

[draft -- do not quote without permission]

On Language and Tools

Jan Koster

1. Introduction

What follows is born out of dissatisfaction with current Minimalism, the received linguistic paradigm since Chomsky 1995. My concerns are not about Minimalism as a program. On the contrary, I subscribe to the overall goal to construct a theory that makes grammar look as perfect as possible and that relegates as much as it can to “third factor” principles. My dissatisfaction is about how this program is carried out in practice. Others disagree, but my personal feeling is that little *theoretical* progress has been made since the 1980s. I emphasize *theoretical*, because empirically speaking the progress has been impressive. One can hardly think of any topic nowadays of which it cannot be said that there is a wealth of literature about it. All of this progress, I claim, is mainly “cartographic” and therefore compatible with pre-minimalist generative grammar and even certain forms of pre-generative structuralism. Part of the theoretical stagnation is due to the fact that some key problems of earlier versions of generative grammar, as they arose for instance in the GB-period, are either unresolved or ignored. But there are deeper problems, it seems, that involve the very foundations of the field.

For the sake of concreteness, briefly consider the following incomplete but representative list of elements of actually existing Minimalism:

(1) **Minimalist theories in practice:**

- a. Extremely derivational (vs. cartography)
- b. Numeration (derivational book-keeping)
- c. Bare phrase structure (Merge)
- d. Internal Merge (copying theory)
- e. Inclusiveness, no-tampering
- f. Feature-driven syntax (ϕ , EPP, edge)
- g. Syntax: optimal way to connect interfaces
- h. Derivation by phase (vP and CP)
- i. Phase-impenetrability (locality)
- j. Agree (probe-goal relation)

It is impossible to completely review these elements here, so, I will just give a few examples. I will briefly criticize Merge, a core notion of the theory, in a minute. Internal Merge (1d) is the successor of “move alpha” in earlier theories and just as problematic as “movement.” Not only is the long tradition of criticism of such notions largely ignored, Internal Merge is even declared to be “perfect” these days, on very shaky grounds (see Koster 2009b for discussion). Similar things can be said about the core principles of the theory, the locality principles (1i) and (1j). It is occasionally observed that the Phase Impenetrability Condition (1i) is very much like the Head-Condition in Van Riemsdijk (1978) (see Boeckx and Grohmann 2004). What remains unsaid, however, is that the earlier condition was --for good reasons-- claimed to be better in the lively debate of 30 years ago because it was also involving the PP. If such issues are not forgotten, they are hardly mentioned anymore. It is difficult to see any progress

here and the conclusion seems justified that no substantial progress has been made as to the nature of locality principles since the 1980s.

This is not a small point because locality principles are by far the most important constraints of the theory that likes to call itself a “principles and parameters” theory. As for parameters, things are perhaps even worse. I cannot get into this topic here, but I believe that the notion “parameter” has hardly been developed beyond the traditional observation that there are “differences” among languages, like with respect to pro-drop or the order of head and complement. In short, the interesting principles were mostly discovered before Minimalism and the notion “parameter” has always remained underdeveloped from a theoretical point of view.

More generally, current generative grammar is often referred to as “a computational theory,” but the current style of the field is so informal that it is practically impossible to find explicit accounts of what exactly is computed and how. With the low level of explicitness considered normal these days, references to “recursive Merge” say little beyond the traditional wisdom that syntax involves hierarchical structures with phrases within phrases of the same kind. There is nothing wrong with that insight but it would be somewhat exaggerated to say that it is revolutionary.

2. What is wrong with the foundations?

It seems to me that generative grammar is ever making the same mistakes leading to the same dead ends over and over again. Most errors are traceable to the beginnings of generative grammar, at a time that the grammar of natural language was approached via the lens of the then very common preoccupation with formal languages. The generative turn in linguistics modified this tradition by interpreting grammars realistically (in terms of individual psychology or biology) but also preserved certain tenets of the formal language tradition. These tenets were not entirely adequate for the study of natural language, as would become clear before long.

The generative turn involved the following paradigmatic notions (FLN standing for “faculty of language in the narrow sense,” see Hauser *et al.*, 2002)):

(2) **Paradigmatic principles**

- a. Methodological individualism: the Faculty of Language (FL) is individual-psychological, ultimately biological (“biolinguistics”)
- b. Theories of language (FLN) must be not word-oriented but syntactocentric
- c. The cultural dimension of language and the role of human agency are beyond the grammatical core of language (called FLN nowadays)

Principles (2a) and (2c) or, of course, two sides of the same coin. Both are errors in retrospect (at least in my opinion) and both were only possible due to the more fundamental error (2b), which in turn was intimately connected to what was inspired by the study of formal languages.

During most of the life-cycle of standard generative grammar, the realistic interpretation was stated in terms of “individual psychology,” although it is true that from the beginning, a connection was made with biology, as in Lenneberg (1967). However, it is only recently that biological terminology became more dominant than references to “psychological reality,” particularly since Jenkins (2000) and Hauser, Chomsky and Fitch (2002). The renewed biological focus took further shape in new journals like *Biolinguistics* (since 2006). It seems to me that this very explicit biolinguistic turn has come down to a *reductio ad*

absurdum of the realistic interpretation of grammar, clearly revealing the weaknesses of the foundations as summarized in (2).

Note that neither rationalism nor the biological foundations of language are at issue here. I assume that all learning is constrained by biological principles, some of them very specific. The days of behaviorism are far behind us and I must confess that I even cannot remember to have ever seen a behaviorist. I consider the view of biologically constrained learning trivially true and from this general perspective language is as biologically based as our ability to play chess or to drive a car. None of that can be done by other organisms after all and at best there is a question as to how specialized or inflexible parts of our brain are with respect to certain tasks. So, the claim that language makes use of biological, innate components is obviously true and therefore trivial as far as I am concerned. This, of course, does not mean that it is a trivial task to find out what is the exact nature of the biological components involved.

The non-trivial claim of current biolinguistics is more specific, namely that grammar is like a specialized organ, or rather like internal computational systems such as the ones found in the mammalian visual system.

