Grammar Came Later:

Triality of Patterning and the Gradual Evolution of Language

Daniel L. Everett

Bentley University

deverett@bentley.edu

Abstract:

This paper argues that language is primarily a tool for communication, rather than principally a tool for thought expression. It makes the case that language has its roots in intentional iconicity of Australopithecines and probably had reached the level of a G1 grammar (linear ordering of symbols + gestures) some 2-2.5 million years ago. Other forms of language, e.g. hierarchical, recursive grammars, are later embellishments that are neither necessary nor sufficient to have human language. The paper looks in detail at the evolution of culture among early hominins and how gap between indexes and icons to symbols might have been bridged. It then discusses the basic composition of phonology, morphology and syntax. The paper rejects the idea of a proto-language, as it also rejects the "X-men" view of language evolution/mutation proposed in Berwick and Chomsky (2016).

1. Introduction

1.1. The cultural-biological-linguistic network

In recent work, Fuentes (2016) argues for an "extended evolutionary synthesis," or EES. He claims that a full picture of our species engages the biological,

the cultural, and the psychological simultaneously as part of this EES-based understanding of the human condition. At the same time, Fuentes acknowledges that current models of what "culture" is and how it interacts with the human psyche and body are poorly developed, at least in the social sense that there is no broad social consensus on what exactly culture is. In Everett (2016) I address the nature of culture and propose a set of fundamental components for it and how my particular notion of culture then interacts with our minds, bodies, and social relationships. That conceptualization of culture plays a role here.

Based on an adequate theoretical foundation, it becomes possible to consider how cultural pressures might have altered our biological development and how our biological development in turn may have altered our cultures, in particular how both led to the evolution of language. The view of what follows below is certainly not completely new. Lieberman (2016), Dediu and Levinson (2013), Luuk and Luuk (2014), Luuk (2013), and many others have addressed similar points and made similar suggestions, long before me.

What is novel about my approach, I believe, is the more articulated notion of culture and a wider understanding of grammar as (at least) a *triality of patterning* (as opposed to, say, Berwick and Chomsky (2016) in particular, but also to even researchers whose work is otherwise compatible with mine, e.g. Dediu and Levinson (2013)). For example, I argue that language (as a triplet of grammar and symbols, along with the combination of intonation and gestures, which mark information structure) evolved by a semiotic progression and that this progression began with discovered icons and moved minimally to non-recursive, non-hierarchical linear

grammars (hierarchical and recursive grammars are secondary embellishments of language and are neither necessary nor sufficient for language). Based on uniformitarian assumptions, I will argue that language likely began with Homo erectus. The idea that language is largely dependent on a particular kind of grammar and that it "popped" into being some 50,000-65,000 years ago, has only weak negative evidence in its favor and fits the facts far less well than the assumption that language evolved slowly and gradually across all Homo species.

As Searle (1972) has pointed out in the past and as Berwick and Chomsky (2016) make clear, a language in the Chomskyan sense is a specific type of grammar, rather than in the sense in which "language" is used by most other researchers and among laypeople. For something to be a human language in their model, it must have a recursive, hierarchical structure-building process (Merge). My view of the grammar is much different, based in part on my own field research, but closely aligned with much other work (Van Valin and LaPolla (1997), Goldberg (1995; 2006), Futrell, et. al. (2016), etc.). and observations. Grammar in the model presented here is a triality of patterning, along with a means (e.g. linearity, hierarchical structures, morphological case, etc.) for interpreting semantic relationships.

Another difference between my view of the evolution of language and that of others, is the need I perceive to see language as a system in which gestures and intonation work in parallel with the *duality of patterning* or *double-articulation* or language that emerges from the work of Hockett (1960), among others., to produce *triple articulation*.

1.2. The emergence of language

Language, whatever its biological basis (on that see section 3 below) is shaped by psychology, history, culture, and so on, as we represented in FIGURE ONE:

Figure One
Language is a nexus

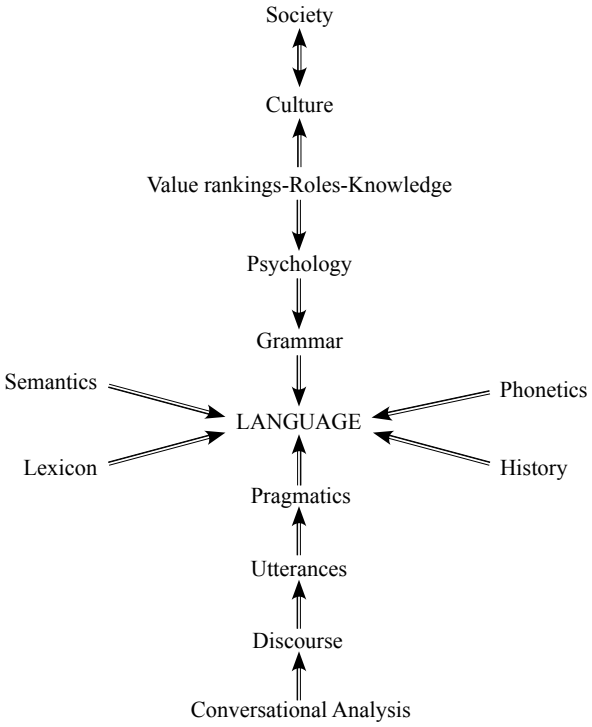


FIGURE ONE summarizes my understanding of language as the confluence of numerous factors external and internal to humans (for a fuller view see Everett 2012; 2016; 2017). At the "north" end of the figure, there are the mental sources of language – the interactions among the societal roles, ranked values, and knowledge structures of culture, as these constrain the development of human psychology and the meaning-computational principles known as grammar. At the south end are the uses of language, from the general set of pragmatic constraints to the outer edges of use in conversation and so on. At the west and east are constraints on physical form and meaning, from semantics to history. Language is that point (or set of points) where all of these factors converge. No one of them is sufficient for language, though all are necessary. If all these factors are what make language possible, then they must have evolved together.

1.3. Uniformitarianism vs. catastrophism

Before answering the "how" question of evolution it is important to make clear one guiding assumption I make. This is that in the absence of evidence to the contrary, a simple assumption throughout scientific disciplines is that the laws that hold today also held yesterday. And the day before. More precisely we assume that the same laws of nature that govern the universe we live in now and its current processes have always worked as we now see them.

In physics, for example, at least there has never been a reason to question the assumption of uniformitarianism (on earth following its formation). Physical laws show no evidence of having changed during the earth's natural history. Lyell's 1833

Principles of Geology develops the reasonableness of uniformitarianism assumptions in earth history studies.

Is there any utility of uniformitarianism in evolutionary, biological, or cultural theories? Well, one useful guideline is that saltations and other catastrophic events should not be proposed unless there is strong evidence to do so (Occam's razor). Otherwise, we can assume that a (mathematically sophisticated, see Berwick and Chomsky (2016, 11ff)) model of evolution accounts for the transformation of ancient life forms into modern life forms via gradual, uniformitarian procedures.

One is free to propose mutations or anything one wants to construct a theory. And of course mutations are among the main drivers of evolution. But the rule of thumb in such proposals must be that "in the absence of evidence, do not assume miracles." Proposals of mutations are otherwise unwarranted and speculative. This is to me the major weakness of the Berwick and Chomsky (2016) "X-men" or "Prometheus" theory of language appearance, a sudden saltation to grammar. There is no need to appeal to mutations when uniformitarian evolution does the job as well or even better.

Moreover, Chomsky and Berwick's (2016) and Tattersall's (2012), among others, speculations on language evolution depend heavily on arguments from the absence of evidence. We know, for example, that current Amazonian communities speak modern human languages. But what record would we have of their languages if they all died out and if we searched for them 500,000 years hence? Forgetting for now that linguists and anthropologists have published grammars, dictionaries, and other studies of Amazonian communication, language, and culture, would they leave

any evidence in their material culture that they were capable of language or symbolic reasoning? Most would not. Arguably even less in fact than what is found for Neanderthals. Unlike Neanderthals, Amazonian art (necklaces, basket designs, carvings, and so on) and tools (bows, arrows, blowguns, darts, poison, baskets, etc.) are biodegradable (aside from the very few cultures that make ceramics). So their material culture would likely leave no trace. We could determine, perhaps, from soil usage that they had villages of a certain size, huts, and so on (Heckenberger, et. al. (2008)), but it would be as difficult to extrapolate from the remnants or other evidence of artifacts that they had language – just as is the case for many ancient hunter-gatherer groups – such as Homo sapiens as it left Africa, or Neanderthals, or Denisovans, Erectus, and so on. Therefore, we must be careful not to conclude, as is the wont of many, based on the absence of evidence about language in the prehistoric record, that earlier hominins lacked language. Uniformitarianism would support the assumption that ancient humans shared all features with modern humans, absent evidence to the contrary.

1.4. Organization of this paper

With the basic view of language as an emergent, culturally-based communication system, ever-evolving for more effective communication, we move on to deliver on the promissory note of offering an account of how this occurred, focusing principally on language's semiotic evolution (but cf. Everett (2017)).

Section two explores in greater depth the idea that language is primarily for communication, also answering the Chomskyan objections to this notion. My thesis

also includes the mundane idea that language developed gradually and that it is not the result of a grammatical saltation for thought-expression, but rather developed symbiotically with the growing complexity of culture and the human brain. In the gradual evolution of language, my main thesis here is that grammar was the latest addition in a three-million year long semiotic progressions and that the more elaborate, structure-dependent, recursive grammar proposed by Chomskyan theory is neither necessary nor sufficient for human language. In my view, again, grammar is a triality of patterning (see below).

Section three focuses on the neurological and physiological adaptations for language in the brain and the vocal tract anatomy. It makes the case that there are no specialized organs or genes for language, though plenty of evidence for evolutionary "recycling" of preexistent structures for human language, likely beginning to appear in their modern forms gradually, as early as two million years ago.

Section four examines the archaeological evidence for the semiotic progression. The different steps in this progression appear in the fossil record exactly as predicted in the Peircean model sketched in the second section of the paper below. We also explore the contributions of art, symbols, and tools, as well as cultural displacement, shared intentionality, and the internal construction of semiotic interpretants to language evolution.

In section five we examine Peirce's notion of infinite semiosis, noting that by virtue of this feature of symbols, the seeds of grammar are inherent in symbols,

where infinite semiosis is one of the precursors of grammatical recursion (other precursors, including gestures and the syllable are discussed in later sections).

This leads naturally into section six, where duality of patterning is discussed.

Section seven looks at a level of semiotic organization that Hockett and traditional linguistics have missed, namely, gestures, which I claim bring language into a triality of pattering, rendering it more complex than either Hockett or Chomsky imagined. (Recent work, such as Floyd (2016) underscores the importance of gestures and the triple articulation of language.)

Section eight takes up in earnest a discussion of the structural components of language – phonology, morphology, syntax, and semantic compositionality, (but see Everett 2017 for more detailed exposition).

The ninth section brings the theory of culture developed in Everett (2016) into the discussion, arguing for the centrality of culture and dark matter to the evolution of language.

In the conclusion to the paper I sum up what we have learned and suggest some new directions for future research.

2. The Nature of Language

2.1. Language is for Communication

In spite of frequent claims to the contrary (Everaert, et. al. (2015), Berwick and Chomsky (2016)), language is primarily a cultural tool for community building (Everett (2012)). Language's computational properties, such as recursive structure-dependency, or even simple linearity, are secondary. Expression of thought,

proposed by Chomskyans as the raison d'etre of language is also a secondary feature, according to the evidence we have in contemporary languages and the evolutionary record. But the uses of grammar for thinking and communication are dependent upon one another and each likely enhances the other.

If communication is the basic function of language, we are not quite so unlike other creatures as some modern linguistics theories would have it. Communication is pervasive throughout the animal kingdom. We are simply the best communicators. And herein lies our distinctiveness.

As stated, an alternative hypothesis comes from the work of Chomsky and his followers. The following quote is typical (from Everaert, et. al. 2015, p729):

> *"... communication, a particular use of externalized language, cannot be the primary function of language, a defining property of the language faculty, suggesting that a traditional conception of language as an instrument of thought might be more appropriate. At a minimum, then, each language incorporates via its syntax computational procedures satisfying this basic property... We take the property of structure dependence of grammatical rules to be central."*[1]

This is a profoundly impoverished and mistaken view of language. It fails to account, among other things for: the evolutionary record; the role of gestures and culture the architectonic shaping of grammar; and redundancy and ambiguity as natural features of language (Piantadosi, et. al. (2011); Everett (2012; 2017)). Yet it

is widespread. Therefore, we must address it. The authors claim in this quote that communication is a "particular use" of "externalized language." This makes it sound rare indeed. They believe that the only kind of language we can actually study is so-called *I(nternal)-language* that is, what the speaker actually knows, rather than an abstract *E(xternal)-language* one might refer to as French, English, Spanish and so on. In fact it is just the opposite. We can *only* study performance, E-language, and the like, in order to draw inferences about the "internal language" of our *theory*. The interactions we observe from speakers are one source of evidence of what speakers know whether via experiments, natural texts, or elicited data. One can embrace the obvious fact that the E-language "English" is an abstraction in order to recognize that recorded utterances in exchanges are *not themselves abstractions*, but the only concrete, empirical sources of what speakers know and cultures produce and what people actually do. Saying that one ignores performance to get at competence (as these authors might assert) is like saying that college exams show nothing because they only measure performance, not competence. Performance just is the only light we have to shed on competence, whether it is performance in dialogue, experiments, story-telling, or individual sentence-elicitation. Moreover, this quote and others like it simply ignore the fossil record.

2.2. Language developed gradually and grammar came last

The principle thesis of this paper is this: *language and its triality of patterning developed symbiotically with culture and the brain gradually, across all hominin*

species. That is, every hominin species directly had some component of human language – from Australopithecines to Sapiens.[2]

The central question to be addressed below is how our species acquired language-quality communication beginning from mere nonlinguistic communication. To find the answer requires us to examine the evolutionary history of Homo sapiens, in order to identify the first clues of language emergence in the evolutionary record. This is a difficult and indirect task, however, because there are no linguistic fossils from our earliest hominin ancestors. No one will dig up the first linguistic sound waves, gestures, or relative clauses. We are forced to use evidence from material culture, migrations, encephalization, and the emergence of cultural practices in order to develop a model of the linguistic abilities of early hominins. We attempt to infer what we cannot observe, as in all of science. Nor does grammar seem to be a mutation, as Berwick and Chomsky (2016) suggest. Humans are not linguistic "X-men," nor was there a linguistic "Prometheus." [3]

One might object that merely focusing exclusively on grammar misses what many researchers consider most interesting about language – things like such as conversational interactional patterns, discourse topic-tracking, metaphor, the usage-based accounts of grammatical forms, cultural effects on grammar and the lexicon, and many others.[4]

Pursuing these ideas, I consider three explicit hypotheses for the origin of human language.

2.3. Three hypotheses

The three principal hypotheses that emerge from the language evolution literature, as well as my own work on the origin and evolution of language are summarized below.

Hypothesis 1, *Grammar Came Last*: The most significant step in the evolution of language was the development of symbols. Grammar is a useful add-on.

Hypothesis 2, *Grammar Came First*: The evolution of language is primarily the about the origin of grammar. Symbols came later (Berwick and Chomsky (2016)). Here grammar is recursive, structure-building Merge.

Hypothesis 3, *Grammar Came Later:* The evolution of language required a symbiosis between grammar and symbols, each one affecting the other. In this view, grammar is naturally a triality of patterning (this paper and Everett 2017).

Each of these hypotheses accord a prominent role to grammar. This is because grammar has clearly been of enormous importance in human communication, thinking, and social structures. The questions are: (i) whether this importance necessarily means that grammar is primary – the definitive characteristic of language (as per Berwick and Chomsky 2016) and (ii) whether a grammar beyond triality of patterning is necessary. I argue that Chomskyan-style recursive grammars are neither necessary nor sufficient for human language (as per Everett 2012; 2017). The idea that grammar is important does not mean that it cannot be secondary. But if it is secondary, there are some issues that we must address. That is because this position entails that there can be language without grammars (of the Chomskyan variety), that grammar is not innate so far as we can

determine, and that we should better appreciate and build upon the tremendous variation in the world's languages in our theories of language evolution.

The hypothesis that one defends will affect one's view of how long language has been evolving among hominins. For example, the grammar-first concept (Berwick and Chomsky 2016) suggests that language appeared suddenly, as little as 50,000 years B.P. That is, not even all Homo sapiens would have had language (since the species is more than 200,000 years old), a bizarre claim that would predict that not all humans could learn language (then or even now, since descendants of non-language bearing sapiens might still be alive).[5] But the semiotic progression/triality of patterning hypothesis suggests that language appeared gradually, over the past 3 million years and that all humans today have, and perhaps all Homo species in the past had, langauge.

Some of the reasons for rejecting the grammar came first hypothesis include: (i) there are languages currently spoken that appear to lack any hierarchical grammar (perhaps Riau (Gil (1994); Jackendoff and Wittenberg (2012)) and Pirahã (Futrell, et. al.; Everett 2005)), their "grammars" being little more than linearity restrictions on words; (ii) there is a good deal of evidence that symbols evolved long before grammar in human linguistic history; (iii) hierarchical grammars are derivative from independent processing advantages of hierarchy in the organization and retrieval of information outside of human languages (Simon 1962); (iv) there are no *well-established* cortical specializations for language or speech, aside from recycling, things like adaptations of the tongue, general fluidity of brain functions (Anderson (2014)), and growth of the prefrontal cortex, itself associated with tool-

making (Morgan, et. al. (2016); (v) there are problems with theories that grammar preceded symbols in language evolution; (vi) nonhuman creatures appear to use syntax (so it – at least in its non-recursive, non-hierarchical form – is not exclusive to humans); (vii) humans have evolved away from cognitive rigidity (as via instincts), to cognitive flexibility (Everett 2016) and learning based on local cultural and even environmental constraints (see Caleb Everett (2014; 2015)). Under our assumptions here any grammatical similarities found across the world's languages would are indicative of either functional pressures on language design or the monogenesis of language, both of which are likely hypotheses.

