

To Appear in Peter Kosta and Katrin Schlund (Eds.)  
*Linguistic Explanations-CEP 2019*

## On Sound-Meaning Correlation

Nirmalangshu Mukherji  
Former Professor of Philosophy  
Delhi University

The intriguing fact about human language is that neither the sound systems nor the thought systems--technically called Sensorimotor and Conceptual-intentional systems--have any significant precedence prior to the emergence of the hominid line. So, a theory of language needs to explain where they came from and how they fell into place to create the conditions for unbounded generativity. In effect, this requirement suggests two novel factors in evolution of human language: word-formation and word-combination. In this context, the recent proposal that the sound component of language is 'ancillary' and that language is basically designed for thought turns the origin of the rich structure of human thought into a mystery since the units of human thought could not have been borrowed from, say, the apes. To look for an alternative explanation, we study Darwin's idea that (a) the sound system might have evolved in part from a pre-existing human music system, (b) the evolving sound systems might have 'acted on' rudimentary units of thought systems to turn them into words. We critically enumerate the validity of Darwin's picture.

## On Sound-Meaning Correlation

It is generally recognised that a comprehensive theory of human languages comprises of three basic questions: (1) what is it, (2) how it is acquired, and (3) how did it evolve (Jenkins 2000). I am setting aside arguably subsidiary questions like (4) what is its neural architecture, and (5) how is it put to use in a population. I am setting (4) aside because the neural basis of language is just an assumption which parasitically depends on answers to (1)-(3); that is, unless we know what language is and how it evolved, we do not know what to look for in the brain (Mukherji Forthcoming). Nothing really is lost in linguistic explanation if we do not have accompanying neural descriptions, if any.

I am also setting (5) aside because language is put to use in a population primarily through sound-meaning correlations, which is the topic for this essay. I do not wish to prejudge the issue at this stage. As we will see, whether the sound component of language is essential for linguistic explanation is currently a contentious issue. We will see that, for some authors, a single person without recourse to the sound system is enough to put language to use for constructing 'inner' thoughts. Notice that, if this view holds, then the discipline of linguistic typology basically disappears or it becomes a discipline of marginal interest, as we will shortly see.

As with the rest of the sciences, especially for the biological sciences, it is plausible to hold that we must reach some stable answer to (1) regarding the nature of human language before we can begin to address the other questions. In that sense, once we have answered (1) in a certain way, some consequences for the other issues follow. For example, at a stage of theory 1, say, suppose we have a corresponding evolutionary picture which also appears to be plausible on independent grounds such

as archaeological records. Now, if there is a major change from theory 1 to theory 2, we may ask whether the changed theory fits the archaeological picture already established, or whether a corresponding change in evolutionary theory is also motivated.

I suggest that such a situation of possible theoretical conflict may be emerging regarding the issue of sound-meaning correlation. In particular, I argue that the recent proposal by Robert Berwick and Noam Chomsky (2016) that the sound component of the design of language is ‘ancillary’ may be questioned on evolutionary grounds. For the modest purposes of this paper, I will keep only to the suggested form of challenge from evolutionary theory; it is not ruled out that the Berwick-Chomsky thesis may be questioned from other directions also, as we briefly indicate near the end of the paper.

### **Mystery of language**

The topic of evolution of human language is particularly intriguing because human languages are unique biological objects in nature. No other nonhuman species is endowed with anything comparable to human language. No doubt, there are some unique features in, say, human limbs and eyes as well. However, it is safe to proceed on the assumption that human limbs and eyes are modifications of biological forms found elsewhere in nature. In fact, it is well-known that the human visual system is one of only five visual systems that have evolved in nature. More specifically, the human three-cone photoreceptors are common among apes and monkeys.

