To appear in *In Search of Universal Grammar: From Norse to Zoque*
Ed. by Terje Lohndal. John Benjamins.

Language as Technology: Some questions that evolutionary linguistics should address[1]

Salikoko S. Mufwene
University of Chicago

Abstract

Over the past two decades, studies of the phylogenetic emergence of language have typically focused on grammatical characteristics, especially those that distinguish modern languages from animal communication. The relevant literature has thus left the reader with the impression that language is either exclusively or primarily mental; in the latter case, its physical features, phonetic or manual, would be epiphenomena that may be overlooked. I argue that language is natural collective technology that evolved primarily to facilitate efficient communication in populations whose social structures were becoming increasingly more complex. It emerged through hominines' exaptation of their own anatomy, thanks to the same mind that was enabling the complex cultures they were producing. Linguistic constraints of various kinds are emergent properties that are largely consequences of the modalities used, a position that does not expect signed languages cum legitimate linguistic systems to replicate the general architecture of spoken languages in all respects. The rest of the paper speculates on how the architecture of spoken languages evolved, gradually, with some features presupposing prior emergence of others, whereas some others conjure up concurrent emergence. The facts suggest a complex non-unilinear evolutionary trajectory with many alternative options, consistent with emergent technologies in which considerations of optimality are absent.

1. Introduction

The title of this essay is a development from my ecological approach to the phylogenetic emergence of language in mankind (Mufwene 2008, 2009, in press). According to it, language evolved gradually as an adaptive hominine response to a life style that was becoming more cooperative and gregarious. Our hominine ancestors then lived in bands that were becoming larger and more complex; they needed more explicit communication to manage their relations

[1] I am grateful to Diane Brentari, Evani Viotti, Terje Lohndal, and Cécile B. Vigouroux for useful comments on the last draft of this chapter. I am alone responsible for the remaining shortcomings.

and communities. The same mind that had led to this new social structure responded to the concomitant ecological pressure for improved communication by domesticating the parts of the hominine anatomy that appeared the most suitable for the purpose. Taking into account communication among nonhuman primates and modern humans' body language today, we may assume that both the hands and the buccopharyngeal structures had then already been coopted for communication by gestures and vocalizations, respectively (Corballis 2002, 2010, 2011). The next step was to co-opt them for richer and more explicit exchanges of information in ways that were less context-dependent in their interpretation.[2] The new ecological social pressures that invited a hominine adaptive response must have included, among other things, the need to talk about past experience, making it inter-individually relevant to solving problems in the present and planning future group activities, as well as explaining consequences of their behaviors.

The position developed here is consistent with current hypotheses about the emergence of modern language, some 50,000 years ago, coinciding with the emergence of rupestrian paintings and the late Homo sapiens' dispersal out of East Africa (Lieberman 2010). I conjecture that "everything" in this phase of the protracted evolution of language was being driven by a mind then endowed with a rich representational capacity and certainly capable of exapting extant means of communication to meet growing social pressures on group members to share their feelings, thoughts, and plans for the purposes of harmonizing their existence.

[2] As is evident from the scholarship in pragmatics, especially Dan Sperber and his colleagues' work on "mind reading" (e.g., Sperber & Wilson 2002), freedom from context in the interpretation of linguistic utterances is relative, not absolute. We cannot deny the reality that some forms of communication are more context-dependent than others. It is in this spirit that my statement about the adaptive emergence of explicit, less context-dependent language must be interpreted.

Thus the emergent means of communication had to meet what Hockett (1959) characterizes as "displacement," a trait apparently unique to human languages, which enables us to talk not only about the present but also about the past and the future, and thus to develop narratives that have world-creating power. Thus, we can not only talk about facts but also engage each other in constructing fictional and mythical universes brought to light not by paintings but with the contents of words. I submit that this is how symbolic communication (Saussure 1916, Deacon 1997) emerged, consisting of arbitrary signified-signifier or meaning-form symbiotic pairs that can be combined in various ways to express diverse information about activities, emotions, thoughts, experiences, conditions of entities, etc., and to situate them in space and time.

Below, I explain why, within this particular approach to the emergence of language, it appears justified to treat both speech and signing as technologies, albeit collective ones that evolved spontaneously (without anticipation of the ultimate systemic outcome) and piecemeal, more by exaptation and expansion of current materials (forms and structures) than by inventions *ex nihilo*.[3] None of these assumptions are at variance with the traditional conception

[3] I started thinking of "language as technology" in spring 2010, while reading Brian Arthur's (2009) very informative book *On the nature of technology: What it is and how it evolved*. In May of the same year, I articulated some of the ideas developed here in a keynote address I gave at the "Conference on Evolutionary Linguistics" at Nankay University, Tianjin, China. I was more encouraged to pursue the idea after noticing that Namhee Lee *et al*. too had thought of language as technology, in their (2009) book *Interactional instinct* in which they focus on language as a complex adaptive system but do not develop the technology idea. As this essay was nearing completion, I decided, out of curiosity, to google "language as technology" to determine whether I was not reinventing the wheel by any chance. My search revealed the following publications, which still do not make the orientation of my endeavor irrelevant or redundant: 1) "Language as a cognitive technology" (2002) by Marcelo Dascal; 2) "Language as a cognitive technology: English-speakers match like Pirahã when you don't let them count" (2008) by Michael E. Frank, Evelina Federenko, and Edward Gibson; and 3) "Ceaseless, unpredictable creativity: Language as technology" (2009) by Jan Koster. Of the three, the last is the closest in spirit to the present essay. However, although we both underscore the significance of phonetics and manual signs as materials that languages are made of, as well as the culture-specific ways in which their architectures vary, we focus on different aspects of the technology. Koster focuses more on the centrality of the word and how grammar evolved because of how it is used

of languages in linguistics as systems consisting of units of various sizes and of rules and constraints governing how they can be used to produce various kinds of meaning thanks to particular morphosyntactic combinations. I argue that in fact the approach makes it easier to also think of languages as cultural artifacts whose protracted emergence can be accounted for on the model of cultural evolution, though the role of biological predisposition for language (evident in hominine mental capacity and particular anatomy) can certainly not be discounted. The approach suggests that typological variation is a natural consequence of languages as cultural inventions, directing attention to parallels between variation in some aspects of linguistic structures with that in the organization of other cultural domains such as kinship, cooking, music, belief "systems," clothing/dress code, and architecture. In fact paleontological evidence suggests that we speak of the evolution of languages, in the plural, rather than just of language, in the non-individuated style of reference (see Section 3.2, below).

2. Do languages really qualify as technology?

The position defended in this essay depends centrally on what is meant by "technology." As explained by Arthur (2009), technology need not be material, it need not have a complex

---

to convey information, whereas I focus on how the overall architecture of the technology evolved in mankind and what particular trajectories the evolution may have followed. This explains also why I will hardly return to his insightful article below, though I share most of his views. The other two essays have to do with how language is coopted to carry out cognitive processes and are therefore peripheral to Koster's and my concerns.

On the other hand, just as I was completing the revisions of this essay, an accident of a completely independent reading led me to Douglas McArthur's (1987) short article "Le langage considéré comme une technologie," in which he sketches a position very similar to mine, but did not elaborate. He states in the English abstract of the paper that language can be considered as "a collection of tools and methods, a 'technology'. Like other technologies[,] it is the product of human invention, is elaborated over a period of time, and continues to change according to needs" (p. 157). He says on pages 159-160 that it can be characterized as "semiotic technology" or "information technology," albeit one with a "biological substratum." I obviously have more space to elaborate the position, including his argument for a polygenetic emergence of languages, in the plural.

structure (though languages do), it need not be monolithic, it need not have been planned, it need not be a conscious invention, it may have acquired this status in an ad-hoc fashion, it need not work perfectly and can be improved several times on different occasions, and it can be adapted to new uses. "Technology (…) is a means to fulfill a purpose: a device, or method, or process" (p. 29). It can be knowledge (when used to explain or handle something else), practice (when extended to the solution of a problem), or science (as in *science and technology*, when scientific knowledge is applied to develop practical things, in order to make our lives easier or to help us solve problems, p. 27).

My position depends largely on what LANGUAGE is understood to be. In fact, from a phylogenetic perspective, I have in mind LANGUAGES, indeed in the plural, consisting of phonologies, vocabularies, and morphosyntaxes that vary from one population/culture to another. I will often downplay the non-individuated abstraction LANGUAGE, conceived of on the basis of common characteristics of languages, especially the most abstract ones that have to do with Universal Grammar (for those who posit it), without any mention of materiality of the sounds or manual gestures engaged in the systems (see also Koster 2009; Evans & Levinson 2009). From an evolutionary perspective discussing languages without any reference to their physical components would be like talking about computers without any reference to their hardware or trying to make sense of how the mind works without any reference to the (activities of the) neurological circuitry in the brain. It is tantamount to ignoring the reason why all the principles that have retained the attention of linguists in such scholarship emerged, just like ignoring the constraints that the hardware imposes on the software of a computer (by way of facilitating or impeding some processes) and the brain on the capacity of the mind.

There would have been no languages for linguists to discuss today if hominines had not coopted vocalizations and gestures to produce speech and signs, respectively, which have made symbolic communication possible. Languages are systems whose elemental units are physical, to which various rules and constraints apply (on various levels of their architecture) and with which semantic interpretations are associated. The elemental units are sounds in the case of speech and handshape, orientation, and movement, among other components, in the case of sign languages.[4] Languages have non-monolithic architectures not only because the scopes of application of their rules overlap in various ways (Mufwene 1992) but also because they consist of material and abstract (computational) parts, indeed like computers, consistent with Arthur's (2009) conception of technology (e.g., p. 31).

