The structure and evolution of symbol

Erkki Luuk*

Institute of Computer Science, University of Tartu, Postimaja pk 149, 51004 Tartu, Estonia

ABSTRACT

The received opinion is that symbol is an evolutionary prerequisite for syntax. This paper shows two things: 1) symbol is not a monolithic phenomenon, and 2) symbol and syntax must have co-evolved. I argue that full-blown syntax requires only three building blocks: signs, concatenation, grammar (constraints on concatenation). Functional dependencies between the blocks suggest the four-stage model of syntactic evolution, compatible with several earlier scenarios: (1) signs, (2) increased number of signs, (3) commutative concatenation of signs, (4) grammatical (noncommutative) concatenation of signs. The main claim of the paper is that symbolic reference comprises up to five distinct interpretative correlates: mental imagery, denotation, paradigmatic connotation, syntagmatic connotation, and definition. I show that the correlates form an evolutionary sequence, some stages of which can be aligned with certain stages of syntactic evolution.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

There is a consensus that two features setting natural language apart from non-human animal communication are symbolic reference and semantically compositional syntax (Deacon, 1997; Hauser, Chomsky, & Fitch, 2002; Hurford, 2004; Nowak, Plotkin, & Jansen, 2000). Thus, minimally, evolutionary linguistics should be concerned with modeling specific stages[1] in the evolution of either one or the other. Elaborations of the evolution of syntax are numerous (e.g. Bickerton, 1998, 2000; Dessalles, 2006; Jackendoff, 1999; Johansson, 2006b; Nowak et al., 2000; Wray, 2000), the evolution of symbol, however, has received much less attention, as the defining characteristics of symbol as well as possible stages in its evolution remain largely vague (but see Deacon, 1997). This paper presents a model of the evolution of symbol. Contrary to the common opinion, symbol is not a uniform monolithic phenomenon. I start from a synchronic viewpoint – from analyzing different cognitive correlates that emerge in interpreting symbols. The claim is that the correlates are mental imagery and different types of mental relations (denotation, connotation, etc.). I show that, among symbol's interpretative correlates, it is possible to separate and identify the more fundamental ones from the relative latecomers in the evolution. The evolution of symbolic reference is then aligned with a simple model of the evolution of syntactic compositionality. The latter is derived from the constraints on the evolution of syntax, outlined by Nowak and Komarova (2001), Jackendoff (1999), Johansson (2006b) and others.

2. Symbol: definition and timeline

As 'symbol' is a pretheoretic term, it has to be defined from the outset. Two features that are usually, if not implicitly, held to be characteristic of 'symbol' are the arbitrary nature of reference (a non-necessary link between a form and its meaning) and a used potential for

[1] I define stages as sets of properties, totally ordered by implication. The definition is inductive – higher stages are defined by certain properties that imply other properties that pertain to lower (i.e. more elementary) stages. The lowest stage is defined by certain properties.

spatiotemporally displaced reference[2] (cf. Hockett, 1960). By contrast, 'index' and 'icon' are defined by a necessary and natural link between a form and its meaning (Peirce, 1998). In icons, the link is based on similarity. In indices, it is based on any other necessary relation (e.g. cause-effect or whole-part relation). In practice, various degrees of iconicity as possible, and symbols (e.g. the Christian cross, or onomatopoetic words like *knock*, *bump*, *crash*) may exhibit iconicity as well. Equipped with these definitions, one notices that the word 'symbol' gets colloquially used in very different senses. For example, status symbols (e.g. expensive clothes) have not much in common with linguistic symbols (e.g. words). By the above definitions, only the latter are symbols. The former are indices of wealth and, more generally, success. Although the difference might seem minor, it has fundamental implications on the archaeological evidence for symbolism. For example, one cannot infer symbolism (and by extension, language) from personal ornaments, as the most parsimonious interpretation of personal ornaments is that they are status symbols (Sterelny, 2008). As personal ornaments are costly, their are indices of success. This observation is archaeologically supported by the fact that the objects used in Pleistocene pendants (mainly predator teeth and suitable seashells) are hard to come by. As killing a predator is evolutionarily extremely costly, it is a bona fide indicator of success. Nevertheless, it is exceedingly common to interpret personal ornaments (and to a lesser extent pigment use) as manifestations of symbolism (e.g. Bednarik, 2008b; Chase, 1994; d'Errico et al., 2003). Observe also that personal ornaments do not imply displaced reference, as they bestow status only to their wearers. Thus, personal ornaments per se are not indicators of symbolism (and, by extension, language).

The same holds for pigment use and art (both representational and abstract). It is very difficult to rule out **all** potential utilitarian uses for Middle Paleolithic pigment pieces (medicine, hide preservative, protection from sun and insects, camouflage, startling of prey and conspecifics etc. – Barham, 2002; Sterelny, 2008; Wadley, 2001) but even where this could be done, there remains the possibility that pigment was used because definitive colors were preferred for esthetic or cognitive (salience) reasons. Even non-human species differentiate between esthetic and non-esthetic stimuli and utilize definitive colors as behavioral cues (Watanabe, 2010) and so do children in their first year (Baldwin, 2006). While coloring is probably uniquely human, there is nothing inherently symbolic about it. For color symbolism to be present, a non-natural, non-random and non-availability-specific link between color and object (or color and figure) has to be evidenced. For example,

a brown, black or white foot figure on a cave wall is probably color iconism (reference by similarity), one purple foot figure is probably a chance but ten purple foot figures suggest color symbolism (except if purple was one of the few pigments available to the artist and the others were equally non-iconic, e.g. crimson and green). Thus, extremely specific configurations of archeological and geological evidence are required to attest color symbolism.

Figurative symbolism is generally easier to establish. Importantly, as figurative paintings and sculptures are at least partly iconic, representational art per se does not entail symbolism. Thus, the vast majority of cave paintings and early sculptures (including the Berekhat Ram figurine from 0.25 mya – d'Errico et al., 2003) have to be excluded from possible indicators of symbolism. However, some early examples of figurative symbolism remain. For example, the two half-lion/half-man figurines from 0.031 mya (Conard, 2003) are symbolic, as they exhibit both spatiotemporally displaced and arbitrary reference (spatiotemporally displaced, because a half-lion/half-man has hardly any potential to refer to anything here and now, and arbitrary because the figurines are not representational **as wholes**). By extension, the figurines are reasonable proxies for language (or at least protolanguage).

As for abstract symbolism, one of the first examples of complex code appears on the La Marche antler from 0.016 mya (d'Errico, 1995). Simple codes are probably in evidence since at least the Ishango bone from c. 0.02 mya (Bogoshi, Naidoo, & Webb, 1987; Brooks & Smith, 1987). Both types of codes imply symbolism. Earlier examples tend to fall into a category of 'abstract art' for which doodling is a more parsimonious explanation than symbolism (Bednarik, 1995a; Halverson, 1995). For example, Bednarik's (1995a) paper on concept-mediated marking in Lower Paleolithic makes no allegations as to the symbolicity or even intentionality of the markings. The markings that Bednarik (1995a) analyzed were abstract, and some of them were very similar to those Middle Paleolithic markings found on bone and ocher pieces from Blombos Cave that are claimed to be "irrefutable evidence of symbolic behavior" (d'Errico et al., 2003, p. 4). The claim relies, of course, on a pretheoretic notion of symbol.

In general, cupules are better candidates for symbols than other forms of 'abstract art' because their manufacture is labor intensive, which rules out doodling as well as coincidental configurations of cut marks (Bednarik, 1995b, 2008a, 2008b). At the same time, it is plausible that some cupules (at least those on horizontal surfaces) were used as containers or were unintentional byproducts of other functional activities (e.g. grinding). Although a functional role does not preclude a symbolic use, it makes cupules' status as an evidence of symbolism ambiguous. Nevertheless, cupules are seemingly a later addition to hominins' behavioral repertoire than personal ornaments (at least 0.15 and 0.3 mya, respectively – Bednarik, 2008a, 2008b).