In contrast, the language system is not replicated anywhere else in nature. Hence, a theory of language needs to explain which components of the language system are truly unique and how they came about. To see the significance of this requirement for language-theory, consider a recent proposal by Berwick and Chomsky (2016, 2). According to them, the ‘basic property’ of language is that it is a ‘finite computational

system yielding an infinity of expressions, each of which has a definite interpretation in semantic-pragmatic and sensorimotor systems (informally, thought and sound)'. This description of the basic property of language has two parts: computational system with unbounded generativity, and the two systems of interpretation.

Arguably, there is little novelty in the first part: computational structures by themselves are found all over nature in the structure of galaxies, continental shelves, sunflowers, zebra stripes and the like (Stewart 1995, Flake 1998, Stewart 2001, Carroll 2005). Even for cognitive systems, even though unbounded generativity is not found anywhere else in the organic world,<sup>1</sup> many human domains seem to have very similar generative power. The human number system is an immediate case in point (Chomsky 2005, Hauser 2008); other human generative systems include the system of music (Katz and Pesetsky 2009, Mukherji 2010, Asano and Boeckx 2015) and the human kinship system (Hale 1966, Jones 2010). Authors have even argued for generativity in early hominid tool-making (Stout 2011, Uomini and Mayer 2013, Joordens et al. 2015, Brooks et al. 2018). In fact, the human form of generativity could well be a general property of human cognition marking the real distinction between human and nonhuman mental systems (Mukherji 2009, Miyagawa et al. 2018, Mukherji Forthcoming).

---

<sup>1</sup> I am setting aside false claims of generativity or computability (including claims about the presence of the operation External Merge in some cases) in animal studies involving insects (Gallistel 1998), bird-songs (Miyagawa 2017), capuchin monkeys (McGonigle et al. 2003), baboons (Seyfarth et al. 2005), Japanese macaques (Schino et al. 2006), chimpanzees (Fujita 2014) etc. The issue is fully discussed in Mukherji (forthcoming).

I mentioned that there are two parts to the Berwick-Chomsky characterisation of language. The real novelty of human language is that its unbounded generativity is geared to interpretation of sound and thought. Berwick and Chomsky only mention the output of the system, that is, the structures generated for interpretation at the two interfaces; they do not mention the input to the system. However, a theory of the specific computational system of human language must explain where the resources for the inputs to the system are coming from such that unbounded structures are generated for the interpretive interfaces of sound and meaning.

What is truly unique about human language is its specific interpretive character: the language system consists of units of sound and units of thought or meaning that are correlated to produce what is commonly called a *word*; words are units of sound-meaning correlation.<sup>2</sup> The computational system of human language implements a combinatorial operation—currently called *Merge*—that puts words together in an unbounded fashion. Even if we think of snowflakes and coastal lines as computational structures, the units for such computations have no parallels with the units of human

---

<sup>2</sup> I am using the notion of a word informally, following common usage. In syntactic theory, what is informally called a ‘word’ is viewed as consisting of phonological, syntactic and semantic features—features for sound, form and meaning respectively. Syntactic features drive computation to generate two structures for sound and meaning at the interfaces. Hence, there is some dispute as to which feature is activated when and where; in Distributed Morphology, semantic features are ‘inserted’ last. Since I am not concerned with computation itself, but only with the input and output of the system, the focus is on features of sound and meaning. In any case, these features need to be assembled from the lexicon. This requires some combinatorial operation. Hence, the word itself is the unit of sound-meaning correlation.

language; neither do the units for music, numbers, kinship, tool-making and the like have parallels with language. In effect, the language system has two aspects of wonder, word-formation and word-combination. So, the big question for the language-theorist is, where are these things coming from?

We have already noted that no nonhuman animal displays generativity. It turns out that no other animal has anything like words as well. Animals do have impressive call systems; some species of birds display pretty complex ‘songs’. However, whatever be the complexity of bird-songs of a species, they are given to the species at once, that is, they are fixed for the species. A given species of birds cannot generate new songs precisely because there are no word-like units in these songs; arguably, there are some ‘syllabic’ units at best. Moreover, there is no evidence that bird-songs correlate with anything in the sense in which human words display sound-meaning correlation. At best, bird-songs relate to some contexts, such as presence of a mate, as a whole (Berwick et al. 2011, Hauser et al. 2014). In contrast, some animal call systems do have sounds/signals that seem to have a ‘referential’ connection with items in the world such as prey, predator, mate, food etc. However, again each call and its ‘referential’ function is fixed for the species and there are just a handful of them. So, these are at best a very finite system of sound-stimulus associations.

