LANGUAGE EVOLUTION AS A CASCADE OF BEHAVIORIAL BIFURCATIONS

WOLFGANG WILDGEN
University of Bremen
wildgen@uni-bremen.de

Resumen

La evolución de la lengua humana probablemente pasó por distintos períodos decisivos o barreras. Nuestro objetivo es mostrar que tienen la forma de una cascada de bifurcaciones, o elecciones binarias. En cada paso, constelaciones de pequeñas causas decidieron del camino a seguir. Esto explica la improbabilitad de un fenómeno como el lenguaje humano en el reino animal (lo cual ha sido el principal argumento de los creacionistas cristianos). La jerarquía y la sucesión de las "decisiones" (dependientes de los principios de la evolución de Darwin) nos proporcionan una serie de indicios sobre la arquitectura del lenguaje humano, en concreto la aparición de la valencia verbal (estructura argumental, construcciones), la polaridad entre el sujeto y el predicado de la frase y otros rasgos sintácticos pueden ser explicados como resultado de estas etapas de la evolución. Asumimos, por último, que la evolución de esquemas básicos de acción (acción manual), como el esquema de EMPUÑAR) tuvieron una importancia fundamental para el desarollo conceptual del lenguaje humano.

PALABRAS CLAVE: evolución, bifurcación binaria, valencia verbal, esq*uema de acción.

Abstract

The evolution of human language probably passed several gateways or decisive periods. Our aim is to show that they have the form of a cascade of bifurcations, i.e. of binary choices. At each step small causal constellations may have decided on which way to go. This explains the overall improbability of a phenomenon like human language in the animal kingdom (which was and is the major argument of creationists). The hierarchy and temporal sequence of "decisions" (Darwinian principles were the deciders) gives us a hint at the architecture of human language. In particular the emergence of verbal valence (argument structure, constructions), of subject-predicate constellations and syntactic devices can be explained as a result of these evolutionary steps. We presume that throughout the evolution basic schemata of (manual) action, like the GRASP-schema were of central importance for the conceptual unfolding of human language.

KEY WORDS: evolution, bifurcation binary, verbal valence, verbal scheme

* This is an elaborated and augmented versiono f a lecture given in Alicante. I thank Prof. Cifuentes Honrubia for his invitation.

1. Introduction

The most astonishing feature of human language and human civilization is that they seem to be (or really are) so dramatically different from anything else in the animal kingdom. Humans are either monsters (or aliens) in relation to other living beings[1] or they are gods (or at least simulacra of God: this is the classical Christian picture which dominated the history of the question until the 18th century). In the perspective of post-Darwinian biology this view is an illusion; everything in the universe, including humans is subjected to general laws of nature and the universal principles of Darwinian evolutionary biology are called mutation (later elaborated in genetics and molecular biology) and selection (at different levels from the molecule to societies). Genetic drift in small populations and morphogenetic fields may complement these fundamental principles. The question arose: What singular event or series of singular events was able to give human evolution a direction and an itinerary which separated it from the predominant success stories in the animal kingdom? How can we explain the huge behavioral differences between a group of chimpanzees in the West African rain forest and the urban societies in some cosmopolite town like New York, Tokyo or Paris?

The line pursued in this article is to postulate a cascade of singularities and decisive moves/changes/catastrophes[2] which were responsible for the dramatic difference between the evolutionary itinerary of humans and other hominids (or mammals).

2. The origin of "Homo loquens": some hypotheses

Scientific research in the last decennia has contributed many facts on the origin of man as "Homos loquens", which make it possible to answer the question: Are there sudden transitions, which created frontier-lines between species which talk, build arte-

[1] A short version of the original lecture in Alicante 2003 has been published in Catalan with the unauthorized prefix ¨Hopefully it is a tale about monsters" ("Una esperançola historia de monstres? cf. Wildgen, 2003). Probably Prof. Garcia Lopez who had invited the contribution took the sympathetic view of our co-hominids who may see humans as monsters of intelligence and technological skill.

[2] In the original meaning in Greek the verb: καταστρέφειν (katastréphein) means *to overturn, to turn around, to reverse.* Catastrophe theory is linked to the work of the mathematicians René Thom and Christopher Zeeman and applies the classification theorem in singularity theory, which gives a list of elementary catastrophes. Applications to linguistics were elaborated in Wildgen (1981, 1982 and 1994). Later work on "semiophysics" is commented in the volumes edited by: Wildgen and Brandt (2010) and Wildgen and Plümacher (2009: in German).

facts, produce art, live in complex societies and those who don't? The transition to *organic* language capacity concerns:

(a) the speech producing organs (control of pulmonary pressure, vocal cords, shape of the larynx, shape of the mouth, lips, teeth);

(b) the auditory system (mainly the inner ear, the cochlea, and the neighboring sense of equilibrium);

(c) the brain (mainly the cortex but also the brain stem) with their capacities of advanced perception (categorization), memory, and self-monitoring (cf. the frontal lobes) which seem to distinguish Homo sapiens from Homo neanderthalensis and last not least, the language centers (Broca and Wernicke-centers).

Dramatic changes in the geometry of the crane, the position of the larynx, the geometry of the inner ear, the absolute and relative volume of the brain (neo-cortex) took place, but these evolutions were so slow, the magnitudes (e.g., of brain volume) so overlapping between species, that a catastrophe like the sudden "creation" of language, of artefacts and art seems impossible. All these continuous changes just contributed to a predisposition for language (cf. Wildgen, 2004: chap. 2). This continuity corresponds to basic principles of Darwinian evolution that rest on the continuous effect of selective features in the environment (selection) and minimal chance deviations in the genetic equipment (mutation). The dramatic behavioral and "cultural" differences between higher primates (e.g., chimpanzees, bonobos, gorillas and orang-utans) and humans ask for other mechanisms explaining rather sudden changes (this means many millennia instead of millions of years). A sudden shift (catastrophe) in a continuum may be caused if a line of equilibrium is crossed (e.g., the model of a balance).[3] This phenomenon is called *symmetry breaking*. Symmetry presupposes at least bimodality, i.e., two modes coexist and there is a path of smooth transition between both in a base space.

[3] Other catastrophes involve the creation/destruction of an attractor, the appearance of bimodality, trimodality, etc.; cf. Wildgen, 1982, 1994, for a description of catastrophe theory and its application to linguistics.

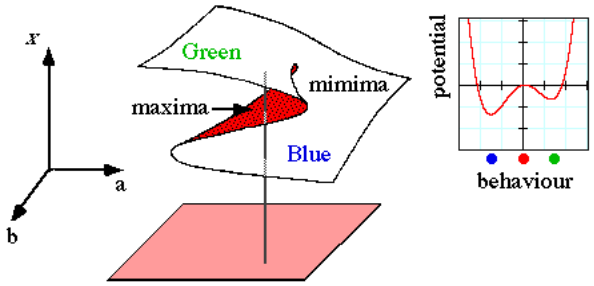

Figure 1: Catastrophic transition between two attractors (Green vs. Blue) given a smooth surface.