As a defender of polygenesis, I submit that languages are technologies that different populations, at different evolutionary stages of the hominine phylogeny, developed piecemeal in response to their social pressures for richer and more explicit communication. They are productive *systems* (a term hardly applicable to animal means of communication) which possess principles and constraints governing how to form larger units from smaller ones and how to form meaningful utterances. Some of the principles are consequences of the particular

[4] I will avoid the mistake of trying to break down sign language in a way that matches speech on every level, because, as explained below, each modality in itself is an inherent ecological factor that constrains the kind of system that emerges. It determines what the most elemental units can be, how they can combine into more complex units and what combinations are not possible. For instance, speech cannot be iconic in the same way that sign language can (as the former depends on hearing and the latter on vision to be perceived and interpreted), nor need the latter be as strictly linear as speech is. On the other hand, the "strict linearity" of speech depends on the articulators involved, since co-articulated sounds (e.g., labiovelar stops) are possible. The co-occurrence of segmental and supra-segmental sequences does not dispute the essence of my claim, because they are produced in parallel and do not engage the same articulators. The supra-segmental elements of an utterance are also structured linearly. For instance, a vowel cannot have a high and a low tone at the same time, though it can have both sequenced, as a falling or rising tone. Much of the discussion below will focus on segmental sequences. The reader should remember that sign languages have their prosodic features too, which are equally ignored in the present discussion.

modalities involved, especially compositionality, which is a consequence of the discrete nature of their elemental units and the linearity they impose on utterances, although linearity is less strictly implemented in sign language than in speech. Culture-specificity is evident in variation within the particular size and composition of phonetic sounds used in a language, though a core of them are attested in all languages (such as the vowels /i, u, a/), in the phonotactic and other phonological principles that have been adopted, in word-formation strategies, and in syntactic constituent order and other conventions. Thus, spoken languages vary typologically not because the essential architectural principles of the linguistic technology (produced by phylogenetically the same human mind co-opting the same anatomy) are different, but because there was latitude for variation in the choice of basic units (perhaps less in sign language, consistent with my explanation below) and in the specifics of the emergent principles. Once the basic units and principles were adopted, it did not matter how a particular population developed the rest of the emergent system. This kind of variation is observable in other cultural domains, such as cooking, music, kinship, and courtship, except where more complex material technology is involved. In modern technological terms, it is like the variation observable in the word-processing industry, as one particular software enables the user to do certain things in ways somewhat different from another software, while they produce comparable results.[5]

[5] Evans & Levinson (2009: 429) state that the putative universals of language are "few and unprofound" and that "languages vary radically in sound, meaning, and syntactic organization." They may thus disagree with my still relatively conservative conclusion. Semantic differences across languages remain what they are, in that words and phrases do not convey identical meanings crosslinguistically, although most of them can translate each other in ways multilingual speakers can live comfortably with. I am focusing here on the simple fact that these differences do not prevent speakers of particular languages from communicating about themselves and their ecologies as adequately as those of others do. In fact, they (can) use their respective, culture-specific strategies to communicate about other cultures without modifying their respective ways of packaging information, which justifies the saying "lost in translation."

Architectural differences between speech and signing deserve some attention, because the modalities used impose their own specific constraints on the architecture of the emergent systems. While the HANDSHAPE, PALM ORIENTATION, and MOVEMENT and POSITION of the relevant articulators make it possible to tell one word-level sign from another, they are comparable as much to sounds (in that, being elemental, the latter help us tell one phonetic word from another) as to morphemes, because the signing parameters often contribute to distinguishing the meaning of one signed word from another in the way spoken morphemes do, when the composition is transparent. For instance, the motion of an entity can be signed differently with the hand(s) assuming the corresponding shape (open or closed), direction (forward or backward), and orientation (up, down, horizontal, etc.), and according to whether or not this articulator is kept steady (rolling, bouncing, weaving, or a steady plane, etc.). To be sure, such iconicity and morphological transparency are not typical of the whole vocabulary of a sign language. In any case, speech is incapable of such iconicity; these are distinctions that are typically expressed linearly rather than synchronously, by agglutinating morphemes to a base in some languages or periphrastically in others, with adverbials or in serial verb constructions.[6]

Assuming that sign languages must have an architecture similar in all respects to that of spoken languages, with phonemes as the elemental units (which combine into morphemes and the latter in turn into words, which combine into sentences), is about as good as expecting the

[6] As we are reminded of by McCleary & Viotti (2010: 198), Charles Hockett was well aware of the fact that the architecture of sign language need not be conceived of on the exact model of speech. Signing is not arbitrary to the same extent that usage of phonetic materials forces speech to be, and in fact it is not strictly linear. "[T]he dimensionality of signing is that of life itself, and it would be stupid not to resort to picturing, pantomiming, or pointing whenever convenient" (Hockett: 1978: 274–275). From the point of view of processing the coded information, visual perception need not be constrained by the same factors as aural perception. These are ecological factors that bore on the evolutionary emergence of linguistic technologies.

motion dynamics of a lizard to be identical to the slithering of a snake or perhaps like expecting a runner and a swimmer to race toward their targets using the same techniques. Because it is based on a different modality than speech, sign language need not be expected to have the same kind of architecture as speech in order to function as adequately as it does as a means of communication.[7] Having different kinds of minimal units and on a level that seems to conflate phonology and morphology in speech does not make signing less linguistic. Actually, the difference should help us articulate more accurately what it takes for a communicative system to be linguistic. This may have to do more with symbolism than with how the symbols are formed, just like a wall in a house remains functionally a wall regardless of whether it is built out of stones, concrete, or wood and the finish consists of plaster or drywall.

There are of course things that one can do with one particular piece of alternative technology but not with the other; and it appears indeed that, evolutionarily, speech has not been more advantageous in all respects. There are situations where it is safer signing than speaking, such as when communicating under water or hiding from a predator (Mufwene, in press), though I focus below on architectural differences that are consequences of modality differences, thus on the fact that different materials have been used in the communication technology. For instance, because hands can imitate shapes and motion better than sounds,

[7] This is not to deny that the two kinds of linguistic systems do also share some features, such as in being modular, using symbolic words/signs (speech perhaps more extensively than signing, which allows more iconicity), and, among other things, in resorting to combinatorics in order to express ideas more complex than those expressed in individual words/signs. The issue that I raise here is about whether "duality of patterning," or "double articulation" (in Martinet's 1960 terminology), applies (exactly) the same way in both kinds of systems. The reader should note that I am not at all questioning the existence of "duality of patterning" in sign language. Even if the latter did not really have a phonological level (a position apparently not so accurate either), the morphological and the lexical levels alone would still justify positing the "duality." This term, like "double" in "double articulation," is a misnomer for a stratification of more than two levels in spoken languages: at least sounds, morphemes, and words.

they can produce more iconic signs than our speech organs (Hockett 1978, Goldin-Meadow 2003, McCleary & Viotti 2010), although even iconic signs reflect some arbitrariness in the form-meaning pairing, as is evident from onomatopoeias in spoken language.[8] Speech would perhaps contain more acoustic icons if it were not phonetic, i.e., consisting of sounds that mean nothing alone and favor arbitrariness over iconicity in the ways they combine into words. A particular limitation with onomatopoeias is that they often consist of sound combinations that are marked compared to those used in the vast majority of words in specific spoken languages, for instance bilabial and dental clicks in several non-click languages. This is not the case with iconic signs in sign language, precisely because signed words are not composed by the same principles as spoken words. The most basic units used in the signing technology are not always the counterparts of phonetic sounds. The same parameters of HANDSHAPE and MOVEMENT, etc. that appear to have a morphemic status in iconic signs are used with no particular meanings in the more arbitrary signs. It appears that the morphological and the phonological levels are conflated into one heterogeneous module of discrete elemental units that do not have a uniform structural status in all words (semantic in some but strictly formal in others). "Duality of patterning" applies in that the elemental units help tell one word from another, regardless of

[8] According to Aronoff *et al* (2005), iconic signs are universal across sign languages and presumably less arbitrary. I think this difference in the faithfulness of iconicity and even in the extents to which it is exploited in the lexicon, not necessarily in the syntax, reflects the greater capacity of signing to encode more transparently the physical properties of the universe of reference than speech. It is consistent with the position being developed in this paper, viz., that some of the architectural properties of languages are consequences of the particular modality that is used.

whether they are interpretable as morphemes or counterparts of phonemes in spoken languages, and then each word helps tell one utterance from another in a discourse.[9]

It is apparently because of the same architectural differences discussed above that signed language is less bound by linearity than speech. The parameters that are combined into iconic signs are capable of expressing meaning, like MOTION UP or MOTION DOWN through the inclination and position of the hand, just as the kind and direction of the motion of a referent can be expressed concurrently by the way the HANDSHAPE (cum classifier) representing the referent is moved within the space used by the signer. Even morphosyntactic incorporation in speech cannot do this, because speech is not produced in a tridimensional space. Besides, as noted above, speech articulators are incapable of producing two sounds concurrently, let alone two morphemes or words, whereas each of the two hands may be engaged in signing two different ideas. This is a significant compensation for signing, because it would otherwise be too slow if it operated on the phonetic, strictly linear model. The size of the articulators and the larger space in which signs are produced (compared to speech) would alone, independent of how syntax is handled, make the technology too energy-consuming and less fluid. (See also Swisher 1988: 350 for a similar position.) So, signing is definitely a different kind of technology. As I show in section 3.1, it is not patterned on speech, though there are similarities in the ways

[9] This difference should not be confused with the fact that spoken languages also have morphemes that consist of one sound only, as at the beginning of ***a****-like* and the end of *fuzz-****y*** in English, or the fact that such words also have a certain amount of transparency that makes them somewhat comparable to iconic words in sign languages. The fact is that TRANSPARENCY and ICONICITY do not mean the same thing and the proportion of iconic words in sign languages is much higher than in spoken languages even if the iconic signs reflect some cultural arbitrariness and constitute only a subset of the lexicon. As noted by Hockett (1978: 274–275), "it would be stupid not to resort to picturing, pantomiming, or pointing whenever convenient." This option makes communication in the manual modality particularly informative and efficient in avoiding the longer time that the linear sequencing of signs would entail.

they relate form to the universe of representation or reference (see section 3.2) and in the general principle of combining smaller signs into larger ones (phrases and sentences), in order to convey more complex meanings.[10] After all, they are emergent inventions by the same mind associated with the same late Homo sapiens.