The human collection of words, called the *lexicon*, is radically different from the nonhuman systems in two fundamental ways. First, although humans do talk about the world in a manner, it is not achieved, contrary to Quine (1960), via direct association between sounds and some stimulus-condition. As Noam Chomsky (2000), and others such as Mukherji (2010), have extensively argued, notions like reference and

denotation do not apply to the human lexicon, even for the so-called proper names.<sup>3</sup> Human ‘reference’ is an extremely complex process that is activated via a vast conceptual network; the network consists of a large store of concepts that are organized at various levels of generality and access. In fact, recent work suggests that large parts of the human referential function are driven by syntax itself (Hinzen 2007, Hinzen and Sheehan 2013); hence, such modes of reference are not available to animals.

Unsurprisingly, very little is known about animal conceptual systems, if any. However, it is most unlikely that animal conceptual systems entered human conceptual system even in part during hominid evolution. For example, it is unclear if chimpanzees, the ‘smartest’ of animals, have any concepts at all that eventually entered into human symbolization. Berwick et al. (2013, 92-3) report Laura Ann Petito’s work to suggest that chimpanzees fail to use anything like the human concept of apple with the label *apple*. It seems to follow that they do not really have the concept of apple the way humans do, apples are just associated with a variety of stimulus-conditions; it could be that they do not have human name-like kind-concepts at all. This result falls in place with Daniel Povinelli’s classic work suggesting that apes cannot demonstrate individual objects (Povinelli 2000). The upshot of this discussion is that the units of human thought—concepts—were most probably not borrowed from non-hominid sources; that is, human concepts have likely to have evolved in the post-chimpanzee hominid line itself.

---

<sup>3</sup> This is not to deny that humans do refer to talk about the world; it is a complicated process some of which is linguistically driven as noted below. The point is, human words, even proper names like *London*, *Joe Six-pack*, *John Doe*, do not have a direct association with stimulus which is the case with animals.

The other fundamental way in which human language differs from animal call systems is even more dramatic. Unfortunately, this aspect of human language is seldom stressed in the biolinguistics literature. The remarkable fact is that human sound-meaning correlations are effectively unbounded themselves, and that they are not fixed for the species. There are thousands of languages and dialects, each with over hundred thousand words in its lexicon. Except for some rare overlap due to historical and geographical reasons, these sound systems are essentially independent of each other; it is hard to locate any commonality in the sound between, say, Khoisan, Pashtu, Bangla, Warlpiri, and Mandarin. No doubt, there are some universal features in the basic phonemic elements of these systems, but the variety and complexity built out of these elements is astronomical. Basically then, human systems of sound-meaning correlations were independently invented and reinvented thousands of times across the world irrespective of geographical location, racial characteristics, and the like.

In fact, there is another related problem with animal call systems which is seldom directly addressed in the biolinguistics literature. Let us call it the *asymmetry problem*. We saw that there is nothing like the human sound-systems in the nonhuman world. However, there are plenty of animal species with rich sound-systems of their own (Hauser 1996), birds and insects for example. Almost invariably in such cases rich conceptual resources are absent. For the great apes like chimpanzees and gorillas, in contrast, there is evidence of rich systems of observation, planning, contemplation and the like, but these are essentially silent creatures (Fitch 2010). Humans are the only species in which rich sound-systems correlate with rich conceptual systems. It is thus reasonable to conjecture that human sound and thought systems are intimately related in their origin and function. I return to the point in the concluding section.