In the empirical application the first step has to consider the existence (or emergence) of the alternatives, i.e. in the above case the two sheets of attractors (stable states) in a continuum (see the unbroken surface in the background; the whole sheet is without limits, only the environment of the instability point is shown in Figue 1). This bifurcation may happen because under different circumstances two types of behavior came up, were in concurrence and finally entered a zone of conflict where one alternative could be chosen (alternatively a situation of ambiguity could exist). The first step concerns the circumstances of change (choice). In human evolution (probably up to now) these choices are not made deliberately by an agent or a group of agents; external forces of nature or internal factors of coherence "decide". In the evolution of man quick climatic changes (in Africa and in other regions of the world) have driven secondary changes in the fauna and flora, they have separated subpopulations for long periods (in East Africa, the cradle of humans, huge lakes with large islands and rivers through the rift valley grew and disappeared again), have driven them to small areas of escape (quasi in a series of quasi-deluve scenarios, followed by extreme drought), where they shrinked under extreme conditions of selection. Thus in the critical period before the appearance of Homo sapiens only a small remaining population was able to survive (e.g. in caves on the shore of the Indian Ocean in Southern Africa; cf. Marean, 2010); before the Homo neanderthalensis died out, this species was driven to the shores of Spain and Italy several times. When Homo sapiens lived in Europe (ca. after 40.000 BP), he suffered under the same conditions, but he was able to survive and to expand again after the end of the last ice-age (ca. 12.000 y BP). The driving mechanism is therefore one of rather quick and dramatic climatic changes. The complex effects on fauna and flora (nutrition and housing) and possibly the enforced mixture with refugees of different provenience in the remaining places triggered the transition which was accompanied by a dramatic reduction of the population (small population enable a genetic drift and allow for quicker changes). In a formal sense one may consider these ecological variations as a heating

up of the process which allowed the transition to new states in the behavioral dynamic system. As soon as the situation cooled down, the new state was stabilized and became the basis of a growing population.[4] This probably happened again and again, whereby humanity (or at least regional groups) were driven towards extinction. Cf. for different scenarios for the evolution of human language Wildgen (2004: chapter 2) and Wildgen (2010b). I will in the following consider the single bifurcations one after the other.

2.1. Towards a dominance of the phonic channel

In higher apes (e.g., chimpanzees) one finds two means of social communication without instrumental or referential functions: grooming (lousing, caring for one another) and social calls. The (manual) mode of grooming dominates and it consumes about 20% of the time budget. The critical transition (from the common ancestors of chimpanzees and men to australopithecines) was probably a dominance shift due to larger groups (cf. Dunbar, 1996) and to stronger social connectivity in groups.[5] Moreover the loss of fur which was probably due to an adaptation to long range running during day time under the sun, reduced the functionality of grooming. This development began already some 4 my BP. Phonic communication was more time-economic and phonic contact with socio-emotional content largely replaced bodily contact. This shift of dominance, which had to cross a point of symmetry, is described in Figure 2.

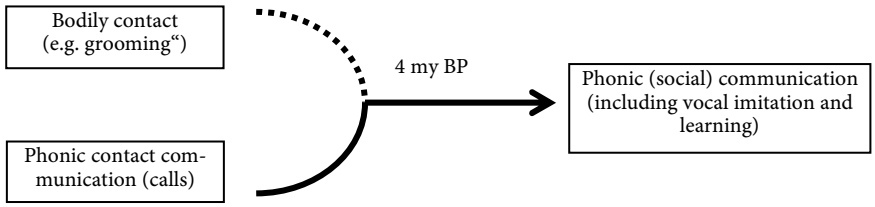

Figure 2: The dominance shift made social phonic communication the central technique of social control and management.

[4] In "harmony theory" the heating and freezing of a dynamical system are considered as generalized features in a decision process (cf. Smolensky, 1986: 236f.)

[5] The absolute size of groups is perhaps not the key (the controversy about the size of hominid groups could impair Dunbar's theses). Thus very large groups of animals may move together and thus fulfil the condition of a "large group"; what counts is the degree of interdependence and reliance on one another. The network should be larger than that of couples, relatives or close "friends" in chimpanzees.

2.2. The expansion of referential signs in the phonic mode

Studies of the behavior of apes in the wilderness have shown that some species have a simple system of calls with referential function. The calls may allow the other apes to distinguish a danger stemming from animals like eagles (attacking from above), carnivores (e.g., lions attacking on the ground), and snakes (creeping in the trees).[6] Parallel to this simple system, bodily motions, gestures and gaze-directions give communicative cues, which allow for a spatial interpretation and can thus be elaborated to a "language" of body postures and gestures. The rich system of gestural signals was, therefore, functionally parallel to a poorer system of calls. If we take the gestures of the hand, it is clear that as soon as hands are fully occupied with other functions like carrying tools and objects, or if communication takes place in the dark or at a distance (with obstacles between), the gestural "language" is ineffective. As such circumstances probably prevailed in the ecology of the australopithecines that lived in the savannah and as the ear had to increase its capacity of discrimination due to the permanent danger of carnivores in the environment,[7] the bimodality between reference by gesture and reference by phonic articulation shifted towards the latter. Figure 3 illustrates this transition.

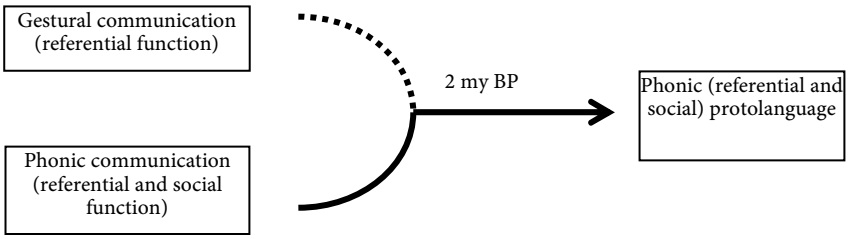

Figure 3: The dominance shift from gestural to phonic in referential communication.

The result of this functional evolution lies midway between biological and cultural evolution insofar as the repertoire of manual behavior in grooming and gestures and

[6] Hauser, Chomsky and Fitch (2010: 31) argue that the calls lack intentionality and that the animals (vervet and rhesus monkeys) are only able to extract information from the acoustic signals. As we cannot interview monkeys, the question of intentionality must remain open. It would, however, be sufficient if a cooperative practice was genetically prepared and quickly developed in these species. Brain scanning results show that learning does not presuppose consciousness, thus even very low levels of consciousness in hominids would not preclude social learning of signs and their meaning; cf. Duss and Henke (2011).