Although emotional attachment is a more parsimonious explanation for burials than symbolism (Sterelny, 2008), grave goods, structures and their configurations can point to symbolism as well. A Neanderthal burial site, La Ferrassie, dated to at least 0.04 mya, suggests a possible symbolic activity (Bednarik, 1995a, 2008a). The find that was

---

[2] The above definition of symbol opens the possibility that, in natural communication, symbolism is uniquely human. For example, while bottlenose dolphins' signature whistles are arbitrary, the whistles' natural potential for spatiotemporally displaced reference is not evident, as each dolphin uses its own distinctive signature whistle (Janik et al., 2006). Although dolphins frequently copy each other signature whistles in the wild, it is not clear whether this qualifies as a reference to third person individuals (which would indicate a used potential for spatiotemporally displaced reference).

perhaps the most implicative of symbolism in La Ferrassie was a limestone slab with 18 cupules (16 of which were arranged in pairs), covering a child's corpse.

By far the oldest evidence for symbolism we have is circumstantial and, ironically, it is evidence for language, the most elusive form of symbolism archaeologically. It is well known that *Homo erectus* crossed at least 19 km of open sea to colonize Flores c. 0.85 mya or earlier (Gibbons, 1998a; Morwood, O'Sullivan, Aziz, & Raza, 1998). It has been (plausibly) argued that watercraft manufacture and navigation entail a level of communication that would be unattainable without (proto)language. Although rafting on flotsam is a possibility (Gibbons, 1998a), it is not a likely one, given *H. erectus*'s ability to manufacture mode 2 tools and successful colonization of much of the Old World (from Africa and Western Europe to Java, China and, possibly, Central Siberia – Ascenzi, Benvenuti, & Segre, 1997; Asfaw et al., 2002; Bar-Yosef & Belfer-Cohenb, 2001; Gibbons, 1998b; Larick et al., 2001; Waters, Forman, & Pierson, 1997). Furthermore, it is now established that *H. erectus* was anatomically capable of (some form of) spoken language, whereas it's evolutionary predecessors australopithecines were not (MacLarnon & Hewitt, 1999; Meyer, Lordkipanidze, & Vekua, 2006). As some of the associated anatomical changes in the vocal tract have evolutionary costs (Fitch, 2000b), it follows that there was selective pressure to evolve enhanced vocalizations. The only reasonable explanation for this pressure is oral communication. Thus, four independent and converging pieces of evidence – *H. erectus*' colonization of Flores (and much of the other Old World); its evolutionarily new anatomic capacity for spoken language; a selective pressure for enhanced vocalizations in *H. erectus*; and oral communication as the most plausible explanation to this pressure – point to some kind of language in *H. erectus* prior to its colonization of Flores. Of course, the evidence is circumstantial but this is always the case in language evolution – the earliest direct evidence for language (an inscription on a wooden tablet – Whitley, 2003) is from less than 0.01 mya.

As *Homo habilis*, the first known *Homo* species (with the earliest occurrence c. 2.3 mya – Spoor et al., 2007), was not scrutinized by MacLarnon and Hewitt (1999), it is possible that this species was anatomically adapted to spoken language as well. According to Tobias (1998), evidence from endocranial casts suggests that *H. habilis* was a speaker. When combined, the archaeological and anatomical evidence indicate that *H. erectus* (and possibly even *H. habilis*) used some kind of (proto)language, the emergence of which could be bracketed between 0.9 and 2.3 mya. At present, this constitutes the earliest evidence of symbolism. Previously, the time when language evolved has been constrained to between 0.5 and 5 mya (Johansson, 2006a).

3. The hierarchical nature of symbolic reference

Terrence Deacon has argued for a hierarchical nature of symbolic reference (Deacon, 1997). Contrary to the commonly held view of symbol (Bickerton, 2003; Christiansen & Kirby, 2003), Deacon's symbols are not

defined by arbitrary reference (in fact, he suggests that indices may be arbitrary as well) but by a closed group of transformations that links them in order to refer:

> "/—/ words need to be in context with other words, in phrases and sentences, in order to have any determinate reference. /—/ Symbolic reference derives from *combinatorial* possibilities and impossibilities, and we therefore depend on combinations both to discover it (during learning) and to make sense of it (during communication)" (Deacon, 1997, p. 83).

> "Thus, not only is symbolic reference a distributed relationship, so is its mnemonic support. /—/ How else could the many thousands of different words we use every day be retrieved so rapidly and effortlessly during the act of speaking or listening?" (Deacon, 1997, p. 97).

Deacon's definition of symbol confines it to human communication systems only. Arbitrary reference, on the other hand, is not uniquely human. Whether vervet monkey alarm calls are arbitrary or not is controversial – at best it makes sense to say that there is an extent to which they are arbitrary (Cheney & Seyfarth, 2005; Deacon, 1997; Owren & Rendall, 2001). But assuming they are arbitrary, there are other aspects that distinguish them from NL symbols (expressions). First, they seem to be produced involuntarily. Second, different sexes produce different alarm calls (Owren & Rendall, 2001, p. 64). Similarly, although European starlings are excellent mimics and can adopt new motifs from their acoustic environment into their songs, the variability is non-arbitrary in the sense that (1) the song structure is preserved (Chaiken, 2004), and (2) it serves to prevent receiver boredom or habituation (cf. Rendall, Owren, & Ryan, 2009). If we generalize the point by comparing the use of (possibly) arbitrary signs in six species (Chaiken, Böhner, & Marler, 1993; Cheney & Seyfarth, 1990; Eens, Pinxten, & Verheyen, 1992; Hurford, 2004; Janik, Sayigh, & Wells, 2006; Pepperberg, 1987, 2000; Savage-Rumbaugh, Shanker, & Taylor, 1998; Seyfarth, Cheney, & Marler, 1980), we get Table 1.

The table informs us that a difference between human and non-human communication is not in arbitrary, learned, or voluntary use of signs.[3] Nor is it the added condition that

---

[3] At this point, it might be useful to compare two best-known models of sign. Peirce's (1955 [1910]) model is Sign Vehicle – Object – Interpretant, Saussure's (1983 [1916]) Signifier-Signified. Given the sequence $O_1–S_1–I_1(O_2)–S_2–I_2$ (S(ign Vehicle), O(bject), I(nterpretant); $I_1(O_2)$ means I_1 viewed as O_2), we have 2 Peirce's signs ($O_1–S_1–I_1$, $(O_2)–S_2–I_2$) but up to 4 Saussure's signs ($O_1–S_1$, $S_1–I_1$, $(O_2)–S_2$, $S_2–I_2$), i.e. any Peirce's sign can be converted to Saussure's signs (with S corresponding to signifier and O and/or I to signified). However, not all Saussure's signs can be converted to Peirce's signs. Thus, Saussure's analysis picks out signs where Peirce's cannot but the inverse does not hold. One reason is that Saussure's sign is more elementary. Another is that Peirce's sign is based on communication, which can be restricted. Some communicative contexts lack objects while others lack interpretants but signifier and signified are always present. For example, the reader of a book (esp. of fiction) has, more often than not, no access to the object, i.e. the author's interpretation. He can interpret it but the result is another interpretant, not the object in the communication that the book (presumably) constitutes. Analogously, an encrypted message lacks an interpretant for a casual observer. In both contexts, we have Saussure's signs (in Peirce's terms, S–I and S–O, resp.) but no Peirce's sign, i.e. the former is more universal.

Table 1
The use of arbitrary signs in six species.

Species:	VM	ES	Gray parrot	Bonobo	BD	Human
Unit of comm.:	AC	Song	Word	Lexigram	SW	NL exp.
Arbitrary	?	?	+	+	+	+
Learned	*	+	+	+	+	+
Voluntary	–	?	+	+	+	+
Natural	+	+	–	–	+	+
Disp. ref.	–	–	?	+	?	+
Phys. uncount.	–	–	–	–	–	+

VM: vervet monkey; ES: European starling; BD: bottlenose dolphin; unit of comm.: unit of communication; AC: alarm call; SW: signature whistle; NL exp.: natural language expression; *: the functions of vervets' calls seem to be learned while the forms are innate; disp. ref.: used potential for spatiotemporally displaced reference; phys. uncount.: physically uncountable number of communication units.