The most interesting aspect of the human lexical items is that somehow two totally disjoint things—sound/gesture and meaning/object—get constantly correlated. In the animal call systems too, the sound, such as the alarm call of a monkey, is arbitrary in that the acoustic shape of the call has no connection with the properties of a tiger. Yet, as the system is (perhaps genetically) fixed for the species, the presence of a tiger does causally trigger the call.

In the human case, the remarkable multiplicity of languages rules out such a ‘genetic’ explanation. As with the animals, the human sound /*dog*/ also has nothing to do with the properties of a dog, not to mention the picture/image of dogs. However, the telling difference from animals is that what is called *dog* in English, is called *kukur* in Bangla, *Kutta* in Hindi, *chien* in French, *maliki* in Warlpiri, *lééchq̣q̣’i* in Navajo, *gǒu* in Mandarin, and so on, in thousands of different variations. All dogs. In the philosophical literature, the relation between sound and meaning is thus viewed as ‘conventional’, but that just labels the problem.

As an aside, let me note that these extraordinary differences in sound between languages is the ground for the discipline of Linguistic Typology. These differences in fact *identify* individual languages for linguistic investigation. The remarkable fact is that there is nothing ‘genetic’, or fixed for individual populations, in this astonishing variety because any human child can pick up any of these languages or any collection of them within the normal maturational cycle. Therefore, in an abstract sense, languages are basically identical insofar as they follow from the human make-up.

For this reason, some biolinguists like Noam Chomsky suggest that the differences between languages are ‘superficial’, languages differ only in morphology. For Chomsky, there is in fact one human language with one lexicon (Chomsky 1993), or at most, ‘there is limited lexical variety’ (Chomsky 1994). Chomsky and others

obviously cannot mean that all languages sound the same; the claim is rather that languages do not vary significantly in syntax and semantics. But this claim simply assumes without argument that human languagehood is to be viewed as essentially represented in its syntax and semantics, not in the sound systems. Yet, once thousands of languages and dialects are identified in terms of how they sound, it becomes a serious research question whether, where and how much languages do in fact share syntax and semantics. That is precisely the domain of inquiry of linguistic typology.

In any case, it is obvious that the morphological differences between languages are neither superficial nor restricted to a limited variety. The extremely complex phenomenon of human sound systems demands systematic explanation, perhaps in tandem with an explanation of human thought, as noted. Research on animal sound-recognition shows that some animals do recognize some of the basic phonemes of human languages (Hauser 2001), some animals are cued to human rhythmic and prosodic patterns (Ramus et al. 2000), and of course parrots etc. can imitate human sounds. But no animal can produce the complex sounds that enter into the formation of words. We may conclude that neither the sounds nor the meanings of human lexicon have a non-hominid origin. Recall that we do not have an explanation of generativity of word-combination unless we have an explanation of word-formation.

### **Evolution of language (Chomsky)**

The preceding considerations on the nature of human language seems corroborated by archaeological record. The leading archaeologist Ian Tattersall (2016) points out that modern humans display two fundamentally novel abilities in connection with the general ability to use language: the ability to ‘process information symbolically’ and the ability to express ideas by ‘using structured articulate language’. Tattersall (2012) held that symbolic reasoning and articulate structuring was absent even in the species

that ‘looked exactly like us’. To emphasize, the emergence of human language required two broad conditions: *symbols* and *structures*. According to Tattersall (2019), these two conditions are related: ‘the elements of symbolic thought map closely onto the vocabularies of words’ to form ‘linguistic building-blocks’ that enter into ‘symbolic mental operations’. Since these conditions are absolutely basic to any conception of language, let us see what narrative is needed to install these conditions.