The trivial claim is true on general grounds and the non-trivial claim is false, or so, at least, I will argue. This has to do with the meaningful distinction that can be made between biologically-based functionality and culturally-based functionality. Consider, as an example, the role of the lungs in respiration and the role of the lungs in the playing of a wind instrument like the trumpet. The former is based on our genetic program and grows automatically, without human interference. Playing a trumpet, however, is based on the same innate structures --the lungs-- but this time the function of the lungs is not formed by our genetic program but by coupling them with an instrument invented by humans in order to achieve human goals. The study of respiration in mammals is generally seen as biology, while playing the trumpet is generally seen as culture. This example illustrates once more that innateness is not the issue. The difference is based on the question whether the use of the same innate structures is mediated by human agency and artifacts (trumpet) or not (respiration). Artifacts, like the trumpet, are invented by humans.

So, the criterion distinguishing the trivial claim from the non-trivial claim is the involvement of artifacts. Language is only possible thanks to artifacts, namely our invented words. This simple observation suffices to refute the non-trivial claim. Whoever rejects this conclusion can choose between two ways out: either it must be shown that organs or internal computational systems (like in mammalian vision) also involve artifacts or it must be shown that the words of our language are not really artifacts. Clearly, the first option is absurd: neither organs like the heart or the kidneys nor mammalian visual computation is based on human inventions. Personally, I think the second option is untenable, too, but in practice it comes closest to a thesis defended in a certain variety of "biolinguistics." But let us first see how the internalist perspective developed in generative grammar.

3. E-language and I-language

Although the internalist perspective was part of generative grammar since the 1950s, it has been characterized with the help of its current terminology since the early 1980s. Particularly in Chomsky (1986), the notions E-language and I-language were discussed, with the further argument that the proper subject matter of linguistics theory (in some reasonably narrow sense) is I-language. The distinction arose in opposition to Quine, who had claimed that languages are characterized by sets of well-formed formulas and that there is no issue as to what is the correct grammar as long as they generate the same set of sentences. Chomsky

rightly rejected this view, characterizing enumerated sets of sentences as E-languages, while the actual mechanisms were characterized as I-languages, objects (“grammars”) selected from a narrow range by children learning their language. The “I” in I-language stands for “internal” (to the human mind) and “individual” (also for “intensional,” a notion to be ignored in this article).

Interestingly, the mistaken preoccupation with E-languages was attributed to the bias created by formal languages. In the study of formal languages, the notion E-language is entirely adequate. As generative grammar was initially introduced as something inspired by the study of formal languages, the confusion was quite understandable, especially since *Syntactic Structures* (Chomsky 1957) explicitly defines a language L as a “set of sentences.” The exact historical context of that is disputed (cf. the introduction of Chomsky 1975 and Tomalin 2006), but that does not really matter for present purposes. What I would like to claim is that it is not only E-language which led generative grammarians astray sometimes, but also something at least as fundamental. What I have in mind is the idea that syntax has an existence in abstraction from the lexicon. The formation rules of formal languages operate on symbols that are fundamentally different from the lexical items of natural languages. The big difference is that the latter have a rich internal structure. This structure is both conceptual-intentional (“semantic”) and syntactic in the sense that knowing a word also means that one knows with what kinds of neighboring elements it can be combined.

Thus, everyone who knows the word *book* knows that it can be combined with a preceding article: *the book* instead of **book the*. Expressing this fact by a rule like NP → Art N misses the point somehow, especially when words are introduced at the very end of a derivation, like in *Syntactic Structures* (Chomsky 1957). In formal languages, however combinatorial properties of symbols are by no means dependent on the internal structure of words. This is the background of the fact that it could ever have been seen as progress that grammar shifted from word-orientation (the Saussurian “sign”) to a syntax-oriented view.

In retrospect, then, the shift away from word-orientation was an error, just like E-language implicitly inspired by the metaphor of formal languages. If linguists had maintained the more traditional word-orientation, it would have been immediately obvious that the notion “I-language” is incoherent. There is a simple reason for that: words are invented cultural objects and are neither internal nor individual. Thus, if a new term, like “I-language,” arises in a language like English, it is invented by some human agent (say, Chomsky) and spreads over a community so that it is nobody’s individual property anymore. It has become part of a public repertoire and is, as such, neither individual nor internal. The *syntactic* properties of this word are also sanctioned, at least partially, by an English-speaking community. There is nothing individual about the fact that the phrase “*the I-language*” is proper English, while the phrase “**I-language the*” is not. Words, in short, are E-elements with properties that include syntactic properties.

The invented, cultural and public nature of words really is the Achilles heel of generative grammar seen as an internalist enterprise. Generative grammar started out on the analogy of formal languages, with residues of the notion E-language and definitively without the idea that syntax is implicit in the internal structure of words. *Syntactic Structures* left the intrinsic combinatorial properties of words (“terminal symbols”) more or less in limbo, while in the same period concern with the lexicon took the form of wrestling with context-sensitive rules. In *Aspects* (Chomsky 1965) the categorial component was kept simple (“context free”) by the introduction of a separate lexicon. The categorial component itself, however, remained independent of lexical items (as in formal languages!), with so-called lexical insertion at the end of a derivation.

Interestingly, introduction of the lexicon revealed a fundamental redundancy in that subcategorization frames seemed to state the same information as was found in the categorial

component. The result was X-bar theory, based on the insight that the redundancy could only be avoided by projecting syntactic structure straight from the lexicon (Chomsky 1970, 1981).

All of this was perceived as steady progress at the time, but in retrospect it was the slow giving up of an error induced by the bias provided by formal languages with their lexicon-independent syntax. In fact, it can be said that it took generative grammar 15 years to return to the structuralist idea that syntax spells out the valency of words. The main lesson learned from this episode was that in a theory in which lexical items have subcategorization properties (like “verb V requires a PP or DP complement”), any form of syntax other than the direct spelling out of these properties, leads to the redundancy problem.