Observe that the word 'uncountable' is not used here in its common set-theoretic sense (to mark the difference between countably and uncountably infinite sets). On the contrary: I claim that the number of possible NL expressions is finite but unbounded – in the sense that one can always add another **up to a certain finite limit that is never reached**. The finity of NL comes from viewing language as a physical computational system rather than a purely mathematical object. As the spacetime that can support physical computational systems is finite (Krauss & Starkman, 2000), physical computational systems cannot, differently from mathematical objects, accommodate infinity. Thus, language is a physically uncountable finite set – it has a finite number of elements (NL expressions) that cannot be physically counted (Luuk & Luuk, 2011). Other physically uncountable finite sets include, e.g. the number of atoms in Jupiter or the Solar System at a given moment etc.

the signs must be part of the species' natural repertoire. Theoretically, a difference could be in naturally occurring spatiotemporally displaced reference (Janik et al. (2006) did not investigate this possibility with bottlenose dolphins).

However, there is a clear-cut difference in the number of signs used. The use of NL expressions by humans is the only natural communication system with a physically uncountable number of signs (Chomsky, 1995; Hauser, Barner, & O'Donnell, 2007; Hauser et al., 2002). (The concession 'natural' has to be made as humans have contrived such artificial systems as well.) This enormous difference between human and non-human communication begs an explanation. While an enhanced mnemonic support, related to the distributed nature of symbolic reference and to almost connectedness in NL signal–signal associations (Deacon, 1997; Ferrer-i-Cancho, Riordan, & Bollobas, 2005) may explain it partly, this does not say much about the conceptual structure required to support these features. One explanation is that non-humans may lack the capacity for mental time travel[4] (Suddendorf & Corballis, 2007; but see Correia, Dickinson, & Clayton, 2007; Osvath, 2009). Other possibly uniquely human prerequisites for this uncountability would be, e.g. relational reinterpretation (representing higher-order relations between perceptual relations – Penn, Holyoak, & Povinelli, 2008) or theory of mind (the ability to attribute mental states to oneself and others – Premack & Woodruff, 1978). It is also proposed that a neurally implemented capacity for syntactic recursion differentiates humans from non-

humans (Chomsky, 2010; Hauser et al., 2002). There are several hypotheses about human cognitive preadaptations for an increased (or even physically uncountable[5]) number of signs, and at least one of them might hold.

The adaptivity of a communication system with uncountable number of meaningful signs lies in detailed context-independent information. Importantly, information about past and future events, other contexts and places, both real and imaginary, is not adaptive per se. However, detailed context-independent information is necessary for the group's elaborate and coordinated behavioral planning, which is adaptive. [6] In addition, a communication system with uncountable number of signs is necessary for recording the group's past experience (e.g. in songs), which in turn supports coordinated behavioral planning. It is not incidental that these two aspects (recording the group's experience, and the group's elaborate and coordinated behavioral planning) together define culture. Boyd and Richerson (1982, 1990) have argued for group selection based on cultural variation over group selection on genetic variation.

Below I attempt to pin down the hierarchical nature of symbolic reference in more linguistic and/or semiotic terms. As mentioned in the Introduction, there is a consensus on two features distinguishing language from non-human communication systems: semantically compositional syntax and learned symbolic reference (e.g. Deacon, 1997; Hauser et al., 2002; Hurford, 2004). Both of these features boil down to connectedness in signal–signal association (Ferrer-i-Cancho et al., 2005). Many non-human species, on the other hand, seem to be close to a perfect communication system – perfect in the sense that every object is mapped to a distinctive signal. Thus, I submit that the main difference between linguistic and non-human signs is that in the latter the referent[7] is pre-determined, whereas in the former it is uniquely (and frequently, in parallel) determined by distinct interpretative correlates. These correlates are: 1. mental imagery, 2. denotation (isolating a possible referent "in the world"), 3. paradigmatic connotation (logico-categorical relations between concepts), 4. syntagmatic connotation (logico-grammatical relations between concepts), and 5. definition (describing the referent with other symbols). All these interpretative correlates have been acknowledged in the literature before but, to my knowledge, they have not been treated together in a systematic fashion. The overall

---

[4] Spatiotemporally displaced reference implies mental time travel (the ability to mentally project oneself forwards or backwards in time) but not vice versa.

[5] Although it may be argued that the uncountability relies not so much on cognitive abilities as on syntax, cognitive processing must support the syntactically induced uncountability.

[6] In language evolution, as opposed to the evolution of cognitive abilities underlying language, some degree of group selection must be assumed. It is not very adaptive to talk to oneself. Thus, the very fact that language has evolved is an evidence of group selection. On the other hand, the underlying cognitive abilities were probably adaptive for the individual (e.g. in behavioral planning) and thus, a target of individual selection.

[7] I use the word 'referent' in the broadest possible sense here: that which is referred to by a sign. I am well aware that this is unorthodox, as philosophers frequently conflate reference with denotation (Hurford, 1998). However, Cambridge Encyclopedia of Philosophy cites reference as a more general term than denotation (Nute, 1995).

amount of attention they have received varies from tremendous (denotation, definition) to scant (syntagmatic connotation). I can only suspect that the authors' frameworks, convictions and/or limited knowledge of the subject have always barred some of them (e.g. mental imagery or syntagmatic connotation) from their sight. For example, formal semantics is programmatically (and probably already terminologically) blind to all the other correlates except denotation (extension) and definition (intension). [8]

With these five interpretative correlates at our disposal, we seldom, if ever, fail to interpret a symbol known to us. The identity of the referent of a symbol X is hardly ever an issue, since a symbol may have different referents under different interpretations. Moreover, being, in each interpretation, **distributed** between a subset of these correlates ensures a degree of precision which is usually sufficient for referent recognition. In this sense, any set of the five interpretative correlates evaluated in an instance of interpretation of a symbol is its reference. This constitutes the vast difference between the human symbolic capacity and the non-human use of symbols (if there is any – this depends on our definition of 'symbol' – cf. Bickerton, 2003; Christiansen & Kirby, 2003; Deacon, 1997). One obvious advantage of the human symbolic capacity is the flexibility it bestows to our understanding (which, however, comes with a tradeoff in interpretation unambiguity). Another, as pointed out by Deacon (1997), is its superior mnemonic support. The produced referents of a symbol are sampled in memory, and serve as aids for future interpretations for not just that one particular symbol but for a set of related symbols. Plausibly, the human symbolic capacity and the syntactic compositionality of language display some dependencies which allow to align the evolutionary scale of one with that of the other. Below I will elaborate on the interpretative correlates in more detail.

3.1. S–S and S–R links

Denotation, connotation and definition are more complex than S–S and S–R links. They subsume S–S and S–R links but the links do not imply denotation, connotation or definition. The main difference between the two groups is that the links do not have to be conceptually mediated. As the links are available to a vast number of species (including humans), they constitute a distinct primitive mechanism that evolution can build on. Thus, the links (both conditioned and unconditioned) predate all the other correlates. For example, the roundworm *Caenorhabditis elegans* has a simple nervous system comprised of only 302 neurons and is capable of several types of associative learning (Ishihara et al., 2002).