[7] Cf. Calvin and Bickerton (2000: 111).

the repertoire of social and referential calls has to be imitated and learned. However, it could together with the evolving tool-use have fastened the growth of the brain, i.e., the selection for bigger brains, which is already a general evolutionary strategy of all species of the lineage. Relatively to higher apes, the resulting protolanguage probably contained a rather large "lexicon" of social and referential calls (e.g. about 30 to 50 patterns), with different types of evaluative modulations (social calls) and categorical distinctions (referential calls). This led to a referentially motivated sign-architecture. The fact that children very often show a vocabulary sprint may point to an innate capacity which given the proper social environment (e.g. a mother who answers questions and gives names) expands and enriches without difficulties the primitive lexicon after the first year by ten to twenty items a day, a rate other hominids would not achieve in their lifes.

2.3. The creation of cultures

It is clear that the capacity of imitation of phonic material, of quick and stable memory entries and corresponding search procedures and semantic net-building faculties presuppose a basically better organized brain (cortex, centers of auditory and visual detection) and enough space (synaptic connectivity) to build a memory which associates phonic patterns with other (visual, olfactory etc.) cues. This leads us to a third bifurcation. It concerns imitation and learning in the case of motor behavior and symbol use. The base-line is defined by the presence of mirror neurons in higher primates and their capability of quick motor learning and motor control (cf. Rizzolatti and Arbib, 1998) and the rise of a theory of mind in late hominid (e.g. chimpanzees). We assume that this base-line was already reached before the lineages of chimpanzees and humans diverged (6-7 my BP). The bifurcation occurs between simple emulation and first "cultures" of tool use in chimpanzees (cf. Boesch and Tomasello, 1998) which is still linked to immediate success (reward) and a more general strategy of imitation and learning, which allows the acquisition of know-how and even symbolic knowledge from others without any immediate pragmatic support (or "grounding"; cf. Cangelosi, Greco and Harnad 2002). In computer simulations this distinction is called "toil" versus "theft" strategy. In the case of symbolic learning a label is either learned in relation to its referent via trial and error or "stolen" from the symbolic behavior of others (the semantics are filled in later). Human infants are systematically "symbolic thieves" in the sense of these experiments. Human cultures accumulate information which can be transmitted without being applied and tested extensively by every single user. Dawkins called this information "memes"; cf. Blackmore (1999).

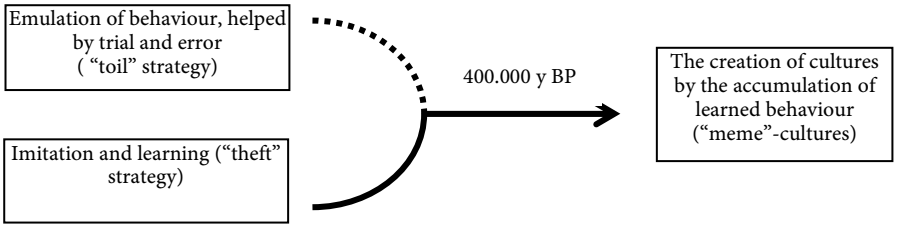

Figure 4: The bifurcation which separates simple cultures based on emulation and "meme"-cultures.

The proto-species Homo erectus lived between 2my BP and 15.000 y BP (Homo florensis recently discovered) and developed a large variety of subtypes which were still able to interbreed if they met. In the late subspecies which had developed on the basis of the Homo heidelbergensis (or Homo antecessor in Spain), a phonic protolanguage that integrated social and referential communication and was able to receive and transmit the accumulated cultural knowledge must have reached a first plateau, which was sufficient for the survival of this new species[8] and allowed its migration into Eurasia and its diffusion in Africa (Homo erectus and Homo ergaster).

Figure 5: Skull of Homo heidelbergensis (cranial volume 1100-1400; ca 500.000 y BP).

Although we know nothing about the communication in these populations, the principle of continuity (on which Darwin founded his theses) makes us postulate a *protolanguage*, i.e. a way of phonic communication which prefigures the language typical for our species. It could correspond to the one and two word utterances of children in their early second year of language acquisition (this implies a mild hypothesis of biogenetic recapitulation) or to very rudimentary pidgins (this implies the capacity of communicative regression).

[8] Recently artefacts also produced by Homo sapiens were found on the Arabian peninsula (but no bones). They were dated 127.000 y BP respectively 95.000 y BP. The relation to other out-of-Africa populations is still unclear. It seems possible that different sub-populations had left Africa in different periods, some of which were isolated and died out due to climatic changes. Cf. Petraglia (2011).

3. Prelinguistic schematizations at the level of Homo erectus

The technical evolution since the pebble culture (2,5 - 1,7 my BP in the Oldowan choppers) had brought about a degree of manual skill and of collective tool production which presupposed complex neural programs for the planning and strategic-intentional processing of tools and materials treated with these tools (from chase and butchering to the shaping of clothes and huts/houses). In this evolution of technical intelligence and its social management, rather complex dynamic and causal schemata had to be cognitively mastered. Figure 6 visualizes several such schemata (cf. for the foundation in catastrophe theory, Wildgen, 1982, 1994).

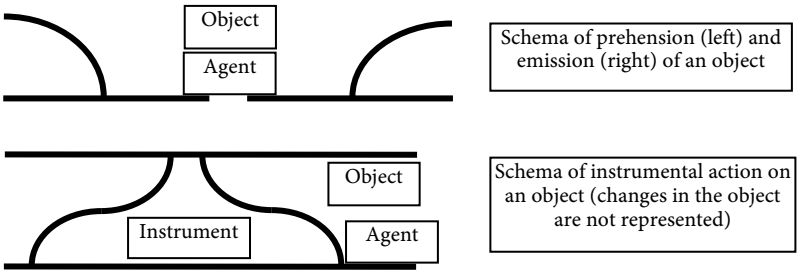

Figure 6: Basic schemata of action cognized in a society with "tool-culture".

The precise shaping of a tool out of a stone kernel asks for the strategically planed excision of material. The excised piece is mostly an intermediate stage of the planned tool, such that after a long series of strokes the final shape is reached. Whereas in Figure 6 one has maximally two or three centers (valence 2 or 3), the excision schema asks for four centers (attractors) of the dynamics, Figure 7 as shows.

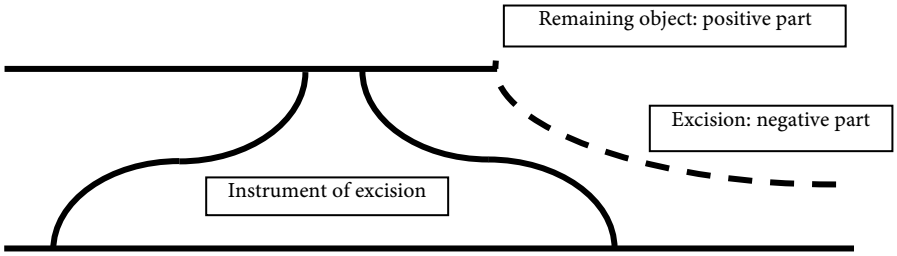

Figure 7: The excision schema (cf. Thom, 1983: 206).