In order to solve the redundancy problem, among other things, it was customary since Chomsky (1970, 1981), to project syntactic structure directly from the lexicon, in accordance with X-bar schemata. Up until the present day, it seems to me, this is the right way to do things. This does not mean, of course, that the specific form of the X-bar theory in Chomsky (1970, 1981) was correct. Like with any other empirical theory, one might hope for steady improvements over the years. In the case of X-bar theory, several modifications were proposed, for instance concerning the superfluousness of intermediate bar levels (Muysken 1982) or the number of Specs. I am not committed to any particular form of X-bar theory, only to the principle that redundancy can only be avoided by projecting syntactic structure straight from the lexicon. I see that insight as the firm conclusion of the first 25 years of generative grammar, even if it means a substantial return to the pre-generative tradition, when syntax was still word-oriented and not seen through the lens of formal languages with their lexicon-independent syntax.

Unfortunately, the introduction of bare phrase structure and Merge in Minimalism has compromised these insights. Bare phrase structure, as I see it, is a matter of notation. It is not substantially different from the versions of X-bar theory that gave up intermediate bar levels or the single-Spec restriction. What is worse, the adoption of Merge *as a sentence generating rule* reintroduced the redundancy problem. What it comes down to is that Merge largely mimics what is already entailed by the subcategorization properties (the “valency”) of lexical items. Both introduce the hierarchical, recursive and binary nature of syntactic structure, showing that something is wrong. All actually existing executions of Merge that I have seen assume the existence of subcategorization features. Chomsky (2000b), for instance, claims that if two elements are merged, a feature F of one of the two elements must be satisfied which is called “the selector.” Similarly, Hornstein *et al.* (2005) use subcategorization information for labeling purposes, where labeling mimics the information (head-complement structure) that was used by X-bar theory to account for endocentricity. But even label-free theories, like the one of Collins (2002), crucially depend on subcategorization. All of this reveals the obvious: even at the most elementary syntactic level (preceding the conceptual-intentional level according to Minimalism), Merge is not sufficient but can only account for endocentricity in combination with lexical information (subcategorization, valency).

In spite of the fact that in practice Merge only exists in tandem with said lexical information, the “recursive Merge only” thesis of Hauser, Chomsky and Fitch (2002), has the flavor of a theory that sees Merge as a computational mechanism completely independent of the properties of our invented words. This is perhaps even necessary for the non-trivial biolinguistics program to make sense. Lexicon-independent computation of this kind evokes reminiscences of the times (pre-1970) when natural language was still approached through the lens of formal languages. Lexicon-independent syntax has been refuted since the 1970s, largely on the basis of the redundancy problem. It is therefore hard to understand why it makes a partial come-back in minimalist theory and its biolinguistic interpretation.

The confusion can be taken away by seeing Merge not as part of a sentence-generating mechanism but as a meta-theoretical characterization of the properties of subcategorization

frames: all involve binary hierarchical structure with recursion. Sentence generation can remain to be seen as the spelling out of subcategorization frames, as in X-bar theory. In a word-oriented syntax of this kind, whatever Merge stands for is *applied* in the creation of lexical items with complex combinatorial properties. In other words, there is a crucial and theoretically important difference between sentence generation *by* Merge and sentence generation as the spelling out of lexical properties *in accordance with* Merge. The former reintroduces the redundancy problem, the latter solves it.

In its application to lexical structures, then, Merge is not something biological but applied biology at best. Of course it is possible to consider Merge in abstraction of its lexical application and to reflect on its biological sources, but at that level of abstraction there is no obvious reason to say that Merge has anything to do with natural language at all. The fact that it is so successfully used in language is not a strong argument for the idea that the biological purpose of Merge is linguistic. Since Gould and Lewontin (1979), the fallacy of reading cultural functions into biological structures is known as “Panglossianism.”

The crux of the matter is that Merge, in abstraction of its lexical application, is not linguistically functional. It is for a reason that before the shift to syntactocentrism, the morpheme was generally seen as the smallest linguistically functional unit. This is the wisdom behind Saussure’s idea that the sign with its dual nature (sound-meaning) is the core element of language. Our combinatorial capacities (as expressed by Merge) are extremely important for language as we know it. No matter how powerful it makes language as a tool, it is only an auxiliary facility. There is just as little reason to call the capacity for Merge the “faculty of language in the narrow sense” (FLN in Hauser, Chomsky and Fitch 2002) as there is reason to call a gasoline engine “a car in the narrow sense.”

4. Words are not labels of concepts

If we accept the fact that words are cultural objects invented by humans, as I think we should, the conclusion seems inescapable that biolinguistics (in the non-trivial sense) is an error. Only the trivial claim can be true: language makes use of innate components but these components are only “linguistic” thanks to their integration via invented signs. The conclusion can only be avoided if it can be shown that words are not really cultural objects but “biological” in some sense after all. In practice, this escape route takes the form of speculations about innate concepts and a natural language precursor known as “the language of thought”(see Fodor 1975). I have argued elsewhere why I find this approach implausible (Koster 2009c). Based on the fact that children are supposed to acquire sometimes a dozen new words a day, a poverty-of-the-stimulus argument can be constructed according to which concepts (as proto-words) are already in place at birth and only wait to be labeled by the arbitrary naming conventions of communities. In this way, words come close to being disguised “biological” entities. Chomsky (2000a: 61) puts it as follows:

Acquisition of lexical items poses what is sometimes called “Plato’s problem” in a very sharp form.

[...]

At peak periods of language acquisition, children are acquiring (“learning”) many words a day, perhaps a dozen or more, meaning that they are acquiring words on very few exposures, even just one. This would appear to indicate that the *concepts are already available*, with much or all of their intricacy and structure predetermined, and that *the child’s task is to assign labels to concepts*, as might be done with limited evidence given sufficiently rich innate structure [emphasis added --JK].

This view is not convincing for a number of reasons. It has been argued in the literature that it is hard to believe that concepts like *carburetor* or *democracy* are innate and foreseen by our primate brain. Analogies have been suggested with the immune system, which selects appropriate antibodies even for antigens to which it has never been exposed before. Unlike chemical substances, however, concepts crucially involve human agency and its record in cultural history. Unlike the class of possible chemical structures (and physical structures in general), the class of *applications* of physical structures is open-ended and unpredictable. It is constrained by physical structure but not determined by it. Nothing can be learned about *democracy* or *carburetors* by inspecting the biological properties of my brain. This is just a special case of the fact that in general function cannot be predicted on the basis of form -- there simply is no logical connection. The function of any form depends on context and the contexts to which brain structures apply are largely created by human agents.