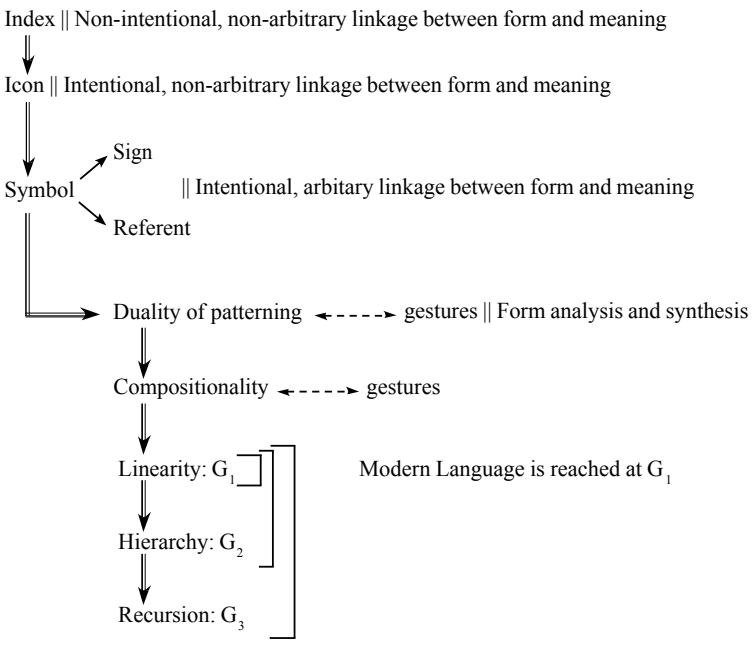
2.4. Language emerged from a semiotic progression and triality of patterning

Thus my thesis in what follows can be labeled *grammar came later* (and this holds especially for recursive, hierarchically-structured grammars) in the evolution of language. This thesis is predicated on the evidence that in human interactions meaning is first and form second. That is, language is primarily a vehicle for meaning. Grammar facilitates meaning transfer, but "grammar" (in the sense of Berwick and Chomsky (2016)) is neither necessary – as shown by Pirahã – nor sufficient for linguistic meanings (as shown by grammatical, structure-dependent organization of information as disparate as DNA organization and computer storage principles). But if grammar came last, what came first? Basically two advances were required of the genus Homo to start it on the road to language: symbols and culture. Grammar would have greatly enhanced any communication system. But it alone cannot have been the source of language or thought (especially if we believe that

animals have thought, since they presumably lack grammar (or grammar is not unique to humans. But this would make it insufficient for language, since creatures could then possess grammar but not language). In what follows I want to review three sources of substantive evidence for this claim. First, there is the fossil record. Icons, indexes, and symbols appear before grammar. Second, the prerequisite of culture is shown partially in the necessity of intentionality and conventionality in this evolutionary semiotic sequencing. Finally, languages without structure-dependent grammars exist. The principal conclusions of this paper are summarized in Figure Two:

FIGURE TWO

The Semiotic Progression



Let's consider the components of Peircean semiotics presented in Figure Two that are crucial to the evolution of language, which is intended to represent a chronological progression. First there are indexes.[6] These are ancient, far predating

humans. Nearly every species uses them. Indexes are non-arbitrary, non-intentional linkages between form and meaning, vital to all creatures. If life forms could not interpret indexes, then lions might never find prey, hyenas would search in vain for carrion, and monkeys would be hard-pressed to avoid snakes and Accipitriformes. One can even cultivate an ability to detect and recognize indexes. In fact, it is advisable to do so. On my walks through the jungle with different indigenous peoples, my companions used indexes to know where we were, what flora and fauna were in our non-visible surroundings, where water was located, and what direction would be best to hunt in. They sniffed, listened, looked, felt, and tasted their way through the forest. I was largely oblivious to the indexes they were detecting, perceiving random smells, sights, and so on, without recognizing what they referenced. The deep knowledge of local index meanings is emic knowledge, as I argue in Everett (2016). Thus indexes are not trivial and are a vital rung in the ladder of human communicative evolution – as soon as they become enriched by culture.

Indexes are in a sense metonymical communication with nature. But though ability to recognize and interpret indexes can be emicized, they are inadequate for language. They lack arbitrariness and intentionality, two crucial components of symbolic language. Their shortcomings for language owe to their very physical link between form and meaning. Arbitrariness is essential to language because it dramatically increases the number of forms that may be used to link to meanings.

But arbitrariness is a later step in the semiotic progression. Before this comes intentionality (Everett (2012)). The signs that entail intentionality in

representations of meanings are the forms *designed* to look like, sound like, taste like, feel like, or represent by convention the things they represent (or acknowledged to have these properties, implicitly or explicitly). Thus icons like the Makapansgat pebble or the Erfoud manuport or early arts, such as the Venus of Berekhat Ram (see below), show us early steps from nonintentional indexes to intentional creation of or use of signs to represent. Even non-iconic art (so far as we know) – e.g. the geometric shell carvings below – show us the merger of intentionality and representation that is vital to all human languages (Everett 2012).[7]

As intentionality meets representation in icons, humans could in principle begin to communicate more. After all, modern day emojis (Evans 2016) are suggestive of an icon-based language. However, this analogy is imperfect because emojis are not purely iconic. They arguably depend the modern grammars from which they emerge for complexity of interpretation and their organization. Nevertheless, emojis do show some of the potential of purely iconic communication (a form of communication that crucially relies on culture, about which more below).

Getting back to our semiotic trail of language evolution, the next step is the symbol. By being both intentional *and* arbitrary, the symbol represents a much longer stride towards modern language than either the index or the icon.[8]

Because of this progression, which is gradual and spread across Homo species, it makes no sense to ask what, say, the first language was. *There was no first language*. Just as there was no first human being. No one woke up one day talking a modern, complex language. Nor, I believe, did anyone wake up one day with a

recursive grammar-building operation in their head, in search of words to supply to such an operation. At least such a proposal is unnecessary when the evidence of the fossil record is carefully evaluated. We can only speculate about the first sound icons. Did an australopithecine try to imitate a bird sound? A chimp call? A falling rock? All languages likely have some form of onomatopoeia, e.g. "whoosh," "crash," "snap, crackle, pop" and so on. But sound symbolism, onomatopoeia, and ideophones, all active in modern languages, show us that such symbols are of wide importance and also likely of early provenance.

According to the theory of Peirce, however, indexes, icons, and symbols are still insufficient for semiotics to get off the ground. We also need something that Peirce referred to as "the interpretant." This is in essence what makes it possible to translate the sign so as to understand its object. More precisely, interpretants are those formal aspects of a sign that make it work qua sign. Thus, for example, the sign "eye" has three components that are separately and together formally combined to refer to our visible organ of vision. These are the conventions on the shape and order of the letters, e, y, and e, that compose it (more generally, its duality of patterning and the units selected to enter the dual pattern). So reverse an /e/ to give its mirror image and the interpretant is lost. But write either a 1cm tall /e/ or a ten-foot tall /e/ and nothing of the interpretant is lost. Thus size is not part of the interpretant of "eye," though directional orientation of the letters is. Thus the symbol is itself analyzed into its meaning-carrying parts.

3. Neurological and physical Platforms for language

Of course language evolution is also about biology, not merely semiotics. It is our biology that in some way underlies our language abilities. Acknowledging this obvious fact, it is perhaps surprising and counterintuitive to some to discover that there is nothing in the body dedicated to language. Not a single organ. Nothing in the brain. And nothing in the mouth (except for the position of the tongue). But this should not be a shock. Evolution always prefers tinkering with or exploiting what already exists over creating the brand-new. What underlies our wonderful human voices is a jury-rigged collection of anatomical parts that we need for other things.

For example, the "vocal apparatus" is a disjointed set of body parts used for other things primarily. The larynx has the primary function (not only in humans but in many other animals) of protecting the trachea from falling food or other foreign objects.[9] Teeth and tongue are obviously multifunctional, the nasal passage as well.

To speak we use our lips, teeth, tongues, nasal passages, vocal cords, and other body bits. The sound ***b***, for example, is made by simultaneously blocking the flow of air from the lungs through the mouth by closing the lips and vibrating the vocal cords, parallel muscles that flank the opening to the trachea, the glottis. The sound ***p*** is identical to ***b*** except that with ***p*** the vocal cords (or folds) do not vibrate. The sound ***m*** is the same as ***b*** except that the air that originates in the lungs also passes through the nose (which is why when you have a cold and your nose is stopped up, ***m***s sound like ***b***s).

Again, every part of the vocal apparatus has a non-speech related function that is more basic from an evolutionary perspective and that is found in other species of primates. Language and speech came later and exploited our bodies and brains as

evolution had produced them, altering them over time. Therefore it is not surprising that mechanisms implicated in human language, like our tongues, teeth, and the rest are not only part of the endowment of modern human biology, but are found in many other animals as well. This is a simple consequence of the continuity of evolution by natural selection. The single unique aspect of the human vocal apparatus that does seem to have evolved specifically for human speech is its *shape*, caused by the position and shape of the tongue (Lieberman (2007; 2016)). This shows adaptation, not complete specialization.

Sign languages, as we see in below, also have much to teach us about our neural cognitive-cerebral platform. Native users of sign languages can communicate as quickly and effectively as speakers using the vocal apparatus. This means that our brain development cannot be so tightly connected to speech sounds that all other modalities or channels of speech are unavailable or less good for language. It seems unlikely that every human being comes equipped by evolution with separate neuronal networks, one for sign languages and another for spoken languages. It is more parisimonious to assume instead that our brains are equipped to process signals of different modalities and that our hands and mouths provide the easiest ones. Sign languages, by the way, also show evidence for syllable-like groupings of gestures so that we know that we are predisposed to such groupings, in the sense that our minds quickly latch on to syllabic groupings as ways of better processing parts of signs. Regardless of other modalities, though, the fact remains that vocal speech is the channel exclusively used by the vast majority of people. And this is interesting, because in this fact we do see evidence that evolution has altered human

physiology for speech.

Although the full inventory of changes to the human vocal apparatus is too large and technical to discuss here, the final result of our tongue placement and shape evolution is good for talking because we can make a larger array of speech sounds, especially the quantal vowels "i," "a," and "u," which are found in all languages of the world. These are the easiest vowels to perceive. We're the only species that can make them well.

On the other hand, although humans can produce a rich array of sounds, including the quantal vowels, they don't need to. By the use of only a small range of consonants, intermixed with three or more vowels, all human meaning can be communicated. Nevertheless there were other modifications in our bodies needed to get us to full human speech capacity.

Now, turning to the brain, we find things much as we did for the vocal apparatus, i.e. preexisting systems exploited for different purposes, though here perhaps there is even less evidence for specific adaptation. Although it is often claimed that there are language-specific areas of the brain such as Wernicke's area or Broca's area, several researchers have shown (see especially Lieberman (2014)) the importance of the subcortical region known as the basal ganglia to the cortical linguistic functions. Basal ganglia are a group of brain tissues that appear to function as a unit and are associated with a variety of general functions such as voluntary motor control, procedural learning (routines or habits), eye movements, and emotional function. The basal ganglia are strongly connected to the cortex and thalamus, along with other brain areas.

The general nature of the basal ganglia, their role in speech and their responsibility for habit formation, teaches us several things. First, these fundamental components of language function are not specifically designed for language. Harm to the basal ganglia can produce a number of aphasic conditions, or language deficits. This means that the responsibility for language lies with various regions of the brain that contribute in multiple ways at a higher level of organization in our mental or cortical life. Second, language is at least partially a series of acquired habits and routines, along with others like skiiing, bicycle-riding, typing, and so on.

In the book, *Rethinking Innateness*, by Jeffrey Elman, Elizabeth Bates, and their colleagues, the authors underscore why we should not jump to conclusions about the significance of the fact that some kinds of knowledge are found in specific regions of the brain:

> "...everything humans know and do is served by and represented in the human brain .... Our best friend's phone number and our spouse's shoe size must be stored in the brain, and presumably they are are stored in nonidentical ways, which could... show up someday on someone's future brain-imaging machine... The existence of a correlation between psychological and neural facts says nothing in and of itself about innateness, domain-specificity, or any other contentious division of the epistemological landscape..." (p241)

They add that ."..*Well-defined regions of the brain may become specialized for a particular function as a result of experience. In other words, learning itself may serve to set up neural systems that are localized and domain-specific, but not innate."* (p242) Therefore, regardless of anyone's findings, or claims about brain specialization for specific tasks, we cannot draw any conclusions about a priori knowledge or universal grammar, say, without first ruling out learning as a source of cortical localization. This is possible, but nothing like this has been done. In fact there is no uncontroversial evidence that brains either do or do not have specialized, hard-wired networks or modules independent of learning. But the evidence for specialization is not strong. In what follows, I discuss some claims for localized brain specialization for language. It is part of urban and academic folklore that language-specialized regions of the brain exist. One of the best known of these regions is "Broca's area."

Broca's area is normally identified as a region of the brain located at the pars opercularis and pars triangularis of the inferior frontal gyrus, that is, on the left side of the brain.

This region of the brain was first identified in the 19th century and became associated with the French researcher and physician, Pierre Paul Broca, working with a patient nicknamed "Tan," because this was the only word that the patient could utter.

As a result of Broca's studies, many people came to think of Tan's affected brain region as specialized for language production. But although everyone agrees that there are regions of the brain involved in language, there is also wide agreement

that Broca's area is not language-specific. For example, recent MRI findings indicate that production of speech is not limited to Broca's area. Also (at least single) lesions to Broca's area alone are not enough to permanently disrupt speech production. A third, quite ironic, problem for claims that Broca's Area is a language area is that the original lesion studied by Broca in Tan's brain seems not to even occur in the same area we currently understand to be "Broca's area." Fourth, Broca's area can be destroyed leaving language intact. Finally, Broca's area is involved in other kinds of brain-activity, such as coordination of motor-related tasks. For example, if you show someone hand shadows of moving animals, Broca's area is activated. The same if someone listens to or performs music. These are not language-specific tasks. Moreover, there seem to be no pure cases of "Broca's aphasia" - aphasics always have other deficits in addition to language. What such results show is that Broca's area could have a more general function, such as the coordination of activities (language production being one such), rather than a dedicated language area. I do not say that we fully understand Broca's area or that we know with certainty that there are no language-specific regions of the brain. I am saying only that we haven't yet discovered such areas (see Friederici (2014) for one review that, while it offers an enlightening discussion of brain areas implicated in language, doesn't show any that have evolved for or are specialized for language). See also Everett (2017) for detailed criticism of Friederici's proposals .

Moreover, new evidence being gathered suggests the opposite - that our brains can use parts of the brain for multiple, diverse functions. Recent findings in research led by Bedny, et. al. 2015 argue that "visual cortex," the region of the brain usually

associated with vision in sighted individuals can be used for non-visual tasks, in particular language, in blind subjects, as well as Braille-reading. This work is extremely important for any attempt to link cognitive functions with specific regions of the brain. And hers is by no means the only research that shows significant brain plasticity. Whatever is responsible for the localization or specialization of different regions of the brain for different cognition functions, it doesn't seem to be the result of specific, genetically-determined links between functions and cerebral topography. The brain seems no more specialized for language than other individual components of the anatomy are responsible for speech. The lungs, larynx, teeth, tongue, nose, and so on are all vital for non-signed language as the hands are crucial for signed languages, but none of these are either individually or collectively language organs.

The same issues arise for any claim of anatomical specialization for language, such as Wernicke's area, located in the posterior section of the superior temporal gyrus in the dominant cerebral hemisphere of a given individual (for right-handed people this is the left hemisphere. For lefties, it is the right hemisphere). At one time, this region of the posterior temporal lobe of the brain was believed to be specialized for understanding written and spoken language.[10]

Unfortunately for anyone interested in marshalling anatomical support for the innateness of language, Wernicke's Area, like Broca's area, isn't exclusively specialized for language. First, just as with Broca's area, there are no agreed-upon definitions of the location and extent of Wernicke's area. There is a general range that people agree on, but no consensus on the exact coordinates of that position in

the brain. That makes it difficult to even say that there *is* such an area with any precision. Second, recent research shows that this region is connected to other regions of the brain that are, as was the case with Broca's area, far more general in their function than language, e.g. motor control (and pre-motor organization of potential activities). Third, as Bedny et. al.'s (2015) research indicates, even if we did find an area specialized for a particular function in one or a million subjects, the next subject we meet could in many cases be using that area of their brain for something else, depending on their individual developmental history.

Then what are we to make of what we currently know about the association between language functions and Broca's and Wernicke's areas? The lesson is that parts of the brain develop in each individual as the homebase for multiple, though related, tasks, such as language, motor control, sequential ordering, and the like. What we see as brain specialization could result from the order in which we experience things in our personal development.[11]

We know now that brain specialization and anatomy can be influenced by culture. This makes it much harder to tease apart evidence for language-specific biology from biological properties resulting from learning or the environment. For example, Keller and Just (2009) argue that reading-challenged children who had experienced as little as six months of intensive remedial reading instruction grew new white-matter connections in their brains. Just's study is one of many that show culture changing the structure and functionality of our brains.[12] Other studies have shown that the connections between portions of the brain can weaken or strengthen over time, based on the cultural experiences of the individual.[13]

Now, since culture can change the form of the brain and since we know of no single cognitive function that must appear in a single location in all brains, then we see once again the difficulty of using any argument from cerebral organization or anatomy for linguistic nativism. We must not forget either that from one vantage point, localization in the brain is often the result of the trivial fact that everything that exists has to exist somewhere. If we know something it will be found some place in our brains. This is no more evidence for innate knowledge than the fact that Bach was born in Eisenach is evidence for predestination. Everyone has to be born somewhere.

Human anatomy and physiology provide no direct support for the proposal that language is innate, aside from the human vocal tract. So let us turn now to archaeology.

4. The archaeological evidence

Having now sketched out the basic types of signs (to which we return later), we are prepared to discuss their archaeological history in more detail.

4.1. Indexes

As we have seen, indexes are ancient, predating the genus Homo. The ability of creatures to connect physical attributes of objects with those objects is likely as old as life itself.

4.2. Icons

4.2.1. The Makapansgat manuport

For more than three million years visual icons have been collected by our hominin ancestors, from Australopithecus to. These icons suggest that the icon-possessor(s) likely grasped a connection between form and meaning – what the icon is a visual representation of. In this light, consider the 2x3 inch stone found in the Makapansgat cave in South Africa.

Makapansgat Manuport/Pebble/Cobble

In 1925, this manuport was discovered in a cave in the Makapan Valley of South Africa, described 49 years later in 1974. This pebble collected by Australopithecus africanus stands out among the Acheulean tools it was found

among, because it clearly is not a tool, but was brought to the cave from elsewhere, almost certainly because it resembled a human face. This manuport indicates that as early as three million years BP early hominins recognized iconic properties in objects around them. Just as I perceive the serpentine iconic properties of tree roots in the Amazon, so the Australopithecines of Makapansgat saw iconicity in a rock with two circular indentations above a groove running transverse to them.

Alternatively, it is possible that this pebble was carried to the cave for reasons unrelated to its resemblance to a human face. It could have become stuck between someone's toes. It wasn't a tool, because the Olduwan tool kit of Australopithecines had no such item. And, again, it didn't match the stone of the other tools.

4.2.2. The Erfoud manuport

Move forward to 300,000 years BP and we find another manuport, this one in modern-day Morocco, picked up by Homo erectus. A cuttle fish bone shaped like a phallus:

Erfoud Manuport

Again, the icon was intentionally appropriated, recognized, and collected, though not intentionally created. What is the significance of such iconic objects for the development of language? Well, the answer depends on who you ask. Berwick and Chomsky (2016)'s view of language (as the output of the Merge operation), rules out any relationship at all between such occasional icons and the development of language. If language is a Merge-constructed set of structure dependencies, then clearly a phallic cuttlefish bone fails to move humans any closer to a computational system. On the other hand, if language is about meaning and symbols are the building blocks of meaning, then icons are vital to the reconstruction of the evolution of language.

4.2.3. Art, tools, and the invention of symbols

Art, tools, and symbols each contribute to our understanding of the other and to the "dark matter" of culture and psychology (Everett 2016) that allows each to emerge. Art, for example is a visual form with shared meaning, the communication of emotions, of cultural moments, of ideas, and so own via shared dark matter. We must learn to see in several different ways to appreciate art. If paintings, we must learn to recognize two-dimensional images of three-dimensional objects (see Everett 2016, chapter four, for further discussion). If sculptures we must learn to see in objects that are iconic or objects that are *not quite iconic* the real-world or imaginary object that the artist (or the collector of an iconic object) intended.