In Figure 8 the transition from simple excision in the case of Oldowan choppers to the manufacturing of the refined handaxe by Homo heidelbergensis is shown. The latter asks for at last 20 to 30 subprocesses as experimental archeologists could demonstrate.

Figure 8: Oldowan choppers from Harar, Ethiopia, ca. 1,7 my and refined hand axe used by Homo heidelbergensis (Acheulian industry).

Another important cultural schema was that of exchange (materials for tools, rare objects for decoration and art, women for extra-tribal marriage). The basic exchange schema consisted of mutual giving and receiving and its formal schema corresponds to that in Figure 6 above, where the role of the instrument is replaced by the exchanged objects given and taken by two agents (the permanent states on top and at the bottom). The question of a fourth force, a value-measure as tertium comparationis of exchange may be asked in this context. This means, if object A is given and object B is received in exchange, do they have equivalent values? This problem is actually solved via the monetary value of objects but originally one has to presuppose some evaluative scale based on cultural conventions. By its conventionality it functions like a code. i.e. a symbolic form not visible but responsible for successful exchange. In catastrophe theory one can use the two-dimensional singularity derived from higher catastrophes (cf. Wildgen, 1982: 62-78 for the mathematical derivation of this schema and Idem, 1994: 129-134 for an advanced interpretation). With the cultural introduction of specific exchange goods (shells, gold, silver, then paper money and bank transfers) this symbolic medium first became touchable (e.g. coins), later it was successively replaced by a procedural technique in a complicated system of banking. In any case the quasi-physical process underlying the use of instruments or the simple exchange is completed by the emergence of a symbolic code which is functionally linked to the quasi-physical processes (enables and controls them). Now, if the proto-cognition and proto-technique of Homo erectus/ergaster and his/her descendants Homo heidelbergensis, Homo neanderthalensis and Homo sapiens had to cope with schemata of this complexity, it had also to organize this complexity in a phonic structure for instruction, report and narrative as a precondition for cultural accumulation of the corresponding competences. A major difference

between Homo neanderthalensis and Homo sapiens concerns the importance of social life (mutual aid and altruism beyond kinship) and as a consequence long range cultural and biological exchange. This may be one of the reasons why Homo sapiens survived the severe ice age period between 30.000 and 13.000 and Homo neanderthalensis did not.[9]

4. Events, actions and time in a protolanguage

Another source-domain of human language is linked to spatial cognition, which may be considered as a kind of "fossil" of protolanguage in modern languages. As Bickerton (2010) suggests, there are two socially organized animals with communicative skills related to space beyond the situation in which interaction occurs: bees and ants, which "transfer information to nest-mates about food sources" (cf. ibidem: 206f. and Wildgen, 2004: 45f.). In the case of humans around 2 my BP, they had to recruit as many companions and as quickly as possible in order to exploit the carcass of a megafauna in competition with other scavengers. Having even simple iconic and indexical signaling techniques they could solve this problem and survive. The selective pressure on this social behavior made spatial orientation by language a central feature in the evolution of language (or some protolangague).The cognitive organization of the ambient world can be deictically structured as *here* (I), *there* (you, objects), and by different steps of *distance: near, far off,* and *midway.* Many languages, e.g., the Hopi-language analyzed by Whorf and later reassessed by Malotki (1979) show a very rich system of categories realized by suffixes (of case) and by pronouns. The basic dimensions of such a system of orientation are:

- The human body (the self) with head (above), feet (below), hands (right and left), face (front), and back (behind); major ego-centric distances: proximate to the trunk (face), in the reach of the hands (of grasping), at a speaking or a sight distance.

[9] There was probably some mixture after 60.000 BC, as 1% to 4% of Neanderthal and Sapiens genomes in out-of Africa populations match closely (beyond the common heritage shared with chimpanzees); cf. Green et alii (2010). The maturation of the brain and the corresponding periods before adulthood were shorter in Homo neanderthalensis, such that only less complex cultural skills could be learned; cf.Gunz et alii (2010). Moreover Homo sapiens may have acquired in the critical period, when small population survived in South Africa in caves (e.g. those near the Pinnacle point, cf. Marean, 2010) the technique of fishing and a diet less dependent on meat. The degree of altruism and social cognition in humans may also have differentiated both populations; its genetic conditions are analyzed in Reuter et alii (2010).

- The landscape and the sky with its orientation: sunrise, sunset, sun (moon, stars), directions of winds (if stable), etc.; distances are those of the skyline, neighboring locations, centre and periphery of the village, etc.

The spatial categorizations, which in the second case depend on the ambient (cultural) ecology, can be translated into gestures and phonetically uttered distinctions. In many languages pre- and postposition, locative cases, verbal affixes contain a closed set of such distinctions. Together with the basic action schemata they make up the (semantic) core of phrasal and sentential syntax that will evolve based on these foundations.

One can distinguish two sub-aspects: processes in space, such as spatial orientation and navigation, and temporal classifications and rhythmical patterns.

4.1. Spatial schemata

The representation of *space* has to do with frontiers (their transition) and perspectives. A first perspective is centrifugal, i.e., starting from the self and its basic bodily motions an *experienced* three dimensional space is cognized: 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 with feet and arms defines the immediate space, where objects may be approached, reached and manipulated. The intermediate space depends on man's ecology; it can be the housing (the cave, abri) or the village; the distal space contains roughly all possible itineraries (of hunting/gathering). The second perspective is centripetal, i.e., the self is seen as the place of effects triggered by external causes. The sky, the horizon (typical points where the sun sets or rises), the favored direction of winds, the ridge of mountains may be the external locus of orientation for the self, who is at the center of a force field or gradient implicit in these delimitations. Many myths and religions refer to this extreme locus of orientation as they interpret the fate of humans as standing under the control of such distant (and often invisible) forces.[10] In Figure 9 the topology and dynamics of such a cognized space are illustrated.

[10] The dependence of religion on models of the universe became evident in the processes of the catholic church against Bruno and Galilee. Although astral deities (the sun, the moon, Venus, etc.) had no explicit function in the Bible and the New Testament, a change in the cosmological world view was understood as a aggression against the foundations of the Christian faith. This shows the continuity of space dependent myths (cf. Wildgen 1998: for the case of Bruno).