Also if we limit ourselves to a child's vocabulary, there is much technology and cultural history, to which the child gradually tunes in with its innate capacities. Word acquisition is life-long enculturation, not some kind of labeling of the biologically given. Take typical words from a child's vocabulary, like *porridge* or *Mickey Mouse*. The word *porridge* cannot be the label of an innate concept because it is based on relatively recent invention and technology: the growing and processing of grains, fire, cooking, etc. Similarly, Mickey Mouse is a 20th-century creation of the studios of Walt Disney. So, it is impossible that even such typical kid words label innate concepts. This does not, of course, exclude the application of innate expectations in general, for instance those of what has been called "naive physics" since the 19th century. Similarly, by seeing Mickey as a kind of person, the child interprets the figure with whatever the child has in terms of innate expectations as to how persons behave, etc.

The labelling view of word acquisition seems to be based on a confusion between concepts and the innate components concepts are applications of. Thus, *porridge* and *Mickey Mouse*, far from being labels of innate concepts, *apply* partially innate components in an assemblage that is provided by, and typical of, the culture the child grows up in. How to apply innate components into concepts is coded in our public cultural record, not in our DNA.

It is also misleading to say that the child learns sometimes as many as 12 word meanings a day. Consider a simple word like *water*. What does a child learn when it acquires this word? That depends on the circumstances under which the child gets acquainted with water. A child might, for instance, find out via a parent's pointing at it that water is a transparent liquid. Does the child acquire THE meaning of *water* this way? By no means! The meaning of a word like *water* is open-ended and depends on the information an individual can associate with a word and how this information is interpreted in context. Thus, we can only talk about H₂O and "heavy water" with some knowledge of chemistry and physics. There is no sharp boundary between "naive physics" and scientific physics and information associated with words grows throughout a lifetime. In general, a common noun like *water* is not the label of a concept but the label of an open-ended memory depository in which information is collected. What a young child initially learns about *water* is not the word's meaning but minimal information for the construction of meaning.

It is my experience with discussions about concepts that the word *concept* is often a source of confusion, since different people use it in different ways. The word *concept*, it seems to me, is ambiguous between "information structure" and "meaning." Information structures can be found not only in human brains, but also in animal brains and in computers. Meanings, however, are the *interpretations* of information structures in context by human agents and are therefore neither found in animals nor in machines. To blur the distinction between information structures and their interpretations (meaning) would lead to absurdities such as saying that smoke detectors possess the concept of "smoke." Perceptual categories

and other complex category detections are, according to this view, about information structures. To call such structures “concepts,” as is often done, would be about zombie concepts, which can be sharply distinguished from human concepts or meanings. In what follows, I will adopt the tradition (going back to Plato) that words and human concepts are tools. The use of cognitive tools is not found in animals and machines, the latter being tools themselves. Machines, like all tools, only have intentionality derived from human intentionality.

With this in mind, it is perhaps clear that if we talk about “innate concepts,” we are not talking about meanings but about zombie concepts, of the kind that can also be found in animals and machines. Meaning involves intentionality and its exercise via human agency, which, as we are assuming here, is like the use of tools. The availability of tools is not innate because tools are cultural objects belonging to a community and its history. We do not enter the world as tool users but only gradually develop our skills as we adopt the culture of our environment.

So, what kinds of tools are words? It has become common in recent years to talk about cognitive or epistemic tools (see Sterelny 2006: 230-233). Most cognitive tools extend memory in one way or another. So, do words. The primary function of words-as-tools, it seems to me, is to maintain addresses for memory depositories. These memory depositories contain zombie concepts and other information structures in accordance with how both our innate capacities and our culture “see” the world. In other words, part of that is natural (like “naïve physics”) and much of it crucially involves the history of a culture (like what we said about *porridge* and *Mickey Mouse*). More advanced information structures associated with words involve combinatorial potential as described by syntax. Part of that is natural as well (perhaps as s-selection), part of it reflects cultural history (c-selection and most things coming under the label of “parameter setting”).

Note that, strictly speaking, words have no inherent meaning under this view. This leads me to the secondary function of words (next to memory storage and organization): the construction of meaning in context. The construction of meaning in context can best be illustrated with the notion of polysemy, introduced into linguistics by Michel Bréal in 1887 but going back to the notions metaphor and metonymy in Aristotle’s *Poetics*) (see Nerlich 2003 for a historical sketch and Pustejovsky 1993 for an account within the generative tradition). Consider a familiar example, the word *book*:

- (3) a. The book weighs a pound
- b. The book is exciting
- c. The book fits on a flash drive
- d. The book only exists in her head
- e. The book was sold to another publisher
- etc.

It is well-known that a word like *book* can be used as something concrete (3a) and something abstract (3b), leading to two different meanings according to any definition of what meaning is. The other examples suggest that the range of possibilities is constrained, but open-ended. A word like *book* is associated with a rich variety of information and the creation of meaning involves, among other things, selection from this wealth in a way appropriate to the context. Meanings, in other word, are “made” by human agents in a creative process. Since the range of contexts is unpredictable, possible meanings are not listed but must be generated in each individual case. This creation of meaning follows certain rules of thumb (like *pars pro toto*, or “object X and representation of X”) but is essentially non-deterministic.

