Language, from its early hominin origin to now, was not primarily being used for practical purposes. We suggest that an essential function of protolanguage was to signal ‘noteworthy’ events, as humans still systematically do. Words could not be so specific as to refer to whole, non-recurring, situations. They referred to elements such as objects or locations, and the communicated event was inferred metonymically. Compositionality was achieved, without syntax, through multi-metonymy, as words referring to elements of the same situation were concatenated into proto-utterances.

1. The limits of protosemantics

Most studies on the origins of language naturally assume that language did not emerge all of a sudden as the full-blown faculty that we enjoy now. Several authors provided arguments in support of the existence of a stable intermediary state in hominin evolution, corresponding to protolanguage (Bickerton 1990; Jackendoff 1999; Nowak & Krakauer 1999; Dessalles 2007a). There is some disagreement, however, about the expressive powers of protolanguage. A crucial issue is to know whether the set of potential meanings was limited, or was open and infinite as for genuine language. In the latter case, protolanguage must have had synthetic power: some cognitive device (protosemantics) was there to assign novel meanings to an unbounded set of word combinations. In this paper, we defend this synthetic option, as opposed to the analytic view in which word combinations emerge through a regularization process and are thus bound to designate recurring meanings (Kirby 2000).

The way people express requests on Web search engines offer a good picture of protolanguage: an order-free concatenation of words, devoid of grammatical indication. To decide how our hominin ancestors attached meaning to proto-utterances, one must first have some idea of what they used protolanguage for. Much depends on whether protolanguage was used to coordinate hunts, to teach children or to establish social bonds, as the corresponding range of meanings would considerably vary. In what follows, we begin with this issue of the function of protolanguage, by first showing the divergence of opinions in that matter. We will show that many of these opinions are in contradiction with current principles of evolutionary theory. We will then provide evidence suggesting that (proto-)human communication is by nature informational and thus necessarily involves an unbounded set of potential meanings. This suggests that protolanguage emerged as a coherent combinatorial device implementing multi-metonymic reference. Lastly, we will briefly address the issue of the transition to language, suggesting that the crucial change was due to the expression of predication, rather than to combinatorial improvements.
2. What was protolanguage used for?

2.1. Protolanguage evolved for a definite reason

For some authors, determining the function of language is a non-issue, principally because language, like any complex biological character, must have multiple functions (Fitch, Hauser & Chomsky 2005:189). According to Lieberman, language enhances virtually all aspects of our practical life, and “it is futile, if not silly, to attempt to find ‘the’ factor that provided the selective advantage for the evolution of human linguistic ability” (Lieberman 2003:19). Yet, because of its complex design, the language faculty (including the protolanguage faculty, as we will suggest) “has all the earmarks of an adaptation” (Fitch 2004). Can complex adaptations result from multiple selection pressures?

Computer simulations have somewhat changed our understanding of evolutionary processes (Dessalles 1996). They have unambiguously confirmed the fact that evolution proceeds in fits and starts (Eldredge & Gould 1972). The reason is that natural selection is an optimizing process that rapidly produces local optima. A species may stay for an indefinite time period in equilibrium on such a local optimum, where all selection pressures cancel each other. It may then ‘discover’ by chance a better optimum in the local fitness landscape, in which case the new selection pressure would rapidly (i.e. in a few dozens or hundreds of generations) lead evolution toward the new local optimum. The high speed of evolution under selection pressure gives the impression of a jump (punctuation) when observed on a geological timescale. The consequence is a crucial distinction to be made between macro-evolution and micro-evolution:

(1) There is no long-term evolutionary trend; selection pressures rapidly balance each other; macro-evolution is unpredictable and non-directed, it has no inertia and no memory (Gould 1996).
(2) Any new complex feature results from some new selection pressure through Darwinian principles; it is rapidly locally optimized, until the selection pressure cancels out or gets balanced; micro-evolutionary processes are directed.

These two principles apply to the evolution of the language faculty (Dessalles 2007a). Together, they make the co-occurrence of several new selection pressures highly unlikely. Contrary to Lieberman’s statement, there must be one reason why protolanguage emerged in the first place, even if subsequent exaptations allowed it to serve secondary purposes.