Vervet monkeys provide another possible example. Although their alarm calls denote four different predator classes (snakes, birds of prey, large cats, primates – Hauser, 1997), this may be obvious to us but not to the monkeys. I suggest that emotional signals getting associated with situation-specific flee responses would be a more parsimonious interpretation of the origin of vervets' alarm calls. The argument is as follows. The only thing that needs explaining is different signals for different predator contexts. As having different signals for different predator contexts enhances vervets' fitness, the differentiation of signals can be attributed to natural selection. For this, prey individuals need to provide only two types of links for natural selection to work upon: the one between predation pattern and flee pattern, and the other between flee pattern and call pattern. As for the link mechanism, conditioned S–R would be sufficient for both. Crucially, a conscious link between predation pattern and call pattern, available to humans, is unnecessary. It has been argued that humans are unique in being able to form 'stimulus equivalence' classes of abstract and unrelated stimuli, e.g. if taught to match A with B and B with C, they will spontaneously match A with C by transitivity (Dickins, 2003; but see Eichenbaum, 2000). I suggest that even if vervets had such a capacity, it would not be required by natural selection for selecting for distinct alarm calls associated with different flee responses. Vervets' putative arbitrary reference to things or events "in the world" (denotation) is thus a non-parsimonious alternative to conditioned response (cf. Soltis, 2009).

3.2. Mental imagery and denotation

Mental imagery is primarily accessible to introspection, an arguably repudiated source of information. However, mental imagery is also a prolific field of study in cognitive science. Despite its name, it is generally agreed upon that mental imagery is not confined to one (i.e. visual) sensory modality (Belardinelli et al., 2004; Cooper, 1995; Pylyshyn, 2003). However, there is some debate on whether mental imagery is per se polysensory or comes in different modalities similar to sensory modalities (visual, auditory, haptic, etc. – Bensafi et al., 2003; Klatzky, Lederman, & Matula, 1991; Reisberg, Wilson, & Smith, 1991).

It is well established that both spontaneous and induced mental imagery have positive effects on text comprehension (Anderson & Kulhavy, 1972; Gambrell & Bales, 1986; Long, Winograd, & Bridge, 1989; Sadoski, 1983, 1985; Thibaut, Rondal, & Kaens, 1995). Language concreteness, its capacity to evoke mental images, has been shown to be one of the most effective determiners of comprehension and learning (Sadoski, Goetz, & Fritz, 1993; Wharton, 1980). There is ample evidence that language comprehension involves the activation of mental imagery of the content of utterances (Barsalou, Simmons, Barbey, & Wilson, 2003; Just, Newman, Keller, McEleney, & Carpenter, 2004; Richardson, Spivey, McRae, & Barsalou, 2003; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). In sum, there is sufficient evidence to include mental imagery among symbol's interpretative correlates.

Moreover, insofar as there is any **substance** to conceptualization, it is probably reducible to an equivalent of mental imagery. This seems necessary, as all the other correlates describe mental relations. As any relation, even an abstract and complex one, would have to be a relation between **something**, it follows that mental imagery is the

---

[8] Alternatively, intension is sometimes defined negatively as "anything referred to by the sign that is not (part of) its extension", which, if conflated with definition, guarantees the blindness.

nearest equivalent for the substance required for conceptualization. It also follows that all conceptually mediated mental relations (denotation, both types of connotations, definition) should presuppose conceptual content, i.e. mental imagery. Only S–S and S–R links, as preconceptual mental relations, can do without. For example, the denotation of the word *rat* requires imagining (or retrieving a memory image of) a specific creature 'in the world', otherwise we could not understand the word in the way we do, irrespective of whether there is a rat in sight or not.

Plausibly, the same holds for all other conceptually mediated relations. Connotations and definitions frequently (although perhaps not necessarily) activate mental images, and even those that do not evoke mental images directly can recruit mental imagery at various stages of their composition. For example, the phrase *furiously sleeping green idea*, standing for a concept without denotation, connotes other concepts, e.g. 'green', which in turn can recruit mental imagery. If the same holds for 'furiously' (which is likely), one can arrive at a crude approximation of the concept of *furiously sleeping green idea* by combining the mental images of these two constituents. The relevance of the concept would be further reinforced by the correct way of combining the mental images, viz. by complementing the semantic frame of the phrase [x y-ing in z way] with mental images of 'green' and 'furiously' for x and z, respectively. [9] Observe that the resulting concept, while crude and subjective, could, if memorized, be used to uniquely identify *furiously sleeping green idea* – an identification motivated by the lack of other unique identifiers, viz. denotation and definition. Moreover, although the resulting concept and its uniqueness are subjective, the concept's derivation is fully objective, i.e. a certain similarity is expected between the concepts of *furiously sleeping green idea* of any two people. Of course, the derivation (and, correspondingly, the concept) would be very similar to that of, e.g. *furiously resting green idea*, *furiously resting green thought*, *furiously reposing green slumber* etc. This should not pose a problem, as the type of combination is sufficiently rare for a subjectively unique identification in normal (i.e. metalinguistically untampered) conditions. Thus, and contrary to what Chomsky (1971 [1957], p. 15) has assumed, the fact that an expression is nonsensical (lacks denotation or extension 'in the world') does not imply that it is meaningless (as meaning can be induced by mental imagery, connotation or definition as well).

Importantly, mental imagery is capable only of approximating objects rather than modeling them exactly. But this is so for a good reason. For if mental imagery were to model objects with perfect accuracy, we would have trouble distinguishing it from them, i.e. we would be prone to hallucinations, which would be very maladaptive. Vagueness, but also extreme flexibility are the prime features of mental imagery. At the same time, mental imagery has also strict limitations, e.g. confinement up to 4-dimensional spacetime in the domain of geometry.

In sum, mental imagery is **autonomous** conceptualization, affected primarily by conceptualization-internal constraints, while the other correlates depend heavily on different external conditions (e.g. the existence of a corresponding entity 'in the world' in the case of denotation). Thus, I hypothesize that an approximation of mental imagery is a prerequisite of conceptual structure, and that the species lacking it are incapable of conceptualization, lack concepts and conceptual structure. Presumably, all the species we have discussed so far have conceptual structure and thus some access to an approximation of mental imagery. While it is true that an approximation of mental imagery in non-humans can be only inferred and is difficult to prove, bonobo's play with an imaginary object (Savage-Rumbaugh et al., 1998) is an example of evidence of the species capacity for mental imagery. More generally, Tomasello and Call (1997) conclude on this topic that evidence for mental imagery in ape play exists but is controversial. However, some mental rotation experiments have provided a more rigorous evidence (Mauck & Dehnhardt, 1997; Vauclair, Fagot, & Hopkins, 1993).

3.3. Connotation and definition

Paradigmatic connotation pertains to lexicon and reflects logico-categorical relations between concepts stored in long-term memory. For instance, *white* is a paradigmatic connotation of *snow*. Paradigmatic connotation relies on a well-known signification phenomenon, metonymy (reference by proximity or whole-part relation). As there cannot be predicates without arguments, paradigmatic connotation pertains also to predicate-argument structure – at least in modern language, where predicate paradigmatically connotes argument and vice versa. [10] This might not have been the case in protolanguage, which, in all likelihood, had initially only arguments, i.e. noun-like entities (Heine & Kuteva, 2007; Luuk, 2009, 2010). However, paradigmatic connotation is a much more general characteristic of language than predicate-argument structure (see Section 5).

In language, syntagmatic connotation pertains to grammar and reflects the logico-grammatical syntagmatic relations a concept steps into. For instance, *if (x) then (y)* or *(y) if (x)* is a syntagmatic connotation of *if*. As Quine (1960) has observed, a word like *sake* can be learned only contextually – i.e. in our terminology, presupposes syntagmatic connotation for a referent. The same holds for all conjunctions and adpositions. Obviously, there cannot be any syntagmatic connotation before the emergence of grammar (e.g. in the form of the first word order constraint). More generally, syntagmatic connotation is a byproduct of fixed sign order. As it is theoretically possible (albeit evolutionarily unlikely) to have, e.g. affixes (morphemes with fixed order) before word order constraints, syntagmatic connotation does not imply word order constraints. However, as morphemes are signs

---

[9] On frame semantics, see Fillmore (1976).

[10] This type of paradigmatic connotation can be more detailed than just described. For example, verbs paradigmatically connote their specific argument roles (Agent, Patient, Beneficiary, Goal etc.).