It is obviously false that anything goes in meaning construction, as the process is heavily constrained by the fact that some parts of the information structure are more unifying than others. With some plausibility, prototypes have been suggested for certain cases, but it is clear that prototypes cannot be seen as THE meanings of the words in question. What is used as unifying the range of interpretations is partially determined by how one gets acquainted with words and their use. As everyone has experienced who tried it, definitions and paraphrases (including sets of markers and meaning postulates) are hopeless as means to give word interpretations a common core. Such means are only one form of getting acquainted with words and their use. In reality, different people get familiar with the use of the same word in different ways and communication is possible to the extent the information structures in question overlap. This can be illustrated with proper names, like *Mozart*. For many, Mozart is primarily known as Europe's prototypical classical composer. That could be the unifying notion for the following examples:

- (4)
- a. Mozart is great (= music)
 - b. Mozart is 30 pages (= score)
 - c. Mozart can be downloaded everywhere (= music files)
 - d. Every pianist played Mozart (= a piece of Mozart)

However, that the Mozart we are talking about here was a composer is by no means the definition of the word *Mozart*, even if applied to a particular man. Others could have known him only as the man next door, without knowing that he was a composer. This could have led to a different set of possible meanings:

- (5)
- a. Mozart is great (= friendly person)
 - b. Mozart is next to the bakery (= his house)
 - c. Mozart will be buried next week (= remains)
 - d. Every neighbor liked Mozart (= the man)

What these examples show is that a (possibly open-ended) class of meanings can be constructed on the basis of the same information associated with some words. As the range of interpretations depends on an unpredictable range of contexts, the total range of meanings of a word cannot be predicted.

Words, then, are addresses of memory depositories which serve as a platform for the construction of meanings in context. In that respect, words are tools in a very literal sense. The notion of words as tools is most familiar from Wittgenstein's *Philosophical Investigations*. The idea goes back, however, to Plato's *Cratylus* and can also be found in the Indian tradition according to the Dutch Sanskritist J.F. Staal. The idea was very common in Wittgenstein's time. In the German speaking world it is found in Karl Bühler's *organon* concept (with reference to Plato) and in the English speaking world it was promoted by the pragmatists, particularly by Charles Sanders Peirce and by John Dewey.

In my view, the notion of words (and concepts) as tools was exactly right and should take center-stage again in theories about the nature of language. Tools are human inventions and fall, as such, within the purview of human agency and cultural tradition. It is a theory that should be preferred over the theory that words are names, either for language-external referents or for innate concepts. Both the idea of a lexicon-independent syntax and the idea that words have an interpretation independent of human agency go back to the Fregean notion of formal calculi and their interpretation. The Fregean paradigm just does not work for natural language and should be radically left behind us, not only in philosophy (as was proposed by the later Wittgenstein) but also in linguistics. Since nothing can be properly called "language"

in abstraction from humanly-invented tools --our words-- the project of “biolinguistics” (in the non-trivial sense) will never fulfill whatever promises it might have for some.

5. Biology, functionality and the extended mind

If we nevertheless want to approach language from a biological perspective, we have to face the problem of the origins of linguistic functionality. I have already hinted at the fact that linguistic functionality, including the functionality of structures compatible with Merge, is neither in Merge itself nor in anything that can be innate. Linguistic functionality crucially depends on our invented lexical items, which belong to a culture shared by a community. Invention is a matter of human agency, the fruits of which are preserved by a public record. There is nothing of that kind in standard biology, which precludes reduction of language, no matter how narrowly conceived, to biology. In standard biology, there are two ways for physical structures to acquire functionality, both depending on natural selection:

- (6) **Function assignment in biology**
 - a. by adaptation
 - b. by exaptation

Adaptation is the gradual adjustment of structures to particular functions by natural selection. Exaptation is a term coined by Gould and Vrba (1982) and was called pre-adaptation in earlier theories. In fact, it is a phenomenon that was discussed by Darwin himself and discussed in the German-speaking world of the 19th century as *Funktionsverschiebung* (see Russell 1916). What it means is that a structure was first selected for one function and was subsequently applied and adjusted to another function, also by natural selection. An example given by Gould and Vrba is the wings of birds that originally evolved as thermo-regulators. I think the notion of exaptation is extremely important because it illustrates the fact that there is no intrinsic relation between form and function. It is not predictable from physical principles which function a given structure will take in the future. Kauffman (2007) rightly takes that as an argument against reductionism, in this case the idea that biology can be reduced to physics.

How does language fit into this picture? Hauser, Chomsky and Fitch (2002) suggest that the faculty of language in the narrow sense (FLN) only includes recursion and that it may have evolved originally for reasons other than language. Possibilities they mention are number, navigation and social relations. Basically, then, they seem to claim that linguistic functionality could be a matter of exaptation. According to standard assumptions in biology, this is a slow process determined by natural selection. Chomsky, at various places, seems to be considering sometimes a more rapid form of evolution, not by natural selection but by “saltation.” That form of evolution is highly controversial among biologists, but that does not really matter from my point of view. Under saltation, there is even less of a transparent relation between form and function than under gradual adaptation of a structure to a function. Ultimately, then, saltational evolution --a much contested option-- would undermine the non-trivial interpretation of biolinguistics rather than support it.

Be this as it may, it seems clear to me that neither adaptation nor exaptation will do for linguistic functionality, for the simple reason that it is based on an invention and therefore on human agency, a phenomenon found nowhere else in the biosphere. What comes to mind here is a distinction that was made by Searle (1995: 20), namely between agentive and non-agentive functionality. Examples given by Searle are the heart and a paperweight. The functionality of the heart developed by natural selection and comes about during ontogeny on the basis of our genetic program, without human interference. When we use a stone as a

paperweight, however, its function is not a matter of natural selection but of human decision. We can even use an object designed for some other purpose as a paperweight, as long as it has the right size and weight. That would show an agentic version of exaptation. Agentic functionality, it seems, is the right notion for language, even in the narrowest sense. This situates language outside the scope of standard biology, which has no equivalent of agentic functionality (apart perhaps from some very rudimentary use of tools among animals).

Words are tools and tools are the prototypical examples of physical structures showing agentic functionality. Moreover, words are not just tools but *cognitive* tools, something unknown even among our closest relatives the great apes. Just as standard tools are an extension of our body, cognitive tools are an extension of our minds. Humans are not just standard biological entities but beings that live in symbiosis with objects that extend the sphere of their intentionality. Also within the human body, we can make a distinction between structures that fall within our sphere of intentionality and structures that do not. The hands, for instance, can be used as tools, while the kidneys cannot. The lungs, as discussed earlier in connection with trumpet playing, are an interesting case in that they have a non-agentic function in respiration but an agentic function in the playing of wind music. The reason is clear: unlike the heart and the kidneys, the functioning of the hands and the lungs is accessible to our control and therefore falls within the sphere of human intentionality.