Those of us working on the evolution of language have often wondered why in all human populations, the majority have settled on speech rather than signing; in fact the latter has typically served as a fallback option for those who cannot produce speech, either because

[10] As explained by Brentari (2011), according to scholars of sign language, "the root node at the top of the structure" where the parameters handshape, place of articulation, movement, orientation, and nonmanual behavior combine "is an entire lexeme, a stem, not a consonant- or vowel-like unit" (691-692). The lexeme is often polymorphemic, though it is typically monosyllabic, because signed language is not strictly linear. The pervasive iconicity enabled at the putative phonological module by parameters such as handshape and movement conjures up morphology rather than phonology in speech, where phonemes help tell morphemes or words apart but do not typically convey specific meaning alone.

However, faithful to her research paradigm, Brentari concludes that the phonologies of the signed and spoken languages differ physically but "[t]he more focus there is on the mind/brain the more sign language and spoken language phonologies will look the same" (719-720). Certainly, like that of spoken language, the architecture of signed language is modular and multistratal (consistent with the "duality of patterning"), especially from the point of view of production, involving combinations of smaller units into larger and larger ones (from the parameters to lexemes and then onto phrases and sentences). However, this architecture need not involve an organization that matches that of speech in all respects. There are alternatives that only one modality can favor but not the other. The fact that signed language (not including fingerspelling, an imitation of English spelling) has a large proportion of elemental units corresponding to morphemes rather than to sounds at the putative phonological level makes all the difference, especially since those corresponding to sounds apparently fall in the minority. Note that the transition variation observed with the sign language parameters could be matched as much with allophonic as with allomorphic variation in speech; it constitutes no compelling evidence for positing an autonomous speech-like phonological module in signed language, at least from the technological perspective developed in this paper. Even the fact that the phonological level in sign language is predicated on analyzing handshape, location, movement, etc. by invoking distinctive features rather than on positing phoneme-like segments (Diane Brentari, p.c., 2/27/2012) suggests that distinctive features need not be associated exclusively with phonetics/phonology. Their function is to enable the distinction of minimal segmental units from each other. Although the elemental units happen to be sounds in spoken language, the monosyllabic units that the distinctive features help distinguish in sign language need not be the counterparts of sounds, as is evidenced by the preceding discussion. Thus, what has been identified as phonology appears to be conflated with morphology. Each modality imposes its own constraints on the development of the technology. The status of signed language as language is far from being diminished by the absence of an autonomous or extensive phonology component. It just makes it different in its architecture, though signed language shares so many other features with spoken language, especially regarding compositionality, "duality of patterning" (but in its modality-specific way), arbitrariness and symbolic nature of denotational signs (to different extents), displacement, etc. It has as much world-creating power in its narratives as speech, though there are certain things that one modality probably expresses more vividly than the other.

they have a hearing impairment (which deprives them of the input) or because they have lost the ability to speak after a particular injury or trauma. The question is especially significant, because, based on the communicative practices of nonhuman primates, it has been hypothesized that intentional communication among hominines probably started with gestures, the distant ancestor of signing (Tomasello & Call 2007, Coballis 2011, Pollick & de Vaal 2007), as explained in the next section.

The reasons why speech has prevailed demographically over signing are numerous and will not be discussed exhaustively here. They include the following: 1) Speech has freed the hands to do other things, such as carrying objects or using tools, concurrently with a communicative activity. 2) Speech does not require the listener to be in the visual fields of the speaker, whereas the addressee must be, in order to see the signs produced by the signer. This also provides the advantage of being able to communicate in the dark or across visual barriers that do not constitute sound barriers. 3) The broadcast capacity of speech has a wider range than that of signing (MacNeilage 2008). The speaker does not have to face the addressee, whereas the signer must. According to Dunbar (1996), this broadcast capacity enables the speaker to "groom" (interpreted here as 'socialize') with several rather than with one other person at a time. 4) Speech can be amplified (as in screaming or shouting) to carry father more successfully than signing can, in order to be perceived from distance.[11]

On the other hand, signing has its advantages: 1) As noted above, it can be more transparent and informative in the description of some events (Goldin-Meadow 2003). 2) In the

[11] Signing can of course be amplified by technology outside the communicator's anatomy, such as projections on a large screen. However, usage of such technology is a much later invention not accessible to all nor on all occasions; it is therefore irrelevant to the present discussion!

proximity of predators, a signer is safer than a speaker who would be giving away information about their location. The same is true of hunters, who would prefer to surprise the game rather than alerting it (John W. Wenzel, p.c., 1/24/2009; Mufwene, in press). Sometimes even whispering can be too loud. 3) Under water, it is certainly smarter to sign than to try to speak (John W. Wenzel, p.c., 1/24/2009). Evidently, we spend much more time out of and on water than under water. Like modern humans, our hominine ancestors must have socialized more frequently than they hunted; and they must have avoided living in the proximity of predators and other dangerous animals. So, ecological reasons must have stacked in favor of speech. 4) Like Swisher (1988), Mufwene (2009, in press) and Corballis (2010) speculate, in addition, that speaking uses less energy, as it depends on compact articulators that move in a much smaller space, and proceeds faster than signing. This is, however, an argument that holds only with regard to linearity. As explained above, signing compensates for this apparent shortcoming in not being absolutely linear, though bilingual people who speak and sign fluently are better placed to decide on the validity of this speculation.

3. How did the linguistic technology emerge phylogenetically?

3.1. The transition from vocalizations to phonetic utterances

The toughest part of this question has always been how speech, a discrete or digital communication system, could have evolved from holistic vocalizations of the kind observable in today's nonhuman primates and certainly in our cries and laughter (Rousseau 1755). The question was poorly answered by such scholars as Carstairs-McCarthy (1999) and Wray (1998, 2000, 2002), as if the transition took place suddenly from holistic vocalizations to modern

phonetic inventories. This explains why their accounts have been disputed, especially by Johannson (2005), Tallerman (2007), and Bickerton (2007), though Carstairs-McCarthy's and Wray's answers are not totally groundless. Likewise, one must account for how signed language could have evolved from gestures (McNeill 2005). The questions are now further complicated by the realization that the evolution to speech, which is intentional, must not have been that straightforward, because nonhuman vocalizations are no more intentional than ours (e.g., cries and laughter), which are innate and realized instinctively. I submit that ape-like vocalizations must first have been co-opted for intentional use before being modified and expanded into speech.

In the case of speech, the question is whether the following position is plausible: as hypothesized by Rousseau (1755), the relevant hominine ancestors would have started to modify their vocalizations into distinctive, recurrent syllable-size chunks, which would have led to the gradual emergence of different vowels and consonants and would have set in place what MacNeilage calls "syllabic variegation." These would have evolved to the capacity to distinguish syllables from each other by some of the segments they consist of and not just by holistic syllable-size vocalizations. Both evolutions, concurrent or successive, would have resulted in the explosion of the vocabulary and the need to develop some syntax of sounds within words (what amounts to phonology), consistent with Hockett's (1959: 38) speculation of why "duality of patterning" is a "design feature" of human languages.

The contrary, "synthetic view" proposed by Johannson (2005), Bickerton (2007), and Tallerman (2007), according to which "protolanguage" would have started with words, which eventually would be combined into sentences thanks to the emergence of syntax," must still

tell us how the words, produced phonetically, started. It is necessary to account for the transition from prephonetic words to the phonetic words, which brings us back to the question addressed not so successfully by Carstairs-McCarthy (1999) and Wray (1998, 2000, 2002). The latter's position does not dismiss the fact that modern speech no longer functions like ape-like and humans' most primitive vocalizations, as it serves more complex communication. However, we must bear in mind that our prelinguistic hominine ancestors did not have the same mental capacity as modern Homo sapiens and that the evolution that innovated phonetic communication did not then necessarily call for the kind of complex communication modern humans engage in today. The evolution that produced the late Homo sapiens also produced more complex minds that must have helped hominines evolve more complex phonetic communication, including various aspects of modern morphology and syntax, as I explain below. Hypotheses about hominine evolution based on paleontological evidence do not support hypotheses of an abrupt jump from ape-like vocalizations to even the kind of protolanguage proposed by Bickerton (1990ff.).

The objections raised by Tallerman (2007) against Carstairs-McCarthy (1999) and Wray (1998, 2000, 2002) include the following: vocalizations are not produced by the same part of the brain as speech (p. 582); they are not intentional/volitional (p. 583); they are produced as much by outbreath as inbreath among nonhuman primates, whereas they are produced only with outbreath among humans (p. 583; actually implosives use inbreath!); they are innate (p. 584); and they lack duality of patterning (p. 584).

However, if every capacity that has been exapted for speech, in the form of domestication and extension of functions, had to remain the same, there would have been no

evolution. Like any other aspect of evolution, something had to be co-opted from somewhere, including the intentional production of vocalizations and their gradual modification into digital sequences. It seems plausible to assume that the mechanisms involved in holistic vocalizations were exapted to do more, although intentionality must have been transferred from gestural behavior.