The next question is to determine which selection pressure gave rise to protolanguage. Answering this question is far from obvious, and there is no consensus concerning this issue. This selection pressure must have been new, as otherwise evolution would have brought other primate species to discover some comparable communication device long before it occurred to our ancestors (Johansson 2005:1994).

2.2. Did protolanguage evolve for practical purposes?

Traditional accounts of language in its early stages inevitably mention its putative role in various practical activities thought to be vital in the daily life of naked hominins. One of these virtues is an increased hunting efficiency, due to the possibility of coordinating actions

---

1 Note that the species is never what natural selection optimizes. Natural selection optimizes relative reproductive success within the species. A feature is locally optimal only as far as no slight change of this feature in some individuals can increase their reproductive success.

2 Evolutionary speed is due to the phenomenon of implicit parallelism (Holland 1965): thanks to crossover, various partial improvements are simultaneously tested at various locations of the genome. Genetic Algorithms take advantage of this principle to solve engineering problems.
Another is the possibility to warn for various sorts of danger like predators or to share opportunities like food sources (Lieberman 1992:23; Bradshaw 1997:100-101; 2001:66; Snowdon 2001:226; Bickerton 2002:209; 2003:84). Another is the didactic virtue of language, in the context of toolmaking (Lieberman 1992:23) or more broadly in the transmission of experience from parents to offspring (Bickerton 2002:221; Fitch 2004; Castro, Medina & Toro 2004:725).

Various critiques can be opposed to such views about the daily use of protolanguage. One is that none of these alleged functions of protolanguage is a response to a new selection pressure, despite the varying ecological contexts in which hominins species were living (Bickerton 2002), so we would expect chimpanzees to talk when hunting or manufacturing tools. Another critique is that coordination in hunting or skill learning is achieved without using language in other species and with hardly any language in our species (Burling 1986:3-4; Dunbar 2003:220). Coordination requires few signs, as illustrated by the rule of the road or the scuba diving code. If selection pressure for such codes existed in hunter-gatherer life, it would have been more likely to produce an innate system than protolanguage.

A widespread assumption is that protolanguage evolved because it made information sharing possible. Hard-won experience, like the location of food sources, could benefit the whole hominin group, leading to its ecological success (Györi 1997:46, 47; Goodson 2003:74; Castro, Medina & Toro 2004:734; Ritt 2004:1-2). The problem is that ecological success has no evolutionary consequences (Williams 1966). Even if some form of “informational collectivism” could improve the ecological success of the first groups that adopted it, it would not create any selection pressure for talking. The problem here is that speakers get no relative advantage from talking, on the contrary! Game-theoretic reasoning shows that the best strategy would be for individuals to take the experience from others and to keep one’s own private (Dessalles 1999). Besides, we would not understand why informational collectivism would be beneficial to hominins and fully uninteresting in other lineages. Note that Fitch’s idea of protolanguage as a “mother tongue” used to teach children (Fitch 2004) solves the first problem (talking behaviour would be advantageous through kin selection), but not the second one (apes are virtually never seen teaching their offspring). Moreover, current spontaneous use of language is very marginally didactic (see below).

### 2.3. The social use of protolanguage

Several authors have put the emergence of language on the ground of social relationships, rather than considering it as a mere tool for improving subsistence. Linguistic abilities have even been claimed to be rooted in primate competence in dealing with social affairs, either directly by importing hierarchical cognitive representations (Worden 1998) or indirectly, by building on the ability to understand others’ minds (Byrne 1995:233). We do not consider these aspects here, as they give no hint of what new selection pressure brought hominins to talk. Various studies consider another social side of language, namely its relation to reciprocal cooperation. Language would have emerged, even in its proto-form, as both a cause and an effect of cooperation (Ulbaeck 1998; Calvin & Bickerton 2000:123; Pinker 2003:28; Nowak & Sigmund 2005:1293; Nowak 2006 :1561). These authors invoke reciprocity to circumvent the fact that information sharing benefits listeners, not speakers, in the first place. We showed, however, that grounding language in cooperation leads to insurmountable difficulties (Dessalles 1999): reciprocal cooperation in the absence of regulatory authority is unstable as soon as the proportion of cooperators increases. Moreover, if language was based on cooperation, speaking should be a cautious act that always brings valuable information to well-chosen ears, and talkative behavior should never surface. Such a picture lies quite far away from spontaneous human language!
A more plausible suggestion is that protolanguage was a new way of establishing and maintaining social bonds (Dunbar 1996). Dunbar’s grooming hypothesis addresses the problem of the new selection pressure: as group size increased, physical grooming was no longer an efficient way to maintain social bonds and language replaced it in that role. Though this hypothesis is highly appealing considering the current use of language, it is obviously incomplete: it does not say anything about what protolanguage should look like, and how it could perform its social function in a reliable way, as compared, say, to synchronized grunts.