(form–function pairs), and all sign order constraints are grammatical constraints in language, there is an equivalence relation between syntagmatic connotation and grammar **in language**. However, syntagmatic connotation is a correlate of fixed order of not just linguistic but any signs. For example, formulas of cause-effect relation, implication ($x \rightarrow y$) and arithmetic ($x + y = z$) exhibit syntagmatic connotation. Of course, formulas like these are derivative of language (as is probably our arithmetic capacity in general – Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Hurford, 1990 [1987]; Nieder, 2005).

Definition is the higher order correlate that must refer to other correlates in order to be understood. Suppose that a definition M of an unknown symbol N contains an unknown symbol O, explained in definition P containing an unknown symbol Q. In order to be understood, M must resort to mental imagery, denotation, syntagmatic connotation, paradigmatic connotation, or any combination of these. In order to be unambiguous, it is necessary (but not sufficient) that M be ultimately reducible to a set of known symbols.

4. The evolution of syntax

In describing the evolution of syntax, the following scenario is widely agreed upon: preadaptations [1] $\rightarrow$ protolanguage($\rightarrow$ preadaptations [2]?) $\rightarrow$ syntactic language.

Certain preadaptations [1] were necessary for protolanguage to emerge.[11] Anatomical preadaptations included changes in the brain anatomy; anatomical preadaptations for speech included changes in middle and/or inner ear anatomy, an enhanced thoracic innervation and a reconfiguration of the tongue and vocal tract (Boë et al., 2007; Fitch, 2000a; MacLarnon & Hewitt, 1999; Martínez et al., 2004; Wynn, 1998). Depending on one's theoretical standpoint, cognitive preadaptations could have been, e.g. theory of mind, joint reference, double scope-blending, episodic memory, conceptual embedding, mental time travel, transitive inference, semantic synthesis, relational redescription and relational reinterpretation (Bruner, 1975; Call & Tomasello, 2008; Dessalles, 2008; Fauconnier & Turner, 2008; Karmiloff-Smith, 1992; Luuk & Luuk, 2008; Penn et al., 2008; Penn & Povinelli, 2007; Piaget, 1999 [1928]; Suddendorf & Corballis, 2007; Tulving, 2005). As protolanguage is, essentially, a language without syntax, it refers to either a holophrastic (Wray, 1998, 2000) or commutatively concatenated (Jackendoff, 1999) language. Although protolanguage is culturally downgraded, both syntactic language and protolanguage are suited for communication.[12] The main difference is that, in syntactic language, linguistic form constrains interpretation better than in protolanguage. As to why protolanguage was eventually substituted with syntactic language, the most plausible explanation is that the transition increased expressivity (cf. Nowak & Krakauer, 1999; Nowak et al., 2000). It is unknown whether it was a solely technological innovation or required some additional anatomical and cognitive preadaptations [2]. However, see Hauser et al. (2002) and Chomsky (2010) for the proposal that the preadaptations included a neurally implemented recursion.

The evolution toward syntax started with an operation on elements (signs) of protolanguage. There are only two possible candidates for this operation: fractionation and concatenation. The hypothesis that posits fractionation as the first operation is that of a 'holistic protolanguage' (e.g. Wray, 1998, 2000). The hypothesis suggests that holistic utterances of protolanguage were, in the advent of syntax, fractured into distinct words. The main counterargument to this hypothesis, supported by Johansson's (2008) calculation, is that the structure of the holistic utterances would have been too ambiguous to yield distinct form-meaning pairs (i.e. words) for the fractioning (Bickerton, 2003; Johansson, 2008; Tallerman, 2007). Thus, the alternative hypothesis, that of concatenation as the first operation, would have to be true. The majority of models of the evolution of syntax subsume this hypothesis (e.g. Dessalles, 2006; Jackendoff, 1999; Johansson, 2006b; Nowak, Komarova, & Niyogi, 2001; Nowak & Krakauer, 1999; Nowak et al., 2000). For example, in Jackendoff's (1999, p. 273) model of language evolution, concatenation precedes the "use of symbol position to convey basic semantic relationships", which implies grammar (cf. Dessalles, 2006; Johansson, 2006b). An increased number of signs are attested as a payoff condition for compound signals (Christiansen & Kirby, 2003; Jackendoff, 1999; Nowak & Komarova, 2001). Thus, one arrives at Table 2.

The table shows the logical and temporal succession of stages of the evolution of syntax. The stages are ordered vertically with each stage describing the achieved state (e.g. 'commutative concatenation of signs' – concatenation of signs irrespective of their order). Table 2 is hierarchical, i.e. at each stage the conditions stipulated by the previous stages (above them) apply as well. This accords with the

---

[11] Alternatively, some of the putative preadaptations could be ascribed to adaptations **for language** in the course of coevolution of language and genetic endowment (cf. Christiansen & Chater, 2008; Fitch, 2008; Hurford, 2008). However, it is still necessary to posit certain preadaptations to get language evolution off the ground.

[12] For example, it has been proposed that the difference between protolanguage and syntactic language is roughly of the order of that between pidgin and creole (Bickerton, 1990; Givón, 1998).

Table 2
The evolution of syntax.

(1) Signs
(2) Increased number of signs
(3) Commutative concatenation of signs
(4) Grammar – noncommutative concatenation of signs

evolutionary principle of building on rather than expunging the earlier stages. The timing of the stages is relative, i.e. the intervals between them might not be equal.

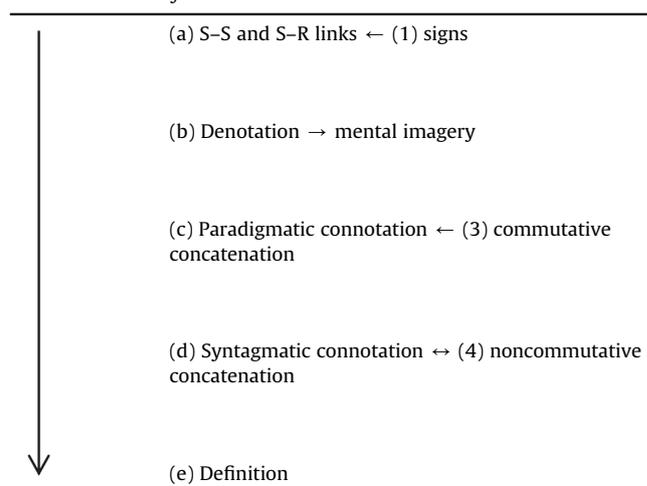## 5. The evolution of symbolic reference

For communication to work, stage (3) (commutative concatenation of signs) must have logico-categorical relationships (i.e. paradigmatic connotation) between the signs. Assuming the availability of suitable stem categories at stage (3), the simplest solution would be to concatenate arguments and predicates, as in [*man go*]. Alternatively, with a categorially uniform stem choice, a solution would be to concatenate different semantic roles, as in [*man forest*], interpreted as 'man go to forest'. In any way, paradigmatic connotation had to evolve sometime before stage (3). At the same time, syntagmatic connotation, the main correlate for interpreting words like *sake, if, which, that* etc., cannot evolve before the first grammatical constraint is at place, i.e. before stage (4). Thus, in deriving the course of the evolution of symbol, we arrive at Table 3.

There are five stages in this scale. As in Table 2, the succession of the stages is logical and temporal; the scale is hierarchical so that at all stages the conditions specified by the previous stages apply as well; the timing of stages is relative, not absolute. As argued at length above, it is parsimonious to assume that the scale begins with an ability to form S–S and S–R links (e.g. conditioned S–R), accordingly, the mental links are the starting point. Denotation depends on mental imagery but perhaps not vice versa (cf. Section 3.2). However, a one-way dependency is difficult to establish here, as little is known about these correlates in non-humans who are the prime source for investigating protosymbolic reference. We know only a few scattered facts, such as that baboons and sea lions have the capacity for mental imagery and that dolphins may be able to denote conspecifics (Janik et al., 2006; Mauck & Dehnhardt, 1997; Vauclair et al., 1993). The scale ends with definition which, normally, should require a full-blown syntax, but at the very least requires a fully developed symbolic reference. As I have argued, paradigmatic connotation is a maintenance condition of stage (3) in Table 2. Thus, at this point, Table 3 can be aligned with Table 2, with stage (3) having stage (c) as its prerequisite. Stage (a), S–S and S–R links are a precondition of stage (1), so at this point Tables 2 and 3 can be aligned as well. Finally, there is an equivalence relation between (d) syntagmatic connotation and (4) noncommutative concatenation. Comparative evidence suggests that certain non-human species have attained at least stage (b) (Janik et al., 2006; Mauck & Dehnhardt, 1997; Osvath, 2009; Savage-Rumbaugh, 1986; Savage-Rumbaugh et al., 1998; Vauclair et al., 1993).