To be sure, as suggested above, the proposed segmentation or "fractionation" need not have started with the gamut of phonetic distinctions now observable in modern languages. It need not have started with Consonant-Vowel sequences such as *pa* and *ma*, and certainly not with the complex hypothetical phonetic sequences proposed by Wray (e.g., *tebima* and *kumapi*); and there was no reason why Tallerman should have been disappointed that Wray's "holistic protolanguage" would not have been able to handle many semantic distinctions. By definition the protolanguage should not be expected to be systemically and semantically as rich as a modern language, especially in its initial stages.

Chances are that the initial variegation that Wray (1998) wanted to illustrate (with the wrong examples!) started with vocalic distinctions within or between vocalizations. Just inserting a glottal stop (or pause) and modifying the vocalic quality would have sufficed to innovate an embryonic digital phonation system from continuous ape-like vocalizations. The evolution must have proceeded gradually, and undoubtedly by trial and error, in attempts to modulate the vocalizations not only melodically (as nonhuman primates do) but also qualitatively (as humans do with vowels). Once the principle of vocalic variegation was in place, facilitated later on by the insertion of a consonant between the early vowels, the phonetic inventory would grow only proportionately to the communicative pressures experienced by the

community of interactants. These pressures themselves were triggered by the increasing cognitive capacities of the relevant hominine ancestors and by the complexification of their social structures.[12]

3.2. The complexification of spoken language

As explained in Mufwene (in press), the phonetic inventories of the emergent languages must have expanded in response to communicative pressures for more and more lexical distinctions. It may have started with contrastive monosyllabic utterances, which may or may not have been symbolic yet. However, the relevant hominine ancestors would have realized soon that more signs could be produced without having to expand the phonetic inventory too extensively if words consisted of more syllables, in alternating and/or (partly) reduplicating sequences. Although many modern languages have managed well with large proportions of monosyllabic words, especially with the help of phonemic tones (which account for semantic distinctions between segmentally identical words), it is equally noteworthy that several, if not most languages, rely heavily on words longer than one syllable. This is especially significant in languages such as Hawaiian with a small segmental inventory (10 vowels and 8 consonants) that are also non-tonal. Variation in the ways that the phonetic sounds have been used in the architecture of different languages, both regarding their specific phonetic inventories and the related combinatorial constraints, underscores the status of languages as cultural artifacts, notwithstanding the significance of humans' biological predisposition for this technology.

[12] I have no expertise on signed language and won't speculate in the dark about how the transition from gestures to signs must have proceeded. We simply must remember that intentional communication appears to have started with gestures earlier than with vocalizations. On the other hand, as argued above, the architecture of signed language need not be patterned on that of speech. Thus the question of transition might not have to be addressed in exactly the same way.

Whenever polysyllabic words emerged phylogenetically, I surmise that, just like children learning the language of their social environment, the hominines that evolved speech did not first spend their time developing a full phonetic inventory and wait until it was in place to start producing words. Johannson (2005), Bickerton (2007), and Tallerman (2007) are correct in drawing attention to the centrality of words in the phylogenetic emergence of language. It is more likely that, as in child language development, the phonetic inventories and the vocabularies grew concurrently, as it became necessary to use more and more units to coin more and more (minimally) contrastive words. This concurrent evolution must have proceeded until the phonetics reached its critical mass and speakers thought that the extant inventories offered them relatively limitless possibilities to produce the new words they needed.

We must bear in mind that there is no standard size that satisfies all populations/ cultures; different language communities have managed equally well with fewer or more segments than others, resorting to combinatorial constraints that are not identical crosslinguistically, although there are some universals both regarding some segments (notably the cardinal vowels /i, u, a/ and voiceless bilabial and coronal oral and nasal stops), as well as particular combinatorics (including CV and NC) that are attested in almost all languages. It is informative that, despite world-wide variation in the contents and sizes of phonetic inventories, most languages cluster in the average middle (Maddieson 2006). Given the large size of possible sounds that humans can produce, it is striking that they avoid more or less the same statistically marked ones and that almost the same sounds constitute the cores of their inventories. There must be a number of ecological constraints stemming from the articulatory and perceptual anatomies involved in the production and processing of phonetic sounds that

account for this state of affairs. Phoneticians are better placed than most of us to explain or investigate them.

As the size of vocabularies expanded, it must have appeared useful to capture formally the semantic relations that obtained between some lexical items in multi-word utterances, i.e., once communication exceeded the one-word utterance stage consisting only of naming and pointing. I am not sure how long the putative holophrastic stage (would have) lasted, because it is difficult to believe that a hominine population would have evolved a rudimentary language just to name entities, and maybe also feelings and conditions, without providing more information. More plausible is the hypothesis of assuming some sort of telegraphic stage perhaps close enough to Bickerton's (1990ff.) "protolanguage" but without all the stipulations he has invoked to distinguish it from modern languages. In its initial stage, it need not have been as complex as a pidgin, produced indeed by modern humans who had already spoken full-fledged languages (if only there were a universal grammar of pidgins at all!). The reason is that, as noted above, our hominine ancestors did not yet have our modern kind of mind, had not yet developed modern kinds of social cultures, and therefore were not under the same ecological pressures for modern style communication, even in pidgin form or child language in its second year, *pace* Bickerton (1990ff.) and Tallerman (2007).[13]

[13] According to Pollick & de Vaal (2007: 8188, col. 1), an embryonic form of predication is attested in bonobos, whose "vocal communication (…) appears more dialogue-like, and includes soft peeps to draw attention to and 'comment' on novel items and environmental events (…), a characteristic shared with early human language development." Unless this characteristic of the bonobos is homologous evolution, it may mean that the ability for predication may have started in an embryonic form long before hominines evolved language. We must remember that there has always been a predisposition for communication among all social animals. It is how hominines diverged from all the others in our evolution that makes all the difference.

Word-formation strategies (themselves also controlled by the relevant emergent combinatorial conventions) must have also started to emerge as alternative, phonetically economical and mnemonically efficient ways of expanding the vocabulary. Both affixation and compounding may have started that early (though practitioners of grammaticization may favor compounding), bearing in mind that subsequent history would not always maintain transparency. These strategies provide the benefit of facilitating both the formation and the recognition of new words, by analogy to previous complex forms, distributing the tasks between memorization and creation. Though the size of the vocabulary used collectively by a language community can become awfully large (such as the over 100,000 words commonly associated with English), a fluent speaker need not know all of them, as long as he/she knows the most common vocabulary and derivational morphemes and can therefore form or interpret some lexical items hitherto unknown to him/her from these materials. So many innovations by different speakers are bound to be the same, as they are constrained by similar combinatorial principles within the same community. Each fluent speaker/signer learns how to package concepts new to them either by exapting extant words (in the right pragmatic context) or by coining new ones from combinations of familiar lexical and derivational materials.

From the point of view of the vocabulary, languages are ingenious technologies that are so adaptive to new communicative needs and must be user-friendly.[14] All users of the

[14] Perhaps languages as technologies can be compared with each other according to how user-friendly they are, though this is also a function of how a particular user is more accustomed to doing particular things. One may also argue that, since variation is a fundamental characteristic of human populations both within and in relation to each other, this kind of inquiry is not particularly fruitful. No population ever aimed at producing an optimal system, despite the preoccupation of some linguists with optimality. It's all a matter of *satisficing* behavior, in which the interactants are happy with satisfying some but not necessarily all conditions for the use of (new) terms or the extension of the usage of extant ones. Adjustments are made ad hoc, often based on analogies, without anticipation of ultimate consequences for the system. (See also Pinker and Jackendoff 2005 for similar

technology will of course not do things exactly the same way, especially when there is choice between affixation and compounding in the same language or even between different affixation strategies (pre-, in-, and suffixing), which generates overlapping and competing strategies and complicates the task for the analyst, but not necessarily for the speakers/signers, who are not in the least interested in figuring out what the overall architecture of their language is or will be like. Languages have definitely not emerged by design; Hockett (1959) was certainly mistaken in speaking of "*design* features." Linguistic "systems" illustrate "emergent systems," which makes it easier to explain why languages change over time, under the pressure of various ecological factors (Mufwene 2001, 2008).[15]

The expansion of the vocabulary appears to have been a consequence of the increase in hominines' cognitive capacity, as pressure increased to name more entities (persons, animals, and objects), activities, and conditions. The same expansion of the cognitive capacity must also have driven the need to comment on particular entities, activities, and conditions. This development marked the emergence of predication strategies in languages, starting with the functional distinction between, on the one hand, SUBJECT or TOPIC and, on the other, PREDICATE. If the vocabulary had had no lexical categories until then, a distinction would have had to start at this point, with NOUNS specializing for the function of ARGUMENT and VERBS for that of PREDICATE.[16]

---

observations. They conclude: "language is useful but imperfect, just like other biological systems" (p. 229), more plausibly, like any system understood to have evolved according to current evolutionary theory, by competition and selection under changing ecological pressures.)

[15] Practitioners of usage-based grammar, such as Croft (2001) and Tomasello (2003), lead to the same kind of position as expressed here. For them grammar is the cumulation of patterns arising from the ways utterances are produced and have no ontological precedence to the latter.

[16] Givón (1979) is particularly informative about which kinds of entities tend to be nouns and which ones tend to be verbs, crosslinguistically, although there are nominalizations of activities (e.g., *construction*, *endangerment*) and of states and conditions (e.g., *insomnia*, *fatigue*, *drowsiness*). It may well be that such nominalizations are

This hypothesis is consistent with Stokoe's (2001: 44) observation on the interdependence of lexical categorization and sentences: "the words in the lexicon must already be sorted by kind, and that sorting had to come from their roles in sentences." However, he thinks that the distinction predates language and must have been in place in Homo erectus, when much of the communication was still gestural, a position apparently shared by Pollick & de Vaal (2007), as shown in note 13.