Tools, especially when they are generalized and found in different places, such as the "kits" discussed below, indicate the existence of shared objectives, problems, and solutions. For example, the 300,000-400,000 year old Schöningen spears are evidence of culture among Homo heidelbergensis (likely a variant – if even that – of Homo erectus) and show that these humans hunted, that they used brute force rather than throwing, and that they hunted as part of their culture. Thus the spears represent cultural objectives, cultural knowledge, cultural techniques.

Symbols are then much like art and tools in having shared objectives and intentionality. Unlike tools, their forms are arbitrarily (in principle) linked to their meanings. But like tools they have a cultural meaning that is not linked directly to an object, but represents an activity, *displaced* from the form and meaning of the spear. A spear means hunting, but the hunting is not necessarily present. Tools also show

an aspect of Peircean signs – the interpretant – in that only certain parts of the tools are meaningfully connected to their tasks, e.g. the edge of the tool, or the handle + the edge, and so on. An axe, for example, can be hollow, or of various different materials. What matters is the quality of the edge for cutting. Thus tools, while unlike symbols in many ways, when examined carefully show similar cognitive thresholds of association.

A Schöningen spear

4.2.3.1. Geometrical sketchings and the Homo mind[14]

Erectus Shell Etchings From Java

The geometric designs above could represent merely a pleasant decoration of a shell, shaped in part by perceptual constraints of the brain that favor geometrical design (Bednarik (1995)). Or these marks could be symbols. Or they could represent something intermediate between icons and symbols, precursors to representation of meaning. I suspect that the former is correct, rather than the latter. Nevertheless, the designer, a Homo erectus man or woman, picked up a shark's tooth and pressed very hard and deliberately to record these shapes. Notice that the lines are solid and continuous, without breaks. To make such marks, our ancestor would have had to have pressed hard enough to cut through both the (now decomposed and missing) brown outerlayer of the shell into the hard white shell proper. He or she would have had to have carved without stopping or the lines would have some visible breaks in them.

Whatever these designs indicate, they are at least a manifestation of intentional activity, perhaps abstract icons, perhaps symbolic (say of the sea voyage from southern, mainland Asia to Java where they were found). And, fascinatingly, they were made some 540,000 years ago.

4.2.3.2. Tool Kits

Morgan, et. al. (2015) assert a strong link between tool development and the emergence of language:

> *"Our results support the hypothesis that hominin reliance on stone tool-making generated selection for teaching and language and imply that (i) low-fidelity social transmission, such as imitation/emulation, may have contributed to the ~700,000 year stasis of the Oldowan technocomplex, and (ii) teaching or proto-language may have been pre-requisites for the appearance of Acheulean technology. This work supports a gradual evolution of language, with simple symbolic communication preceding behavioural modernity by hundreds of thousands of years."*

As the references to Morgan et. al. (2015) illustrate, this is a growing area of research, linking tool-making to language evolution, via brain development. Thus the presence of tools in a society are, on my reading, not as strong as evidence for language as symbols, but closely related.

But in discussing tools relative to language, we are also after the qualities that both share of culture, shared intentions, and the ability to *match form and*

function, the conceptual basis of signs. The Oldowan tool kit below shows stones crudely shaped to work as weapons and tools, such as a hammer or a hand-axe. They would have not been precision instruments compared to later tools, but they represent a tremendous step forward in human technology, serving as precursors to culture.

Oldowan tools are the earliest known. They were used from roughly 2.6 million years ago, perhaps by Australopithecines but certainly by Homo erectus (or Homo habilis, if one accepts that as a separate non-erectus species). Oldowan tools included choppers, scrapers, and pounders, shown in the image below:

Oldowan Tools

To produce an Oldowan tool, a "core rock" is struck on its edge by a round "hammerstone." The striking produces a sharp, thin flake, leaving conchoidal fractures on the core rock, as seen in the image. The flakes are often reworked for

other purposes. There are implications in tool-building for the cognitive development of the species.

In the first place, making tools requires planning, imagination (having an image of what the final tools should look like) and communication of some sort for instructing (even if only by example) others, e.g. people from a younger generation, in how to make tools. The sequential operations call upon the prefrontal cortex and produce *cultural* selectional pressure for more cortical horsepower. However, this pressure might have worked, the larger prefrontal cortex of earlier Homo, relative to Australopithecene tool makers, responded to this need. Not surprisingly, therefore, at about 1.76 million years ago, about 300,000 after the rise of Homo erectus, Oldowan tools were joined by other erectus-manufactured tools, a new type called the Acheulean.

But why would it take 300,000 years for erectus to develop the technological advance of Acheulean tools? Was this related to an absence of language, as Morgan, et. al. (2015) suggest? Possibly. Some certainly claim that the long period without innovation is evidence for a lack of language among the tool users. But this does not necessarily follow.

We know, for example, that human cultures, even in the 21st century, are resistant to change. We know that imitation is favored strongly above innovation when what is being imitated still works fine (Boyd and Richerson (1988; 2005)). So the 300,000 year lag might have been a result of a lack of cognitive or linguistic development, but just as likely it is explained by the nearly universal principle of

"satisficing" (Simon (1962)), i.e. nature is satisfied with what is with what is "good enough," not striving for the best, as well as the valuing of imitation over innovation.

Acheulean tools:

Though Oldowan and Acheulean tools overlapped in their use by earlier hominins, Acheulean tools were more advanced. They were carried from Africa to Europe by Homo erectus, with Spain being their earliest destination, about 900,000 years ago. Acheulean tools were not created exclusively by striking stone upon stone as were Oldowan tools. They also involved shaping after flaking with bone, antlers, wood and other tools, which provided more control for the toolmaker. Also, Acheulean tool-makers preferred to use the cores over the flakes as the primary tools. So they were an advance over as well as a complement to Oldowan tools.

From the Acheulean industry improvements resulted in the closely-related Levallois technique (ca. 500,000 years ago). In the spread of all of these tools,

however, we see communication, if not in explicit instruction or linguistically, then in the revelation of the tools themselves to other hominins, as they spread and their utility became known.

The Levallois technique required fine work along the edges of a core, followed by a final blow that lifted the flake, presharpened by the earlier striking. These tools were often made of flint, a more workable material, and thus had finer edges, as seen in the photo below:

Levalloisian Tools

The complexity and uniformity of Levalloisian tools lead some to argue that language is implicated in their manufacture, in order to account for error-correction which is assumed to have been necessary. But speaking is not absolutely required since, as I describe in Everett (2016), learning is often a matter of observation,

followed by trial and error under a watchful eye, with very little verbal communication required, even in modern societies. However, some form of communication does seem to be necessary for feedback, even in language-minimal training. Moreover, though, there is no doubt that making tools together and the correction of flawed techniques by learners would have favored language development for instruction. And this was occurring with the first hominins, Homo erectus, to actually produce intentionally iconic and geometrical art. The idea that erectus was capable of some sort of sophisticated communication, e.g. a G_1 language (figure two) is supported not only by their art and tools but by the fact that erectus was the first hominin species to leave Africa and sail the ocean, from mainland Asia to Java. It stretches credulity to believe that they could do this without symbolic communication.

4.2.3.3. Representational Art

Bednarick (2008) suggests that icons like the Makapansgat pebble and the Erfoud manuport may have led to new neural pathways for recognizing that one thing may stand for another. I am not sure exactly what he intends by this, though I do agree that new ways of thinking can lead to new evolutionary pressures on the brain, enhancing the ability to comprehend representations more complex than mere indexes. And Morgan, et. al. (2015) do offer a more concrete proposal along these lines.

An interesting further bit of evidence in the evolution of the symbolic comes from early art, such as the 250,000 year old Venus of Berekhat Ram, shown below:

Venus of Berekhat Ram

Some deny that this is art, claiming that is nothing more than rock bearing a human resemblance, the same as the Makapansgat manuport. However, it does show some evidence upon closer examination, of being slightly shaped to take on more of a "venus likeness." And there is some suggestion of red ochre added to the

stone as a form of decoration. It is not a straightforward *objet d'art* in this sense. But still the evidence suggests strongly that this is the oldest extant work of art in the world.

4.2.3.4. Mistakes, stylizations, and the invention of symbols

From my very first encounter with the rain forest, I was on the lookout for snakes. Every tumescent root that "slithered" across my path, partially covered in leaves, appeared to me first as a writhing, threatening serpent and only secondarily as inert flora. I perceived parts of the root and they represented in my mind parts of a snake. My mistake, provided by my emotional state, led me to, in a sense, see the roots of the jungle as snakes, only later to reinterpet them as colubrinus icons, eventually as snake symbols.

Symbols arise naturally within minds and cultures able to learn, retain, and integrate knowledge into a personal sense of personal and group identity. One example, just given, is how the mind makes use of errors, perhaps moving moving from misperceptions to icons to symbols, one image "standing for" another.

But they also arise from adaptation of the natural to the conventional in culture. One treatment of this route towards symbolization is proposed in Urban (2010). In his work on ritualized lamentation in Ge languages Urban argues that a particular metasignal – signal relationship (ritualized lamentation derived from natural crying) illustrates a common journey from "strategic vocal manipulation, to cultural metasignals, to language." Further, he claims that "strategic vocal deceptions in nonhuman primates are possible precursors of true socially

constructed, socially shared metasignals, which in turn may be ancestors of modern human language." Though I do not think that these metasignals actually get us to language, I do believe that Urban has put his finger on a fecund source of symbol evolution.

Another area in which symbols arise is in tracking social relationships. Most primates, among many other creatures, have elaborate social organizing principles, e.g. polyandry, polygyny, dominance relations, cross-cousins, parallel cousins, and so on. These concepts are learned via interactions, based initially on physical opposition, e.g. male vs. female, strong vs. weak, mateable vs. nonmateable, mother vs. child, and so on. As we use concepts we come to understand them (Brandom (1998)). So one can say, I believe accurately, that even without language, many animals use something like concepts as they negotiate their ways through social relationships (see for example Safina (2015)). Keeping track of such relationships would have increased the cultural/cognitive selectional pressures for symbols (see also Barnard (2012, 89ff) and Dunbar (1998)).

To be certain, numerous researchers have written on the evolution of the symbolic (Luuk 2012; Cangelosi 2001; Sereno et. al. 2014, inter alia). However, as illuminating as these discussions are, and all have been helpful to my own thinking on these matters, they share a common lacuna, i.e. the connection of symbol-evolution and grammar to a well-developed theory of culture. Thus, for example, Luuk (2012, 88) claims that "... status symbols (e.g. expensive clothes) have not much in common with linguistic symbols (e.g. words)... one cannot infer symbolism (and by extension, language) from personal ornaments, as the most parsimonious

interpretation of personal ornaments is that they are status symbols (Sterelny 2008)." Luuk rejects the linguistic significance of status symbols because he claims they fail to show Hockett's (1960) feature of "displacement," reference to something absent from the situation of utterance. As he says "... personal ornaments do not show displaced reference, as they bestow status only to their wearers." But this is false, and displays a lack of appreciation for culture. Status is not inherent in ornaments nor in individuals, nor do ornaments bestow status. If I find the crown of the King of England, putting it on my head not only does not make me a king, it opens me up to the status-lowering charge of being an imposter. Status in fact derives from the network of culture, as discussed in Everett (2016) and below, i.e. from abstract value-rankings, hierarchical knowledge structures, and social roles. Status symbols are social indexicals (Silverstein (2003; 2004)), dependent for their meaning on abstract, and thus displaced, culture, as I define it in my work. Thus although Luuk is correct that status symbols are not per se linguistic symbols, both classes of symbols are arbitrary, *socially* indexical, and displaced. Therefore, they are conceptual kin. To have one in fact is not entirely unrelated to having the other. In fact, they would, ceteris paribus, be expected to occur together in the same society at the relevant level of conceptual complexity or simplicity.

Displacement, the element which Luuk finds missing in status symbols, is in fact, as Everett (2005) argues, itself subject to cultural constraints. So the crucial components from my perspective are the more general ones of arbitrariness and intentionality, which are crucial not only to symbols but to language proper.

It is perhaps worth adding that displacement and imagination, the latter not found among Hockett's design features, are both crucial to language and both depend on both increased intelligence and cultural values and standards. Imagination takes place within a cultural context and is spurred by that context. Some cultures value innovation and imagination more than others, for example (such as Silicon Valley culture vs. Catholic Church culture, idealizing somewhat). But, once again, for the evolution of symbols and communication, culture is vital – without imagination and displacement, there is no language.

So what are some of the possible paths to the development of symbols? To refer back to my example of roots across the jungle floor from earlier, when I see logs at times walking through the jungle, I jump back if I haven't had a good look, worried that one might be a snake. This mistaken association of one thing for another leads naturally to intentional use of that mistaken impression to represent. As early hieroglyphic writing systems from different parts of the world show, this use of representations based on resemblance, though not quite iconic, can serve as the basis for reference to one thing by another not only with intentionality obviously present, but also as the form is modified to be more general, with arbitrariness, creeping in as it moves from the status of icon to symbol. And just as this has happened in writing systems, so it is likely that it has happened in spoken systems.

In Dunbar's (1998) research and later work by Barnard (2012), it is argued that the development of kinship relationships would have created concepts in need of forms, i.e. this would have exerted pressure on the communication system to go

beyond icons to symbols. Concepts go looking for forms to serve as cultural exchange. On the other hand, if many researchers are correct and animals do indeed have concepts (Safina (2015), Panksepp and Biven (2012)) then this begs the question, why don't animals develop symbols? One easy answer is that animals lack the language gene. But this is not terribly insightful, pushing the explanation back one level to the evolution of the gene rather than the evolution of symbols.

Moreover, while we know little about the evolution of a mysterious language gene (FOXP2 is certainly not that gene (Diller and Cann (2012, 170ff)), we do know a great deal more about the evolution of human intelligence and we know that humans, as they define intelligence, are more intelligent than non symbol-using creatures.[15] Thus a richer array of concepts requiring symbols and a richer, more inventive intelligence would have both been under pressure to find a joint solution to concept communication. Symbols would have been a rather easy step, especially as culture developed and status symbols, burial symbols, and the like evolved, as per above.

5. Infinite Semiosis

In his work on indexicality and metaindexicality, Silverstein (2003; 2004) analyzes the recursive properties of human thinking as applied to the use of language in representing not only social-cultural meaning (values, local knowledge structures, social roles, etc) but the recursivity of pragmatics. This theme is further explored, and more explicitly linked to recursion, in Levinson (2013).

Peirce (1977; 1992) anticipated both Levinson and Silverstein, however, in proposing that symbols are recursively constructed of other symbols (some overtly, some covertly – the latter as part of what I refer to as *dark matter of the mind* (Everett 2016). In Peirce's writings, infinite semiosis is the idea that signs are multifunctional, each sign also function as an *interpretant* at the same time that they are composed of interpretants. For example, a sign has an interpretant, but an interpretant is also a sign, so the interpretant also is a sign. This is a kind of conceptual, though not completely formal, recursion. And conceptual recursion is a huge step forward in human communication. This means that a string of signs is never finite. For example, according to Peirce we see infinite regress even in a sequence like:

$$\text{SIGN}_1/\text{INTERPRETANT}_1 \rightarrow \text{SIGN}_2/\text{INTERPRETANT}_2 \ldots \rightarrow \text{SIGN}_N$$

This representation looks finite until we realize that Sign_n cannot be the end because if it lacks an interpretant it is not a sign. Likewise, Sign_1 cannot really be the beginning, because it is by definition an interpretant of an earlier sign. So there is no beginning nor end to semiosis. It is infinite.

In Peirce's reasoning, symbols alone produce infinite strings of form and meaning, so that recursion, such as it is here, i.e. largely conceptual, is a property of signs necessarily but, as we see directly, not necessarily of grammar (unless grammar is looked at as a form of sign, which formal linguists deny, though researchers in Construction Grammars happily accept).

6. Duality of Patterning

There has been a revival of interest of late into the idea of duality of patterning. See in particular Ladd (2012 and references cited therein). Many of the new works focus on either Martinet's concept of double articulation, Hockett's idea of duality of patterning, or some conflation of the two. It is common to find confusion in discussions of these ideas, because they are easy to confuse with other concepts such as compositionality and what Hockett called "productivity." I want to make two novel points in this section and then sketch how duality of patterning might have evolved. The first point is that the idea of duality of patterning in fact predates both Martinet and Hockett, going farther back than even Saussure's original distinction of form vs. meaning, the two sides of the sign. We find the intellectual roots of duality of patterning, in my opinion, in the Peircean notions of the sign, the interpretant, and infinite semeiosis.

The initial idea of duality of patterning is to highlight the supposition (not shared by all) that language builds meaningful units out of meaningless units. The meaningless units of a language are, depending on one's theory, its phonemes, syllables, phonological phrases, and other sound-based units. These combine to form meaningful units like morphemes, words, phrases, sentences, stories, and conversations. Martinet's concept of double articulation is a bit different and resembles more the "three hierarchies" model of Kenneth L. Pike below.

Is duality of patterning then just another name for combinatoriality or compositionality? No. After all, compositionality is the ability to have open-class elements and one can in fact observe duality of patterning without open-endedness. As Ladd (2012) shows, one can have a series of colored lights in which the colors are

meaningless while their combinations are meaningful. That is duality of patterning. But if there are only three colors, say, then the number of messages is not open-ended and so although the messages result from duality of patterning, they lack an ability to produce a non-bounded set of messages. Moreover, bees have, according to Hockett (1960) dances that seem to be open-ended in the analog representation of distance to honey sources, but still lacking duality of patterning. And compositionality is largely a function of building meaning from smaller parts where an open-ended semantics meets an open-ended syntax. And duality of patterning in its basic form is therefore not compositionality. Yet I believe that these are ultimately distinctions without a difference. If a human mind can come up with duality of patterning, it can extend this to general open-ended compositionality and grammar straightforwardly, especially if we return to the basics of Peirce's semiotics.

Icons, indexes, and symbols are constructed on a simple semiotic property, namely, that one thing can indicate or represent or mean another. Thus each is a composite of a form and a representation or meaning, and each also entails Peirce's *interpretant*. Signs in all their forms are therefore steps towards duality of patterning. And they take us closer to another crucial conceptual step up the ladder from communication to language, the etic vs. the emic (the perspective of the outsider vs. the perspective of the insider). Signs alone do not get us all the way to either. But by associating meaningful representations with otherwise meaningless forms (as with symbols), the distinction between form and meaning is highlighted and the possibility of manipulating the forms separately, analyzing them then

synthesizing resultant parts into other forms we move towards duality.[16] And because symbols are interpreted by members of a particular group, they lead to the insider vs. outsider contribution to language evolution, the etic. vs. the emic (see 8 below).

The semiotic progression we have examined, Indexes → Icons → Symbols, overlain with gestures, gets us part way towards language as widely understood. But other steps are necessary. Following symbols, the most important invention for language, we need a way to make more complex utterances out of symbols. We need organizing principles, or *grammar*.

There are two major organizing principles of language – the syntagmatic and the paradigmatic. These underlie all grammars and discourses (Saussure 1983, 122, 123). Both of these facilitate communication by allowing more information to be packed into individual semantic or semiotic constituents of language than would be possible without their basis in duality and triality of patterning. Both syntagmatic and paradigmatic organization follow from the nature of symbols and the transmission of information.