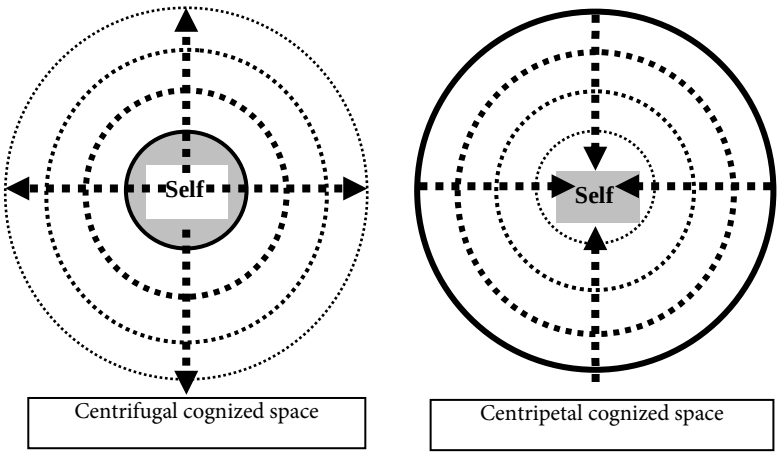

Figure 9: Force fields of centrifugal versus centripetal orientation.

The cognizing of such schemata for orientation may only show up in behavior (as it does in many animals), it may be gestured or it can be deictically organized in a phonic language (cf. Levinson, 2001: 317 ff.). For the Homo erectus the cognized space seems clear. The inner space is defined by the use of hands and instruments, the medium space by the choice or construction of dwelling-places (to which the group could return). The centrifugal organization is involved in long-range excursions and migration. As the orientation system cannot be genetically coded it has to be learnt, adapted to changing contexts and socially shared. Language is one possible solution to this problem, be it gestural (behavioral) or phonic and as humans have chosen the path of phonation it is plausible that our ancestors began to proceed further into this direction (cf. Figure 2).

4.2. Temporal schemata

The representation of *time* is rooted in the classification of multimodal sensory inputs using specific temporal rhythms (clocks). Pöppel (1997, 2010) proposed two temporal windows for multimodal integration:

- The window of 30 msec. Only after a stability of 30 msec does an event become an object of (multimodal) perception; it can be classified, labeled, compared, i.e., further processed.

- The window of 3 sec. A sequence of events can be understood as a structure. In this window the smaller units (30 msec) are correlated as: before – after, cause – effect, etc. This is the point where a notion of structured temporality is born.

In a similar vein (but without reference to neurobiology) Bickerton (2000: 275) mentions the „higher level signal coherence" as a precondition for hierarchical structuring and tries to explain the "catastrophic" transition to syntax along this line. A protolanguage must categorize events and actions (by proto-verbs) and must discriminate stable entities (by proto-nouns). The question arises as to whether temporal, dynamic, quantitative, 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 discussed by Bickerton: phrase structure (X-bar-structures) and government (argument clusters). These conditions may include the use of recursive rules in the neural "machinery" (Chomsky's metaphor) retrieving and producing language, but this link is controversial; cf. Fitch (2010). I will argue in the next pages that there are intrinsic complexity barriers which could have blocked the further elaboration of a protolanguage for a long (evolutionary) time-span (cf. Wildgen 2004: Chapter 8 for a fuller treatment).

In order to have access to a complexity measure I shall introduce a model of event-schemata. One can take grasping (with the hand) as the basic scenario. The action-concept GRASP involves two stable entities: the body (the hand) and the object. Every point on the lines in Figure 7 is an attractor, i.e., the perception of a stable entity in the 30 msec window (cf. above). The whole schema should fit into the 3 sec window, e.g., in the sentence:

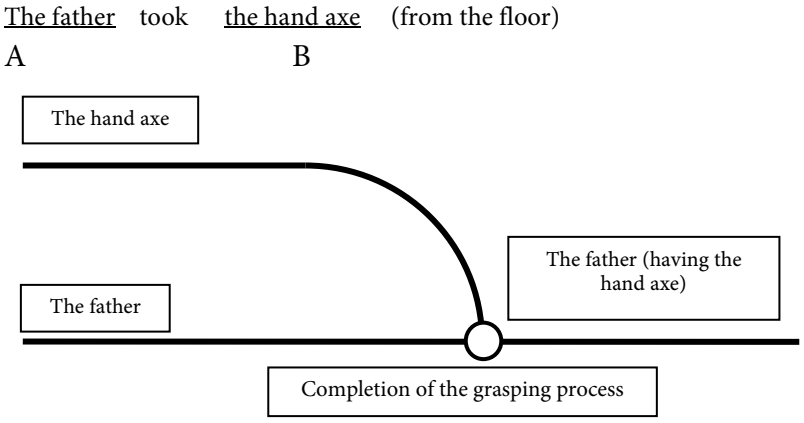

Figure 10: Process (catastrophe) schema of GRASP.

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 which was adapted to strong grasping (as places on the bones, where muscles are attached, show, cf. Piveteau, 1991: 74 f.). One may distinguish three ways of grasping:

- the force grip (e.g., of a branch)
- the precision grip (e.g., of a small tool)[11]
- the refined grip (e.g., of a needle)[12]

The refinement refers to the topology of the capture. These distinctions which have a long evolutionary history constitute a kind of manner specification in relation to the schema in Figure 7. As the distinction between several 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 can be cognized does not necessarily mean that it could be transformed into phonic signals. If we assume a frequent vocalization (inferred from the evolution of the sub-laryngeal tract; cf. Lieberman, 1989) and a steady increase of memory (due to the growth of the brain) linked to an advance in social cognition (cf. Dunbar, 1996) it becomes clear that this cognitive schema and subsequent ones are a preadaptation 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 be used as a kind of ground for iconic/metaphorical transfer to all kinds of manipulations on objects. As soon as instruments were used this schema could be iterated.

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

The schema of grasping assembles causal/enabling/intentional meaning components, which are necessarily present in the purposeful shaping of a tool and it also lays the groundwork for force-dynamics (cf. Talmy, 1988) in human language. In this development a first barrier of complexity appears. While the schemata shown in Figure 6 are

[11] 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 and Reynolds, 1993: 123). It had as consequence the "conversion of active behavior to crepuscular and diurnal phrases of activity" (*ibidem*).

[12] Cf. Piveteau, 1991: 29, who calls it: "préhension de délicatesse".

dynamically and topologically simple, the composition of such schemata needs specific restrictions to guarantee stability for such a composition.[13]

A second restriction concerns the manner component. The evolutionary old distinctions between forms of grip and manners of locomotion (related to the dynamics of the legs; e.g. go, *run, jump, climb* etc.) are topologically basic and could 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-complex (T = Time, M = Mode, A = Aspect). They are the next step which could have "evolved" after the protolanguage. The evolutionary period associated with these developments is probably linked to the final speciation of Homo sapiens between 200.000 and 100.000 y. BP. The order of emergence of grammatical features could have been:

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

I have started from the grasp schema, but there are simpler schemata. The dynamically simplest schema is that of stable existence. If we apply the 3 sec-window, any entity not changing in this window is a candidate. As the inputs of classification or labeling-reaction are not only spatio-temporal events but also qualities, one can assume the slow increase of quasi nominal/adjectival labels as soon as memory capacities enabled it and social demands asked for it. One could imagine that labels for other people, animals, plants, and artefacts were the first candidates for a growing lexicon. This development is also the natural continuation of classificatory capabilities of other mammals (even birds and fishes) and the differentiated warning calls of specific apes[14]. The cries of alarm, disturbance and food constitute a basic lexicon with reference to specific situations and they have distinctive pragmatic values, e.g., as asking for, responding to, informing about, etc.