2.4. The political origin of protolanguage

An easy way to determine the biological function of human communication is to slightly impair it. The experiment is easy to perform: one should just stop talking during one week, or utter only trivial statements like “The door is grey”, “There are eight lamps in this room”, “My sister owns two bikes”. The negative consequences on one’s social network will be immediate. As the study of spontaneous conversation suggests (Dessalles 2007a), language is a competitive display in which participants, in turn, demonstrate their ability to be relevant. This competition for relevance makes sense in a political species like ours, where individuals seek the best possible coalition partners. The basic assumption, in line with Dunbar’s grooming theory, is that with the increase of coalition size, individuals changed their criteria when choosing allies; qualities like physical strength became less valued, whereas the qualities required for being relevant became essential. Those who failed to demonstrate their relevance lost any chance of establishing valuable alliances.

This political theory of the origin of language has several strong points. First, it is consistent with evolutionary principles: new selection pressure (due to new coalition size), advantage for speakers (who get opportunities to get into new coalitions or to secure their position in current ones), advantage for listeners (who can appraise potential coalition partners, and also benefit from the information given). Second, it is consistent with current spontaneous use of language (see below) and therefore avoids hypothesizing biological function change between protolanguage and language (Calvin & Bickerton 2000:123; Bickerton 2002:222; Burling 2005:181-182). Third, it avoids any bootstrapping paradox, as displaying one’s ability to be relevant can start with mere pointing gestures. Fourth, it accounts for various, otherwise mysterious properties of protolanguage, such as open learned vocabulary, talk about seemingly futile matters, and ‘omnipotence’ (ability to talk about anything that can be experienced, instead of being bound to expressing food location (like bees) or to coordination signals like orders and requests). Moreover, it suggests a plausible transition from protolanguage to language (section 5). On the negative side, let us mention that the hypothesized increase in coalition size is left unexplained and remains an irreducible assumption of the scenario.

3. Proto-pragmatics and proto-semantics

At this point, we suppose that protolanguage, like language now, was used by our ancestors to competitively display their ability to be relevant. What can we know about relevance in hominin life? The observation of small children and of spontaneous conversations gives us a hint. Before reaching twelve months of age, children spontaneously point towards unusual stimuli, trying to attract the attention of the parent (Carpenter, Nagell & Tomasello 1998). Apes are never observed to perform such declarative pointing (Tomasello 2006). The systematic communication of unexpected events, almost a reflex in adults, is a

3 We consider coalition size, instead of group size, to be the crucial factor (Dunbar 1996:66-67). The fundamental distinction is that coalition mates choose each other, instead of merely happening to end up together for historical or ecological reasons.
universal and distinctive feature of our species. On July 6, 2005 14:00, in my research department, several colleagues rushed out of their office to announce the highly unexpected victory of London against Paris to be the host city for the 2012 Olympic Summer Games. If the purpose was merely to share information, there was no point to hurry. But if giving news is just a means to demonstrate one’s ability to get good information, as the political scenario suggests, then showing that one is the first to know is more essential than the news itself. The following conversation took place between two children aged eight and ten (original in French):

M: Did you see there’s again balloons up there this morning?
Q: Yes, I know.
M: You, be quiet! I’m not talking to you, I’m talking to the others. [To his father] Did you see there’s balloons up there this morning?