If one has symbols then there is no huge breakthrough entailed in placing the symbols of an utterance in some linear order. Once this order becomes conventionalized, through culturally sanctioned, frequent practice, linearity of symbols, i.e. their syntagmatic organization, is largely completed. And as the syntagmatic organization so imposed uses multiple symbols, it will not be a huge leap to use symbols for events and symbols for things initially, to fill analogous

positions in different syntagmemes. This would stimulate the paradigmatic organization of language in TABLE ONE:

TABLE ONE- DUALITY OF PATTERNING

Paradigmatic	**Syntagmatic**		
$Symbol_{filler\ 1}$	$Symbol_{slot\ 1}$	$Symbol_{slot\ 2}$	etc.
$Symbol_{filler\ 2}$			
$Symbol_{filler\ 3}$			
etc			

Instructions: *Select one paradigmatic filler and place it in appropriate syntagmatic slot.*

From the idea that we have an inventory of symbols to be placed in a specific order, not a huge jump cognitively, we derive the notions of *slot* and *filler* which are the basis of all grammar (and much more important than operations over them, e.g. recursion or "Merge").

According to figure two above, once we have symbols and duality of patterning, we are moving closer to a full human language and have a linear grammar (if such can be called a grammar). Nothing else is necessary for grammar and only the "highlighters" are necessary to have the full form of language, able to mark and manipulate information structure. And as we organize our symbols we might next begin to break the symbols themselves down further into duality-patterned units. Thus a word such as "cat," a symbol, can have the syntagmatic organization of a syllable, c-a-t, and the (idealized) paradigmatic organization of phoneme fillers for slots, fillers drawn from the inventory of sounds the language

chooses and constructed according to paratactic rules. The combination of meaningless/phonological units to form grammatical/meaningful units can thus be bidirectional, going from phonemes --> syllables --> words and down from words --> syllables --> words.

The syllable is vital to the development of duality. It is a natural organizing constraint set on the arrangement of phonemes (individual speech sounds) that works to enable each phoneme to be better perceived. It has other functions, but the crucial point here is that it is primarily an aid to perception, arising from the matching of our ears to our vocal apparatus over the course of our evolution, rather than a prespecified mental category.[17] A very simple, even though technically inaccurate account of syllabic organization is that speech sounds are arranged in order, though most languages diverge from this to greater or lesser degree based on their individual histories, by a property called *sonority*, roughly corresponding to amplitude or loudness.[18] Thus syllables can be viewed for our purposes here as syntagmatic units in which their individual paradigmatic slots (onset, nucleus, coda, in that order) produce a crescendo-decrescendo effect, where the nucleus is the most sonorous element and the onset and codas are the least sonorous elements. For example, the (word and) syllable *bad* is well-formed in English because *b* and *d* are less sonorous than *a*, and thus are found in the onset and coda positions respectively while *a* is in the nuclear position. The syllable *bda* would be ill-formed because a minimally sonorant element, *b*, is followed by another element of the same sonority, *d*, rather than immediately by an increase in sonority. (I ignore the

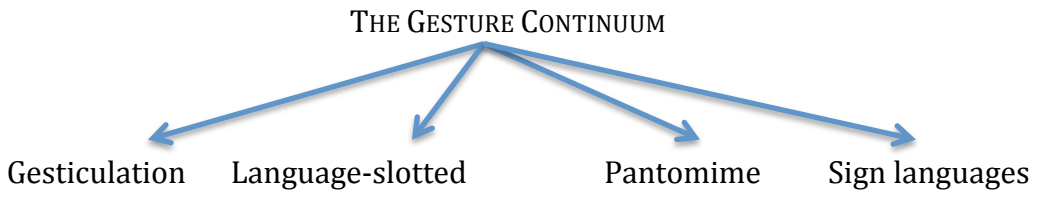
tremendous amount of cross-linguistic variation in syllable structure for purposes of exposition).

This perceptual and articulatory organization of the syllable brings duality of patterning, and its extension to grammar and compositionality, to language almost for free. It is possible that syllables came later in language evolution than, say, words and sentences, but any type of sound organization, whether in phonemes of Homo sapiens or other sounds by erectus or neanderthalensis, would entail a form of duality of patterning. Thus speech would have been a stimulus to syntagmatic and paradigmatic organization in syntax, morphology, and elsewhere in language.

7. Gestures and Triality of patterning[19]

7.1. Introduction

Language is even richer than duality of patterning, however. The relationship of gesture to language is crucial for an understanding of language evolution. Since many non-specialists, including many linguists, are unaware of the different kinds of gestures in language, I repeat below what McNeill (2005; 2000) refers to as the "gesture continuum:"



7.2. Intentionality, pointing, and the origin of gestures as directedness[20]

Gesticulation is the core of the theory of gestures. It involves gestures which intersect grammatical structures at the "Growth Point" (see below) and which are practiced by all humans so far as we know, even the blind and others with different types of cognitive disorders, such as proprioceptive deficiencies.[21]

Language-slotted gestures are those which can actually replace a word. For example, imagine that you tell someone "He (use of foot in a kicking motion) the ball" where the gesture replaces the verb "kicked" or "She (use of open hand across your face) me (for "She slapped me")." These gestures occupy grammatical slots (see also in an utterance and replace the grammatical units, usually words, that we otherwise expect there. They are improvised and used for particular effects in particular circumstances. They reveal speakers' understanding of the positions, words, and structures of their syntax. Since they interact with already developed language, these are not crucial to the evolution of language.[22]

Pantomime: Pantomime gestures simulate an object or an action without speech. Like gesticulations and language-slotted gestures, they are also not conventionalized, meaning that their forms may very widely. But pantomimes are not a possible source for language because the supplant language, they do not supplement it.

Emblems: These are conventionalized gesture-symbols that function as isolated "signs," such as the forefinger and the thumb rounding and touching at their tips to form the "OK" sign or "the bird," the upraised, solitary middle finger. They are noniconic symbols.

Sign-Languages are gesture-based languages that use gesture in a static, rather than dynamic, way. That is, there are gestural morphemes, sentences, and other grammatical units that are tightly conventionalized. Sign languages replace spoken language. *Sign languages do not enhance or interact with spoken language*. This is why McNeill argues that spoken languages did not, and could not, have begun as sign languages. On the other hand, sign languages themselves make use of gesticulation, in addition to the conventional signs that form the lexical items of the grammar.

7.3. The role of gestures in the evolution of grammar

As we move to consider gesture theory in more detail, the bedrock concept is the "Growth Point." This is the core unit of language in the theory of McNeill and it (or something like it) must be understood if we are to better understand how gesture and nongestural language combine. According to McNeill (2012, 24), the growth point is that point where "... speech and gesture are (a) synchronized, (b) co-expressive, (c) jointly form a 'psychological predicate,' and (d) present the same idea in opposite semiotic modes." This description of the growth point contains several technical terms. By "co-expressive" McNeill means that (p12) for the symbols used simultaneously in gesture and speech "... each symbol, in its own way, expresses the same idea..." By "psychological predicate," terminology deriving from the work of Lev Vygotsky (1978), McNeill (2012, 33) intends the moment of the expression when "... newsworthy content is *differentiated from a context*." By "semiotic opposites" McNeill means that gestures are dynamic, created on the spur

of the moment and, though influenced by culture and convention, are not themselves lexical or conventionalized units. Speech on the other hand contains grammar which is highly conventionalized (grammatical rules, lexical items, etc) and is thus "static communication."

In short, gesture studies entail a conception of language as a dynamic process of communication rather than a static knowledge of grammar. Gestures are manufactured by speakers in real time, following a culturally articulated unconscious (Everett 2016). They are dynamic. They are not merely applications of uni-modal rules, but is a multimodal holistic event. Gestures are actions and processes par excellence. McNeill (2005, 31) defines a gesture unit as "... the interval between successive rests of the limbs." Like most units of human activity (Pike 1967, 82ff; 315ff), it is useful to break down gestures into components. Thus, McNeill (and others) argues that gestures should be analyzed in terms of *prestroke*, *stroke*, and *poststrokes*. And just as onsets, codas, and nuclei in syllables may be long or short, gestures may be lengthened in their different constituents, what McNeill calls "holds." In the prestroke, which like the other constituents of a gesture, may be "held" to better syncronize timing with the spoken speech, the hands move from their rest position in anticipation of the gesture. The stroke is the meaningful core movement of the hands. The poststroke is the beginning of the retraction of the gesture. The work on gesture is full of rich illustrations of these gestural constituents and how they at once synchronize with and add dynamically to speech.

These constituents and holds are strong indications that utterances of human languages are tacitly designed in real time, again supporting the dynamic cognitive

and communicational nature of human language. As Kendon (2004, 5) says, "Gestures are part of the 'design' of an utterance." One of the clearest ways in which gestures show design are in the constituents of gestures (prestroke, stroke, and poststroke) and how they are often held to synchronize precisely with spoken speech. The question of how gestures are learned and controlled is every bit as interesting as those few aspects of language that linguistics as a discipline currently addresses.

Gesture also brings culture and grammar together. For example, David Efron's (1941) pioneering work under Boas represents the first modern study to examine the link between culture and gesture. But it is not the only one. There is a now a sizeable literature on such effects. To take one example, de Ruiter and Wilkins (1998) and Wilkins (1999) discuss the case of Arrernte in which the connection or "binding" of speech and gesture is overridden by culture and dark matter. According to de Ruiter and Wilkins, the Arrernte regularly perform gestures *after* the co-expressive speech. The cultural reason the authors suggest is that the Arrernte make much larger gestures physically than are found in many other cultures, using movements of the entire arm in gesturing. Thus, as they interpret the phenomena, the larger gestures and space required by the Arrernte demand more planning time, favoring the performance of gestures following the relevant speech. Alternatively, as suggested by McNeill (2004, 28ff), the Arrernte may simply prefer the gestures to follow the speech. The lack of binding and different timing would simply be a cultural choice, a cultural value. Gestures for the Arrernte could then be interpreted similarly to the Turkana people of Kenya, in which gestures function in part to echo

and reinforce speech, other potential cultural values, and functional enhancements of language communication, expressed through gestures. Whatever the analysis, one must appreciate the relevance and significance accorded to culture in McNeill's and other researchers modern analyses of gesture, following on in the Boasian tradition inaugurated by Efron.

In his work McNeill introduces the vital term, "equiprimordiality" into the discussion of gestures and their relationship to the evolution of human speech. By this he intends that gestures and speech were equally and simultaneously implicated in the evolution of language. To understand this, we must ask how the Growth Point and the imagery-language "dialectic" evolved. Here McNeill (2012, 65ff) relies on George Herbert Mead's (1974) seminal work of on the evolution of the mind as a social entity, with special attention to language and gestures. Mead's claim on gestures (1974, p47ff) is that "Gestures become significant symbols when they implicitly arouse in an individual making them the same response which they explicitly arouse in other individuals" (this was probably written in the 1920s). McNeill's insight is to take Mead's conjecture and tie it in with Rizzolatti and Arbib's (1998) discussion of the involvement of mirror neurons in language. What McNeill claims is that Rizzolatti and Arbib missed a crucial step which he refers to as "Mead's Loop," wherein one's own gestures are responded to by one's own mirror neurons in the same way that these neurons respond to the actions of others, thus bringing one's own actions into the realm of the social and contributing crucially to the development of a theory of mind – being able to interpret the actions of others under the assumption that they have minds like we do and think according to

similar processes. Thus McNeill at once links his research program and the evolution of language more generally to the brain and society in an interesting and unique way, also highlighting the ineffable cerebral, and social connections in the formation of language, culture, and dark matter.

According to McNeill (2012, 69), Mead's Loop entails that "speech and gesture had to evolve together." "There could not have been gesture-first *or* [emphasis McNeill's, DLE] speech-first.[23]" This follows, he claims, because Mead's loop creates a "dual semiotic." "To create the dual semiotic of Mead's Loop, they [speech and gesture, DLE] had to be equiprimordial." Mead's Loop made possible the dynamic aspects of speech as well as the analysis of otherwise holophrastic constructions into parts, such as words, phrases, sentences, morphemes, phonetic segments, and so on. McNeill explains this (2012, 67) by claiming that;

> *"Semiotically, it* [Mead's Loop, DLE] *brought the gesture's meaning into the mirror neuron area. Mirror neurons no longer were confined to the semiosis of actions. One's own gestures... entered, as if it were liberating action from action and opening it to imagery in gesture. Extended by metaphoricity, the significance of imagery is unlimited. So from this one change, the meaning potential of language moved away from only action and expanded vastly."*[24]

I notice a couple of things in this quote relevant to our present concerns. First, language expresses action and meaning, with structure as an aid, which sets it off from a great deal of linguistic analysis, e.g. the tradition exemplified in Berwick and Chomsky (2016). Second, Mead's Loop and the Growth Point place compositionality

in a somewhat different light in the evolution of human language. Most linguists, myself included, would have likely answered, when asked what the great quantum leap in the evolution of language was, "compositionality."[25] But if McNeill is even partially right here, the Growth Point's evolution from Mead's Loop is the prerequisite for compositionality. Compositionality emerges from the act of communication, not from a genetic saltation. In fact, in Everett (2012a) I allude to the possibility that compositionality relies on non language-specific cognitive abilities. Interestingly and unfortunately, this possibility is ignored in recent works on the evolution of language, e.g. Fitch (2010), which prefer to see language as cognitively unique.

Once we get past this initial hurdle of how gestures become meaningful for humans, other notions arise to fine-tune the evolutionary story of the gesture-speech nexus. McNeill's theory, (e.g. 1992, 311ff) takes a perspective similar to Construction Grammar (Goldberg 1995) in claiming that utterances – gesture/speech wholes – are initially "holophrastic," used as single words or unanalyzable wholes, and later, through reuse and gestural focusing on specific components of the holophrastic construction, analyzed in more detail, leading to grammatical rules in a way reminiscent of the discovery methods of Harris (1947, Longacre (1964), and others, i.e. distributional isolability and recombination).

As gestures and speech join as signs in the social space, gestures take on one of two perspectives (McNeill 2005, p34). They either represent the viewpoint of the observer/speaker, *OVPT*, or the viewpoint of the person being talked about, or *Character View Point*, *CVPT*. Thus as we practice language and culture we learn these

things – different viewpoints, different ways of highlighting content and attributing ownership of content.

For example, McNeill gives an example of one person retelling what they say in a cartoon of Sylvester the cat and Tweetie-bird. When their hand movements are meant to duplicate or stand for Sylvester's movements, then their perspective is CVPT. But when their hand movements indicate their own perspective, then their perspective is OVPT. The ability to see events from different perspectives is "imaginative displacement" and is the crucial cognitive-cultural foundation for symbols and grammar.

Another interesting component of McNeill's theory of language evolution concerns his own take on recursion. Recursion is (see inter alia Everett 2010) a tool for more tightly packing information into single utterances.[26] Thus he independently arrives at an important conclusion found in recent debates on recursion by providing a model of language evolution and use in which recursion is useful but not essential, a very similar point to Everett (2005, 2008, 2009, 2010, 2012a, 2012b, and many others).

Some disagree with McNeill and argue instead that language did evolve from gestures. For example, in recent years, Tomasello (1999, 2008), Corballis (2002), Hewes (1973), and Arbib (2005), among many others, have argued that "... language evolved, not from the vocal calls of our primate ancestors, but rather from their manual and facial *gestures*" (Corballis 2002, ix), McNeill argues that there are two insurmountable problems with the "gesture-first" theory of language evolution. In the first place, speech did not supplant gesture. Rather, his work as well as that of

his students and many, many others show, gestures and speech form *an integrated system*. The gesture-first origin of language predicts asynchrony between gesture and speech, since they would be separate systems. But they are synchronous and parts of a single whole. Further, code-switching between gestures and speech is common. Why, if speech evolved from gestures, would the two still have this give-and-take relationship? Moreover, if the gesture-first hypothesis is correct, then why, aside from languages of the deaf, is gesture never the primary "channel" for any language in the world?

The second major problem with the gesture-first theory is that if gestures could be substituted by speech, they would not then be of the right type to form a language. This follows because in the absence of language, the available communicative gestures would have to be pantomimes. But, as McNeill makes clear throughout his work, *pantomime repels speech*. Pantomime does not accompany speech – it fills in missing values or gaps in speech. It is used in lieu of speech.

Also, as McNeill makes clear throughout his work, speech is built on a stable grammar. The only gestures which provide stability are the conventionalized and grammaticized gestures in sign languages. In this case again, however, gestures are either used instead of or to supplement speech and thus are unlikely to be the source of language. Summing up, then, had sign language or other gestures, e.g. pantomimes or language-slotted gestures, preceded speech, then there would have been no functional need for speech to develop. As McNeill (2012, 60ff) puts it "First, gesture-first must claim that speech, when it emerged, supplanted gesture; second, the gestures of gesture-first would [initially, DLE] be pantomimes, that is, gestures

that simulate actions and events; such gestures do not combine with co-expressive speech as gesticulations but rather fall into other slots on the Gesture Continuum, the language-slotted and pantomime."

McNeill's work further develops a specific proposal for the role of gestures in the evolution of grammar. As he (2012, 77) suggests, "an area of life where a syntactic ability could evolve is the cultural and social encounter." Here he cites the work of Freyd (1983) on "shareability" – the idea that structures and meanings must come to be shared among individuals if we are to say that they speak the same language, i.e. are utilizing the same outputs of conventionalization (another instance of the "actuation problem" (Weinrich, Labov, etc). In particular, McNeill appeals to Freyd's "discreteness filter," an idea akin to the generative notion of "discrete" in the phrase "discrete infinity (for criticisms of the latter see Everett 2010)." The idea is that our utterances are initially holophrastic, noncompositional. Then as humans began to learn a repertory of such utterances, these would begin to change via the GP, such that gestures would highlight some portions of the previously unanalyzable whole, leading to an analysis of the holophrastic into component parts – top-down parsing that eventually results in compositionality. This fascinates me because it presents a picture of how learning and emicization of the relationship between gestures and grammar can drive development of the very evolution of language.

This points to a stark difference between McNeill's theory and other theories of language evolution (such as Hauser, et. al. (2002)). In McNeill's theory the compositionality of syntax arises from actual language *use* via GPs, not from a sudden appearance of compositionality via, say, recursion. And, once again, it

identifies the particular function of language as communication, without which language, including gestures, would not arise. In fact, in McNeill's theory (as in Kinsella's (2009) and Everett's (2012a), among many others), compositionality *precedes* recursion. And this is just as the dark matter model (Everett 2016) of language emergence (see MacWhinney and O'Grady 2016) predicts – language emerges from a process of emicization, reanalysis, and re-emicization in order to better satisfy communication needs (Hopper (1988); MacWhinney (2006); Steels (2005); Rosenbaum (2014), etc). McNeill expresses this well: "Contrary to traditions both philological and Biblical, language did not begin with a 'first word. Words emerged from GPs. There was an emerging ability to differentiate news worthy points in contexts; a first 'psychological predicate' perhaps but not a first word." By demonstrating how compositionality could have come about by use and thus entered all human languages from early human interactions, McNeill undermines the need to appeal to genetics or biology to account for language evolution, instead supporting the account I develop here. In the context of his discussion of compositionality, moreover, McNeill offers an extremely interesting discussion of how recursion itself might have entered grammar, one quite compatible with my own (Everett (2010), Everett (2012a), Everett (2012b)).