---

[13] The elementary catastrophes on which the schemata are based are local evolutions, i.e. in the neighbourhood of a point (the singularity). In the case of composition one enters the domain of global analysis which goes beyond catastrophe theory. One has to integrate the local fields to a global map using techniques of gluing together or overlapping, as if a set of snapshot is used to build an overall image.

[14] Cf. Fischer and Hammerschmidt (2001) for a critical discussion and experiments with Barbary macaques.

5. The transition to full-fletched languages: a first sketch

The increase of the lexicon and the availability of case-frames (action – schemata) and spatial categorizations establish the source-domain, in which a very complex system of grammar could emerge by self-organization. The compositionality of human languages is the result of such a self-organization as Kirby (2000) showed by computer simulations. In this process less general (idiomatic) rules give way to more general rules through learning and in the transmission sequence between users (cf. Hurford, 2000). The basic capacity for the production and understanding of a rapid sequence of phonic events is primarily due to short- and long-term phonetic memory. Studdert-Kennedy (2000: 17) says: "Without a pre-adapted system for storing phonetic structure independently from its meaning, syntax could not have begun to evolve." The basic syntactic capacities of man are better demonstrated by morphological and phrase compositionality than by sentential syntax (word order, movements, trace, etc.). The basic capacity is "a way of systematizing existing vocabulary items and being able to create new ones", rather than some "fancy recursive generativity of syntax" (Jackendoff, 2002: 244).[15] This is also demonstrated by the extreme diversity of syntactic devices in the languages of the world and the impact of historically recent (cultural) evolutions on features like dominant word order.

The valence patterns found in verbs (basic sentences) and nouns/adjectives (basic noun phrases) could be the results of a fourth bifurcation which has split the holophrases of a protolanguage into different functional subparts.[16] As the holophrase (compare the

[15] If recursion means that an open field of possible utterances which are either referentially very global or very precise is enabled, as Fitch (2010) concludes after a thorough analysis of a set of definitions for recursion and the criteria which allow us to say that a behaviour follows an underlying recursive process, then one may accept this as a description of a question to be solved by an evolutionary theory of language.

[16] The question at what date this change showed up is part of a controversy. Some authors believe that only the rich symbolic products demonstrated by cave painters (after 35.000 y BP) can be considered as evidence for the existence of "homo symbolicus" (Cassirer). This means that only at a stage when Cro Magnon man entered Europe and replaced slowly the Homo neanderthalensis who had occupied this region for almost 200.000 y the human capacity for symbolic thinking and symbolic language was there. In this case Homo sapiens originated in Eastern Africa, living parallel to Homo neanderthalensis in the Near East (Palestine) for almost 50.000 y and after his expansion to southern Asia and Australia had acquired this capacity on his return to Europe (and Africa). This hypothesis is, however, incompatible with the acceptance that all present humans have the same (inborn) language capacity. One could argue that the species which was biologically shaped in East Africa around 200.000 y BP had the latent (biologically present but behaviourally uncovered) faculty of language. This dormant faculty finally surfaced some 50.000 y ago under specific environmental conditions or it was just discovered

holophrastic units of child speech) is split up into two (or three) constituents, a "memory" of its unity is left; in Tesnière (1959: 11) it is called "connexion", an invisible link between separated parts of the sentence. The intuition about such a dynamical split is very old. In Wildgen (2002a, b) the Platonic theory of predication has been shown to underlie this idea. It seems that there exists an intuitive (silent) knowledge accessible to reflection about the evolutionary bifurcation and the force field it created in basic sentences. The bifurcation goes, however, in the opposite sense compared to the bifurcations mentioned before. It does not eliminate or reduce a former opposition but creates new oppositions and these are the germs for further structural bifurcations responsible for the dynamic complexity of human languages. The splitting of holophrastic units has two effects. On one hand the dynamic tension between the two different parts remains active and drives the dynamics of sentence formation/understanding; on the other hand new levels of complexity replace the holophastic gestalt; these are conversational turns, islands of monologue in the conversational dynamics, such as narratives, descriptions or argument sequences. They will be discussed in the next section. Figure 11 illustrates the basic split and the dynamic configuration it has created. This evolutionary step was decisive for the emergence of the complete language capacity in Homo sapiens in his pre-out-of-Africa stage (ca. 100.000 to 60.000 y BP). Cf. also Wildgen (2009) for biolinguistic aspects of grammar.

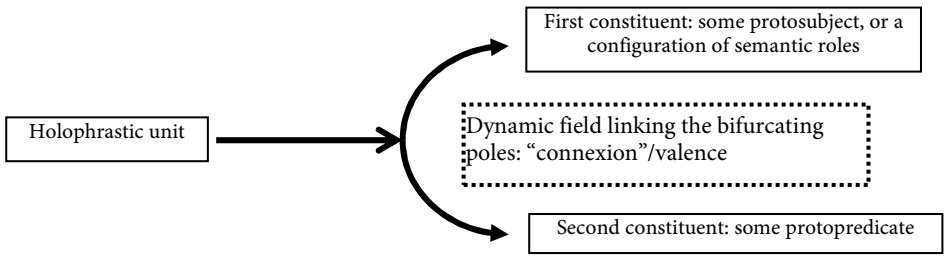

Figure 11: (Productive) Bifurcation and the emergent dynamic field (valence).

by some individuals and groups. In this case, language would not be a biologically (genetically) fixed faculty but a cultural innovation which rapidly spread after 50.000 in the community of Cro Magnon man. The Homo neanderthalensis was for some reason or other not able to adopt this cultural innovation and died out. In a certain sense this would close the domain: evolution of language and go back to hypothetical constructions in the 18th century: How did (single) humans discover the possibility of using language efficiently (cf. Herder and others; the most prominent anthropologist arguing for such a late origin of human language is Tattersall; cf. Tattersall, 2010).

The basic question is: How can such a tension between the simultaneously specified poles of subject (thema) and predicate (rheme) not only be conserved but also used for a growing lexicon of constructions (cf. Wildgen, 2008: chapter 6 for "construction grammars")? It probably concerns at the same time the massive ambiguity of syntactic constructions and lexical items used in their frame. Theoretically most sentences are multiply ambiguous, but normal speaker are almost not concerned by the effect of this ambiguities. The answers to such questions and challenges will probably come from neurosemantics, a field beyond our concern in this article.

