

Under the developmental perspective, the content of UG might be identified with category- or structure-sensitivity inherited to a certain collection of neurons that integrate the input from several fields. What is innate to the child is the need to develop his or her neural architecture by searching all given linguistic input for items of minimal size that can be given structural generalizations in terms of categories concerning reference and predication, as well as for operations that combine these items.

Most of what we know about languages developed after the separation into language families. The general picture is determined by three factors: (i) *Rapid enrichment of the vocabulary*: An increasing collection of items put under the given structural choices gives rise for more expressivity, and allows for more explicit and less vague lexical items. (ii) *Structural sensitivity*: An inherent structural (categorical) feature is more important for categorization and computation than a purely cognitive or semantic feature; hence, there is a tendency for generalization. (iii) *Extension of the category system*: The invented lexical categories are complemented by functional categories that express certain specific dimensions of variation. Based on the categorical distinctions, the set of combinatorial templates and constructions is extended.

The later developments within language families include the following: (i) Completion of clausal grammar: introduction of an argument linking device which allows to realize *all* arguments, including the less salient ones. (ii) Introduction of syntax proper: topicalization, decomposition of complex predicates into independent clauses, argument linking by syntactic position such as SVO, case and agreement, interaction of morphology and syntax. (iii) Development of complex syntax (such as relative constructions, control structures), introduction of a grammatical subject as the primary target for argument sharing in these structures. (iv) Grammaticalization of functional categories such as complementizer, determiner, auxiliaries.

7. Conclusion

I have argued in this paper for the following three claims. (i) In the evolution of language several independent sources are integrated. In particular, symmetry and off-line processing may originate from the evolution of the hand, while discreteness and double articulation may originate from the evolution of the vocal apparatus. Both lines of evolution probably took place in the period of hominids and were integrated in the period of homo. (ii) The major step towards protolanguage, defining the common ancestor of all human languages, was the invention of the lexical categories noun and verb, by itself an important generalization with structural consequences. (iii) Syntax proper, allowing the notion of movement, was a later

development, mainly triggered by the need for topicalization and taking place independently in various ways in the several language families.

Clearly, such a stand is speculative, only based on considerations of what is plausible for theoretical and typological reasons. It offers a program to start from rather than some specific results. Taking into account also aspects of later linguistic diversification helps us to restrict our assumptions on UG, so that it becomes conceivable that UG emerged by a 'small' step compared with the rich complexity of modern grammars. In the course of time, subsequent generations produce an increasingly richer input, which enables following generations to generate more sophisticated grammars. Some features that are often thought to be universal by syntacticians need not have been present in protolanguage.

One argument I have been confronted with states that considering the development of languages implies nothing about the underlying language capacity (Wolfgang Klein, Manfred Bierwisch, p.c.). People may have possessed the language capacity (UG) for a long time before they had the chance for sufficiently rich input. I agree that this argument, as it stands, is correct. Yet, it makes it almost impossible to ever indicate the relevant step. If UG contributes anything to human fitness, it must have given an advantage at the time when it was invented. Such a view enforces us to consider the invention of UG to be a small step based on the preceding evolution of all the necessary preconditions. Therefore, I propose that UG reduces to the capacity of making structural generalizations in view of categorial distinction. Hence, the invention of the first categories (which, as I believe, were verb and noun) is a good approximation to the starting point of UG. Moreover, a simple feature such as structure-sensitivity can explain why all the individual languages indeed progressively generalize by way of tradition and contact, that is, why grammaticalization is correctly assumed to be unidirectional.

References

Carstairs, MacCarthy, Andrew. 1999. *The origins of complex language. An inquiry into the evolutionary beginnings of sentences, syllables, and truth*. Oxford University Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1998. Minimalist inquiries: the framework. *MIT Occasional Papers in Linguistics* 15, 1-56. Dept. of Linguistics, MIT.

Crawford MA, Bloom M, Broadhurst CL, Schmidt WF, Cunnane SC, Galli C, Gehbremeskel K, Linseisen F, Lloyd-Smith J, Parkington J. 1999. Evidence for the unique function of docosahexaenoic acid during the evolution of the modern hominid brain. *Lipids* 34 Supplement: 39-47.

Crawford, Michael and Marsh, David. 1989. *The driving force*. Heinemann.

Gallese, V., Fadiga, L., Fogassi, L. and Rizzolatti, G. 1996. Action recognition in the premotor cortex. *Brain* 119, 593-609.

Hardy, Alister 1960. Was man more aquatic in the past? *The New Scientist*, March 17: 642-645.

Heine, Bernd and Kuteva, Tania. 2002. On the evolution of grammatical forms. In Wray, Alison (ed.) *The transition to language*, 376-397. Oxford University Press

Hurford, James R. 2003. Language beyond our grasp: what mirror neurons can, and cannot, do for language evolution. In Iller, K., Griebel, U. and Plunkett, K. (eds.) *The evolution of communications systems: A comparative approach*. MIT Press.

Indefrey, Peter. 2002. *Listen und Regeln. Erwerb und Repräsentation der schwachen Substantivdeklination des Deutschen*. Ph.D dissertation, University of Düsseldorf

Krywaniuk, Andrew. 1996. Adventures in pseudoscience: The aquatic ape theory. *Science* 19, Number 12, Imprint.

Liberman, A.M. 1957. Some results of research on speech perception. *Journal of the Acoustical Society of America* 29, 117-123.

Morgan, Elaine. 1990. *The scars of evolution: what our bodies tells us about human origin*. Souvenir Press. [Penguin 1991]

Morgan, Elaine. 1997. *The aquatic ape hypothesis*. Souvenir Press.

Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. 1996. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3, 131-141.

Rizzolatti, Giacomo and Arbib, Michael A. 1998. Language within our grasp. *Trends in Neuroscience* 21, 188-194.

Van Valin, Robert D. Jr. To appear. Minimalism and explanation. In M. Polinsky and J. Moore (eds.) *Explanation in Linguistics*. Proceedings of a conference held at UCSD. Stanford: CSLI

Verhaegen, Marc J. M. 1988. Aquatic ape theory and speech origins: a hypothesis. *Speculations in Science and Technology* 11, 165-171.