Gesture is the "third eye," the *triality of language* production patterning. Because it cannot be removed from language – not even in thinking – we must conclude that language evolved primarily for communication, a communication that engages the entire body. It is interesting that theories which argue that language did not evolve for communication ignore gesture entirely. [27]

8. The Structural Components of Language

If this model of the semiotic progression is on the right track, duality of patterning, along with gestures, triality of patterning, are the foundational organizational principles. However, after that, we would expect that many languages will discover the utility of hierarchy, which Simon (1962) has argued in detail to be extremely useful in the transmission or storage of complex information.

The relationship of these grammars to the Chomsky-hierarchy (Chomsky (1963)) is not a simply mapping. Because this hierarchy is unrelated to our central purpose, I ignore it.

8.1. Phonology

Phonology is, like all other forms of human behavior, constrained by the memory-expression tension: the more units that a language has, the less ambiguously it can express messages, but the more there is to learn and memorize. It is also concerned by the preference of humans to simplify and generalize. And it is also constrained by what we can hear, which gets us to the co-evolution of the articulatory and auditory apparatuses. The latter constraint gives us the basic units of perception. The former two introduce variability into the system, itself an emergent property of the culture from which the phonology arises.

The articulatory apparatus of humans is of course interesting because no single part of it – other than its shape (which includes the position of the tongue) – is specialized for speech. Philip Lieberman has written extensively on the evolution

the human vocal apparatus, so I refer the reader to that work (2013, 2007) for lengthy and insightful discussion of the details.

The human vocal apparatus has three basic components – moving parts (articulators), stationary parts (points of articulation), and air flow generators. It is worth underscoring the fact that (as many have argued, but see Everett (2012; 2017)) that the evolution of the vocal apparatus for speech is likely to have followed the beginning of language. Though language can exist without well-developed speech abilities (since it could be, as many modern languages are, whistled, hummed, or signed), there can be no speech without language. Neanderthals, for example, did not have speech capabilities like those of Sapiens. But they most certainly could have had a working language (at the G_1 level easily, but perhaps also G_3, see below for more discussion of these language types) without a Sapiens-like vocal apparatus. For example, the inability of Neanderthals to produce /i/, /a/, and /u/ (Lieberman (2007)) would be a handicap for speech, but these cardinal vowels are neither necessary nor sufficient for language (not necessary because of signed languages, not sufficient because parrots can produce them).

Speech is enhanced, as has been mentioned, when the auditory system co-evolves with the articulatory system. We become able to hear and use most easily the sounds we naturally can produce.

Individual speech sounds, phones, are produced by the articulators – tongue and lips for the most part – meeting or approximating points of articulation – alveolar ridge, teeth, palate, lips and so on. Some of these sounds are louder because they offer minimal impedance to the flow of air out of the outh (and for many out of

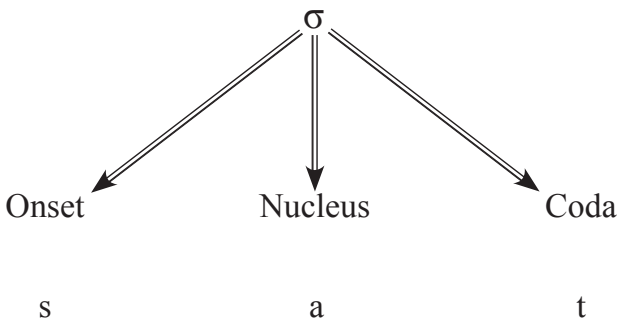
the nose). These are vowels. No articulator makes direct contact with a point of articulation in the production of vowels. Other phones completely or partially impede the flow of air out of the mouth. These are consonants. With both consonants and vowels the stream of sounds produced by any speaker can be organized so as to maximize both information rate (consonants generally carry more information than vowels, in information-theoretic terms) and perceptual clarity (consonants are easier to perceive in different positions of the speech stream, such as immediately preceding and following vowels and the beginnings and ends of words). Vowels and consonants, since speech is not digital in its production, but rather a continuous stream of articulatory movements, assimilate to one another slightly in some contexts (not always the same contexts in every language). Additional modifications (such as aspiration, voicing, and syllable structure, among others) enhance perception of speech sounds. These enhancements are ignored often by native speakers when they produce speech since the enhancements are, depending on the language, add-ons and not part of the target sound (e.g. [p] vs. [p^h]). Speakers are unaware of such enhancements usually and only with special effort can they learn to hear them (showing that these "allophones" are part of their dark matter and that the relationship of the enhancements to the target sounds are part of the speakers' emic knowledge – see section 8 below). The study of the physical properties of sounds, regardless of the speakers' perceptions and organization of sounds, is phonetics. The study of the emic knowledge of speakers, i.e. what enhancements are ignored by native speakers and what sounds they target, is phonology.

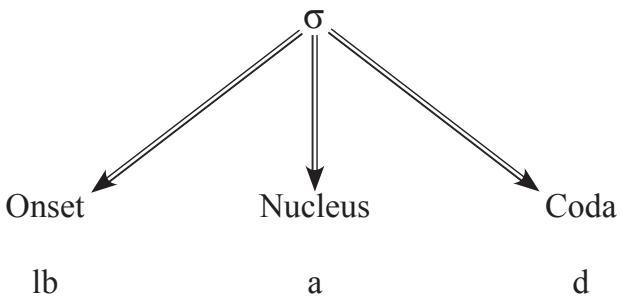
The study of measurable physical entities observed in a culture is what Kenneth Pike (1967) referred to as the *etic*. How those entities are perceived, used, and classified by members of a culture is the *emic* (see Everett 2016 for a detailed exposition of etics and emics in the construction of culture).

There is a long tradition that breaks these basic sounds, vowels and consonants, down further into phonetic features, e.g. [+/- voiced], [+/- advanced tongue root], and so on. But no harm is done to our exposition here of language evolution if we ignore such finer details.

Thus the lowest unit on the evolving phonological hierarchy is the individual speech sound, perceptually enhanced and contextually altered to unconsciously aid native speaker perception and categorization of their target speech sounds. A phonological hierarchy independent of any specific language is an etic hierarchy. The actual hierarchy utilized or organized by native speakers, while never deviating radically from the etic, is the emic hierarchy.

Moving up the phonological hierarchy we arrive once again at the, which introduces duality of patterning into speech sound organization. To elaborate slight further on what was said earlier about the syllable, consider the syllables below:


σ
Onset
Nucleus
Coda
s
a
t



As per our earlier discussion of sonority, we expect that the syllable [sat] will be well-formed, ceteris paribus, while the syllable [lbad] will not be because the

latter violates sonority sequencing (see footnote 10), while the former respects the principle.

The syllable is thus a hierarchical, non-recursive structuring of speech sounds. It functions to enhance the perceptibility of phones and often works in languages as the basic rhythmic unit. One can imagine that given its extremely useful contributions to speech perception, that syllables began to appear early on in the linking of sounds to meaning in language. It would have been a very useful and easy add-on, improving dramatically the perceptibility from a hypothetical system that might merely place speech sounds in an unrestricted linear arrangement of just vowels or random vowels and consonants. The natural limitations of our auditory and articulatory systems would have exerted pressure for speakers to hear and produce syllabic organization early on.

However, once introduced, syllables, segments, and other units of the phonological hierarchy would have undergone culture-based (Everett (2012)) elaborations, in order to make sounds that identified one group as the source of those sounds, because speakers of one culture would have preferred some sounds to others, some enhancements to others, and so on. The inventory of sounds would also have been culturally limited (which of the sounds that can be produced and perceived relatively easily does one culture choose to use?). After selection, these sounds will change over time, again subject to articulatory, auditory, or cultural pressures, or via contact with other languages (Thomason and Kaufman (1989)).

Other units of the phonological hierarchy include phonological phrases, groupings of syllables into phonological words or units larger than words. These

phrases or words are used to distribute pitches for additional information-marking (see Everett (1979)) on phonological words and Pierrehumbert and Hirschberg (1990) on the functions of phonological phrases). And above these phrases there are what Pike (1967) refers to as "contours" or "breath groups" – groupings marked at their boundaries by breathing and intonational end points. Pitch and other prosodies are also used to mark information structure (predicates vs. topics, old vs. new information, emphasis, hand so on). All of these uses of phonology emerge gradually as humans move through the semiotic progress from indexes to grammar. And every step of the way they would have likely been accompanied by gesture.

The phonological hierarchy that emerges from natural, perceptual-production enhancing processes:

Phonological Hierarchy

Phonemes

Syllables

Phonological words

Phonological phrases

Phonological paragraphs

Phonological texts

Conversational features

Phonology is also constrained by two other sets of factors. The first is the environment. Sound structures can be significantly constrained by average

temperatures, humidity, atmospheric pressure, and so on where the language arose (Caleb Everett (2014: 2015)). Thus to understand the evolution of a specific language, we must know something about its original culture and ecological circumstances.

But phonological hierarchies are also affected by pressure to mutually encode or decode a message with other components of grammar. The need for syntax, phonology, morphology, meaning and gestures to work together to mark the information structure (VanValin and LaPolla (1997), intentions (Searle (1983)), and so on being communicated is a natural external pressure that brings them into some alignment (though as Pike (1967) says at many times in his work, the "hierarchies are skewed." That is, it is not always the case that the exactly match – phonological phrases can fall across syntactic boundaries (as in Steedman (2001) who argues for distinct phonological and syntactic phrasing), gestures can follow the entire utterance (as McNeill (2000; 2004) describes for Arrernte, based on the work of de Ruiter and Wilkins (1998) and Wilkins (1999)), and semantics can cross sentence boundaries (as in phrase sequences like "You drink. You drive. You go to jail." These are three syntactic units interpreted semantically as a single conditional). The interaction of gestures, phonology, and information structure with syntax and grammar more generally offers more strong evidence that the primary purpose of language is communication, contra the hypothesis of Berwick and Chomsky (2016), Everaert, et. al. (2015), and so on.

8.2. Morphosyntax

Like McNeill, I think it makes more sense to suppose that early language was holophrastic (though this is not essential to the theory here). That is, the first attempts at communication were unstructured utterances that were neither words nor sentences, simply interjections or exclamations. They were simply polyphonal things with gestures that could be contextually interpreted in many cases.

How expressions catch on is a puzzle known as the "actuation problem" (Weinrich, Herzog, and Labov 1968). Just as new words or expressions or jokes today, several possible enabling factors might be involved. Speakers might like the sounds of certain components of a holophrastic utterance more than others, a pitch and a gesture occurring simultaneously might have highlighted one portion of the utterance to the exclusion of other parts. As highlighting is also picked up by others and begins to circulate, for whatever reasons, the highlighted, more salient portion becomes more important in the transmission and perception of the utterance being "actuated." Gestures, with functions that overlap intonation in some ways, contribute to this, either reinforcing the highlighted portions or by marking other portions of the utterance as of secondary importance, but still more important than other portions. Along with prosody (pitch, loudness, length), gestures and other markers of salience (body positioning, eyebrow raising, etc.) have the joint effect of beginning to *decompose* the utterance, breaking it down into *salient*, *less salient*, and *non-salient*. Members of a community then take such etic components of etic utterances, gradually associating meaning with the different parts. Once the utterances are decomposed (analyzed), and only then, they can be (re)composed (synthesized), i.e. used to build additional utterances. And this leads to another

necessary property of human language, semantic compositionality – being able to encode or decode a meaning of a whole utterance from the individual meanings of its parts.

Thus from natural processes linking sound and meaning in utterances, there is an easy path via gestures, intonation, duration, and amplitude to decomposing an initially unstructured whole into parts and from there recomposing parts into wholes.

Pike (1967) placed morphology and syntax together in another hierarchy, worth repeating here (though I use my own, slightly adapted version):

Morphosyntactic Hierarchy

Morpheme

Word

Phrase

Sentence

Paragraph

Monologic Discourse

Conversations

It is likely that the first utterance was made to communicate to someone else. Of course, I wasn't there. Nevertheless the prior and subsequent history of the semiotic progression strongly supports this.

The evidence suggests that there is no innate "phonological mind" (Berent 2013; Everett 2016). That is, it seems that none of phonology is innate. The simplest hypothesis is that along with the co-evolution of the vocal apparatus, the hearing apparatus, and linguistic organizational principles led to the existence of a well-organized sound-based system of forms for representing meanings as part of signs. There are external, functional, and ecological constraints on the evolution of sound systems.[28]

Syntax develops as duality of patterning becomes more intricate over time, subject to culturally significant communicational objectives and conventions, along with computational constraints (see Shannon (1948)).

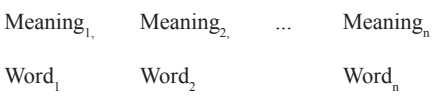
Morphology comes in different varieties, falling into a small set of categories that exhaust both the empirical possibilities and likely probabilities of organization of form and meaning in word structure.

Here are the choices in building morphological (word) structures. We can compose words from parts or not. If not, we have an isolating language, such as Chinese. If so, then we can combine multiple meanings in each part, just one meaning per part, or one meaning per multiple parts. In the first case we have a fusional language, such as Spanish. If the second possibility we have an agglutinative language like Turkish. If the third, we get special kinds of morphemes called circumfixes (e.g. in German past tense of the verb *spielen* 'to play' we get the past participle form, ge*spiel*t, where *ge-* and *–t* jointly express the past tense, circumscribing the verb they affect. We could instead use pitch or another prosody to add meaning to word, deriving tonal *simulfixes* (e.g. Pirahã words **ʔáagá**

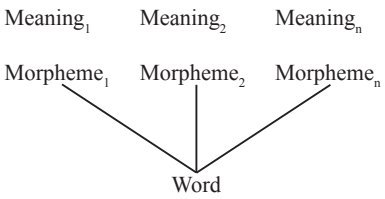
"permanent quality" vs. **ʔaagá** "temporary quality"). Or one could express some meaning via consonants and other meanings via vowels, in which case we have a nonconcatenative system such as in Arabic languages.

It is unlikely that any language has a pure system. But what this brief summary of morphology shows is that if we take all the morphological systems of the world in consideration, morphology is not hard in its basic forms (later cultural decisions on what to represent and how to associate one meaning with others can make learning a specific system extremely difficult). This is summarized below:

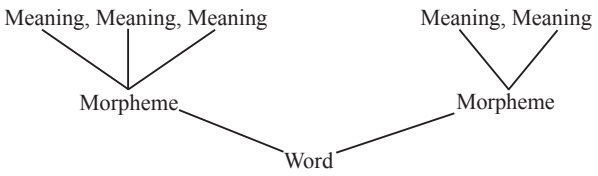
1. Isolating:



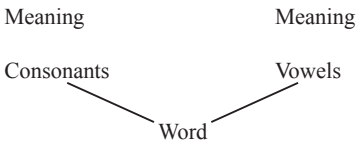
2. Agglutinative



3. Fusional



4. Nonconcatenative



These are the choices that every culture has to make. It could choose multiples of these strategies, but simplicity (for memory's sake) will favor less rather than more mixing of the systems. From the beginnings of the invention of

symbols, at least 550,000 years ago by Homo erectus, there has been sufficient time to discover this small range of possibilities and, via triality of patterning, to build morphological systems from these possibilities.

8.3. Syntax[29]

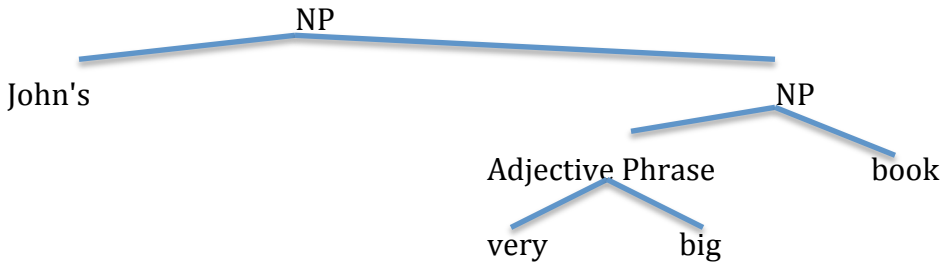
8.3.1. Introduction

Arguably Chomsky's greatest contribution to the understanding of human languages is his Emile Post/Marcel Schützenberger – inspired classification of different kinds of grammars, based on their computational and mathematical effects.[30]

Yet while the Chomsky hierarchy is insightful and useful to computer scientists, I ignore it here in order to discuss a less formal, but perhaps more useful system for this particular discussion of the evolution of language.

There are linear grammars, hierarchical grammars, and recursive hierarchical grammars available to cultures, though neither Chomsky nor anyone else categorizes grammars in exactly this way. It is not a set of mathematically distinct options. They are empirical choices. A linear grammar would be an arrangement of words left to right. A word ordering in a linear grammar is not likely random. Throughout all languages heretofore studied, grammatical strategies are used to keep track of which words are more closely related. One common strategy is to place words closer to words whose meaning they most affect (Bybee (2010)). Another is to place words in phrases hierarchically, keeping track of them

configurationally (assuming that such configurations are represented somehow in the minds of speakers and hearers), as in:



The final common way to keep track of which words are most closely related is to mark words with *case* or *agreement.* Greek and Roman allow words that are related to be separated by other words in a sentence, so long as they are all marked with the same case (genitive, ablative, accusative, nominative, and so on). Agreement, as in "He likes John" Where "he" is third person singular and the "s" on "likes" is also third person singular, is another way to keep track of word relationships in a sentence.

8.3.1. Linear grammars

Placing words together in linear order without hierarchical structure is a logical option and would likely have been selected first, since it is the easiest for short utterances in intimate societies (Everett (2005), (2008), (2012)). Futrell, et. al. (2016) argue that the modern language Pirahã (and see Jackendoff and Wittenberg (2012) for similar arguments for the Indonesian language, Riau) can be analyzed as having a linear grammar.

In a linear grammar, semantic compositionality operates largely independent of the syntax, as it does in English sentences such as "Eat. Drink. Man. Woman." and similar perfectly intelligible, but non-syntactically compositional speech. Though many languages, if not all, have examples of compositional semantics independent of syntax (another example from English would be "Bill has gray hair." "Really?" "Or so John says." See Everett (2010) for more detailed discussion), the closer the syntax matches the semantics, the likely easier the semantic interpretation in more complex utterances. But complex utterances are a choice, not genetically predestined. And the fact that there are such languages, at least given the best data currently available (see Futrell, et. al. 2016), suggests not only that linear grammars were the likely initial stage of syntax, but that such concatenative approaches are still viable strategies for modern grammars and that attaining such a grammar is to have attained full human language.

8.3.2. Hierarchical non-recursive grammars

In an important paper, Karlsson (2009) argues that "Standard Average European" (SAE) languages are hierarchical but non-recursive, with strictly circumscribed levels of embedding. To cite him here (Karlsson 2009, 193):[31]

> *"(1) I2max: the maximal degree of initial embedding is two (some 100 instances found, as in (6)).*
> *(2) Qualitative I2-constraints: Double initial embedding strongly prefers a) written language, b) an if-clause as higher embedding, c) a sentential subject,*

i.e. a what-clause as lower embedding, and d) finiteness. Cf. (6).

(3) C3max-w: in written language, the maximal (utterly rare) degree of multiple center-embedding is three (thirteen instances retrieved, (7)).

(4) C2max-s: in spoken language, the maximal (utterly rare) degree of multiple center-embedding is two (less than five instances retrieved, (8)).

(5) Only-postmodifying-self-embedding: only postmodifying (i.e. typically relative) clauses allow self-embedding (7, 8), i.e. repeated embedding of precisely the same type of clause.