6. Further steps of complexification

Sentential patterns may be very simple even in human languages, e.g. in pidgins, in learner languages and even in normal languages with a broad usage (e.g. the so-called minimal languages analyzed by Gil 2009). Human utterances are, however, not restricted to isolated sentences. To the contrary, the natural units are sequences of sentences, so called turns in conversation, adjacent pairs as in: question-answer and rather often monologue sequences as in narratives, descriptions and arguments. Recurrent patterns may be due to ritual behavior; rule governed sequences in relation to major functions occur in fairy tales, jokes and songs. The every day myths commented by Barthes (1957) and political rhetoric in general demonstrate the ongoing relevance of ritualized symbolic behavior. In terms of Peirce's classification of signs the textual domain belongs to the sign level of *argument* (the two others are *rhematic sign* and *dicent sign*). The argument exemplified by a syllogism consisting of three propositions is based on logical coherence (consequence, deduction). In the case of the narrative the coherence is rather a temporal and a causal (motivational) one. In the case of a descriptive text (in a tourist guide or in the oral description of a route to follow to reach a specific place) the coherence is spatial and temporal. A similar "coherence" of sequential actions is asked for in tool manufacturing, where up to 30 steps have to be followed to produce a perfect hand axe out of a piece of flint-stone. Thus the capability to control far reaching coherence patterns in action (and later in words) is surely a central feature which distinguishes the language of our species from a protolanguage. The question if recurrent actions embedded into a holistic gestalt (manufacturing the hand axe) apply some recursive rule is as open as the question if human utterances use such recursive rules (the fact that a model reproducing human sentences can use such rules is not decisive insofar as the same sentences can also be produced by some non-finite state program without recursive rules; cf. Fitch, 2010).

A more fundamental problem concerns the level of language to which a selective process could apply. As this is normally the phenomenal level of holistic behavior, we presume that the textual behavior is the proper level on which selection effects play a role. Therefore human evolution must have selected for the effective use of language in social communication and not at the level of sentences or words. These levels are only selection relevant insofar as they allow the construction of coherence in narratives, description or arguments. This line of research must be extended instead of the misuse of evolutionary arguments to foster the position of paradigms established in mid of the last century, such as the Chomskyan one.

7. Conclusions

The *biological* evolution of human language is basically a continuous process in which the bodily preconditions were provided. In the further *social* evolution that started with the increase of group size (or group organization) and new forms of symbolically ruled social behavior three bifurcations with symmetry-breaking and convergence occurred: from manual to phonic contact management, from gestural to phonic deixis and referential location and from context dependent emulation to cultural learning. After these dramatic shifts towards a complex phonic system of communication with socio-emotional *and* referential function, and cultures based on symbolic transmission, the centrality of language for human survival and expansion was firmly established. Language capacity became a species-defining character of humans (and had dramatic selection consequences). The migration of the Homo erectus to Europe and Asia and its survival for almost 2 my demonstrated this selective advantage at the level of the protolanguage.

The next stage had to cross a barrier of complexity linked to multi-valued valence patterns. The stone-age industries and the communicational consequences of it, prehistoric art and myths, show that this barrier had been crossed when Homo sapiens began to migrate out of Africa (between 100.000 and 70.000 y BP); cf. Wildgen 2004: Chapter 4 and 5 for a treatment of stone age techniques and Paleolithic art.

The major effects on language were:

- a larger and steadily growing lexicon;
- the mastery of rapid and complex strings of phonic signals and corresponding functional-semantic patterns as shown in lexical innovation, composition, grammaticalization, and complex phrasal syntax;

- a new level of creativity in language and art linked to a growing complexity and diversity of languages.

Finally with the rise of agriculture and the breeding of animals, metallurgy, ship building, towns and large civilizations in Egypt and Mesopotamia a new era of symbolic communication arrived which lies beyond the scope of his paper.

Received: 14.04.2011
Accepted: 14.11.2011

Bibliography

Ballmer, Th. T. (1982): *Biological Foundations of Linguistic Communication*. Amsterdam, Benjamins.

Barthes, R. (1957): *Mythologies*. Paris, Seuil.

Bickerton, D. (2000): "How Protolanguage Became Language". En Knight, Ch., M. Studdert-Kennedy, J. R. Hurford (drs.): *The Evolutionary Emergence of Language. Social Functions and the Origins of Linguistic Form*. Cambridge, Cambridge U.P., págs. 264-284.

Bickerton, D. (2010): "On two Incompatible Theories of Language. Evolution. En Larson, R. K. *et alii* (drs.): *The Evolution of Human Language. Biolinguistic Perspectives*. New York, Cambridge U.P., págs. 199-210.

Blackmore, S. (1999): *The Meme Machine*. New York, Oxford U.P.

Boesch, Ch. y M. Tomasello (1998): "Chimpanzee and Human Cultures", *Current Anthropology*, 39(5), págs. 591-614.

Calvin, W. H. y D. Bickerton (2000): *Lingua ex Machina. Reconciling Darwin and Chomsky with the Human Brain*. Cambridge (Mass.), MIT-Press.

Cangelosi, A., A. Greco y S. Harnad (2002): "Symbol Grounding and the Symbolic Theft Hypothesis". En Cangelosi, A. & D. Parisi (drs.): *Simulating the Evolution of Language*. London, Springer, págs. 191-209.

Dunbar, R. (1996): *Grooming, Gossip, and the Evolution of Language*. Cambridge (Mass.), U.P. Harvard.

Duss, S. y K. Henke (2011): "Bewusstsein für Lernen unnötig", *Spektrum der Wissenschaft*, January 2011, págs. 16-18.

Fischer, J. y K. Hammerschmidt (2001): "Functional Referents and Acoustic Similarity Revisited: the Case of Barbary Macaque Alarm Calls", *Animal Cognition*, 4, págs . 29-35.

Fitch, W. T. (2010): "Three Meanings of "recursion": Key Distinction for Biolinguistics". En Larson, R. K. *et alii* (drs.): *The Evolution of Language. Biolinguistic Perspec-*

tives. Cambridge U.P., Cambridge, págs. 73-90.

Gil, D. (2009): "How Much Grammar Does it Take to Sail a Boat?". En Sampson, G., D. Gil y P. Trudgill (drs.): *Language Complexity as an Evolving Variable*. Oxford U.P., Oxford.

Green, R. E. *et alii* (2010): "A Draft Sequence of the Neandertal Genome", *Science*, 328, págs. 710-722.

Gunz, Ph., S. Neubauer, B. Maureille y J.-J. Hublin (2010): "Brain Development after Birth Differs between Neandertals and Modern Humans", *Current Biology*, 9. November 2010.

Hurford, J. R. (2000): "Social Transmission Favours Linguistic Generalization". En Knight, Ch., M. Studdert-Kennedy, J. R. Hurford (drs.): *The Evolutionary Emergence of Language. Social Functions and the Origins of Linguistic Form*. Cambridge U.P., Cambridge (chap. 19), págs. 324-352.

Jackendoff, R. (2002): *A Theory of Language*. Cambridge, Cambridge U.P.