M urged to signal the unexpected spectacle of dozens of hot-air balloons drifting overhead. Again, the point is not information pooling, but rather to show that one knew first. Q’s reaction is quite revealing in this respect.

A reasonable assumption is that this ‘first-to-know’ display, which we do not share with apes, was what motivated the emergence of protolanguage. Though modern humans have other sources of relevance (see below), signaling unexpected events seems to appear first in the ontogeny of human referential communication. In adulthood, individuals devote a considerable amount of time reporting noteworthy events: conversational narratives may represent more of one half of the total speaking time (Dessalles 2007b) (see Norrick 2000 for a comprehensive description of this behavior). Drawing attention to unexpected events can be achieved through several means, ranging from mere pointing accompanied by some unspecific interjection: “Look!”, to complex narratives involving several temporal embeddings. Only the simplest are relevant to proto-pragmatics.

We may conjecture a two-step transition to protolanguage. In a first step, larger coalition size led to the emergence of a new selection pressure for the ‘first-to-know’ display. Readiness to signal unexpected states of affairs through deictic pointing and exclamation was immediately selected. At this stage, a single referential word could be uttered (Jackendoff 1999:273) together with deictic gestures (Bickerton 2002:219). This mode of communication could be used to refer to immediate events (the single-word stage is also to be found in human ontogeny, around twelve to eighteen months of age; see Bloom 1993:29).

The next step is where protolanguage strictly speaking enters the scene. Protolanguage is characterized by meaningful word combinations. Our ancestors, at this stage, were able to refer to situations that were beyond the reach of immediate perception. For instance, “strangers-fire-plain” could mean that there were strangers on the plain, as evidenced by the smoke that the witness could see over there. The transition from the one-world stage to the multi-world stage required a non-trivial cognitive adaptation, as we will see. For now, we can draw important conclusions from the preceding observations.

If protolanguage was used to refer to unexpected situations, then from its very beginning, the number of potential meanings was huge, virtually infinite. The set of unexpected situations is of course not the complement of one’s past experiences. Unexpectedness is more restrictively defined as the property of being simpler than anticipated: a rare event occurring in close vicinity, a situation made unique in its own kind due to a deviant characteristics, a coincidence, etc. (Dessalles 2007b). For practical purposes, it can be modeled by Shannonian information log \(1/p\), where \(p\) is the perceived probability of the event. Curious species like

---

4 ‘Simpler’ here means that the minimal cognitive characterization of the situation is more concise than expected (Dessalles 2007b).
chimpanzees are sensitive to certain dimensions of unexpectedness (Nishida et al. 1999), even if they don’t communicate about it. We may suppose that hominins’ sense of unexpectedness was intermediary between what triggers ape and human curiosity. The fact remains that the set of situations that are worth signaling in the hominin world, because of their unexpectedness, is unbounded. This has definite consequences on the semantics of protolanguage.

4. Reference through metonymy

Several studies on the emergence of referential communication aim at showing that compositional languages systematically emerge when expressive power is rewarded, under constraints that limit the efficiency of non-compositional languages. Such constraints are, typically, the risk of mistaking one word for another, due to the limited available sound space (Nowak & Krakauer 1999), or the limited number of signal-meaning associations that can be transmitted from one generation to the next (Kirby 2000). Figure 1 shows how communication efficiency is computed. Typically, the ability for agent B to correctly understand agent A’s intended meaning i is \( \sum_j p_{ij}q'_{ij} \), where the sum is computed over all possible signals j that A may have possibly chosen to designate object i (with probability \( p_{ij} \)). \( q'_{ij} \) stands for B’s probability of decoding signal j as i.

![Figure 1: protolanguage as an encoding-decoding device](image)

The preceding sum, averaged over all possible meanings i, provides an estimate of the language’s efficiency. Efficiency, here, amounts to avoiding decoding mistakes. All these models unsurprisingly presuppose that the set of meanings that can be unambiguously referred to is of the same order of magnitude as the set of available (possibly composed) signals. Successful languages are those which establish a one-to-one mapping between the two sets (Brighton & Kirby 2006).