(6) [Main [I-1 If [I-2 what he saw through security] did not impress him] Tammuz ...] (BNC)

(7) [Main Der Ritter von Malzahn, [C-1 dem der Junker sich als einen Fremden, [C-2 der bei seiner Durchreise den seltsamen Mann, [C-3 den er mit sich führe,] in Augenschein zu nehmen wünschte,] vorstellte,] nötigte ihn ...] (von Kleist, Michael Kohlhaas)

(8) [Main A lot of the housing [C-1 that the people [C-2 that worked in New Haven] lived in] was back that way.]"

Further clarifying, Karlsson goes on to say (p193) *"No genuine triple initial embeddings nor any quadruple center-embeddings are on record ('genuine' here means sentences produced in natural non-linguistic contexts, not sentences produced by professional linguists in the course of their theoretical argumentation)."*

That is, there is no evidence from Karlsson's thorough research that *any* SAE language is recursive in a practical sense. Of course, one could argue that they are

recursive in an abstract sense or generated by a recursive process. But this more abstract perspective is largely irrelevant to the claim here that fully modern human languages need not show evidence for recursion. Moreover, Karlsson's work demonstrates that it is possible in principle to have languages that are hierarchical (at a minimum embedding structures are prima facie evidence for hierarchy in a language) but not recursive.

8.3.4. Hierarchical recursive languages

The final type of language would be any language that shows both hierarchy and recursion. There are many claims in the literature that there are such languages. As we see in the next section, there is even a claim that all languages are built by a recursive process, Merge, and that this process represents the dividing line between "proto-languages" and modern languages. However, the fact remains that if no language has been documented in which sentences generated by speakers are found to be actually unbounded, then recursion is a property of the grammar-writing process, not necessarily of grammars themselves. There may be theoretical reasons for claiming that Merge or some such recursive process underwrites all modern human languages, but it has no empirical significance in the understanding of the evolution of languages.

8.4. Semantics and compositionality

As we said earlier, there are many examples in all languages that show non-direct mapping from syntax to semantics. In discourses and conversations,

moreover, the meanings are composed by speakers from disconnected sentences, partial sentences, and so on. This supports the idea (see also Everett 2012 and, especially, Everett 2016 on the "dark matter of texts") that compositionality is a function of matching syntactic forms to meanings through the medium of culture, with culture constructing meaning from utterances. Many examples of this are given in the works cited. Also, there are theories that allow semantic compositionality to match to the syntax by "linking rules" that allow for much more cultural guidance and independence of the syntax and semantics than some more formal theories (e.g. Chomskyan Minimalism). See VanValin and LaPolla (1997) for one approach.

In addition, some scientists have claimed that all semantic compositionality is done from linear, paratactic grammars, regardless of the superficial syntactic structures of the language in question, further strengthening the proposal that semantic meaning is not isomorphic with syntactic structure. (see Hobbs (2008) and Language Log – http://itre.cis.upenn.edu/~myl/languagelog/archives/005380.html for such a paratactic account of semantics).

8.5. Pragmatics

The general principle of pragmatics is that "*Language is not saying all you mean and not meaning all you say*." While we lack space to provide any account of the evolution of pragmatics, I refer the reader to works such as Sperber and Wilson (1996), Grice (1991), and many others that show that the pragmatics, what I (Everett 2012 and Everett 2016) interpret as cultural constraints, fill in gaps and alter literal meanings to get a speaker meanings in local contexts of utterances. This

means that neither semantics nor syntax, etc. are sufficient to guarantee meaning transfer between interlocutors. And this must have especially been the case during the evolution of language when syntax did not yet even exist.

9. The cultural foundation of language

9.1. The etic vs. the emic

Indexes are readable by members of any culture, in fact by members of most species. They are clues to the environment necessary for survival across many lifeforms. And thus we know that the ability to connect a representation to a form (in indexes) is an ancient ability of the genus Homo. We were never without it. Icons, on the other hand, require more. First, they are *intentional*. Whether making an icon or simply collecting one, the reader of the icon must understand that it physically resemble what it represents. Understanding, whether indexes, icons, or symbols is an intentional act, directly or indirectly (because of unconsciously looking for and recognizing connections). An index is non-intentional.

But though icon creation requires intentionality, it is non-arbitrary. Arbitrariness is crucial for language because it affords access to an unlimited number of symbols, subject to local duality of patterning constraints. Thus we cannot have a full language of emojis because they are non-arbitrary (at least for now). Symbols have no physical connection to what they represent. So they are not indexes. Symbols further do not necessarily resemble what they represent so they are not icons. Symbols are at once intentional and arbitrary.

However, because they are arbitrary, they require agreement by members of a culture to function. For any arbitrary representation to serve more than one person, those who use the symbol must agree that the symbol means the same to all. They must agree to both its form and meaning. Symbols are like jokes in this sense – they have a first use, but subsequent uses depend on whether the symbol in question is accepted by the people who decide to use them.

But convention on any large scale sufficient for the emergence of language requires culture. The role of culture in language, values, knowledge, social roles, and array of other central features of human experience is taken up in Everett (2016). There I offer the following definition of culture:

> *Culture is an abstract network shaping and connecting social roles, hierarchically structured knowledge domains, and ranked values. Culture is dynamic, shifting, reinterpreted moment by moment. Culture is only found in the bodies* (the brain is part of the body) *and behaviors of its members.*

Everyone is born outside of culture and language. We are all aliens as we emerge from our mothers' wombs (though not completely because learning about our new culture and language begins in the womb). We are outside the culture of our own birth looking in. Our senses deliver physical impressions to us but only after much exposure to both these sensations and our and others' reactions to them (verbal, bodily, social, etc.) can we begin to interpret those impressions. The sensations of our senses at this outsider stage of our lives are *etic* sensations – i.e.

merely forms, lacking culture-based interpretations. As we learn to interpret such sensations, especially as they become automatized, our understanding and experience of them becomes *emic*, experiences interpreted from within a culture.

Although this definition emerges from a theory of culture, it also interrogates the construction of such a theory. What should, for example, a "theory of culture" do exactly? Well, it should enable us to understand. It should remove or radically lessen our surprises regarding human social behavior and indexical values. It should predict or at least explain individuals' behaviors in specific cultural contexts. It should offer an understanding of the major institutions of a society. It should provide a guide and methodology for investigating specific cultures.

9.2 Conclusion

Language is a nexus, a point of intersection and interaction of meaning (semantics), conditions on usage (pragmatics), the physical properties of its inventory of sounds (phonetics), a grammar (syntax (sentence structure), phonology (sound structure), morphology (word structure)), discourse and conversational organizational principles, and so on. And even though the components of language are numerous, the whole is a gestalt entity – it is more than merely the sum of its parts.

There are entire communities of linguists that identify themselves by the different subareas of the diagram. There are pragmaticists, conversational analysts, syntacticians, morphologists, phoneticians, semanticists, and so on. But none of them is studying language as a whole, only the parts they are interested in

professionally. A syntactician is to language as an opthomologist is to the body. Both are necessary, but are (understandably) tackling very small pieces of the pie.

The ultimate accomplishment of language is building relationships – cultures and societies (societies precede and envelop cultures, as I explain in Everett 2016). We build these relationships through stories and conversations, even written ones, that establish and justify shared value-rankings (all of our values are hierarchical, as we see, for example, in the fact that for soldiers patriotism is valued above the commandment not to kill, etc), knowledge-structures (such as that red and blue belong to the set of colors and that colors to the set of qualities, and so on), and social roles (author, editor, teacher, laborer, father, mother, etc).

Grammar is a tremendous aid in this task of language to build relationships. It is also a help in thinking (see Everett 2012). But in spite of the focus of many linguists on grammar as the core of language (as Chomsky and his followers arguably do), grammar itself in fact tends not to be any more important than other components of language in our diagram. Again, grammar is extremely useful, but there could be languages without grammars.

10. Who had language?

10.1. Why this is an important question

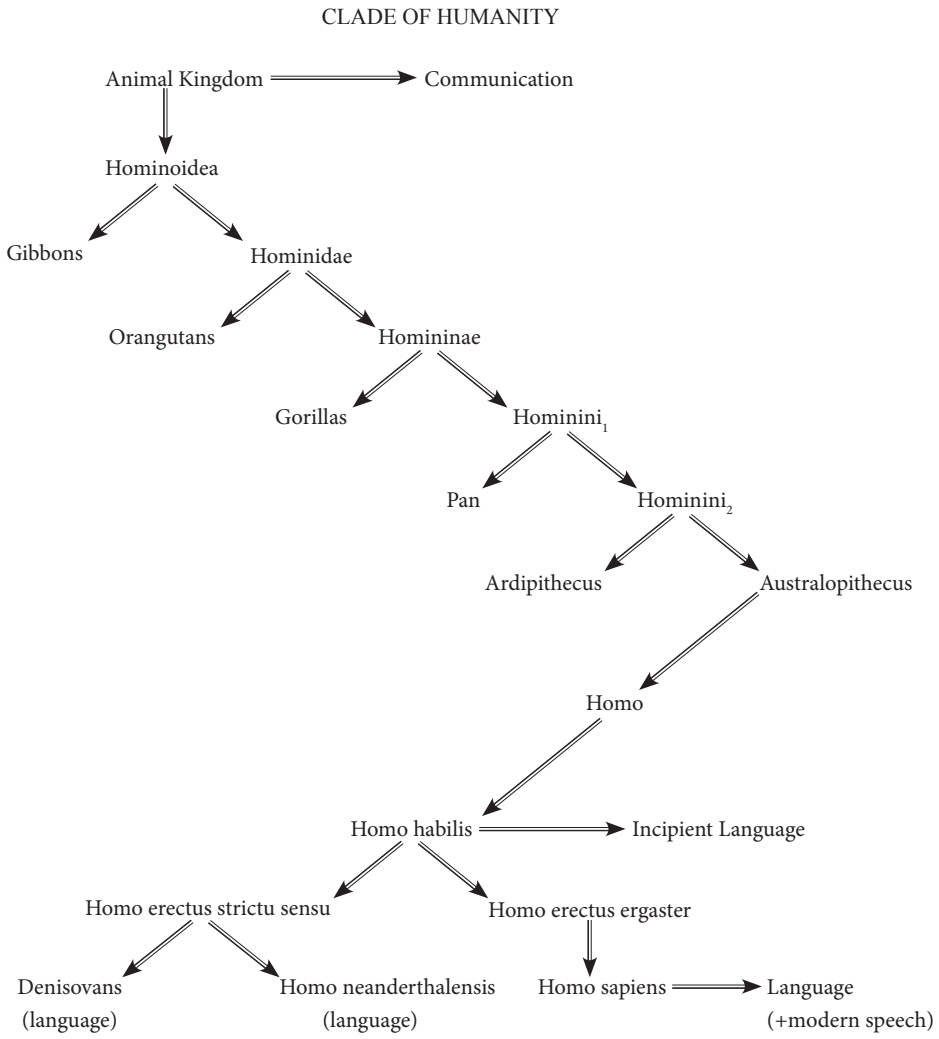
It is essential to consider carefully when language (symbols in triality of patterning linear, hierarchical, or recursive grammars) first emerged. The earlier the date supported by the evidence, the more likely the process is to follow from uniformitarian assumptions rather than those of catastrophism. And this tells us

whether language is only for X-men or that we are creatures like any other, resulting primarily from gradual natural selection.

10.2. The evidence summarized

All animals communicate, hence the arrow at *Animal Kingdom*, in Figure Three below. Not all animals have language. We can only talk about the evolution of language in the context of the evolution of our species. Each one of the early Hominini will be discussed in varying detail in this chapter. I have listed Hominini_1 and Hominini_2 in order to represent my view that the genus Homo likely descended from and is not merely another branch with Australopithecus.

FIGURE THREE[32]



Safina (2015) has argued that animal communication far exceeds what researchers have commonly noticed in the past. Panksepp and Biven (2012) have

also argued that animal emotions, often the basis of utterance interpretation, are likely very similar to those of humans, in feeling if not in interpretation (but cf. LeDoux 2016 for an alternative account). Still, although animals make use of indexes regularly and perhaps some make sense of icons (e.g. the dog barking at the television screen when other dogs are "on" the screen), there is no evidence of animals using symbols in the wild (I do not discount this possibility tout court or a priori, but I am not aware of any well-supported/accepted claims to the contrary).[33]

Among humans, however, as we have seen, there is evidence that both Homo erectus and Homo neanderthalensis used symbols. And with symbols + concatenation plus highlighters (intonation and gestures), there is language, according to my theory here. Thus possession of symbols, especially in the presence of evidence that culture existed, evidence strong for both erectus and neanderthalensis, it is highly likely that language was in use in their communities. In fact, it is implausible in the extreme to assert that Homo erectus's cultural accomplishments were brought about by, say, unstructured grunting.

10.3. Any need for proto language?

Under my theory, there is no need for a concept of proto-language as a special theoretical construct. There are symbols and they are placed in a culturally-agreed upon order. And they are interpreted by rich cultural mechanisms that may or may not be reflected in distinct components of the language proper (such as recursion). Once we have order, symbols and culturally-enabled interpretative schemes, we have language. Not proto-language. Thus I find little use for this notion.

Even if there are simpler languages, they are languages, quantitatively not qualitatively different from more complex languages, as I have attempted to show by reference to some modern languages.

11. The Grammar Came First Alternative

As we have seen several times above, there are two opposing views of the nature of language. One is Chomsky's, which Searle (1972) describes as follows:[34]

> *"The syntactical structures of human languages are the products of innate features of the human mind, and they have no significant connection with communication, though, of course, people do use them for, among other purposes, communication. The essential thing about languages, their defining trait, is their structure. The so-called "bee language," for example, is not a language at all because it doesn't have the right structure, and the fact that bees apparently use it to communicate is irrelevant. If human beings evolved to the point where they used syntactical forms to communicate that are quite unlike the forms we have now and would be beyond our present comprehension, then human beings would no longer have language, but something else."*

Searle concludes that*"It is important to emphasize how peculiar and eccentric Chomsky's overall approach to language is."*

A natural reply to this would be that what is one person's "peculiar and eccentric" is another person's "brilliantly original." The best work often is eccentric and peculiar. But I want to argue that Chomsky's view of language evolution is not

to be rejected because it is original, but because it is wrong. He has continued to double-down on this view for decades. For example, Berwick and Chomsky (2016) present a clearly written, cogently argued and otherwise convincing theory of language and its evolution, which furthers the 60 year-old syntacto-centric program of linguistic theorizing that Chomsky initially placed on the global intellectual map, the one that Searle questions above. Chomsky's view was in fact so novel and shocking that it was initially thought by many to have introduced a revolution into linguistic theory and been the first shot fired in the "cognitive revolution" that some date to a conference at MIT on September 11, 1956.

But it was neither a linguistic revolution nor a cognitive revolution. In the 1930s Chomsky's predecessor, Leonard Bloomfield, developed a theory of language remarkably like Chomsky's, in the sense that structure rather than meaning was central and communication was also considered secondary to structure. And another predecessor of Chomsky's, Edward Sapir, had since the 1920s argued that psychology (what some today would call cognition) interacted with language structures and meanings in profound ways. Still, Chomsky has staked out his claims clearly over the years, reiterating them in his new work on evolution, namely, that "language" is a computational system, not a communication system. More specifically, it is a set of endocentric, binary structures created by a single operation, Merge, and only secondarily used for story-telling, conversation, sociolinguistic interactions, and so forth. This led Hauser, Chomsky, and Fitch (2002) to propose recursion as the special feature that underwrites human language and renders it

superior to other animals' communication systems.[35] And in Chomsky's proposal, recursion is both the foundation of language and a mutation.

Yet according to my theory (and McNeill's), recursion would have begun to appear in language, as we saw earlier, via gestures, prosodies, and their contributions to the decomposition of holophrastic utterances. This is because gestures, unlike the eventually compositional static outputs of grammar, are gestalt units (though not all gesture researchers accept this). This is a fundamental difference between these dynamic units vs. static syntax. Gestures are wholes without meaningful parts. And the meaning of the whole is not derived from the meaning of the parts. Thus although we can observe several submovements in the larger gesture, none of these smaller acts has any meaning apart from the gesture as a whole. Gestures are in this sense *anticompositional*.

As auditory symbols developed they would have been used in strings of symbols. Gestures and intonation, whether precisely aligned or only perceived to be aligned with specific parts of utterances, would have led to a *decomposition* of symbols. Other symbols could have been derived from utterances that had little internal structure initially, but were then likewise broken down via gestures, intonation and so on (McNeill (1992; 2002; 2005; 2012)).

But syntax became analyzable and *following this* recursion was able to play a role in the grammar. In this sense, recursion for McNeill, as for me (inter alia Everett 2005, 2008, 2009, 2010, 2012a, 2012b) is a nonessential, yet extremely useful component of language evolution (contra Hauser, et. al.(2002)). Recursion is used to render the syntagmatic (string) paradigmatic (a slot), enabling speakers to pack

more information into single utterances and, as I point out in Everett (2012b), making it easier to follow complex events via oral discourse. McNeill (2012, 223) cites Shelley's "The Masque of Anarchy" to illustrate the syntagmatic to paradigmatic shift:

"His big tears, for he wept full well,

Turned to millstones as they fell."

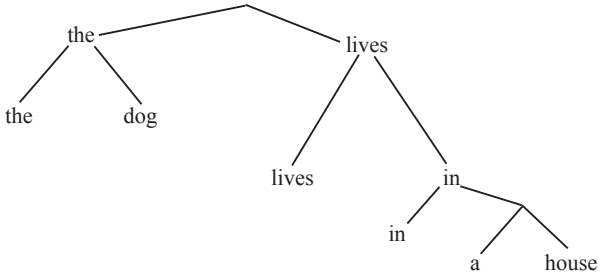
"The rhyming '-ell's, on the axis of combination, project a new semantic opposition..." From the rhyming the words are highlighted – potentially leading to their separate storage and analysis as words, *parts of utterances*, introducting compositionality into grammar. This opposition is between paradigmatic parts of larger sentences (syntagmemes) that themselves derive from syntagmemes. Thus having provided a plausible scenario for the evolution of syntax, McNeill turns to consider the resultant spread of static grammar.

Chomsky's grammar-first-theory (GFT) is the polar opposite of what we have seen above. It throws out gradual evolution, ignoring the evolution of icons, symbols, gestures, langauges with linear grammars, and so on, in favor of genetic saltation in the form of a sudden ability to do recursion. Again, according to GFT, communication is not the principal function of language. While all creatures communicate in one way or another, only humans have anything remotely like language because only humans have structure-dependent rules.[36] According to Chomsky language is not defined by its many functions or by an evolutionary process that built its properties slowly and messily over millennia. It is, rather, a perfect computational system that appeared quite suddenly, probably about 50,000

years ago, as the human brain passed a certain threshold of organizational complexity.