There is some independent evidence as to the reality of these stages. In modeling a general communication framework based on Zipf's law (which holds for NL) with graph theory, and without referring to the interpretative correlates (aside from distinguishing denotation from paradigmatic connotation), Ferrer-i-Cancho et al. (2005) arrived at the model that corresponds exactly to stage (c) in Table 3, i.e. a concatenable set of signs with denotation and paradigmatic connotation but without syntagmatic

Table 3
The evolution of symbolic reference.

(a) S–S and S–R links ← (1) signs
(b) Denotation → mental imagery
(c) Paradigmatic connotation ← (3) commutative concatenation
(d) Syntagmatic connotation ↔ (4) noncommutative concatenation
(e) Definition

→, ←: implication; ↔: equivalence.

connotation and noncommutative concatenation. At the time of its publication, this result was totally unexpected and theoretically unexplained. As Ferrer-i-Cancho et al. (2005, p. 564) wrote: "While researchers are divided when considering syntax /−/ or symbolic reference /−/ as the essence of human language, we hypothesize that syntax and forms of reference higher than mere signal-object associations are two sides of the same coin, i.e. connectedness in signal-signal associations" (cf. Sole, 2005). The authors described their model as "a proto-grammar, from which full human languages can easily evolve" (Ferrer-i-Cancho et al., 2005, p. 562).

6. Conclusion

This paper attempted to gain an interdisciplinary insight into the evolution of symbol. It arrived at the following, rather eclectic but apparently coherent results. Of these results, 2, 5 and 6 are almost entirely new. The rest build more on pre-existing, sometimes implicit notions in arriving at the formulations (in Tables 1–3 and the definition of symbol) that are at least partly novel.

1. Symbol is defined by the arbitrary nature of reference (a non-necessary link between a form and its meaning) and a used potential for spatiotemporally displaced reference.
2. Archaeological and anatomical evidence brackets the earliest evidence for symbolism between 0.9 and 2.3 mya.
3. The comparison of the use of arbitrary signs in six species (vervet monkeys, European starlings, gray parrots, bonobos, bottlenose dolphins and humans) shows that a difference between human and non-human communication is not in arbitrary, learned, voluntary or natural signs but in the number of signs used. Language is the only natural communication system with a physically uncountable number of signs (Table 1).
4. Contrary to some claims (Cheney & Seyfarth, 2005; Hauser, 1997; Seyfarth et al., 2010), reference to distinct predator classes is not required to explain the emergence of differentiated natural alarm signal systems (e.g.

vervets' calls). Conditioned response and natural selection are sufficient to explain it.

5. Symbolic reference can be analyzed in terms of five interpretative correlates: mental imagery, denotation, paradigmatic connotation, syntagmatic connotation, and definition. I hypothesize that the correlates form a derivational and evolutionary sequence, given in the five-stage evolutionary scale of symbolic reference (Table 3): (a) S–S and S–R links, (b) denotation and mental imagery, (c) paradigmatic connotation, (d) syntagmatic connotation, (e) definition. I also hypothesize that denotation implies (an approximation of) mental imagery.

6. At three stages, the evolution of symbol (Table 3) can be aligned with the evolution of syntax (Table 2): (1) signs, (2) increased number of signs, (3) commutative concatenation of signs, (4) grammatical (noncommutative) concatenation of signs. This scale of syntactic evolution is compatible with several well-known scenarios of the evolution of syntax (e.g. Dessalles, 2006; Diessel & Tomasello, 2005; Jackendoff, 1999; Johansson, 2006b; Nowak & Komarova, 2001; Nowak & Krakauer, 1999; Nowak et al., 2000). I hypothesize that the evolutions of syntax and symbol are related in the following way: (1) signs imply (a) S–S and S–R links, (3) commutative concatenation implies (c) paradigmatic connotation, and (4) noncommutative concatenation and (d) syntagmatic connotation are equivalent.

Acknowledgements

References

Anderson, R. C., & Kulhavy, R. W. (1972). Imagery and prose learning. *Journal of Educational Psychology, 63*, 242–243.

Ascenzi, A., Benvenuti, A., & Segre, A. G. (1997). On the paleopathologic findings exhibited by the late *Homo erectus* of Ceprano, Italy. *Human Evolution, 12*(3), 189–196.

Asfaw, B., Gilbert, W. H., Beyene, Y., Hart, W. K., Renne, P. R., WoldeGabriel, G., et al. (2002). Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature, 416*, 317–320.

Baldwin, J. M. (2006). *Handbook of psychology: Senses and intellect* (2nd [orig. 1890] ed.). Whitefish, MT: Kessinger Publishing.

Bar-Yosef, O., & Belfer-Cohenb, A. (2001). From Africa to Eurasia – early dispersals. *Quaternary International, 75*(1), 19–28.

Barham, L. S. (2002). Systematic pigment use in the Middle Pleistocene of South-Central Africa. *Current Anthropology, 43*(1), 181–190.

Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*, 84–91.

Bednarik, R. G. (1995a). Concept-mediated marking in the Lower Palaeolithic. *Current Anthropology, 36*(4), 605–616.

Bednarik, R. G. (1995b). Reply to comments on Bednarik's "Concept-mediated marking in the Lower Palaeolithic". *Current Anthropology, 36*(4), 626–630.

Bednarik, R. G. (2008a). Cupules. *Rock Art Research, 25*(1), 61–100.

Bednarik, R. G. (2008b). The origins of symboling. *Signs, 2*, 82–113.

Belardinelli, M. O., Matteo, R. D., Gratta, C. D., Nicola, A. D., Ferretti, A., Tartaro, A., et al. (2004). Intermodal sensory image generation: an fMRI analysis. *The European Journal of Cognitive Psychology, 16*(5), 729–752.

Bensafi, M., Porter, J., Pouliot, S., Mainland, J., Johnson, B., Zelano, C., et al. (2003). Olfactomotor activity during imagery mimics that during perception. *Nature Neuroscience, 6*(11), 1142–1144.

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

Bickerton, D. (1998). Catastrophic evolution: the case for a single step from protolanguage to full human language. In J. R. Hurford, M. Studdert-Kennedy, & C. Knight (Eds.), *Approaches to the evolution of language: Social and cognitive bases* (pp. 341–358). Cambridge: Cambridge University Press.

Bickerton, D. (2000). How protolanguage became language. In C. Knight, J. R. Hurford, & M. Studdert-Kennedy (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 264–284). Cambridge: Cambridge University Press.

Bickerton, D. (2003). Symbol and structure: a comprehensive framework for language evolution. In M. H. Christiansen, & S. Kirby (Eds.), *Language evolution: The states of the art* (pp. 77–93). Oxford: Oxford University Press.

Boë, L.-J., Heim, J.-L., Honda, K., Maeda, S., Badin, P., & Abry, C. (2007). The vocal tract of newborn humans and Neanderthals: acoustic capabilities and consequences for the debate on the origin of language. A reply to Lieberman (2007a). *Journal of Phonetics, 35*, 564–581.

Bogoshi, J., Naidoo, K., & Webb, J. (1987). The oldest mathematical artifact. *Mathematical Gazette, 71*, 458.

Boyd, R., & Richerson, P. J. (1982). Cultural transmission and the evolution of cooperative behavior. *Human Ecology, 10*, 325–351.

Boyd, R., & Richerson, P. J. (1990). Group selection among alternative evolutionary stable strategies. *Journal of Theoretical Biology, 145*, 331–342.