Within such a framework, compositionality would have emerged through a regularization process, as a structure-preserving mapping from meanings to signals would be more likely to be understood and learned (Kirby 2000). A signal like a-bc is better to refer to a situation AB if there is a semantic mapping f such that, for instance, \( f(a) = A \) and \( f(bc) = B \), since a may be advantageously reused in a-d to mean, say, DAE. Signals like abc and ad would have been used ‘holistically’ first, in the absence of any semantic mapping. Then semantic mappings like \( f(a) = A \) would have been abstracted from fortuitous regularities like the co-occurrence of a in both strings abc and ad, and of A in both situations AB and DAE. Protolanguage would have started with a limited inventory of holistic signals until compositionality would have taken over. For this process to work, however, the set of target meanings must be kept reasonably small, at least at the beginning, for individuals or populations to notice regular patterns in recurring signal-meaning pairs and to abstract partial mapping from them.

To meet this constraint, protolanguage is supposed to have been used for basic needs such as commands, greetings, requests or threats (Wray 1998:52). As we saw in section 2.2, the
protolinguistic faculty is unlikely to have been selected for such limited function (already present in apes). If, as we claimed, protolanguage was used (as, in part, language now) to refer to unexpected situations, then holistic signals were useless. Almost by definition (see end of section 3), unexpected situations are not recurrent. It would be complete waste to devote specialized signals to refer to situations of interest which, for the most interesting of them, would occur only once or twice in a lifetime. The fact that Mary is back after a two-year absence probably belongs to the kind of events our hominin ancestors were competing for announcing first. They must have had other means to say it than a specialized word meaning ‘Mary-is-back-after-a-two-year-absence’.

A reasonable alternative to holistic messages is metonymy. In the appropriate context, uttering ‘Mary’ together with a pointing gesture is sufficient to mean that Mary is coming back after a two-year absence. The word “Mary” is metonymic because it refers only to part of the communicated situation. The pointing gesture is metonymic as well, as the indicated direction is just one of the many properties of the communicated situation. Metonymy is, by essence, highly ambiguous. How can communication rely on it? The word “Mary” and the pointing gesture could have meant ‘Mary climbed up this tree a long time ago’ or ‘Mary found mushrooms over there’. The answer is that most of these alternative meanings are not unexpected, whereas Mary’s return is.

Human communication is, as presumably protolinguistic communication was, inherently an inference-based process and not a mere coding-decoding procedure (Sperber & Wilson 1986). Individuals use relevance as a criterion to guide interpretation. Unexpectedness provides a constraining relevance criterion to make metonymy unambiguous in most concrete situations. The ‘algorithm’ runs like this: perform associations from what is said or shown until they produce an unexpected event. In most situations, a single word and a pointing gesture will be sufficient for hearers to get the point.

5. Compositionality without syntax

The usual notion of protolanguage is not limited to single-word utterances (Bickerton 1990; Jackendoff 1999). We may understand a proto-utterance like “strangers-fire-plain” as a multi-metonymic indication worked out by the speaker to make interpretation easier for listeners. This form of expression implements compositionality (as there is a structure-preserving semantic mapping), though not syntax (Bickerton 1990:122). Its clear advantage is to reduce ambiguity (Nowak & Krakauer 1999). It is tempting, then, to say that protolanguage was compositional from its very beginning.

Such a conclusion would ignore the complexity of multi-word semantic interpretation. Listeners must integrate the different associations triggered by the different words, ‘stranger’, ‘fire’, ‘plain’ into one single states of affairs, instead of imagining several disconnected situations. We humans do have this semantic synthesis ability. It is not obvious that we share it with chimps (Dessalles 2007a:181). The synthetic power of protolanguage relies entirely on semantic synthesis: when speakers take the trouble of using several words, unexpectedness is likely to become apparent only when the meanings associated with these words have been successfully combined. The non-trivial character of semantic synthesis may justify the existence of a single-word stage to be distinguished from protolanguage (Jackendoff 1999; Dessalles 2007a:183).