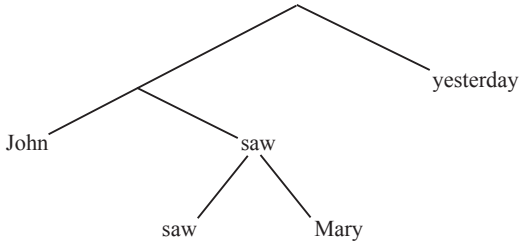
On the surface, this is a logically reasonable theory – clear, strong, and interesting. It simply has no connection to the facts discussed above. To see why, consider that Chomsky's conceptualization of grammatical structure has gotten more restrictive over the years (potentially a good thing, of course), by eliminating entire classes of potential structure-dependent rules in favor of a single such rule. The class of structures that Chomsky admits are generated by a recursive rule he calls "Merge." Merge (which comes in two varieties, "external" and "internal") generates only binary, endocentric (headed) structures of the type in (1) or (2a), but never structures like those in (2b):

1. The dog lives in a house.

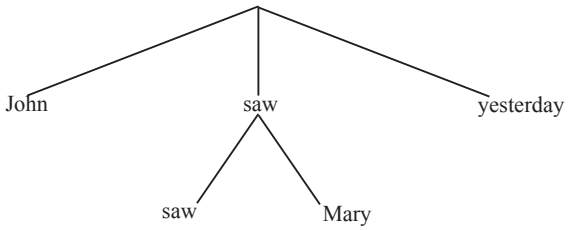


2. John saw Mary yesterday.

a.



b.



Example (2b) is a ternary-branching tree which has no head, i.e. it is exocentric. To understand how Merge expresses these restrictions, consider the definition of Merge offered by Everaert et. al. (p number 2016):

"Merge is a (dyadic) operation that takes two syntactic objects, call them X and Y, and constructs from them a single new syntactic object, call it Z. X,Y can be building blocks that are drawn from the lexicon or previously constructed objects. Put simply, Merge (X,Y) just forms the set containing X and Y. Neither X nor Y is modified in the course of the operation Merge."

Further, the authors add:

"If X and Y are merged there are only two logical possibilities. Either X or Y are distinct, and neither one is a term of the other, or else one of the two elements X or Y is a term of the other, where Z is a term of W if it is a subject of the other or the subject of a term of the other. We can call the former operation 'External Merge': two distinct objects are combined.

(i) Merge (read, that book) ==> {read, that book}

If alternatively X is a term of Y or vice versa and X and Y are merged, we call this 'Internal Merge'. So for example, we can (Internal) Merge which book and John read which book, yielding the following:

(ii) Merge (which book, John read which book) ==> [which book, John read which book}

In this case the result of merging X and Y contains two copies of Y. Following further operations, this structure will surface as in (iii), under a constraint to externalize ('pronounce') only the structurally most prominent copy of which book:

(iii) (Guess) which book John read."

Now the thesis is not merely that Merge is an operation on grammar. It is rather the "basic operation" that underlies *all* of language, conceived of as a grammatical core with add-ons (e.g. conversation structure, constraints on appropriate usage, metaphor, story-telling, idiophones, idioms, and so on).

Let's assume for the moment that Merge is correct in the sense that, as the authors would have it, it accurately describes the syntax of every language that has ever been spoken, is currently spoken, or ever will be spoken (if a communication system evolved that lacked Merge, it would not be a language according to Berwick and Chomsky (10ff, 53, elsewhere in their book)).

What could be the source of such an amazingly perfect computational system? One could say, as Berwick and Chomsky do (2016; 70ff), that Merge came into being as a "minor mutation." The simple operation Merge would thus have "popped" into being and is part of the human cognitive production and perception apparatus, no more subject to widescale variation across human populations than our visual, auditory, or olfactory systems. It is how we *must* organize our grammars. There could no more be grammars without Merge than healthy human eyes that lack stereoscopic vision. This is an interesting idea. And its claim that Merge did not evolve insulates it from the demand that an evolutionary account of Merge be offered. No such story is required for something that just appeared suddenly. [37]

But what is the evidence for the sudden emergence of Merge in the evolutionary, pre-historical record of Homo sapiens? In the main this evidence is asserted to be the mysterious appearance of complex tools, symbolic, representative art, evidence for culture (e.g. burial rites, homogeneity of tool and home designs,

etc), as well as the extinction of competing species, e.g. Homo neanderthalensis. Many of these things happened simultaneously about 50,000 years ago according to some (but not all) paleoanthropologists, e.g. Ian Tattersall (2013). Barring the possibility that aliens came to earth and taught these things to our ancestors, it is possible that these achievements did in fact appear simultaneously via the rapid appearance of language and culture among humans. Or, in Chomsky's terms, if Merge suddenly appeared. However, as I have argued, based on several current Amazonian material cultures, as well as the anthropological record, this is a fantasy. There are at least three serious problems with this view. First, it ignores the first "cognitive revolution" of Homo erectus, which could have only been underwritten by language. Second, it ignores the potential explanation for the second revolution (the appearance of complex sapiens tools and art) to be found in cumulative cultural knowledge. Third, it is unnecessary, as the model above is able to account for both languages with and without recursion without sudden appearance or mutations that we otherwise have no evidence for.

On the other hand, as one reviewer for this journal points out, it might appear that a mutation is not incompatible with the view of language evolution proposed above. There are two major cognitive revolutions in the history of our species – one that occurred nearly two million years ago with erectus and the other that occurred roughly 50,000 years ago with the proliferation of cultural artifacts of Homo sapiens. It is possible that the second revolution was the result of a mutation. But there is no evidence for that. What there is evidence for is gradual evolution and the appearance of grammar, compositionality, and recursive thought (tools that

contain parts within parts, such as erectus rafts as proposed by Bednarik (2003) long before

12. Conclusion

In what preceded I defended the grammar came last hypothesis. We discussed many reasons for rejecting the grammar came first hypothesis. Some of these are: (i) there are languages currently spoken that appear to lack any hierarchical grammar (e.g. Riau – Gil 1994 and Jackendoff and Wittenberg (2012)) and Pirahã (Futrell, et. al.; Everett 2005)), their "grammars" little more than linearity restrictions on words; (ii) there is a good deal of evidence that symbols evolved long before grammar in human linguistic history; (iii) hierarchical grammars are derivative from independent processing advantages of hierarchy in the organization and retrieval of information outside of human languages; (iv) there are no *well-established*; (v) nonhuman creatures appear to use syntax (so it is not exclusive to humans); (vi) humans have evolved away from cognitive rigidity (as via instincts), to cognitive flexibility (Everett 2016) and learning based on local cultural and even environmental constraints (see Caleb Everett (2015)). Under this latter assumption any grammatical similarities found across the world's languages would tell us more about the communicative constraints on grammars than anything about human evolution per se.

Moreover, if the this paper is on the right track, limiting our understanding of "grammar" to the sentence and its constituents, the long-time position of Chomsky

(1963 and many others) is severely misguided, for leaving out conversation, discourse, gesture, and myriad other components that make sentences possible.

Language is not an organ of the brain, at least there is no compelling reason to propose this given current knowledge. It is a cultural tool (Everett (2012)) evolving over time and altering its users as little as possible in the process.

References

Anderson, Michael L. 2014. *After Phrenology: Neural Reuse and the Interactive Brain*. MIT Press.

Arbib, Michael A. 2005. "From Monkey-Like Action Recognition to Human Language: An Evolutionary Framework for Neurolinguistics." *Behavioral and Brain Sciences* 28 (2): 105–24.

Arbib, Michael A. 2012. *How the Brain Got Language: The Mirror System Hypothesis*. Oxford University Press.

Barnard, Alan. 2012. Gene*sis of Symbolic Thought*. Cambridge University Press.

Bednarik, Robert G. 2013. "On the neuroscience of rock art interpretation. Time and Mind." *The Journal of Archaeology, Consciousness and Culture* 6(1): 37–40.

Bednarik, Robert G. 2008. "Beads and the origins of symbolism." Time and Mind: The Journal of Archaeology, Consciousness and Culture 1(3): 285-318.

Bednarik, Robert G. (2003). "Seafaring in the Pleistocene". *Cambridge Archaeological Journal*. **13** (1): 41–66.

Bednarik, Robert G. 1995. "Concept-mediated marking in the Lower Palaeolithic." **Current Anthropology**, 605-634.

Bedny, Marina, Hillary Richardson, and Rebecca Saxe. 2015. ""Visual" Cortex Responds to Spoken Language in Blind Children," **The Journal of Neuroscience**, August 19, 2015 • 35(33):11674 –11681.

Berent, Iris. 2013. *The Phonological Mind*. Cambridge: Cambridge University Press.

Berwick, Robert C. and Noam Chomsky. 2016. *Why Only Us?* MIT Press.

Bolhuis, Johan J., Martin Everaert, Robert Berwick, and Noam Chomsky. 2016. *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain.* MIT Press.

Boyd, Robert, and Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

Boyd, Robert, and Peter J. Richerson. 1988. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Brandom, Robert B. 1998. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Bybee, Joan L. 2010. *Language, Usage, and Cognition*. Cambridge University Press.

Cangelosi, Angelo. 2001. "Evolution of communication and language using signals, symbols, and words." *IEEE Transactions on Evolutionary Computation*, 5(2):93-101.

Chomsky, Noam. 2010a. “Minimal Recursion: Exploring the Prospects.” In *Proceedings of the Conference on Recursion: Structural Complexity in Language and Cognition*, edited by Margaret Speas and Tom Roeper. Amherst: University of Massachusetts. In press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origins, and Use*. New York: Praeger.

Chomsky, Noam. 1984. “On Language and Culture.” In *Contrasts: Soviet and American Thinkers Discuss the Future*, edited by Wiktor Osiatyński, 95–101. New York:

MacMillan.

Chomsky, Noam. 1972. *Language and Mind*. Enlarged edition.

Chomsky, Noam. 1963. "Formal properties of grammars." In: Luce - Bush - Galanter (eds) *Handbook of mathematical psychology*. Wiley, New York

Corballis, Michael C. 2007. "Recursion, Language, and Starlings." *Cognitive Science* 31 (4): 697–704.

Corballis, Michael C. 2002. *From Hand to Mouth*. Princeton, NJ: Princeton University Press.

Dediu, Dan and Steven C. Levinson. 2013. "On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences" Front. Psychol., 05 July 2013 | http://dx.doi.org/10.3389/fpsyg.2013.00397

De Ruiter, Jan P. and David Wilkins. 1998. "The synchronization of gesture and speech in Dutch and Arrernte (an Australian Aboriginal language)." In S. Santi, I. Guaïtella, C. Cavé and G. Konopczynski (eds.), Oralité et Gestualité, pp. 603-607. Paris: L'Hamattan.

Diller, Karl C. and Rebecca L. Cann. 2009. "Evidence against a genetic-based revolution in language 50,000 years ago." In: *The Cradle of Language*. R Botha, C Knight (eds). New York: Oxford University Press. pp. 135–149.

Dunbar, Robin. 1998. *Grooming, Gossip, and the Evolution of Language*. Harvard University Press.

Efron, David. 1941. *Gesture and Environment*. King's Crown Press. New York.

Evans, Vyvyan. 2015. "Beyond words: How language-like is emoji?" Oxford Dictionaries blog--20th November 2015

Everaert, Martin B.H. Marinus A. C. Huybregts, Noam Chomsky, Robert C. Berwick & Johan J. Bolhuis. 2015. "Structures, Not Strings: Linguistics as Part of the Cognitive Sciences." *Trends in Cognitive Sciences* 19 (12):729-743.

Everett, Caleb D. 2015. “Climate, Vocal Folds, and Tonal Languages.” (with D. Blasi and S. Roberts) *Proceedings of the National Academy of Sciences*.

Everett, Caleb D. 2014. “Evidence for Direct Geographic Influences on Linguistic Sounds: The Case of Ejectives.” *PLOS ONE* 8 (6): e65275. doi:10.1371/journal.pone.0065275

Everett, Daniel L. 1979. “Aspectos da Fonologia do Pirahã.” Master’s thesis, Universidade Estadual de Campinas. http://ling.auf.net/lingbuzz/001715.

Everett, Daniel L. 1982. “Phonetic Rarities in Pirahã.” *Journal of the International Phonetics Association* 12:94–96.

Everett, Daniel L. 1983. “A Lingua Pirahã e a Teoria da Sintaxe.” PhD diss., Universidade Estadual de Campinas. Published as *A Lingua Pirahã e a Teoria da Sintaxe*, Editora da UNICAMP, 1992.

Everett, Daniel L. 1985. “Syllable Weight, Sloppy Phonemes, and Channels in Pirahã Discourse.” *XI Annual Meeting Berkeley Linguistics Society, University of California Berkeley, Parasession on Metrical Phonology*.

Everett, Daniel L. 1986. “Pirahã.” In *Handbook of Amazonian Languages I*, edited by Desmond Derbyshire and Geoffrey Pullum, 200–326. Berlin: Mouton de Gruyter.

Everett, Daniel L. 1988. “On Metrical Constituent Structure in Pirahã Phonology.” *Natural Language and Linguistic Theory* 6: 207–46.

Everett, Daniel L. 1994. “The Sentential Divide in Language and Cognition: Pragmatics of Word Order Flexibility and Related Issues.” *Journal of Pragmatics and Cognition* 2 (1) : 131–66.

Everett, Daniel L. 2001. “Monolingual Field Research.” In Fieldlinguistics, edited by Paul Newman and Martha Ratliff, 166–88. Cambridge: Cambridge University Press.

Everett, Daniel L. 2004. “Coherent Fieldwork.” In *Linguistics today*, edited by P. van Sterkenberg, 141–62. Amsterdam: John Benjamins.

Everett, Daniel L. 2005a. “Periphrastic Pronouns in Wari.” *International Journal of American Linguistics* 71 (3): 303–26.

Everett, Daniel L. 2005b. “Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language.” *Current Anthropology* 76:621–46.

Everett, Daniel L. 2008. *Don’t Sleep, There Are Snakes: Life and Language in the Amazonian Jungle*. New York: Pantheon.

Everett, Daniel L. 2009a. “Wari’ Intentional State Construction Predicates.” In*Investigations of the Syntax-Semantics-Pragmatics Interface*, edited by : Robert Van Valin, 381–409. Amsterdam: John Benjamins.

Everett, Daniel L. 2009b. “Pirahã Culture and Grammar: A Response to Some Criticisms.” *Language* 85:405–42.

Everett, Daniel L. 2010a. “You Drink. You Drive. You Go to Jail. Where’s Recursion?” Paper presented at the 2009 University of Massachusetts Conference on Recursion. http://ling.auf.net/lingbuzz/001141.

Everett, Daniel L. 2010b. “The Shrinking Chomskyan Corner in Linguistics.” Response to the criticisms Nevins, Pesetsky, and Rodrigues raise against various papers of Everett on Pirahã's unusual features, published in *Language* 85 (3). http://ling.auf.net/lingbuzz/000994.

Everett, Daniel L. 2012a. *Language: The Cultural Tool*. New York: Pantheon Books.

Everett, Daniel L. 2012b. “What Does Pirahã Have to Teach Us about Human Language and the Mind?” *WIREs Cognitive Science*. doi:10.1002/wcs.1195.

Everett, Daniel L. 2013a. “A Reconsideration of the Reification of Linguistics.” Paper presented at The Cognitive Revolution, 60 Years at the British Academy, London.

Everett, Daniel L. 2013b. “The State of Whose Art?” Reply to Nick Enfield’s review of *Language: The Cultural Tool* in *Journal of the Royal Anthropological Institute* 19 (1).

Everett, Daniel L. 2014. “The Role of Culture in Language Emergence.” In *The Handbook of Language Emergence*, edited by Brian MacWhinney and William O’Grady, TK–TK. Hoboken, NJ: Wiley-Blackwell.

Everett, Daniel L. 2014. “Concentric Circles of Attachment in Pirahã: A Brief Survey” In *Different Faces of Attachment: Cultural Variations of a Universal Human Need*, edited by Heidi Keller and Hiltrud Otto, TK–TK. Cambridge: Cambridge University Press.

Everett, Daniel L. 2016. *Dark Matter of the Mind: The Culturally Articulated Unconscious*, University of Chicago Press.

Everett, Daniel L. 2017. *How Language Began* (working title). New York: Liveright-

W. W. Norton.

Everett, Daniel L., and Keren Everett. 1984. “On the relevance of Syllable Onsets to Stress Placement.” *Linguistic Inquiry* 15:705–11.

Fitch, W. Tecumseh. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.

Fitch, W. Tecumseh. 2008. Nano-intentionality: a defense of intrinsic intentionality. Biology & Philosophy March 2008, Volume 23, Issue 2, pp 157–177

Floyd, Simeon. 2016. "Modally hybrid grammar?: Celestial pointing for time-of-day reference in Nheengatú." **Language,** 92: 1, pp. 31-64 | 10.1353/lan.2016.0013

Freyd, Jennifer. 1983. "Shareability: The Social Psychology of Epistemology." **Cognitive Science** 7: 191-210.

Fuentes, Augustin. 2016. "The Extended Evolutionary Synthesis, Ethnography, and the Human Niche: Toward an Integrated Anthropology." **Current Anthropology** Volume 57, Supplement 13, June 2016

Futrell, Richard, Laura Stearns, Steven T. Piantadosi, Daniel L. Everett, and Edward Gibson. 2016. ‘A Corpus Investigation of Syntactic Embedding in Pirahã’, PLOS One. Co- authored by Richard Futrell, Laura Stearns, Steven T. Piantadosi, and Edward Gibson, March 2.

Gil, David. 1994. “The Structure of Riau Indonesian.” *Nordic Journal of Linguistics* 17:179–200.

Goldberg, Adele. 1995. *Constructions: A Construction Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Grice, Paul. 1991. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Harris, Marvin. 1999. *Cultural Anthropology*. Boston: Allyn and Bacon.

Harris, Marvin. 2001. *Cultural Materialism: The Struggle for a Science of Culture*. Walnut Creek, CA: AltaMira.

Harris, Zellig. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Hauser, Marc, Noam Chomsky, and Tecumseh Fitch. 2002. "The Faculty of Language: What Is It, Who Has It, How Did It Evolve?" *Science* 298: 569–1579.

Heckenberger, Michael J., J. Christian Russell, Carlos Fausto, Joshua R.Toney, Morgan J. Schmidt, Edithe Pereira, Bruna Franchetto, Afukaka Kuikuro. 2008. "Pre-Columbian Urbanism, Anthropogenic Landscapes, and the Future of the Amazon," *Science* 29 August 2008: Vol. 321. no. 5893, pp. 1214 – 1217, DOI: 10.1126/science.1159769.

Hickok, Gregory. 2014. The Myth of Mirror Neurons: The Real Neuroscience of Communication and Cogni"tion, W.W. Norton Hewes 1973

Hobbs, Jerry R., 2008. "Deep Lexical Semantics," *Proceedings*, 9th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2008), Haifa, Israel, February 2008.

Hockett, Charles. 1960. "The Origin of Language." *Scientific American*, 203, pp. 89–97.

Hopper, Paul. 1988. “Emergent Grammar and the A Priori Grammar Postulate.” In *Linguistics in Context: Connecting Observation and Understanding*, edited by Deborah Tannen. Ablex, New York.

Hurford, James R. 2011. *The Origins of Meaning: Language in the Light of Evolution*. Oxford University Press.

Jackendoff, Ray. 2003. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray, and Eva Wittenberg 2012. *Even Simpler Syntax*: https://depts.washington.edu/lingconf/abstracts/JackendoffandWittenberg.pdf.

Jackendoff, Ray, and Eva Wittenberg. in preparation. *A Hierarchy of Grammatical Complexity*.

Karlsson, Fred. 2009. "Origin and maintenance of clausal embedding complexity." In: Geoffrey Sampson, David Gil and Peter Trudgill, eds., *Language Complexity as an Evolving Variable*. Oxford University Press, pp192-202.