Kirby, S. (2000): "Syntax without Natural Selection: How Compositionality Emerges from Vocabulary in a Population of Learners". En Knight, Ch., M. Studdert-Kennedy, J. R. Hurford (drs.): *The Evolutionary Emergence of Language. Social Functions and the Origins of Linguistic Form*. Cambridge U.P., Cambridge, págs. 303-323.

Levinson, S. C. (2001): *Space, Language and Cognition. Explorations in Cognitive Diversity*. Cambridge U.P., Cambridge (Mass.).

Liebermann, Ph. (1989): "The Origins of Some Aspects of Human Language and Cognition". En Mellars, P. y Ch. Stringer (drs.): *The Human Revolution. Behavioural and Biological Perspectives on the Origins of Modern Humans*. Edinburgh, Edinburgh U.P., págs. 391-414.

Malotki, E. (1979): *Hopi-Raum. Eine sprachwissenschaftliche Analyse der Raumvorstellungen in der Hopi-Sprache*. Tübingen, Narr.

Marean, C.W. (2010): "Coastal south Africa and the Modern Human Lineage and the Coastal Adaptation". En Bicho, N. *et alii* (drs.): *Changing Coastlines and the Antiquity of Coastal Settlement*. New York, Springer.

Petraglia, M. D. (2011): "Trailblazers across Arabia", *Nature*, 470, pgs. 50–51 (03 February).

Piveteau, J. (1991): *La main et l'hominisation*. Paris, Masson.

Pöppel, E. (1997): "A Hierarchical Model of Temporal Perception", *Trends in Cognitive Science*, 1(2), págs. 56-61.

Pöppel, E. (2010): *Der Rahmen. Ein Blick des Gehirns auf unser Ich. München*, dtv (cf. chapter 20).

Quiatt, D. y V. Reynolds (1993): *Primate Behaviour. Information, Social Knowledge, and*

the Evolution of Culture. Cambridge, Cambridge U.P.

Reuter, M. *et alii* (2010): "Investigating the Genetic Basis of Altruism: the Role of the COMT Val158Met Polymorphism", *Social Cognitive and Affective Neuroscience Advance Access (SCAN)*, October 28, págs. 1-7.

Rizzolatti, G. & M. A. Arbib (1998): "Language Within Our Grasp", *Trends in Neuroscience*, 21(5), págs .188-194.

Smolensky, P. (1986): "Information Processing in Dynamical Systems: Foundations of Harmony Theory". En Rumelhart, D. E. *et alii* (drs.): *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, vol. 1: Foundations, págs. 194-281.

Studdert-Kennedy, M. (2000): "Evolutionary Implications of the Particulate Principle: Imitation and the Dissociation of Phonetic Form from Semantic Function". En Knight, Ch., M. Studdert-Kennedy, J. R. Hurford (drs.): *The Evolutionary Emergence of Language. Social Functions and the Origins of Linguistic Form*. Cambridge, Cambridge U.P., págs. 161-176.

Talmy, L. (1988): "Force Dynamics in Language and Cognition", *Cognitive Science*, 12(1), págs.. 49-100.

Tattersall, I. (2010): "A Putative Role for Language in the Origin of Human Consciousness". En Larson, R. K. *et alii* (drs.): *The Evolution of Human Language. Biolinguistic Perspectives*. New York, Cambridge U.P., págs. 193-198.

Tesnière, L. (1959): *Éléments de syntaxe structurale*. Paris, Klincksieck.

Thom, R. (1983): *Mathematical Models of Morphogenesis*. New York, Horwood (Wiley).

Wildgen, W. (1981): "Archetypal Dynamics in Word Semantics. An Application of Catastrophe Theory". En Eikmeyer H. J. y H. Rieser (drs.): *Words, Worlds and Contexts. New Approaches to Word Semantics*. Berlin, de Gruyter, págs. 234-296.

Wildgen, W. (1982): *Catastrophe Theoretic Semantics. An Elaboration and Application of René Thom's Theory*. Amsterdam, Benjamins.

Wildgen, W. (1994): *Process, Image, and Meanings. A Realistic Model of the Meanings of Sentences and Narrative Texts*. Amsterdam, Benjamins.

Wildgen, W. (1998): *Das kosmische Gedächtnis. Kosmologie, Semiotik und Gedächtnistheorie im Werke von Giordano Bruno (1548-1600)*. Frankfurt, Lang.

Wildgen, Wolfgang (2002a): "Dynamical Models of Predication", *STUF (Sprachtypologie und Universalienforschunmg)*, 4, págs. 403-420.

Wildgen, W. (2002b): "Modelos de predicación y clasificación dinamica oracional de los verbos y los papeles configuracionales", *E.L.U.A*, 16, págs. 219-253.

Wildgen, W. (2003): "L'evolució de les llengües: continuïtat i catàstrofe", *Mètode. Rivista de Difusió de la Investigació*, 39, págs. 73-75.

Wildgen, W. (2004): *The Evolution of Human Language: Scenarios, Principles, Cultural*

Dynamics. Amsterdam, Benjamins.

Wildgen, W. (2005): "Le problème du continu/discontinu dans la sémiotique de René Thom et l'évolution des langues", *Cahiers de Praxématique*, 42, págs. 121-144.

Wildgen, W. (2008): *Kognitive Grammatik. Klassische Paradigmen und neue Perspektiven. Berlin, de Gruyter.* [Cognitive Grammar. Classical Paradigms and New Perspectives].

Wildgen, W. (2009). "Sketch of an Evolutionary Grammar Based on Comparative Biolinguistics". In: Röska-Hardy, L. S. y E. M. Neumann-Held (drs.): *Learning from Animals? Examining the Nature of Human Uniqueness*. Hove and New York, Psychology Press, págs. 45-59.

Wildgen, W. (2010a): "Evolutionary Pragmatics, (augmented and revised version)". En Fried, M., J.-O- Östman y J. Verschueren (drs.): *Variation and Change. Pragmatic Perspectives [HoPH= Handbook of Pragmatics Highlights 6]*. Amsterdam: Benjamins, págs. 81-99.

Wildgen, W. (2010b): "Sprache und Selbstorganisation: Von der biologischen zur kulturellen Evolution der Sprache". En Stroh, C. (ed.): *Von Katastrophen, Zeichen und vom Ursprung der menschlichen Sprache. Würdigung eines vielseitigen Linguisten. Wolfgang Wildgen zur Emeritierung*. Brockmeyer, Bochum, págs. 1-19.

Wildgen, W. and M. Plümacher (drs.) (2009): *Prägnanter Inhalt - Prägnante Form. Thematic volume of: Zeitschrift für Semiotik*, 31 (1-2). Tübingen, Stauffenburg.

Wildgen, W. and P. A. Brandt (drs.) (2010): *Semiosis and Catastrophes. René Thom's Semiotic Heritage*. Bern, Lang.