Instead of explaining how the human-specific sound and thought systems came about, Chomsky simply asserts, in what he calls a ‘fairy tale’, that ‘there were primates with pretty much our sensorimotor (SM) and conceptual-intentional (CI) systems, but no language faculty, and some natural event took place that brought about a mutation that installed a language faculty’ (Chomsky and McGilvray 2012, p.14-5). Clearly, by ‘language faculty’ Chomsky basically means the generative system that constructs interpretive structures for the two interfaces—the basic property. In effect, Chomsky assumes what needs to be explained: how did the human-specific SM and CI systems become available in the first place for the subsequent insertion of the ‘language faculty’?

In this connection, it is a puzzle for me that the archaeologist Ian Tattersall, who emphatically pointed out the two novelties of word-making and word-combination in the evolution of human language, appears to approve of Chomsky’s picture. Thus, he suggests that the evolution of human language was ‘underpinned by an algorithmically simple interface between sensorimotor and conceptual-intentional systems that were co-opted from pre-existing functions or potentials’ (Tattersall 2019). Like Chomsky, Tattersall didn’t explain which ‘pre-existing potentials’ led to the SM and CI systems.

In any case, it is obvious that at least one ‘pre-existing’ function could not have been directly ‘co-opted’ from our last non-hominid ancestor, the chimpanzee, because these great apes are strikingly silent creatures (Fitch 2010). So, at least the sensorimotor part—the sound part, in particular—needs to be co-opted either from other post-chimpanzee progenitors in the hominid line, or from chirping birds and tweeting insects from remote evolutionary branches. While the appeal to progenitors just pushes the problem backwards to the earlier progenitors themselves, the appeal to birds and insects requires incredible evolutionary narratives linking humans with songbirds, notwithstanding interesting evidence from nonhuman vocal learning activated by a few hundred genes (Fitch 2010; Miyagawa et al. 2014; Berwick and Chomsky 2016). After all, nonhuman vocal learning just suggests how patterns of sound are formed in a species and, let us grant, there could have been ‘deep convergence’ for that purpose between birds, humans and other noisy species. Nonetheless, the suggestion by itself says nothing about the specific character and complexity of human sound-systems and how they became correlated with the CI systems.

Furthermore, given the serious limitations of the conceptual repertoire of chimpanzees, as noted, it is extremely unlikely that human CI systems were co-opted from some pre-existing nonhuman function. So, they must have evolved largely during hominid evolution itself prior to the emergence of humans, thereby solving the asymmetry problem specifically in the human case. I am unable to locate any significant response from Chomsky and his colleagues on how ape-limitations were

overcome, even though Chomsky has repeatedly asserted the vast differences between animal and human CI systems.<sup>4</sup>

Insofar as the sound side of the asymmetry problem is concerned, Chomsky's response is to set the problem aside for explaining the origin of language because, according to Chomsky and his colleagues, sound is ancillary, it is an evolutionary afterthought for communication (only). As Marc Hauser (2009) puts it, 'language evolved for internal thought and planning and only later was co-opted for communication, so this sets up a dissociation between what we do with the internal computation as opposed to what the internal computation actually evolved for'. It is important to note that neither Chomsky nor Hauser provides any archaeological or genetic evidence for the claim that 'internal computation' for thought evolved *earlier*, and was 'co-opted for communication' *later*.

As an important aside, notice that the 'sound-is-ancillary' hypothesis (SAH), if valid, casts doubt on the linguistic significance of much evolutionary evidence focused on the anatomy of speech. Speech is something that can be proxy studied by fossil evidence of growth in anatomical structures related to vocal abilities, as well as by direct behavioral evidence of vocalization in extant species (Tattersall 2019). In fact, as Miyagawa et al. (2018) suggest, 'the often-stated idea that "language does not

---

<sup>4</sup> Most recently, Chomsky (pc) writes: 'There seem to be sharp differences between human concepts and anything in the animal world. The components of animal systems seem to have a one-one association with identifiable "mind"-external events: rustling of the leaves elicits a "warning cry." Hormonal changes elicit some vocal noise. For this and many other reasons human concepts seem to have evolved independently of animal concepts.' I am citing this correspondence at length because Chomsky has written these things in published material for many years.

fossilize” is not quite true: pieces of *externalized* language may turn out to be hidden among the art forms produced by our early modern human ancestors’ (emphasis added). Perhaps this is the reason why, as we will now see, Darwin the scientist urged and made an attempt to derive as much as he could from such slender and indirect evidence.