It's not evident yet that the category ADJECTIVE, which can also be used predicatively, would have started yet (in languages that have it, as it is not universal). Chances are that it would not have emerged before pressure mounted to articulate more explicitly or specifically how a noun picks out its referent out of a wide range of denotata. Once they emerged, adjectives would specialize for modifying nouns but would also be candidates for predicating states or conditions about the subject. Languages would then vary typologically regarding whether a copula is required when adjectives are used predicatively.

Otherwise, predication may have started simply with utterances in which a lexical item is used as an argument and another as a predicate. The emergence of the lexical distinction between NOUN and VERB would facilitate a response to another need for a more informative communication system, viz., to specify reference for the subject/topic and time reference for

---

consequences of speaking about activities, states, and conditions as events or entities that are identified or identifiable as arguments of predicates in utterances. On the other hand, Evans & Levinson (2009: 434-435) draw our attention to the fact that in languages such as Straits Salish the fundamental distinction remains PREDICATE vs. ARGUMENT, which is indicated syntactically by the affixation of function-specific markers to the same lexical items for 'man', 'big', 'run', etc. In the lexicon, there is no particular way to identify them as either nouns or verbs. For the sake of not complicating the exposition, I will leave this typological option alone, though it can be used to support the position that it is perhaps more accurate to speak of the evolution of languages, in the plural. From the perspective developed in the present essay, communication problems were not solved in identical ways from one population/culture to another.

the predicate, though one is reminded now of languages such as Mandarin in which ASPECT is expressed grammatically but not TENSE. Modern languages suggest that these two time notions need not have arisen concurrently in the emergent grammars, as the time at which a state of affairs occurred can be expressed with a time adverb such as 'now', 'yesterday', and 'tomorrow'. Besides, modern languages also suggest that the categories need not have been structured identically in different communities. It is also likely that the relevant distinctions did not all emerge at the same time, though this raises the question of figuring out what the most fundamental distinctions are in different systems.

As with word-formation strategies, specifying reference need not have started as an affixation process either; the strategy may have been periphrastic, as in languages with isolating morphosyntax, though we have no evidence for speculating, like Humboldt (1836) and Schleicher (1853), that the morphosyntax of languages has evolved from isolating to fusional systems, going through agglutinating and polysynthetic systems. The evolution of languages from holistic vocalizations to polysyllabic utterances and from holophrastic to multi-word utterances need not have favored the earlier emergence of isolating morphosyntax, though this was one of the possibilities. It is noteworthy that a number of philologists up to the 19th century thought that the earliest forms of speech involved agglutination, perhaps because, in elliptical constructions, they are reminiscent of holophrasis, as most of the relevant information is agglutinated to the verb, which is uttered as one phonetic word.

One particular phenomenon that is relevant to the issue of the evolutionary order of morphosyntactic structure, especially for those who will not consider the option of alternative choices by different populations of speakers, is quantification and demonstratives as part of

specifying reference with nouns. All languages of the world have these strategies, and they are executed periphrastically, whereas markers of nominal NUMBER and GENDER, the latter of which is an optional grammatical category (in the sense that a language may choose not to include it) can be expressed either periphrastically or by affixation, displaying typological variation across languages.[17]

It is also noteworthy that even incipient pidgins, which are associated with reduced morphosyntax, preserve numerals and demonstratives, which suggests that, from a phylogenetic perspective, they are deeply entrenched in the cognitive substratum of modern human capacity to communicate linguistically and must have emerged early, just like the organization of the first telegraphic utterances into subject/topic and predicate. Demonstratives and quantifiers such as 'little' and 'much' are also acquired early in child language, just like the words for 'no' and 'yes', which, according to Jackendoff (1999), must also have emerged early in the protracted phylogeny of hominine languages. We may thus infer that along with, or soon after, the emergence of predication, our hominine ancestors developed strategies to express quantity (including basic numerals) and learned to situate the referents of nouns used in utterances spatially with demonstratives. Markers of NUMBER, and GENDER, and

[17] This is a most interesting case of the emergence of typological diversity, as some languages have semantic gender (associated with denotational properties, as in English), whereas some others have (survivals of) formal or arbitrary gender (as in Latin and the Romance languages), while others have developed noun classes (with more numerous distinctions and originally also associated with denotational properties, as in the Bantu languages), and still others have numeral classifying systems (also with more numerous distinctions and even more transparently associated with denotational properties, as in Sinitic and Austronesian languages). But variation in this domain does not stop here, as languages within these types also vary according to the specific parameters along which the distinctions are articulated and the specific number of distinctions made. While they also underscore the fact that languages are cultural artifacts, such patterns of crosslinguistic variation call for a clearer articulation of the concept of "parameter-setting" (as opposed to pattern-identification, as one may expect, from an emergentist perspective), in ontogenetic language development.

CASE, which are easily dispensed with in pidgins, appear to be useful accessories that may have developed later, as the communicative technology became more sophisticated.[18]

The question of time reference appears to be more difficult. Creoles tell us that ASPECT must be distinguished from TENSE, and the former may be more critical than the latter, as is confirmed by Sinitic languages, for instance. The reason probably is that, as noted above, deictic adverbs such as 'now', 'today', 'yesterday', 'tomorrow', and 'later', which emerge early in ontogenetic language development and probably did in language phylogeny, can be used to situate states of affairs in time, just like demonstratives can situate entities spatially. It is not evident either when markers of MOOD may have evolved, to express the speaker's/signer's epistemic stand toward their utterances, thus making it possible to distinguish what is factual from what is hypothetical or conditional, what is known from observation or direct experience from what is inferred or has been learned by hearsay, etc.. Though MOOD can be expressed periphrastically (albeit through modal verbs/predicates), combinations of verbs for this purpose (as also in the case of periphrastic aspectual constructions) presuppose the emergence of embedding, a strategy that enables a verb/predicate to function as the complement/argument of another. Is mood the kind of nuancing strategy in language that may have emerged later

[18] The more involved reader will have noticed by now that gradualism is generally assumed in this paper, because, given the complexity of the architecture of modern human languages, any abrupt transition from no language or a protolanguage (Bickerton 1990ff.) to a language would have been too overwhelming to handle. It would have caused the collapse of the attempts to communicate explicitly and more efficiently, as the emergent technology would have been overwhelming and too difficult to adapt to. Although hominine evolution involved mutations, the ensuing population dis- and replacements did not occur in one swoop, as is evident from, for example, the protracted coexistence of the Neanderthal and Homo sapiens over tens of thousands of years. The genetic changes associated with evolution to modern languages needed to be reproduced and spread, gaining the critical mass for population-wide changes to take effect. It is also unlikely that the necessary communicative needs themselves all arose around the same time, because the complexity of hominine cognitive capacity must itself have evolved gradually. The architecture of language itself also suggests that the emergence of some structural features presupposed that of others before them, as I try to show in the rest of this essay.

than the distinction between statements, questions, and orders/requests? The latter are fundamental aspects of social interaction and communication, which the linguistic technology was being developed to serve, for the purposes of facilitating cooperation! I suspect that strategies for these illocutionary distinctions must have emerged early in communication, just as they also do in child language. They are not lost in pidgins, though the specific strategies used may not be those in currency in the languages they evolved from, the so-called "lexifiers."

These considerations raise the important question of whether all aspects of predication started at the same time. What was said about predication above applies without any obvious problem to intransitive, mono-argument utterances of the SUBJECT/TOPIC + PREDICATE kind. It says nothing about larger predicate phrases involving a second and possibly a third argument (as with the verbs *tell*, *take*, and *put* in English) and even adverbial adjuncts. Future research will have to address the question of whether the earliest cases of predication were restricted to SUBJECT/TOPIC + INTRANSITIVE PREDICATE constructions (expressing what the referent of the subject/topic is doing/experiencing or has done/experienced) or also involved constructions expressing a relation between two or more arguments in terms of what one is doing, or how one is related, to the other.[19] Is the evolution from holophrastic utterances (simply naming entities and states of affairs) to those involving predication likely to have involved a shift to the predicate as the central element of an utterance, in which, when more than one argument is involved, it specifies what relation obtains between them?

[19] Just in order to keep the exposition manageable, I will ignore typologically different, unaccusative languages, as well as polysynthetic ones whose elliptical utterances may not even contain an independent, non-agglutinated subject or topic.

This may be a case where ontogenetic development may not be informative about phylogenetic development, because children are exposed to both intransitive and transitive constructions, aside from being born with a late Homo-sapiens mind that is predisposed to any linguistic distinction that is consistent with his/her cognitive maturation. Our hominine ancestors need not have developed both types of constructions (transitive and intransitive) concurrently. We need different kinds of evidence to answer the question.

Be this as it may, multiple-argument constructions would have called for strategies (argument order) and/or devices (case makers or adpositions) that identify the arguments' thematic roles. Is it possible that, as suggested by Jackendoff (1999), our hominine ancestors would have first relied on constituent order (such as in di-transitive constructions, in which the dative argument precedes the direct object)? The hypothesis is not implausible if one agrees with students of grammaticization such as Heine & Kuteva (2007) that inflections and other grammatical morphemes are erstwhile periphrastic markers, consistent with the alternative of the evolution of non-isolating morphosyntax discussed above.[20]

The emergence of predication obviously involved quite a number of details that our hominine ancestors had to attend to. Short of involving a miracle, it is not evident that the emergence of the morphosyntactic devices and strategies used in modern languages could have occurred in the abrupt, catastrophic style of evolution hypothesized by Bickerton (1990ff.),

[20] I will leave alone the question of when adjuncts may have emerged, especially those that have nothing to do with situating states of affairs in place and time. The fact that they are optional suggests that they may have emerged later than the basic strategy of predication. On the other hand, it is also conceivable to distinguish markers of time (and place) from the other adjuncts. Adverbial clauses are of a different order of complexity that must be discussed separately from manner and instrumental adverbs, for instance. They presuppose prior development of clause formation strategies, thus of predication, and of devices that connect adjuncts to the main clause.

endorsed by Tallerman (2007), and even still assumed by those such as Wray (2002) who offer alternative hypotheses.