These considerations also apply to the brain. Much of the processes going on in the brain is not accessible to us and therefore falls outside the sphere of human intentionality. Words and their use, however, are accessible to consciousness and willful arrangement and therefore do fall within the sphere of human intentionality. This does, of course, not mean that we have conscious access to *how* the brain manages the use of words, but that is true for all tool use, including the use of a hammer or a computer. I have no idea what happens in my brain when I use a hammer. In order to be used as a tool, it suffices that a structure is under our control, which requires physical access in the case of non-cognitive tools and access to consciousness and working memory in the case of cognitive tools. This conclusion, incidentally, entails that consciousness is a necessary condition for language.

The upshot of these considerations is that we are beings not only defined by our biological body but also by the sphere of our intentionality, which partially extends beyond the body and inalienable possession and is therefore shareable with others. In other words, part of our intentional sphere lies within the body, like when we use our hands, and part of it falls outside of the body, when we use structures as tools that are not biologically given. The same applies to cognitive tools, which are partially confined within the brain (like when we use visual imagery for planning) and partially also outside of the brain, like the words of language.

The most important preliminary conclusion at this point is that human agency and intentionality cannot be characterized by the standard forms of functionality known from biology (*viz.*, adaptation and exaptation). One missing concept was Searle's agentic functionality. Recent developments in neurobiology give further substance to this key notion for the understanding of the biological foundations of culture. In a very important new book, Dehaene (2009) gives a neurobiological account of another form of cognitive technology, our use of writing systems. Writing and reading are interesting because, unlike in the case of spoken language, it is entirely uncontroversial that writing is a relatively recent invention, say, of 5000 years ago. That is way too short a period for a capacity to have developed by natural selection. Nevertheless, as Dehaene shows, our use of graphic systems (of all cultures!) is governed by very specific areas of the brain, reminiscent of the classical areas of Broca and Wernicke. Dehaene and others have identified a region in the occipito-temporal sulcus of the left hemisphere that Dehaene has dubbed "the letterbox" [From: Stanislas Dehaene, *Reading*

in the Brain: The Science and Evolution of a Human Invention. New York: Viking.
 Illustration given here pending permission of use from the publisher]:

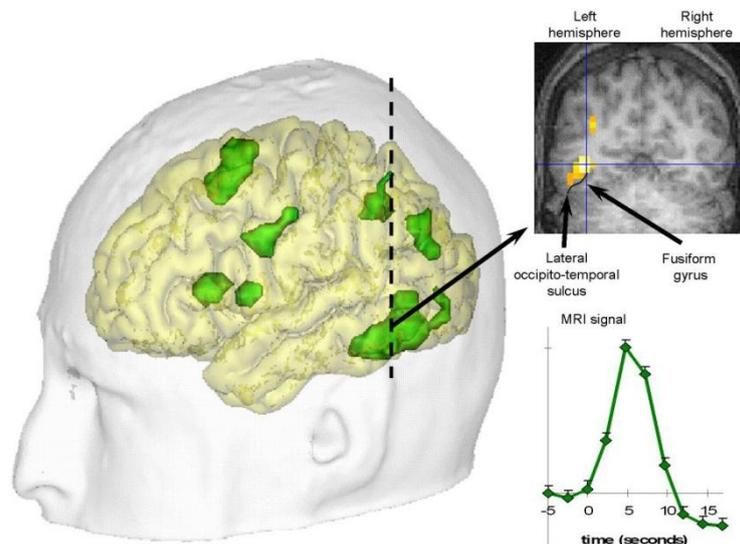


Fig. 1

This is a very important discovery because it shows that even uncontroversially cultural phenomena, like writing and reading, are not controlled by some mechanism of general intelligence, but by very specific areas evolved for very specific purposes. In this case, the area in question is specialized for the recognition of shapes and for object invariance (i.e., the capacity to recognize something as the same object when seen from different angles and perspectives). This is largely an evolved, innate capacity, with obvious survival value for apes and monkeys as well (hence the finding of homologous structures in their brains). However, we humans have been able to give these innate structures a new function, namely the function created by the invention of writing systems. This is the agentive form of function assignment discussed earlier and Dehaene calls it our capacity for “recycling:” our capacity to give very specific, innate brain structures a new function by using them in a new, culturally invented context.

Recycling, in Dehaene’s sense, is precisely the notion we are looking for and I think it sets a standard for all cognitive phenomena with a cultural, agentive component, including natural language. So, instead of two forms of functionality (see (6) above), we have to make a three-way distinction:

- (7) **Function assignment**
- a. by adaptation
 - b. by exaptation
 - c. by recycling

Recycling (7c) via cultural invention sets humans apart from the rest of the biosphere, which establishes form-function connections mostly by (7a) and (7b). Note also that Dehaene’s

notion opens the way to a non-Panglossian approach to the relation between culture and biology. What I mean is that in the case of reading an intrinsic connection between brain structure and function is excluded, so that nobody would, in a non-metaphorical way, talk about “a reading organ,” “the faculty of writing” or “innate alphabets.” Dehaene speaks about “the letterbox” but, clearly, that is a metaphor describing an end result and not some genetically given structure destined for reading and writing. It is my belief that the same approach should be adopted for linguistics and for how we see the relation between language and biology. Use of innate structures: yes, biological structures with a fixed cultural purpose: no. Linguistic functionality is not based on adaptation or exaptation, but on an invention of a cultural context that recycles neural machinery originally evolved for other purposes. Exaptation and recycling both “redefine” the function of existing biological structures, but unlike exaptation, recycling does justice to the role of human agency as shown in cultural invention.

I see both spoken language and writing systems as cognitive technology, primarily memory technology. Derived from that, we have a set of linguistic tools for the support of thinking and for communication. If both are cognitive technologies, an interesting question arises as to the differences. Spoken language differs from writing, after all, in that it is universal and acquired much earlier in life and practically without explicit instruction. It seems to me that none of that makes a difference for the logic of the problem. Nothing in the recycling hypothesis says anything about the fact that some forms of recycling are easier to obtain than others. Singing songs, for instance, is universal and acquired early in life, while playing the piano or the veena is far from universal, learned later in life and with much more effort and explicit instruction. Nevertheless, both are uncontroversially cultural activities making use of innate capacities. It is just that some activities lean more heavily on easily accessible innate capacities than others, as can be demonstrated with numerous examples.