### **Evolution of language (Darwin)**

Charles Darwin suggested a very different narrative to account for the evolution of language. Darwin’s story is almost entirely speculative and based on common sense since whatever knowledge we now have about hominid evolution, architecture of the brain, structure of language etc. were obtained much after Darwin’s time. Yet, it is of great interest that Darwin’s narrative broadly matches the archaeological picture sketched above. In a famous passage in his *Descent of Man*, Darwin (1871, 57) conjectured about an ‘early progenitor of man’ as follows.

As the voice was used more and more, the vocal organs would have been strengthened and perfected through the principle of the inherited effects of use; and this would have reacted on the power of speech. But the relation between the continued use of language and the development of the brain has no doubt been far more important. The mental powers in some early *progenitor of man* must have been more *highly developed than in any existing ape, before even the most imperfect form of speech* have come into use; but we may confidently believe that the continued use and advancement of this power would have reacted on the mind itself, by enabling and encouraging it to carry on long trains of thought. A long and complex train of thought can no more be carried on without the aid of words, whether spoken or silent, than a long calculation without the use of figures or algebra. (Emphasis added)

There is much discussion of this passage in the recent literature on evolution of language. For example, Fitch (2010, 472-4) also engages in a detailed analysis of this seminal passage. However, Fitch also adds sundry recent ideas to Darwin's story: vocal learning in animals such as songbirds, gestures of apes, phonological properties of speech, cultural distribution of music, and the like. Fitch needs all this to prepare the basis for his theory of 'musical protolanguage'. In my view, in doing so, he missed Darwin's sharp and focused narrative in the passage. Nothing in the passage itself suggests anything about sexual selection, animal vocalisation, musical protolanguage and other irrelevant topics.

In this very carefully thought out passage, Darwin attempted a thought experiment on how the different elements that are needed for the eventual emergence of language might have been harnessed from within the rather restricted resources available in the hominid line. More specifically, Darwin adopts something like the familiar Aristotelian view of language as a system of sound-meaning correlations to show how the elements of speech and thought developed to get correlated in a hypothetical evolutionary narrative.

The first step of the narrative raises the obvious need for postulating some *post-ape but pre-linguistic mental powers* for explaining the evolution of thought; human language and thought could not have directly emerged from the chimpanzee brain. It needed a much larger brain size and significant evolutionary time to attain the powers of a species that eventually spawned structured thought. A significant property of the initial mental power was that it was independent of and perhaps prior to speech. Beyond this, Darwin did not say what constitutes these powers. However, his subsequent use of the expressions *mind* and *long train of thought* suggests that he viewed these mental powers as closely related to thought. Suppose then by *mental*

*powers* Darwin meant a rich conceptual repertoire that must have been vastly more developed than the last apes to provide for ‘words’, as we will see.

The second step of the narrative suggests that speech—that is, *language* in Darwin’s terminology—was already developing independently of mental powers as a result of the strengthening of vocal cords. It is obvious that, since mental powers were thought to be already in place outside the realm of speech, Darwin could not have thought of the strengthened vocal cords as *mental* power. His only requirement was that strong vocal cords laid the basis for continued use of speech with subsequent effects on the mind. Needless to say, continued use of speech required strong vocal cords, just as the ability to play cricket requires strong arms, but having strong arms does not guarantee a cricketing mind. Unfortunately, in the passage, Darwin did not say where speech came from; he only said that the vocal cords grew stronger over time due to singing. Obviously, he assumed that singing (= music) was already in place; in fact, for Darwin it must have been going on for a while independently of both speech and mental powers. The assumption might have led recent authors to launch their own fables about musical protolanguage and Caruso theories; I am setting these proposals aside (But see Mukherji, Forthcoming).