In any case, predication brought along with it the possibility of what Hockett (1959) identified as "displacement," which enables interactants to exchange information not only about the *hic et nunc* of their interactions (which is more typical of holophrastic statements) but also about the past and the future, as about places other than where the speaker/signer is. Thus predication would have exerted a lot of cognitive pressure to specify time reference, in order to situate states of affairs in time. Assuming that modern languages' specifiers of the verb are the counterparts of those associated with the noun (although modern generative syntax identifies them as heads of projections), the emergence of this capacity to situate states of affairs in time may have emerged concurrently with the capacity to specify where the referents are situated spatially (with demonstratives) and whether one or more than one is involved (with number markers), or even whether reference is generic rather than to particulars (with articles or some other devices).

Generic reference for the noun may also have been a concomitant of the expression of habits and typical conditions. Now, the question is whether, at some point in the evolutionary trajectory of language(s), there could have been an explosive evolution in the hominine mental capacity that caused all these strategies to emerge concurrently in language(s), while allowing for typological variation from one to the other. Or may the evolution have been as gradual as in child language (Mufwene 1999), though the time between the emergences of the different strategies/devices may also have involved short phylogenetic distances? I am not sure that the question can be answered conclusively at this point.

All languages make distinctions along the lines expressed above, despite the obvious typological variation that linguistics has articulated since the 19th century. They all also involve negation of the predicate phrase or constituents within the predicate phrase, which must have emerged early too, as is also suggested by child language development, though here too, there is typological variation regarding the position of the relevant marker in an utterance and whether, in the first place, the marker is the same for wide-scope and narrow-scope negation.[21] I am inclined to consider typological variation as evidence for polygenesis in the phylogenetic emergence of language, since the geographical distribution of paleontological evidence, dispersed over the huge area of East(ern) Africa, defies the possibility of monogenesis. Besides, modern humans continue to show that there is usually more than one way of solving the same problem, notwithstanding the fact that no global typology has emerged that shows what particular parametric combinations are not permitted in a language. Greenberg (1966) identified only common tendencies and some implicational universals among the combinations that are attested.

Before we proceed with the larger aspects of syntax, we should discuss a couple of things about the morphosyntactic delimitation of the noun, viz., not all languages have articles, no more than those who have them have a uniform way of articulating their distinctions morphosyntactically. (For instance, the functional distribution of the definite and zero articles is not identical in English and in French.) On the other hand, it appears that definite articles

[21] The negation discussed here is different from what is the counterpart of 'yes' and can occur elliptically. As is evident from French, which has *non* and *ne … pas* as basic negation markers, the two kinds need not have the same form; nor need they have emerged around the same time. (French also uses *pas* alone for narrow-scope negation.) Child language development suggests that the independent negation is evolutionarily anterior to negation with scope over or within the predicate phrase.

usually evolved from (distal) demonstratives and indefinite ones from the quantifier 'one' in many languages. So, it may be assumed that grammaticization was involved in this case, which suggests later, optional emergence of these markers compared to quantifiers and demonstratives. These considerations too lead to the conclusion that the linguistic technology evolved gradually and incrementally, with different hominine populations innovating similar but not identical strategies at the same stages of cognitive evolution, a normal state of affairs in cultural evolution. Similarities in evolutionary trajectories do not necessarily reflect constraints exerted by UG. They may also reflect system-inherent limitations that offered no other alternatives but to exapt those materials that are the most adaptive for the new functions. The fact that similar grammaticization processes have been repeated in the histories of modern languages may reflect a bias that the modern problem-solving mind has undergone in human evolution. Perhaps this is the best evolutionary account there is for the concept of UG. The concurrent variation regarding quantifiers and demonstratives would have involved word order (notwithstanding the internal structure of the domains), about which nothing has been said so far. [22]

Whether the emergent grammar opted for SVO, SOV, or some other kind of major-constituent order in the sentence, and for head first or head last in the arrangement of constituents within the noun phrase appears to be as inconsequential to the emergent technology as the design of cars with the steering wheel on the left or the right side of a four-

[22] An interesting analog to the emergence of typological variation in the evolution of languages, assuming polygenesis, is the evolution of desktop computers, with the Mac and PC engineers developing similar but not identical technologies at more or less the same times with different degrees of success with certain tasks. Despite differences, they satisfy the needs of their customers, in their own respective cultures, and they do not fail in their most basic tasks. Each technology is also ready to adapt to their customers' new needs.

wheel vehicle, though there are secondary consequences, such as the location of the ignition key and dashboard commands.

The same variation also emerged regarding whether or not a language would resort to case markers, postpositions, or just word order to express thematic relations in an utterance.[23] However, all languages have developed markers of relations between some constituents, to express joint participation (conjunctions), accompaniment (comitative relations), and instrumental, manner, and temporal adjuncts to the core predicate phrase, among others. Since the direction and origination of motion is, more often than not, not incorporated in the meaning of the verb, languages usually include adpositions and/or special case markers to mark these verb complements. The relevant systems are sometimes elaborate in some languages, such as *to, into, onto, out of,* and *from*, in English. In some languages, they are not, especially if the direction and origination of motion are incorporated in the relevant verbs. Some languages have opted for redundancy, expressing with the adposition what is incorporated in the meaning of the verb, such as *extract* NP *from* and *import* NP *(in)to*.

Also, as Heine & Kuteva (2007) show, the history of many languages suggests that many adpositions, if not all of them, have evolved from erstwhile nouns or verbs (the case of prepositions such as *down*, *inside*, *behind*, *while*, *including*, *during*, *except(ing)*, and *regarding* in English). Thus, they are apparently later developments, by exaptation of extant materials, than the fundamental categories of nouns and verbs. How does their emergence bear on that of predication? Can we assume that, like many other aspects of language, the internal structure of

[23] Escure (2009) is perhaps the latest creolist discussing serial verb constructions to point out that serial verbs often express thematic relations that some languages express only with adpositions. Serial verb constructions are indeed an important typological option for sentence expansion, to which I return below.

the predicate phrase evolved incrementally? At what stage of language phylogeny are the syntactic structures requiring adpositions likely to have emerged? May those marking complements of motion verbs, or transfer of possession for some verbs (like *take* and *give*), have emerged earlier than those marking adjuncts?

It seems to me that the question is really that of when these devices emerged in language, rather than that of whether they emerged later than the categories of nouns and verbs. On the other hand, we cannot overlook the fact that these functional categories (prepositions and conjunctions) are not as extensive in some languages than others, in which combinations of nouns and a connective functionally similar to *of* in English are used to express the same grammatical relations (in fact as in *in front/back of*, *on top of*, *aside from*, *except for*, and *along with* in English itself).[24]

These facts suggest that perhaps just a small core of "all-purpose" connectives were used, aside from verbs, to mark grammatical relations in the early stages of the phylogenetic emergence of language. The derivative prepositions and conjunctions that have extended the inventories of functional categories must have emerged later, apparently in order to make communication either more precise or less vague, depending on one's idea of what precise communication is. Note that, even during this particular phase of the evolution of language, there is also room for typological variation even within the same language. While there are many cases where new unitary prepositions have been grammaticized from nouns and verbs,

[24] This kind of vocabulary expansion is not restricted to prepositions only. It is also attested in the category of quantifiers, with composite expressions such as *a lot of*, *a number of*, *an amount of*, *a multitude of*, and *a bunch of* in English. That their lexical status is somewhat different from quantifiers such as *many*, *several*, and *few* is made evident by the fact that they are intensified differently, for instance, *a whole lot of*, *a (very) large number of*, and *a certain amount of*, as opposed to *very many* and *very few*, with the latter behaving like adjectives and the former like nominal phrases.

there are also many cases of composite prepositions. Some languages have chosen this particular option. French has a few, such as *au sujet de* 'about', *autour de* 'around', *en provenance de* 'from', and *au dessus de* 'on top of, above'. The transparency of the prepositional phrases (not *preposition phrase* as P + NP) is as striking as that of the semi-auxiliary *be going to* (+ infinitive) in English or that of prepositional phrases such as *on top of, in back of, in front of*. Again the evidence speaks for a gradual, perhaps protracted evolution of language toward more and more productive and increasingly complex syntax, reflecting how extant materials were coopted toward the production of new expressions. As suggested above, we have no reason for assuming that at the initial stages of the emergence of language our hominine ancestors' minds and their communicative needs were just like ours today. Consistent with Darwinian evolutionary theory, the gradual complexification of the language technology suggested by the data may reflect the gradual complexification of the hominine cognitive capacity and need for more sophisticated communication.