Another popular possibility is saying that language involved a certain measure of coevolution (Deacon 1997). That, too, is entirely compatible with the recycling hypothesis and I even consider it likely that the speed of access and use of language was highly favored by natural selection. However, that would be facilitating evolution, not function-creating evolution. Recall that the function of organs like the heart is created by evolution, while linguistic functionality is created by human invention. Nothing of principle excludes the possibility that the use of inventions is rewarded by natural selection.

In sum, then, I think that Dehaene’s notion of recycling is the missing link and the right concept for human capacities that integrate biological and cultural phenomena. More in particular, Dehaene’s approach is compatible with the following assumptions about language that I partially share with Chomsky:

- (8) **Recycling is compatible with:**
- a. poverty-of-the-stimulus arguments
 - b. rationalism
 - c. the role of culture and agency
 - d. the rapid evolution of language

Recycling is compatible with poverty-of-the-stimulus arguments because it makes use to a significant degree of innate structures. It is compatible with rationalism, because, being based on innate structure, it claims that learning involves not general but specific mechanisms within a narrow range. Unlike the standard interpretation of “biolinguistics,” it also does justice to the cultural and agentic aspects of language.

Last but not least, it makes the relatively recent emergence of language plausible without recourse to extreme forms of saltationism. Nobody knows when natural language first

appeared, but a fairly common guess is some 200,000 years ago. That is too short a period for normal, gradual evolution by natural selection. One assumption is that some kind of saltational mega-mutation took place, leading to language almost instantaneously. Although the days of extreme gradualism and exclusive point mutation are behind us in biology, it still makes good biological sense to see evolution as a process always building on existing structures, without transitions that are all too radical. The mystery is perhaps solved if we assume that language is, at least partially, a matter of cultural evolution: something we owe to new inventions --words-- that integrated and recycled non-linguistic capacities that had evolved for other reasons in the brains of our primate ancestors. In the spirit of my instrumental view of language, the key to studying the origins of language might be the evolution of tool use and its extension to the use of cognitive tools to support and organize memory.

Altogether, it seems to me that the recycling approach is the only one available at the moment compatible with all considerations summarized in (8).

6. Consequences and conclusions

Since the 1980s, I have rejected the formulations of linguistic theory in terms of individual psychology, which are currently more commonly referred to as “biolinguistics.” My original objections (Koster: 1988 [1993], 1989) were inspired by a modified version of Karl Popper’s World 3 concept and also, more at the background, by the opposition against sociobiology. This led to a rejection of brain-mind identity, i.e., the thesis that theories of the mind are theories of the brain at a certain level of abstraction. The myth of the isolated mind (as the subject matter of cognitive science) was also rejected by others, like Donald (1991, 2000), leading to the concept of “the extended mind” (see also Malik 2000 and Logan 2007). Recently, the idea of the extended mind has gained an unexpected popularity, particularly thanks to Clark and Chalmers (1998) and books like Clark (2009) and “The Cambridge Handbook of Situated Cognition” (Robbins and Aydede 2009, with much relevant discussion). Jerry Fodor (2009) has criticized the idea of the extended mind, mainly on the basis of the claim that brain-external cognition is a matter of “derived content,” while brain-internal cognition is based on “underived content.” By and large, this is applying double standards, because much of what is going on within the brain involves derived content as well.

As we saw in our discussion of tool use, the body (including the brain) is just not an absolute boundary in defining the sphere of human intentionality. As I see it, the human mind is crucially defined by the sphere of its intentionality. On the “hardware” side, this sphere includes not only the brain but also brain-external extensions in the form of tools, including cognitive tools like the words of language and writing systems. Since our intentional sphere is partially external to the body and therefore not an inalienable possession, we can share it with others. All tool use sets us off from primates and other animals, but, more than anything else, the cognitive tools of our language make all the difference. It is particularly thanks to our language that our mind extends beyond the brain, even if all “content” is ultimately derived from unknown sources from within the brain. We have an extended mind thanks to our invented tools, which in turn are the fruits of our agency as preserved by our culture. Cognitive science may abstract away from ill-understood notions like agency and culture for practical reasons, but in that way it will definitely miss the human essence insofar as it extends beyond the cognition of animals and zombies. Being in denial of the “human difference” runs counter to the spirit of the Enlightenment (see Malik 2000). There is much to learn about the biological sources of human nature, but ultimately human life, including

language, can only be understood at a level beyond biology, namely the level including agency and its record in our cultural history.

It is from this point of view, supported by argument I hope, that I am not convinced by biolinguistics (in the non-trivial sense) and its precursors in terms of individual psychology and I-language. I-language, furthermore, was concluded to be a residue of the study of formal languages with their lexicon-independent syntax. According to the view that syntax is about properties projected from lexical elements, notions like I-language and biolinguistics can eventually not be formulated, as, clearly, lexical items are non-individual inventions. They belong to a cultural tradition with an existence external to the individual. I believe that the paradigmatic metaphor language-as-a-biological-organ should be replaced by the idea of language as a set of publicly available tools. In particular, these tools must be seen as cognitive tools, which are of a kind unique to humans. Naturally, this idea is compatible with the presence of rich internal structure that we are unaware of, possibly even computational structure. The point to keep in mind here is that the use of *all* cultural objects involves unconscious, inner computational dimensions, for instance when we interpret a two-dimensional painting as representing a three-dimensional reality and even when we try to hit a nail with a hammer.

I am not only skeptical about the notion I-language, but also about the idea that syntax is a computational system of the same internal kind as the computational systems involved in mammalian vision. I find this latter view unintelligible, as, obviously, mammalian visual computation does not depend on external, publicly available cultural objects, like the words of language.

Luckily, there is an alternative to biolinguistics (in the non-trivial sense). I propose that Dehaene's notion of "recycling" is seen as paradigmatic for the study of the biological foundations of language. Last but not least, I reject brain-mind identity in favor of some version of the theory of "the extended mind." A human being cannot be defined in terms of the body and the brain alone, but only as a being living in symbiosis with cognitive and non-cognitive tools, which extend our mind-defining sphere of intentionality. Most important of all, thanks to having an extended mind in this sense, we are not the lonely prisoners of our brain but we live in a cognitive world shared with others.