Brooks, A. S., & Smith, C. C. (1987). Ishango revisited: new age determinations and cultural interpretations. *African Archaeological Review, 5*(1), 65–78.

Bruner, J. S. (1975). From communication to language – a psychological perspective. *Cognition, 3*(3), 255–287.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

Chaiken, M. (2004). Starling song: an open-ended learner. In P. Marler, & H. Slabbekoorn (Eds.), *Nature's music: The science of birdsong* (pp. 83). Amsterdam etc.: Elsevier.

Chaiken, M., Böhner, J., & Marler, P. (1993). Song acquisition in European starlings, *Sturnus vulgaris*: a comparison of the songs of live-tutored, tape-tutored and wild-caught males. *Animal Behaviour, 46*, 1079–1090.

Chase, P. G. (1994). On symbols and the Palaeolithic. *Current Anthropology, 35*(5), 627–629.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.

Cheney, D., & Seyfarth, R. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review, 22*(2–4), 135–159.

Chomsky, N. (1971 [1957]). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1995). *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (2010). Some simple evo-devo theses: how true might they be for language? In R. K. Larson, H. Yamakido, & V. Deprez (Eds.), *Evolution of human language: Biolinguistic perspectives* (pp. 45–61) Cambridge: Cambridge University Press.

Christiansen, M. H., & Chater, N. (2008). Brains, genes, and language evolution: a new synthesis. *Behavioral and Brain Sciences, 31*(5), 537–558.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: consensus and controversies. *Trends in Cognitive Sciences, 7*(7), 300–307.

Conard, N. J. (2003). Palaeolithic ivory sculptures from southwestern Germany and the origins of figurative art. *Nature, 426*, 830–832.

Cooper, L. A. (1995). Varieties of visual representation: how are we to analyze the concept of mental image? *Neuropsychologia, 33*, 1575–1582.

Correia, S. P., Dickinson, A., & Clayton, N. S. (2007). Western scrub-jays anticipate future needs independently of their current motivational state. *Current Biology, 17*(10), 856–861.

d'Errico, F. (1995). A new model and its implications for the origin of writing: La Marche antler revisited. *Cambridge Archaeological Journal, 5*(2), 163–206.

d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., et al. (2003). Archaeological evidence for the emergence of language, symbolism, and music – an alternative multidisciplinary perspective. *Journal of World Prehistory, 17*(1), 1–70.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: W.W. Norton.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science, 284*(5416), 970–974.

Dessalles, J.-L. (2006). Du protolangage au langage: modèle d'une transition. *Marges Linguistiques, 11*, 142–152.

Dessalles, J.-L. (2008). From metonymy to syntax in the communication of events. *Interaction Studies, 9*(1), 51–65.

Dickins, T. E. (2003). General symbol machines: the first stage in the evolution of symbolic communication. *Evolutionary Psychology, 1*, 192–209.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language, 81*, 1–25.

Eens, M., Pinxten, R., & Verheyen, R. F. (1992). Song learning in captive European starlings, *Sturnus vulgaris. Animal Behaviour, 44*(6), 1131–1143.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience, 1*, 41–50.

Fauconnier, G., & Turner, M. (2008). The origin of language as a product of the evolution of double-scope blending. *Behavioral and Brain Sciences, 31*(5), 520–521.

Ferrer-i-Cancho, R., Riordan, O., & Bollobas, B. (2005). The consequences of Zipf's law for syntax and symbolic reference. *Proceedings of the Royal Society of London B Biological Sciences, 272*(1562), 561–565.

Fillmore, C. J. (1976). Frame semantics and the nature of language. InHarnad, S., Steklis, H. D., & Lancaster, J. (Eds.), (1976). *Annals of the New York Academy of Sciences: Conference on the origin and development of language and speech, Vol. 280* (pp. 20–32).

Fitch, W. T. (2000a). The evolution of speech: a comparative review. *Trends in Cognitive Sciences, 4*(7), 258–267.

Fitch, W. T. (2000b). Vocal production in nonhuman mammals: implications for the evolution of speech. In *Paper presented at the evolution of language conference, Paris.*

Fitch, W. T. (2008). Co-evolution of phylogeny and glossogeny: there is no "logical problem of language evolution". *Behavioral and Brain Sciences, 31*(5), 521–522.

Gambrell, L. B., & Bales, R. J. (1986). Mental imagery and the comprehension-monitoring performance of fourth- and fifth-grade poor readers. *Reading Research Quarterly, 21*, 454–464.

Gibbons, A. (1998a). Ancient island tools suggest *Homo erectus* was a seafarer. *Science, 279*(5357), 1635–1637.

Gibbons, A. (1998b). In China, a handier *Homo erectus. Science, 279*(5357), 1636.

Givón, T. (1998). On the co-evolution of language, mind, and brain. *Evolution of Communication, 2*, 45–116.

Halverson, J. (1995). Comment on Bednarik's "Concept-mediated marking in the lower Palaeolithic". *Current Anthropology, 36*(4), 621.

Hauser, M. D. (1997). *The evolution of communication.* Cambridge, MA: MIT Press/BradfordBooks.

Hauser, M. D., Barner, D., & O'Donnell, T. (2007). Evolutionary linguistics: a new look at an old landscape. *Language, Learning and Development, 3*(2), 101–132.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction.* New York: Oxford University Press.

Hockett, C. F. (1960). The origin of speech. *Scientific American, 203*(3), 89–96.

Hurford, J. R. (1990 [1987]). *Language and number: The emergence of a cognitive system.* Oxford: Basil Blackwell.

Hurford, J. R. (1998). Review of "The symbolic species: the co-evolution of language and the human brain", by Terrence Deacon, 1997. *The Times Literary Supplement, 34.*

Hurford, J. R. (2004). Human uniqueness, learned symbols and recursive thought. *European Review, 12*(4), 551–565.

Hurford, J. R. (2008). Niche-construction, co-evolution, and domain-specificity. *Behavioral and Brain Sciences, 31*(5), 526.

Ishihara, T., Iino, Y., Mohri, A., Mori, I., Gengyo-Ando, K., Mitani, S., et al. (2002). HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in *Caenorhabditis elegans. Cell, 109*(5), 639–649.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences, 3*(7), 272–279.

Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences of the United States of America, 103*(21), 8293–8297.

Johansson, S. (2006a). Constraining the time when language evolved. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The evolution of language: Proceedings of the 6th International Conference (EVOLANG6)* (pp. 152–159). Singapore: World Scientific.

Johansson, S. (2006b). Working backwards from modern language to proto-grammar. In A. Cangelosi, A. D. M. Smith, & K. Smith (Eds.), *The evolution of language: Proceedings of the 6th International Conference on the evolution of language (EVOLANG6)* (pp. 160–167). Singapore: World Scientific.

Johansson, S. (2008). Seeking compositionality in holistic proto-language without substructure: do counter-examples overwhelm the fractionation process? In A. D. M. Smith, K. Smith, & R. Ferrer-i-Cancho (Eds.), *The evolution of language: Proceedings of the 7th International Conference (EVOLANG7)* (pp. 171–178) Singapore: World Scientific.

Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: an fMRI study. *NeuroImage, 21*(1), 112–124.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science.* Cambridge: MIT Press/Bradford Books.

Klatzky, R. L., Lederman, S. J., & Matula, D. E. (1991). Imagined haptic exploration in judgments of object properties. *Journal of Experimental Psychology Learning Memory and Cognition, 17*(2), 314–322.

Krauss, L. M., & Starkman, G. D. (2000). Life, the universe, and nothing: life and death in an ever-expanding universe. *The Astrophysical Journal, 531*, 22–30.

Larick, R., Ciochon, R. L., Zaim, Y., Sudijono, S., Rizal, Y., Aziz, F., et al. (2001). Early Pleistocene 40Ar/39Ar ages for Bapang Formation hominins, Central Jawa, Indonesia. *Proceedings of the National Academy of Sciences of the United States of America, 98*(9), 4866–4871.