Multi-metonymic communication, even guided by the constraint of unpredictedness, remains somewhat ambiguous if we compare it with the precision that can be achieved through language. Hence the common idea that protolanguage must have been bound to express concrete and immediate meanings. Was the evolutionary emergence of syntactic abilities a response to this ambiguity problem? Indeed, our syntactic language is far more
expressive and precise than protolanguage. On closer examination, however, syntax appears to fulfill an entirely different function.

Syntax is, by nature, a tool devoted to the expression of predicates. Both phrase structure and marking (cases, classes, agreement, ...) are well-designed to express ‘who did what to whom’ (Jackendoff 1999; Dessalles 2007a:224). Simulations have shown that whenever meanings have predicate structure, syntactic language inevitably emerges to express that structure (Batali 1998; Kirby 2000). Conversely, syntax is useless if there is no predicate structure to express. If protolanguage ever existed as a multi-metonymic device, then there was no other structure to convey than the co-occurrence of the relevant concrete objects and locations within the situation to be signaled. Protolanguage was locally optimal (see section 2.1) to express this co-occurrence. The great divide between our species and the preceding one, as we suggested elsewhere, is the ability to form predicates5 (Dessalles & Ghadakpour 2003, in response to Hurford 2003). This ability emerged, not for its referential expressive power, but to express logical processes, and especially negation. This new system was a protection against lying (Dessalles 2007:330). It motivated the emergence of syntax (to express predicates) and led to the argumentative use of language (to negotiate logical consistency). Relevance has now two independent dimensions: narrative (event-based) relevance dates back to the time of protolanguage and depends on unexpectedness, whereas argumentative relevance is properly human.

6. Summary

If protolanguage ever constituted a long-standing step in the evolution of the language faculty, it must have been the consequence of a new selection pressure. We suggested that it evolved as hominins engaged in ‘first-to-know’ competitive displays. The function of proto-utterances was not practical, but was referential: individuals strived to attract others’ attention to unexpected events, as they do now, from early infancy on, with language.

In Molière’s Le Bourgeois Gentilhomme (1670), M. Jourdain wonders how a short phrase in (supposedly) Turkish, “bel-men”, could mean such a complex idea as: “you should go quickly with him to prepare yourself for the ceremony so as to see your daughter right soon and draw up the marriage settlement” (see Dessalles 2007a:233). By the time of our hominin ancestors as well, extremely unambiguous signals were useless, especially to refer to rare, unexpected events. Conversely, a high level of ambiguity was tolerable, and indeed it significantly increased expressive power, since interpretation was guided by the constraint of unexpectedness. Lexical abilities never evolved to remove ambiguity, just like hand-pointing never evolved to reach laser precision. A vague understanding of direction is most often sufficient to get the point.

In our account, compositionality is already achieved through multi-metonymy, which requires listeners to combine the meanings associated with words into a coherent situation. Syntax evolved in a later stage, to express predicates. The consequence of this stepwise evolution is that protolanguage was not a mere rough draft of human language, but a genuine faculty, qualitatively different from language, which was locally optimal for its function.

---

5 Some authors would consider that a word like “mammoth” uttered together with a pointing gesture would constitute a form of predication (Bickerton 2002:219-220; Hurford 2003). We suggested that there is no continuity with human predicates (Dessalles & Ghadakpour 2003).

6 Unexpectedness encompasses abstract dimensions in our species (Dessalles 2007b). Emotional change also play a definite role in narrative relevance (Dessalles 2007c).
7. References


http://www.enst.fr/~jld/WWT/


http://www.ling.ed.ac.uk/~jim/BBSNEURO/dessalles.html


http://www.ling.ed.ac.uk/~jim/newro.htm


http://www.ling.ed.ac.uk/%7Esimon/Papers/Kirby/Syntax%20without%20Natural%20Selection%20How%20compositionality.pdf


http://biologybk.st-and.ac.uk/cultures3/articles/download/Ethogram.doc


http://www.ped.fas.harvard.edu/people/faculty/publications_nowak/Nowak_Science06.pdf


http://www.ped.fas.harvard.edu/people/faculty/publications_nowak/nature05c.pdf


http://pinker.wjh.harvard.edu/articles/papers/Language_Evolution.pdf