Bibliography

Boeckx, Cedric and Kleantes Grohmann

2004 Putting phases into perspective. Paper presented at *Tools in Linguistic Theory*, Budapest, May 2004.

Chomsky, Noam

1957 *Syntactic Structures*. The Hague: Mouton.

1965 *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.

1970 Remarks on nominalization. In R. Jacobs and P. Rosenbaum, Eds., *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn and Company: 184-221.

1975 *The Logical Structure of Linguistic Theory*. New York: Plenum Press.

1981 *Lectures on Government and Binding*. Dordrecht: Foris.

1986 *Knowledge of Language*. New York: Praeger.

1995 *The Minimalist Program*. Cambridge, Mass.: MIT Press.

2000a *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

- 2000b Minimalist inquiries. In R. Martin *et al.*, Eds., *Step by Step*. Cambridge: MIT Press: 89-155.
- Clark, Andy
2008 *Supersizing the Mind: Embodiment, Action and Cognitive Extension*. Oxford: Oxford University Press.
- Clark, Andy & David Chalmers
1998 The extended mind. *Analysis* 58: 7-19.
- Collins, Chris
2002 Eliminating labels. In Samuel David Epstein and T. Daniel Seely, Eds., *Derivation and Explanation in the Minimalist Program*. Malden, etc.: Blackwell: 42-64.
- Deacon, Terence W.
1997 *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W.W. Norton & Company.
- Dehaene, Stanislas
2009 *Reading in the Brain: The Science and Evolution of a Human Invention*. New York: Viking.
- Donald, Merlin
1991 *Origins of The Modern Mind*. Cambridge, Mass.: Harvard University Press.
2000 The central role of culture in cognitive evolution: a reflection on the myth of the 'isolated mind'. In L. Nucci, Ed., *Culture, Thought and Development*. Philadelphia: Lawrence Erlbaum Associates: 19-38.
- Fodor, Jerry
1975 *The Language of Thought*. Hassocks, Sussex: Harvester Press.
2009 Where is my mind? review of Clark (2008). *London Review of Books*, 12 Feb 2009.
- Gould, Stephen Jay and Richard C. Lewontin
1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London*. B 205 (1161): 581-598.
- Gould, Stephen Jay and Elisabeth Vrba
1982 Exaptation: a missing term in the science of form. *Paleobiology* 8: 4-15.
- Hauser, Marc, Noam Chomsky & Tecumseh Fitch
2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* 298: 1569-1579.
- Hornstein, Norbert, Jairo Nunes & Kleanthes Grohmann.
2005 *Understanding Minimalism*. Cambridge: Cambridge University Press
- Jenkins, Lyle
2000 *Biolinguistics*. Cambridge., Mass.: MIT Press.
- Kauffman, Stuart
2007 Beyond reductionism: reinventing the sacred. *Zygon* 42, 903-914.
- Koster, Jan
1988 Language and epistemology. *Groningen Papers in Theoretical and Applied Linguistics TENK Nr. 6*. Published in French as Koster (1993). (Also at: <http://www.let.rug.nl/koster/old%20papers.html>)
1989 How natural is natural language. In J.E. Fenstad *et al.*, Eds., *Logic, Methodology, and Philosophy of Science VIII*. Amsterdam: Elsevier.
1993 Langage et épistémologie [Translation of Koster 1988 into French by Nicolas Ruwet]. *Recherches Linguistiques* 22, 59-74.

- 2009a Ceaseless, unpredictable creativity: language as technology. In *Biolinguistics* 3: 61-92.
- 2009b IM not perfect: the case against copying. Ms., University of Groningen.
- 2009c Word meaning and the preformationist fallacy. MS., University of Groningen.
- Lenneberg, Eric
1967 *Biological Foundations of Language*. New York: John Wiley & Sons, Inc.
- Logan, Robert
2007 *The Extended Mind: The Emergence of Language, the Human Mind, and Culture*. Toronto: University of Toronto Press.
- Malik, Kenan
2000 *Man, Beast, and Zombie: What Science Can Tell Us About Human Nature*. New Brunswick, N.J.: Rutgers University Press.
- Muysken, Pieter
1982 Parametrizing the notion "head." *Journal of Linguistic Research* 2: 57-75.
- Nerlich, Brigitte
2003 Polysemy: past and present. In Nerlich et al. 2003: 39-79.
- Nerlich, Brigitte, Zazie Todd, Vimala Herman and David Clarke, Eds.,
2003 *Polysemy: Flexible Patterns of Meaning in Mind and Language*. Berlin: Mouton De Gruyter.
- Popper, Karl R.
1972 Epistemology without a knowing subject. In Karl R. Popper, *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press: 106-152.
- Pustejovsky, James
1993 *Semantics and the Lexicon*. Dordrecht: Kluwer.
- Riemsdijk, Henk van
1978 *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Dordrecht: Foris.
- Robbins, Philip and Murat Aydede, Eds.
2009 *The Cambridge Handbook of Situated Cognition*. Cambridge: Cambridge University Press.
- Russell, Edward Stuart
1916 *Form and Function: A Contribution to the History of Animal Morphology*. London: Murray
- Searle, John R.
1995 *The Construction of Social Reality*. London: Penguin Books.
- Sterelny, Kim
2006 Cognitive load and human decision, or, three ways of rolling the rock uphill. In Peter Carruthers, Stephen Laurence and Stephen Stich, Eds., *The Innate Mind: Volume 2: Culture and Cognition*. Oxford: Oxford University Press: 218-233
- Tomalin, Marcus
2006 *Linguistics and The Formal Sciences: The Origins of Generative Grammar*. Cambridge: Cambridge University Press.
- Wittgenstein, Ludwig
1953 *Philosophical Investigations*. Oxford: Basil Blackwell.
- Jan Koster
Department of Linguistics
University of Groningen
Postbus 716,

9700 AS Groningen
The Netherlands
J.Koster@rug.nl