Keller, Timothy A. and Marcel Adam Just. 2009. "Altering Cortical Connectivity: Remediation-Induced Changes in the White Matter of Poor Readers." **Neuron**, Volume 64, Issue 5, p624–631, 10 December 2009.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kinsella, Anna R. 2009. *Language Evolution and Syntactic Theory*. Cambridge: Cambridge University Press.

Kirby, Simon, Mike Dowman, and Thomas L. Griffiths. 2007. "Innateness and culture

in the evolution of language." ***Proceedings of the National Academy of Sciences***. www.pnas.org_cgi_doi_10.1073_pnas.0608222104

Kirby, Simon, Hannah Cornish, aand Kenny Smith. 200 Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. ***Proceedings of the National Academy of Sciences.*** www.pnas.org_cgi_doi_10.1073_pnas.0707835105

Thompson, B. Kirby, S & Smith, K. 2016. "Culture shapes the evolution of cognition," *Proceedings of the National Academy of Sciences* of the United States of America - PNAS., 10.1073/pnas.1523631113

Labov, William. 2010. *Principles of Linguistic Change, Volume III, Cognitive and Cultural Factors*. Wiley-Blackwell.

D. Robert Ladd. 2012. "What *is* Duality of Patterning Anyway?" Language and Cognition 2012; 4(4): 261 – 273

LeDoux, Joseph. 2015. *Anxious: Using the Brain to Understand and Treat Fear and Anxiety*. Viking.

Levinson, Stephen C. 2006. "On the Human 'Interaction Engine.'" In *Roots of Human Sociality: Culture, Cognition, and Interaction*, edited by Nick J. Enfield and Stephen C. Levinson, 399–460. New York: Berg.

Levinson, Stephen C., and Asifa Majid. 2014. "Differential Ineffability and the Senses." *Mind and Language* 29 (4): 407–27.

Levinson, Stephen C., and Pierre Jaisson. 2005. *Evolution and Culture: A Fyssen Foundation Symposium*. Cambridge, MA: MIT Press.

Levinson, Stephen C. 2013. "Recursion in Pragmatics." *Language*, 89:1, pp. 149-162

Lieberman, Philip. 2013. *The Unpredictable Species: What Makes Humans Unique*. Princeton, NJ: Princeton University Press.

Lieberman, Philip. 2007. "The Evolution of Human Speech: Its Anatomical and Neural Bases." *Current Anthropology* 48 (1): 39–66.

Longacre, Robert. 1964. *Grammar discovery procedures: a field manual.* Janua Linguarum, series minor, Academic Training.

Luuk, Erkki. 2013. "The structure and evolution of symbols." *New Ideas in Psychology*, 31: 2, 87-97.

Luuk, Erkki, Hendrik Luuk. 2014. The evolution of syntax: Signs, concatenation and embedding. Cognitive Systems Research. doi:10.1016/j.cogsys.2013.01.00

Lyell, Charles. 1833. *Principles of Geology*. John Murray. London.

MacWhinney, Brian. 2004. "A Unified Model of Language Acquisition," In; *Handbook of Bilingualism: Psycholinguistic Approaches*, edited by J. Kroll and A. De Groot, TK–TK. Oxford: Oxford University Press.

MacWhinney, Brian. 2006. "Emergentism—Use Often and With Care." *Applied Linguistics* 27 (4): 729–40. doi:10.1093/applin/aml035.

MacWhinney, Brian and William O'Grady, eds. 2016. *The Handbook of Language Emergence*, Hoboken, NJ: Wiley-Blackwell.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.McNeill, David. 2005. Gesture and Thought. Chicago: University of Chicago Press.

McNeill, David. 2012. How Language Began: Gesture and Speech in Human Evolution. Cambridge: Cambridge University Press.

McNeill, David, ed. 2000. *Language and Gesture*. Cambridge: Cambridge University Press.

Mead, George Herbert. 2015. *Mind, Self, and Society.* University of Chicago Press.

Morgan. T.J.H, N. T. Uominic, L.E. Rendella, L. Chouinard-Thuly, S. E. Street, H. M. Lewis, C. P. Cross, C. Evans, R. Kearney, I. De la Torreg, A. Whitene, K.N. Lalanda. 2015. Experimental Evidence for the Co-Evolution 4 of Hominin Tool-Making Teaching and Language." NATURE COMMUNICATIONS | 6:6029 | DOI: 10.1038/ncomms7029 | www.nature.com/naturecommunications

Müller RA, Basho S. 2004. "Are nonlinguistic functions in "Broca's area" prerequisites for language acquisition? FMRI findings from an ontogenetic viewpoint." Brain Lang. 2004 May;89(2):329-36.

Panksepp, Jaak, and Lucy Biven. 2012. *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. New York: W. W. Norton.

Peirce, C. S. 1977. *Semiotics and Significs*. Edited by Charles Hardwick. Bloomington: Indiana University Press.

Peirce, C. S. 1992. *The Essential Peirce: Selected Philosophical Writings (1867–1893)*. Indiana University Press. Bloomington, IN.

Peirce, C. S. 1998. *The Essential Peirce, Volume 2: Selected Philosophical Writings, 1893–1913*. Indiana University Press. Bloomington, IN.

Pepperberg, Irene M. 2012. "Evolution of Communication and Language: Insights from Parrots and Songbirds." In: Tallerman, Maggie M. and Katherine R.K. Gibson. 2013. *The Oxford Handbook of Language Evolution*. Oxford University Press.

Piantadosi, Steven T., Harry Tily, and Edward Gibson. 2011. "The Communicative Function of Ambiguity in Language." **Cognition**: Volume 122, Issue 3, March 2012, Pages 280–291.

Pierrehumbert, Janet and Julia Hirschberg. 1990. "The meaning of intonational contours in the interpretation of discourse." In *Intentions in Communication,* edited by P.R. Cohen, J. Morgan and M.E. Pollock. Cambridge, MA, MIT Press: 271-311.

Pike, Kenneth L. 1967. *Language in Relation to a Unified Theory of the Structure of Human Behavior*. 2nd rev. ed. Janua Linguarum, series maior, 24. The Hague: Mouton.

Rizzolatti, Giacomo, and Michael A. Arbib. 1998. “Language within Our Grasp.” *Trends Neuroscience* 21:188–94.

Rosenbaum, David A. 2014. *It’s a Jungle in There: How Competition and Cooperation in the Brain Shape the Mind*. Oxford: Oxford University Press.

Safina, Carl. 2015. *Beyond Words: What Animals Think and Feel.* Henry Holt and Company.

Saussure, Ferdinand. 1983. *A Course in General Linguistics*. Open Court Publishing.

Searle, John. 1972. “Chomsky’s Revolution in Linguistics.” *New York Review of Books*, June 29, 16–24. [Reprinted in On Noam Chomsky: Critical essays, ed. by Gilbert Harman, 2–33. Garden City, NY: Anchor Books, 1974.1

Selkirk E. 1984. “On the Major Class Features and Syllable Theory.” *In Aronoff and Oehrle*.

Sereno, Martin I. 2014. "Origin of symbol-using systems: speech, but not sign,

without the semantic urge." *Philosophical Transactions of the Royal Society B* **369**(1651):20130303. (doi:10.1098/rstb.2013.0303

Shannon, Claude E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: 379–423, 623–56.

Silverstein, Michael. 2004. "Cultural Concepts and the Language-Culture Nexus." *Current Anthropology* 45 (5): 621–52.

Silverstein, Michael. 2003. "Indexical Order and the Dialectics of Sociolinguistic Life." *Language and Communication* 23:193–229.

Silverstein, Michael, and Greg Urban, eds. 1996. *Natural Histories of Discourse*. Chicago: University of Chicago Press.

Simon, Herbert A. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106(6): 467–82.

Slater, Peter. 2012. "Bird Song and Language." In: Tallerman, Maggie M. and Katherine R.K. Gibson. 2013. *The Oxford Handbook of Language Evolution*. Oxford University Press.

Sperber, Dan, and Deidre Wilson. 1996. *Relevance: Communication and Cognition*. Wiley-Blackwell.

Steedman, Mark. 2001. The Syntactic Process. Bradford Books, MIT Press.

Steels, L. 2005. "The Emergence and Evolution of Linguistic Structure: From Lexical to Grammatical Communication Systems." *Connection Science* 17:213–30.

Sterelny, Kim. 2008. *Thought in a Hostile World: The Evolution of Human Cognition*. Wiley Blackwell.

Sterelny, Kim. 2014. *The Evolved Apprentice: How Evolution Made Humans Unique*.

Bradford Books. Cambridge, MA.

Tallerman, Maggie M. and Katherine R.K. Gibson. 2013. *The Oxford Handbook of Language Evolution*. Oxford University Press.

Tattersall, Ian. 2013. *Masters of the Planet: The Search for Our Human Origins.* St. Martin's Griffin.

Thomason, Sarah Grey and Terrence Kaufman. 1992. *Language Contact, Creolization, and Genetic Linguistics*. California University Press.

Tomasello, Michael. 2001. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2010. *Origins of Human Communication*. Bradford Book, MIT Press.

Tomasello, Michael. 2014. *A Natural History of Human Thinking*. Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Harvard University Press.

Urban, Greg. 2010. "Metasignaling and Language Origins." **American Anthropologist**, New Series, Vol. 104, No. 1 (Mar., 2002), pp. 233-246 2010

Van Valin, Robert D., and Randy LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.

Vygotsky, Lev S. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Edited by Michael Cole. Harvard University Press. Weinrich

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. “Empirical

Foundations for a Theory of Language Change." In *Directions for Historical Linguistics*, edited by W. Lehmann and Y. Malkiel, 95–189. Austin: University of Texas Press.

Wilkins, David P. 1999. "Spatial Deixis in Arrernte Speech and Gesture: On the Analysis of a Species of Composite Signal as Used by a Central Australian Aboriginal Group." Paper 6 in *Proceedings of the Workshop on Deixis, Demonstration and Deictic Belief in Multimedia Contexts, held on occasion of ESSLI XI*, edited by Elisabeth André, Massimo Poesio, and Hannes Rieser, 31–45. Workshop held in the section 'Language and Computation' as part of the 'Eleventh European Summer School in Logic, Language and Information', August 9–20, 1999, Utrecht, The Netherlands.

Notes

[1] Searle's (1972) review in NYRBs of Chomsky's revolution.

[2] A secondary thesis is that *language was likely achieved by Homo erectus*, though this is not essential to the principal thesis and conclusion.

[3] "It is not likely that there was any single mutation causing the origin of language or even speech, as seen by the complex relationship between $FOXP_2$ and $CNTNAP_2$, and by the fact that $FOXP_2$ regulates several hundred genes, including many that are non-language related." (Diller and Cann (2012) p175ff)

[4] In section __ we consider an influential perspective on language as manifested as *I(nternal)-Language* vs. *E(xternal)-language*. I will argue that this distinction, while conceptually reasonable, has little impact on a theory of the evolution of language.

[5] If there could have been Homo sapiens without language 50,000 years ago, the hypothesized time of the "Merge leap" for Berwick and Chomsky, i.e. more or less when the ability to do recursion entered human minds/brains and languages, then there is no reason that there still could not pockets of humans without language, i.e. without recursive thought or expression. This seems like a strange prediction, but it should be easily enough to verify. Finding humans that not only lack but who are completely unable to understand or produce recursive language would be striking support for the UG/recursion theory of language origins.

[6] ***INDEXES****: Common to perhaps all animals. Thus indexes are extremely ancient, perhaps from the first single-celled creatures.*

ICONS*: At least three million years old.*

***SYMBOLS**: 250,000-540,000 years ago.*

There is no language which has meaning without convention. Yet conventions imply culture (Everett (2016)), conventions being general cultural agreements, e.g. word meanings. Finally, there is animal thought – if we say that language is structure-dependent and that it is a necessary condition for thought (as opposed to merely enhancing thought), then we are not only claiming that other hominins did not think, but that no other creature thinks, since other creatures lack the structure-building operation Chomsky proposes as the foundation of language, i.e. *Merge*.

In what follows we examine the arguments for and against the two main positions just mentioned and to develop a prolegomenon for a more empirical (and empiricist) view of language than is found in much of the formal linguistics literature (e.g. Berwick and Chomsky (2016); Everaert, et. al. (2016) and so on), a prolegomenon which, I hope, complements other empirical work (e.g. Tallerman & Gibson (2012); Thompson, et. al. (2016); Hurford (2011)).

[7] Tecumseh Fitch's 2008 paper on nano-intentionality offers an account of the rise of intentionality. There are many good ideas in this paper. But what it fails to explain is the role of culture in severely constraining what is attended to (see Everett 2016).

[8] On the other hand, both indexes and icons are likely still found in all languages.

[9] The remainder of this section borrows heavily, often verbatim, from Everett (2012).

[10] There is a great deal of discussion in the literature these days about birdsong and its possible connection to language (Slater 2012; Pepperberg 2012; Bolhuis, et. al. (2016). However, one thing that these discussions all miss is the contribution of culture to human language. As Kirby, et. al. (2008), Thompson, et. al. (2016) and Kirby, et. al. (2007), indicate, and as I argue in Everett (2012), birdsong lacks the component of culture, in the human sense (Everett 2016). And culture can, as Kirby and his co-authors point out, override innate biases, such as they might exist.

[11] "There is incomplete consensus on the anatomical demarcation of Broca's area in the left inferior frontal gyrus and its functional characterization remains a matter of debate. Exclusive syntactic specialization has been proposed, but is overall inconsistent with the neuroimaging literature. We examined three functional MRI (fMRI) datasets on lexicosemantic decision, tone discrimination, and visuomotor coordination for potential overlap of activation. A single site of convergent activation across all three paradigms was found in the left inferior frontal lobe (area 44/45). This result is discussed in the context of animal and human studies showing inferior frontal participation in visuomotor and audiomotor functions as well as working memory. We propose that Broca's area involvement in lexical semantics and syntax emerges from these nonlinguistic functions, which are prerequisites for language acquisition." Müller and Basho (2004), 329ff.

[12] White-matter is named for the white (because fatty) material (technically, myelin sheaths) that surrounds the nerve fibers connecting parts of the brain used for higher cognitive functions.

[13] A similar kind of study was published as a letter to the journal Nature in October of 2009, by a group of scientists showing that brain structure is affected by literacy and that literacy changes the anatomy of their brains.

[14] http://www.zmescience.com/science/archaeology/homo-erectus-shell-04122014/

[15] As Diller and Cann (2012: 175ff) put it "It is not likely that there was any single mutation causing the origin of language, or even speech, as seen by the complex relationship between FOXP2 and CNTNAP2and by the fact that FOXP2 regulates several hundred genes, including many that have non-language related functions ..."

[16] As we see in ___ below, gestures are also crucial to understanding how duality and compositionality happen.

[17] Unfortunately, there is no space in this paper available for a discussion of the evolution of speech. But see, inter alia, Everett 2017 for a discussion.

[18] Everett 2016 offers a sustained discussion of phonology related to Universal Grammar, and criticizes severely the notion that either sonority or phonology is an innate property of human minds. Selkirk (1984) proposes the following "sonority hierarchy": [a] > [e o] > [i u] > [r] > [l] > [m n ŋ] > [z v ð] > [s f θ] > [b d g] > [p t k]

But, as discussed in Everett (2016a and 2016b), this hierarchy is in fact of little use.

[19] Much of the following is drawn directly from Everett (2016).

[20] This section borrows heavily from Everett (2016).

[21] Some gesticulations, like language-slotted gestures and pantomimes, are not conventional – they may vary widely and have no societally fixed form (though they are culturally influenced).

[22] In a fascinating new paper, Floyd (2016) discusses a rich set of gestures and their implications for linguistic theory in the Nheengatu language of Brazil.

[23] As Sascha Griffiths pointed out to me (email May 04, 2014) the gesture/speech equiprimordiality hypothesis avoids the nagging question of how long after we had gesture languages the human larynx would have waited to descend (see Lieberman (2007)).

[24] Mirror neurons have been all the rage in recent years for helping to explain language evolution, but most accounts strike me as still too speculative to be of much use See Hickok (2014).

[25] Compositionality bubbles throughout the system, once we get culture. Icons, for example (as a referee for this journal rightly pointed out to me) are compositional, just as the bodies they represent are. It is possible, therefore, that Australopithecines saw iconic elements, such as the Makapansgat Pebble, as containing parts. But we have no evidence to support this at present. We do know that language-bearing modern humans see the compositionality in icons, but culture is crucial for any animal to make much of the observation that wholes have parts.

[26] It is also worth noting that if McNeill's story is on the right track, the contribution of Merge to the evolution of language was neither necessary nor sufficient.

[27] Interesting new work by Floyd (2016) on Nheengatu gestures underscores the triple articulation of language, as well as the purpose of language for communication.

[28] I do not believe that sign languages are the same as phonologies, though they arise when phonologies are unavailable (e.g. deafness or lack of articulatory ability) or when other cultural values render gestures preferable. See above for a discussion of gestures. Since gestures are related to the eyes rather than the ears, their organizing principles are different in some ways. Of course, because both phonological and gestural languages are designed by cultures and the minds, subject to similar constraints of computational utility, they will also both share features, as is often observed in the literature.

[29] There is some overlap between the account presented here and an independently developed set of proposals by Luuk and Luuk (2013), a very important article in the literature on the evolution of syntax. Like the account below, Luuk and Luuk argue that syntax develops initially from the concatenation of signs, moving then from mere concatenation to embedding grammars. My differences with their proposals, however, are many. For one thing, they seem to believe, which is common enough, that compositionality depends on syntactic structure, failing to recognize that semantic compositionality is facilitated by syntactic structure, but not dependent on it in all languages. Further, they fail to recognize the *triality of patterning in a cultural context* and that modern languages do not need embedding. They do seem to embrace the view of Everett (2012) that language is a cultural tool, though they do not cite that work.

[30] Futrell, et. al. (2016) argue that there exist modern human languages that come in lower in the Chomsky hierarchy than Chomsky would have predicted.

[31] Karlsson (2009, p192-193) explains his notation as: *"'I' stands for initial clausal embedding, 'C' for clausal center-embedding, 'F' for final clausal embedding, and the raised exponent expresses the maximal degree of embedding of a sentence, e. g. I2 is double initial embedding as in sentence (6). Expressions like C-2 indicate type and embedding depth of individual clauses; e. g. C-2 is a center-embedded clause at depth 2."*

[32] Communication – which precedes language – is about as old as life. Communication is found in every species of the animal kingdom. I define it as the successful exchange of information.

[33] I certainly accept the idea that gorillas and other creatures have been taught to use symbols in the lab.

[34] http://www.nybooks.com/articles/1972/06/29/a-special-supplement-chomskys-revolution-in-lingui/

[35] Though the authors use the term "recursion" in their 2002 article, they really do not mean recursion as understood by people doing research outside of Chomsky's Minimalism, but they actually intend Merge, as discussed in the text. This has caused tremendous confusion, though, ultimately, the issues do not change and Merge has been falsified in several modern grammars (see Everett 2012, among many others).

[36] This is circular in the sense that Chomsky takes a feature that only humans are known to have, structure-dependency, and claims that this defines language so therefore only humans have language.

[37] Chomsky often mentions favorably Wallace's well-known skepticism that natural selection could possibly explain the mental life of humans. And yet, this is unfortunate, because Wallace had strong theistic views that manifest themselves from time to time. Wallace had no scientific reason to reject natural explanations for the mental. He had only a religious bias. It is unfortunate that Chomsky choose to perpetuate this.