The gradual complexification of the grammar is also evident in the use of a clause as the complement of a higher predicate, as the modifier of a noun (e.g., relative clauses), or as an adjunct adverbial clause. Made possible by the principle of recursion,[25] this expansion of the

[25] Although recursion has properly been associated with the repetition of particular combinatorial patterns in syntax, typically with complement and modifying clauses and with multiple embeddings of preposition and noun phrases, one may also argue that it is deeply rooted in the strategy of compositionality itself, which allows longer and longer expressions. Originally suggested by Martinet's (1960) principle of "double articulation" and Hockett's (1959) and Hockett & Ascher's (1964) "duality of patterning," compositionality applies in more than one module of the architecture of language, viz., in combinations of sounds into morphemes and words, as well as in combinations of the latter into sentences, eventually in combinations of the latter into larger discourse. Syntax has focused on a specific kind of it, whereby some patterns for expanding constituents are repeated within themselves to produce even larger phrases. We also know from modern grammars that constraints emerged that restrict particular combinations and various syntactic processes from occurring under particular conditions. The syntactic constraints may have a cognitive basis, while the phonetic ones may have acoustic and/or articulatory motivations. We will not deal with them in this essay.

sentence also led to the emergence of complementizers, among other alternative devices (such as the conversion of the verb into a gerund, i.e., verbal noun), in systems where particular subordination strategies emerged (as in Western European languages), as opposed to verb serialization.[26]

Givón (2009) hypothesizes that subordination evolved from serialization. A problem with this hypothesis is that we have no evidence that all subordinators evolved from verbs, although there are some languages where the serial verb 'say' has evolved to function as a complementizer. Although this does not rule out the possibility that the first instances of clause-embedding involved no subordinators, it is also noteworthy that some serial verbs express thematic relations that are expressed with adpositions in some other languages. This suggests that serial verb constructions are not simply the counterparts of subordinate clauses. One might otherwise argue that coordinate clauses and preposition phrases all evolved from serial verb constructions. The evidence for the position is just not compelling; it is typically based on idiomatic translations of serial verb constructions into English or French. The typological evidence simply suggests that different populations have chosen to organize their grammatical systems differently, consistent with the earlier observation that the human mind is capable of producing alternative solutions to the same problems. This is a central point in Evans

[26] We must bear in mind that the term *serial verb construction* was coined in reference to the fact that in languages producing such constructions verbs or verb phrases are sequenced without conjunctions or subordinating morphemes between them. The syntactic relations between the serialized verbs are not uniform: the verbs may be considered as conjoined in some constructions, whereas in some others the serial verb is the object of the head verb, as in the Gullah construction *Uh tell um come home* 'I told him/her [to] come home'. The mischaracterization of some serialized verbs as "instrumental" and "dative" also reflects the fact that their semantics are not uniform from one construction to another. As argued in Mufwene (1989, 1990), serialization is just an alternative strategy, in opposition to morphosyntactic subordination and conjunction (with overt markers for specific syntactic functions), for the formation of complex sentences. See also Evans & Levinson (2009: 442) for a similar position.

& Levinson's (2009) argument for not subordinating linguistic diversity to language universals. From the perspective of the present essay, the relevant problem is how to pack various information hierarchically or otherwise in a sentence.

A less controversial hypothesis is that subordinators emerged after the application of recursion to clause-size units made it possible to expand a clause with another clause that functions as a complement of a verb, as the modifier of a noun, or as an adverbial adjunct. Languages vary regarding the strategies they evolved to mark such clauses. More research is needed regarding what structures had to be in place before others that possibly coopted them emerged. This will then help us determine with some certainty at what particular phase of the phylogeny of languages some devices and structures that do not appear to be fundamental evolved. The fact that the primary characteristic of serial verb constructions is not to have any function words between the sequenced verbs and for the verbs to be uninflected may help us sort out what are some of the non-fundamental devices in the architecture of modern languages. Candidates include not only complementizers but also the distinction between finite and nonfinite forms of verbs, as well the nominalization of the subordinate construction as done in Japanese and Korean before a postposition is attached to it to mark its syntactic function.[27]

The above observation is not intended to deny the usefulness of complementizers and other subordinators in facilitating sentence-processing. It simply highlights the fact that some

[27] The way complement clauses in Japanese and Korean are bounded morphologically with a nominalizer followed by a postposition suggests in fact that speakers had to realize at some point that the noun or noun phrase could designate not only an entity (material or abstract) but even a state of affairs. What Japanese and Korean do so transparently is reminding the hearer that the complement clause must be interpreted as a noun phrase. We can thus be certain that the capacity for a verb to have a nominal object provided the substrate for the emergence of complement clauses.

devices in the architecture of languages may be treated as accessories that enhance how they functions but are not essential or critical parts (Mufwene 1999). This could also have been said of the copula, gender, nominal number, and agreement strategies (between an overt subject and its verbal predicate and between a noun and its adjectival modifier), the kinds of things that pidgins dispense with (not inevitably!) and whose functions children understand later during child language development.

Recursion has also enabled the modification of a noun by another noun as in the English *my teacher's book's cover* and its French translation *la couverture du livre de mon professeur*, as well as the modification of a preposition phrase by another preposition phrase as in *the wine in the glass on the table by the window*. Whenever these kinds of constructions emerged during the phylogenetic emergence of language(s) (indeed when English had not emerged yet!), they presuppose the earlier emergence of the simpler constructions that could be coopted into the production of the larger ones.[28] Exaptation was thus a key strategy in the phylogenetic emergence of language(s).

4. **Conclusions**

[28] The reader should remember that the emergence of modern languages such as English and French did not follow the same evolutionary trajectory discussed here. They reflect historical change from ancestor languages and speciation under particular contact conditions, although some of the processes involved in the phylogenetic emergence of language have been repeated, such as grammaticization. Some typological shifts have also taken place during the relevant changes such as from VSO to SVO in English. All this is only part of the complex story showing that the linguistic technology continues to be adapted to its users' communicative needs under particular ecological pressures, including language contact (Mufwene 2009). This all means that what language history can tell us about the phylogenetic emergence of language may be very limited, especially since historical changes have not always proceeded from less to more complex structures. Often, just the opposite has been true, such as in the morphosyntax of Indo-European languages such as English and the Romance languages, from the point of view of distinctions within particular domains or of constraints on the application of particular rules.

The primary goal of this essay was to show how languages can be thought of as emergent technology whose function is to enable sophisticated and more explicit communication than, as far as we know, nonhuman primates' vocalizations and gestures. In the way that Pinker & Jackendoff (2005: 231) explain it, "it is an [evolutionary] adaptation for the communication of knowledge and intentions." Though it is more fashionable to think of writing as technology, that of putting on material objects (clay, papyrus, paper, and now the computer screen) what would otherwise be spoken, language is itself technology developed through the domestication by hominines of their own anatomy to express their thoughts and feelings, to describe various states of affairs around them, to relate past experiences, to plan future states of affairs, and, as claimed by several students of the evolution of language, to cooperate toward the sustenance of their societies.

The second goal was to speculate on how this technology evolved, albeit in an unplanned way, by exaptations by hominines of extant structures, in response to growing communication pressures arising from their expanding mental and cognitive capacities. It seems obvious that, while domesticating appropriate anatomical parts of their bodies, the technology started with the capacity to name entities and states of affairs, with the phonetic inventories increasing up to a critical mass (variable from one population to another) to satisfy the continuous expansion of the vocabulary. This evolution already set hominines apart from nonhuman primates, whose vocabulary of calls is considered limited to innate needs and does not expand. The next step was the emergence of predication and strategies to specify reference

(for both nouns and verbs[29]), a capacity that linguistic apes have hardly acquired (at least not to the extent observable in modern human languages), which underscores the evolution of a mind particularly hominine as a critical evolutionary prerequisite to the emergence of the linguistic technology. So what sets us apart from other animals is not so much language as the mind that made it all possible and also accounts for why our cognition and cultures are also so different from those of other animals.

Past these initial stages, the story of the phylogenetic trajectory of language has been more difficult to construct, defying any unilinear and perhaps also rectilinear account. My admittedly unfinished story shows how partial accounts such as Jackendoff's (1999), Heine & Kuteva's (2007), and Givón's (2009) are. My speculations agree only partially with theirs, certainly in showing that the evolution was incremental and that some strategies and devices presuppose prior emergence of some others that served as their substrates. This conclusion underscores the need to better understand the ways in which the different aspects of the architecture of language are interconnected. We must sort out what particular components have evolved concurrently and which ones presuppose the earlier emergence of others. We must also think hard over what counts as evidence for our hypotheses.

I have also highlighted the relevance of typological variation to the question of whether we should conjecture monogenesis or polygenesis in the evolution of languages. Taking into account the geographical distribution of paleontological evidence across East(ern) Africa, it is difficult to dismiss the possibility of polygenesis as what made typological variation possible.

[29] I am aware of the tradition that associates reference only with nouns, but linguists also speak of "time reference." It is both meanings that are intended here.

Different hominine populations need not have solved their communication problems in identical ways, though they were developing technologies that co-opted the same anatomical structures to produce speech or manual signs. Uniformity in the phylogenetic emergence of languages seems to end with the choice of the same materials in the development of the technology and the common, universal principles that apply because they may be consequences of both the materials used and the mind producing the technology.[30]

These considerations suggest that we could as well be speaking of the evolution of languages, indeed in the plural, with emphasis of diversity, at least sometimes, rather than just speaking of the emergence of language, in a non-individuated and unified way. If the commonalities are largely consequences of both the materials used to produce the technology and the mind driving the emergence of the technology, then scholars such as MacNeilage (2008), Koster (2009), Lieberman (2010), Jackendoff (2010), and others are justified in questioning the postulation of a language organ or Universal Grammar as THE central factor accounting for how languages have evolved and for their common properties. If anything, the name *language organ* itself must be a misnomer, especially because it corresponds to no anatomical part of hominine structure, nor to any combination of organs specializing just for language.[31] Universal Grammar, as the sum of mental specifications and/or constraints

[30] Evans & Levinson (2009) are among those that are reluctant to invoke UG (under whatever name) to account for the common properties, especially constraints, shared by most, if not all, languages. They include a relatively detailed discussion on p. 439 that I will not try to summarize here.