Long, S. A., Winograd, P. A., & Bridge, C. A. (1989). The effects of reader and text characteristics on imagery reported during and after reading. *Reading Research Quarterly, 24*(3), 353–372.

Luuk, E. (2009). The noun/verb and predicate/argument structures. *Lingua, 119*(11), 1707–1727.

Luuk, E. (2010). Linguistic arguments probably antedate linguistic predicates. In A. D. M. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language* (pp. 230–237). Singapore: World Scientific.

Luuk, E., & Luuk, H. (2008). Evolutionary framework for the language faculty. In A. D. M. Smith, K. Smith, & R. Ferrer-i-Cancho (Eds.), *The Evolution of language: Proceedings of the 7th International Conference (EVOLANG7).* Singapore: World Scientific.

Luuk, E., & Luuk, H. (2011). The redundancy of recursion and infinity for natural language. *Cognitive Processing, 12*(1), 1–11.

MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: the role of enhanced breathing control. *American Journal of Physical Anthropology, 109*, 341–363.

Martínez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., et al. (2004). Auditory capacities in Middle Pleistocene humans from the Sierra de Atapuerca in Spain. *Proceedings of the National Academy of Sciences of the United States of America, 101*(27), 9976–9981.

Mauck, B., & Dehnhardt, G. (1997). Mental rotation in a California sea lion (*Zalophus californianus*). *Journal of Experimental Biology, 200*(Pt 9), 1309–1316.

Meyer, M. R., Lordkipanidze, D., & Vekua, A. (2006). Evidence for the anatomical capacity for spoken language in *Homo erectus. American Journal of Physical Anthropology Supplement, 42*, 130.

Morwood, M. J., O'Sullivan, P. B., Aziz, F., & Raza, A. (1998). Fission-track ages of stone tools and fossils on the east Indonesian island of Flores. *Nature, 392*, 173–176.

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*(3), 177–190.

Nowak, M. A., & Komarova, N. L. (2001). Towards an evolutionary theory of language. *Trends in Cognitive Sciences, 5*(7), 288–295.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of universal grammar. *Science, 291*(5501), 114–118.

Nowak, M. A., & Krakauer, D. C. (1999). The evolution of language. *Proceedings of the National Academy of Sciences of the United States of America, 96*(14), 8028–8033.

Nowak, M. A., Plotkin, J. B., & Jansen, V. A. (2000). The evolution of syntactic communication. *Nature, 404*(6777), 495–498.

Nute, D. (1995). Intension. In R. Audi (Ed.), *The Cambridge encyclopedia of philosophy* (pp. 379). Cambridge: Cambridge University Press.

Osvath, M. (2009). Spontaneous planning for future stone throwing by a male chimpanzee. *Current Biology, 19*(5), R190–R191.

Owren, M. J., & Rendall, D. (2001). Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology, 10*, 58–71.

Peirce, C. S. (1955 [1910]). Logic as semiotic: the theory of signs. In J. Buchler (Ed.), *Philosophical writings of Peirce* (pp. 98–119). New York: Dover.

Peirce, C. S. (1998). *The Essential Peirce, Vol. 2.* Bloomington, IN: Indiana University Press.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31*(2), 109–130, discussion 130–178.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 731–744.

Pepperberg, I. M. (1987). Acquisition of the same/different concept by an African grey parrot (*Psittacus erithacus*): learning with respect to categories of color, shape, and material. *Animal Learning & Behavior, 15*, 423–432.

Pepperberg, I. M. (2000). *The Alex studies: Cognitive and communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.

Piaget, J. (1999 [1928]). *Judgement and reasoning in the child*. London: Routledge.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515–526.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, Mass.; London: MIT Press.

Quine, W. V. (1960). *Word and object*. Cambridge, MA: The MIT Press.

Reisberg, D., Wilson, M., & Smith, J. D. (1991). Auditory imagery and inner speech. In R. H. Logie, & M. Denis (Eds.), *Mental images in human cognition* (pp. 59–81). Amsterdam: Elsevier Science Publishers B.V.

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour, 78*, 233–240.

Richardson, D. C., Spivey, M. J., McRae, K., & Barsalou, L. W. (2003). Spatial representations activated during real-time comprehension of verbs. *Cognitive Science, 27*, 767–780.

Sadoski, M. (1983). An exploratory study of the relationships between reported imagery and the comprehension and recall of a story. *Reading Research Quarterly, 19*, 110–123.

Sadoski, M. (1985). The natural use of imagery in story comprehension and recall: replication and extension. *Reading Research Quarterly, 20*, 658–667.

Sadoski, M., Goetz, E. T., & Fritz, J. B. (1993). Impact of concreteness on comprehensibility, interest, and memory for text: implications for dual coding theory and text design. *Journal of Educational Psychology, 85*, 291–304.

Saussure, F.d. (1983 [1916]). *Course in general linguistics* (R. Harris, Trans.). La Salle, Illinois: Open Court.

Savage-Rumbaugh, S. (1986). *Ape language: From conditioned response to symbol*. Oxford: Oxford University Press.

Savage-Rumbaugh, S., Shanker, S., & Taylor, T. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.

Seyfarth, R., Cheney, D., Bergman, T., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour, 80*, 3–8.

Seyfarth, R., Cheney, D., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science, 210*(4471), 801–803.

Sole, R. V. (2005). Syntax for free? *Nature, 434*, 289.

Soltis, J. (2009). What do animal signals do? *Animal Behaviour, 78*, 1485–1486.

Spoor, F., Leakey, M. G., Gathogo, P. N., Brown, F. H., Antón, S. C., McDougal, I., et al. (2007). Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature, 448*, 688–691.

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12*, 153–156.

Sterelny, K. (2008). *What is behavioural modernity?*. Paper for the Nicod Lecture. http://www.institutnicod.org/Sterelny-WhatisBehaviouralMordernity.pdf.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: what is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*(3), 299–313.

Tallerman, M. (2007). Did our ancestors speak a holistic protolanguage? *Lingua, 117*, 579–604.

Thibaut, J. P., Rondal, J. A., & Kaens, A. M. (1995). Actionality and mental imagery in children's comprehension of declaratives. *Journal of Child Language, 22*(1), 189–209.

Tobias, P. V. (1998). Evidence for the early beginnings of spoken language. *Cambridge Archaeological Journal, 8*(1), 72–78.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford: Oxford University Press.

Tulving, E. (2005). Episodic memory and autonoesis: uniquely human. In H. S. Terrace, & J. Metcalfe (Eds.), *The missing link in cognition: Origins of self-reflective consciousness* (pp. 3–56). Oxford: Oxford University Press.

Vauclair, J., Fagot, J., & Hopkins, W. D. (1993). Rotation of mental images in baboons when the visual input is directed to the left cerebral hemisphere. *Psychological Science, 4*(2), 99–103.

Wadley, L. (2001). A general view and a South African perspective from Rose Cottage Cave. *Cambridge Archaeological Journal, 11*(2), 201–221.

Watanabe, S. (2010). Pigeons can discriminate "good" and "bad" paintings by children. *Animal Cognition, 13*(1), 75–85.

Waters, M. R., Forman, S. L., & Pierson, J. M. (1997). Diring Yuriakh: a lower Paleolithic site in Central Siberia. *Science, 275*, 1281–1284.

Wharton, W. P. (1980). High imagery and the readability of college history texts. *Journal of Mental Imagery, 4*, 129–147.

Whitley, J. (2003). Archaeology in Greece 2003–2004. *Archaeological Reports, 50*, 1–92.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. *Language and Communication, 18*(1), 47–67.

Wray, A. (2000). Holistic utterances in protolanguage: the link from primates to humans. In C. Knight, J. R. Hurford, & M. Studdert-Kennedy (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form* (pp. 285–302). Cambridge, UK: Cambridge University Press.

Wynn, T. (1998). Did *Homo erectus* speak? *Cambridge Archaeological Journal, 8*(1), 78–81.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.