[31] David Lightfoot, who has perhaps written the most on the subject matter, is not particularly informative. To wit, Lightfoot (2005) characterizes the "language organ" variously as "biological," "psychological," "bodily," and yet not anatomical (in that it cannot be localized in the brain), quite unaware of the confusion his inconsistent characterization produces. He also argues that the "linguistic capacity [or grammar, also identified with the "language organ"] is better understood in functional rather than anatomical terms" (57) and later compares the development of grammar in the child to the growth of hair (58), subject to particular ecological conditions, the

putatively determining the properties shared by all languages, would be a consequence of the evolution I have attempted to account for here. The biological foundation of languages, more accurately interpretable as cultural artifacts, appears to lie in the way that biological evolution endowed hominines with a particular evolving mind capable of producing the sophisticated modern communication technologies we now produce and use. Because of other cultural artifacts and conventions that only humans produce and use today that other animals do not, I submit that what distinguishes humans from other animals is not so much language as the mind that made it all possible. Of the lessons I have learned from this writing this essay, perhaps the most significant is how complex the subject matter of the phylogenetic emergence of language(s) is and what a particularly rich perspective one gains from thinking of language(s) as technology, and especially how many questions arise, the answers to which are not obvious.

References

Aronoff, Mark, Irit Meir, & Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81: 301-344.

Arthur, Brian W. 2009. *The nature of technology: What it is and how it evolves*. New York: Free Press.

Bickerton, Derek. 1990, *Language and species*. Chicago: University of Chicago Press.

Bickerton, Derek. 2007. Language evolution: a brief guide for linguists. *Lingua* 117: 510-26.

"trigger." This is at best an elusive characterization of UG that can hardly be operationalized to explain either how language evolved or how specific languages are learned by their speakers.

Bickerton, Derek. 2010. *Adam's tongue: How humans made language, how language made humans.* New York: Hill & Wang.

Brentari, Diane. 2011. Sign language phonology. In *The handbook of phonological theory*, John Goldsmith, Jason Riggle, & Alan Yu (eds.), 691-721. Malden, MA: Blackwell.

Carstairs-McCarthy, Andrew. 1999. *The origins of complex language: an inquiry into the evolutionary beginnings of sentences, syllables, and truth*. Oxford: Oxford University Press.

Corballis, Michael C. 2002. *From hand to mouth: the origins of language.* Princeton: Princeton University Press.

Corballis, Michael C. 2010. Did language evolve before speech? In Larson *et al*. (eds), 115-123.

Corballis, Michael C. 2011. *The recursive mind: The origins of human language, thought, and civilization*. Princeton: Princeton University Press.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Dascal, Marcelo. 2002. Language as a cognitive technology. *International Journal of Cognition and Technology* 1: 35-61.

Deacon, Terrence R. 1997. *The symbolic species: the co-evolution of language and the brain.* New York: W. W. Norton and Company.

Dunbar, Robin. 1996. *Grooming, gossip, and the evolution of language.* Cambridge, MA: Harvard University Press.

Escure, Geneviève. 2009. I verb serialization simple? Evidence from Chinese Pidgin English. In *Simplicity and complexity in creoles and pidgins*, Nicholas Faraclas & Thomas B. Klein (eds.), 109-123. London: Battlebridge.

Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32: 429‑492

Frank, Michael C., Daniel L. Everett, Evelina Federenko, & Edward Gibson. 2008. Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108: 819-824.

Givón, T. 1979. *On understanding grammar*. New York: Academic Press.

Givón, T. 2009. *The genesis of syntactic complexity*. Amsterdam: John Benjamins.

Goldin-Meadow, Susan. 2003. *The resilience of language: what gesture creation in deaf children can tell us about how all children learn language.* New York: Psychology Press.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies.* The Hague: Mouton.

Heine, Bernd & Kuteva, Tania. 2007. *The genesis of grammar: A reconstruction.* Oxford: Oxford University Press.

Hockett, Charles F. 1959. Animal "languages" and human language. *Human Biology* 31: 32-39.

Hockett, Charles F. 1978. In Search of Jove's Brow. *American Speech* 53: 243–313.

Hockett, Charles F. & Robert Ascher. 1964. The human revolution. *Current Anthropology* 5: 135-168.

Humboldt, Friedrich Wilhelm Christian Karl Ferdinand Freiher von (1836-1839), *Ueber die Kawi-Sprache* (Munich: Saur). Transl. by Peter Heath as *On language: The diversity of human language-structure and its influence on the mental development of mankind* (Cambridge: Cambridge University Press).

Jackendoff, Ray. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3: 272-279.

Jackendoff, Ray. 2010. Your theory of language evolution depends on your theory of language. In Larson *et al*. (eds.), 63-72.

Johannson, Sverker. 2005. *Origins of language: Constraints on hypotheses*. Amsterdam: John Benjamins.

Koster, Jan. 2009. Ceaseless, unpredictable, creativity. *Biolinguistics* 3: 61-92.

Larson, Richard K; Déprez, Viviane; & Yamadiko, Hiroko (eds). 2010. *The evolution of human language: Biolinguistic perspectives.* Cambridge: Cambridge University Press.

Lee, Namhee; Mikesell, Lisa; Joacquin, Anna Dina L.; Mates, Andrea W.; & Schumann, John H. 2009. *The interactional instinct: The evolution and acquisition of language*. Oxford: Oxford University Press.

Lieberman, Philip. 2010. The creative capacity of language, in what manner is it unique, and who had it? In Larson *et al*. (eds), 163-175.

Lightfoot, David. 2005. Plato's problem, UG, and the language organ. In *The Cambridge Companion to Chomsky*, James McGilvray (ed.), 42-59. Cambridge: Cambridge University Press.

MacNeilage, Peter. 2008. *The origin of speech.* Oxford: Oxford University Press.

Maddieson, Ian. 2006. In search of universals. In *Linguistic Universals*, R. Mairal & J. Gil (eds.), 80-100. Cambridge: Cambridge University Press.

Martinet, André. 1960. *Elements de linguistique générale*. Paris: Armand Colin.

McArthur, D. 1987. Le langage considéré comme une technologie. *Cahiers de Lexicologie* 50: 157-164.

McCleary, Leland & Evani Viotti. 2010. Sign-gesture symbiosis in Brazilian sign language narrative. In *Meaning, form, and body*, Fey Parrill, Vera Tobin, & Mark Turner (eds.), 181-201. Stanford: CSLI Publications.

McNeill, David. 2005. *Gesture and thought.* Chicago: University of Chicago Press.

Mufwene, Salikoko S. 1989. Equivocal structures in some Gullah complex sentences. *American Speech* 64: 304-26.

Mufwene, Salikoko S. 1990. Serialization and subordination in Gullah. In *When verbs collide: Papers from the 1990 Ohio State Mini-Conference on Serial Verbs*, Brian Joseph & Arnold Zwicky (eds.), 91-108. (*OSU Working Papers in Linguistics* 39).

Mufwene, Salikoko S. 1992. Why grammars are not monolithic. In *The joy of grammar: A festschrift in honor of James D. McCawley*, Gary Larson, Diane Brentari, & Lynn McLeod (eds.), 225-50. John Benjamins.

Mufwene, Salikoko S. 1999. The language bioprogram hypothesis: Hints from Tazie. In *Creolization, diachrony, and language acquisition*, Michel DeGraff (ed.), 95-127. Cambridge, MA: MIT Press.

Mufwene, Salikoko S. 2001. *The ecology of language evolution.* Cambridge: Cambridge University Press.

Mufwene, Salikoko S. 2008. *Language evolution: contact competition, and change.* London: Continuum Press.

Mufwene, Salikoko S. 2009. The evolution of language: Hints from creoles. In *Language, Evolution, and the Brain*, James Minett and William Wang (eds.), 1-33, Hong Kong: City University of Hong Kong Press.

Mufwene, Salikoko S. in press. The origins and the evolution of language'. In *The Oxford handbook of the history of linguistics*, Keith Allan (ed.). Oxford: Oxford University Press.

Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: What's special about it?', *Cognition* 95: 201-236.

Pollick, Amy S. & Frans B. M. de Vaal. 2007. Ape gestures and language evolution. *Proceedings of the National Academy of Sciences* 104: 8184-8189.

Rousseau, Jean-Jacques. 1755. *Discours sur l'origine et les fondements de l'inégalité parmi les hommes.* Amsterdam : M. M. Rey. Translation of "Essai sur l'origine des langues," published in Moran & Gode (1966).

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*, ed. by Charles Bally & Albert Sechehaye in collaboration with Albert Riedlinger. Paris: Payot. Trans. Wade Baskin. 1966. *Course in general linguistics.* New York: McGraw-Hill.

Schleicher, August. 1853. Die erstern Spaltungen des indogermanischen Urvolkes. *Algemeine Monatschrift für Sprachwissenschaft un Literatur*, September 1853, 786-787.

Sperber, Dan & Deidre Wilson. 2002. Pragmatics, modularity and mind-reading. *Mind and Language* 17: 3-23.

Stokoe, William C. 2001. The origins of language. In *The Routledge companion to semiotics and linguistics*, Paul Cobley (ed.), 40-51. London: Routledge.

Swisher, M. Virginia. Similarities and differences between spoken languages and natural signed languages. *Applied Linguistics* 9: 343-356.

Tallerman, Maggie. 2007. Did our ancestors speak a holistic protolanguage? *Lingua* 117: 579-604.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition.* Cambridge: Harvard University Press.

Tomasello, Michael & Joseph Call. 2007. Ape gestures and the origins of language. In *The gestural communication of apes and monkeys*, Joseph Call & Michael Tomasello (eds.), 221-239. London: Lawrence Erlbaum & Associates.

Wray, Alison. 1998. Protolanguage as a holistic system for social interaction. *Language and Communication* 18: 47-67.

Wray, Alison. 2000. Holistic utterances in protolanguage. The link from primates to humans. In *The evolutionary emergence of language: Social function and origins of linguistic form*, Chris Knight, Michael Studdert-Kennedy, & James R. Hurford (eds.), 285-302. Oxford: Oxford University Press.

Wray, Alison. 2002. Dual processing in protolanguage: performance without competence. In *Transition to language*, Alison Wray (ed.), 113-37. Oxford: Oxford University Press.