Anyhow, in the third step of the narrative, Darwin stated that, at some point during the growth of the brain accompanied by the continued use of speech that ‘acted’ on mental powers, ‘words’ occurred, both spoken and silent. The mention of ‘silent’ words suggests that Darwin did not identify speech entirely with articulation, or what is now called ‘externalization’. In that, his idea of speech was akin to Descartes’ idea of sign which Descartes thought may occur in humans even without the ‘organ’ of



speech.<sup>5</sup> It is plausible to suppose that both Descartes and Darwin wished to emphasise the *symbolic* character of words. In due course, these words contributed to long trains of thought just as long calculations require the aid of figures and algebra. Except for saying that somehow words resulted out of interaction between speech and mental powers, Darwin did not elaborate on what constituted the interaction. He also did not give any hint of how the ‘algebra’ for ‘long trains of thought’ evolved when the words became available. There are other infirmities in the narrative which I set aside for now.

So, the fable is that a progenitor of man started out with richer conceptual resources than the apes due to larger brain size. The progenitor was also endowed with strong vocal tracts that grew out of prolonged singing over generations, an inherited trait according to Darwin. The strong vocal cords led to speech which also grew as the brain grew. As speech grew, conceptual growth and growth of speech converged to give rise to words which eventually contributed to long trains of thought ‘algebraically’, so to speak.

Darwin’s fable needs at least two distinct phases, perhaps separated by evolutionary time, for distinct operations: speech acting on mind to form words, and

---

<sup>5</sup> The suggested effect of speech on thought is very different from what is known as *inner speech* (Moro 2016, Chapter 8). Moro makes the interesting but not entirely surprising claim that a ‘sound envelope’ accompanies the process of thinking even when someone is reading something mentally, in silence. Once words are learned as units of sound-meaning correlations, inner images of sound are likely to accompany words even in writing. The phenomenon is like the reading of music from written scores. In my view, Descartes and Darwin were making the more fundamental point that sound enters the very formation of units of thought.

algebraic construction of train of thoughts. For the origin of symbols, Darwin did not explicitly mention CI systems, but he did require that some mental powers had developed much beyond apes. Needless to say, everyone needs to explain how, beginning with the limited conceptual repertoire of chimpanzees, the mental powers of immediate ancestors of humans reached near-human proportions. I touch on this critical point in the concluding section.

So, what did Darwin assume when he remarked that, several million years after the departure from the primate line, his progenitor of man had rich mental powers? Suppose, with advanced sensory systems, increased memory, emergence of digital manipulation involving demonstration, ability for imitation and vocal learning and the like, some of the growth in mental powers amounted to growth in individual stock of concepts: RED, WOOD, RIPE, BANANA, HOLE, WATER, etc. I am following the standard convention of mentioning concepts in capitals, sounds in slashes, and mention of words in italics. Of course, people like Berwick, Chomsky and Hauser, who adopt SAH to deal only in concepts, need to explain how the concept of banana, namely BANANA, gets to have the shape BANANA; I am just assuming that it happens.

Darwin seems to have indirectly appealed to such an operation of labelling concepts *via symbols* with his suggestion that the power of speech ‘acted’ on the mental power to create ‘words’. Suppose this simply means that the sound /banana/ was linked up with the concept BANANA, whatever it is. In effect, it means that a combinatorial operation forms a pair-list (/banana/, BANANA), represented by the word *banana*. In other words, the combinatorial operation enables humans, and *only* humans, to symbolize a concept with the mark of a sound; needless to say, sound is not the only marker. In this way, the ‘sign’ signals the presence of ‘hidden’ thought,

as Descartes put it. I wish to stress the point that symbolization (=marking) requires a combinatorial operation which is a uniquely human endowment since no other animal displays symbols, as Tattersall emphasised.

With the introduction of words, Darwin's narrative satisfied the condition of availability of symbols because one of the ways of marshalling a symbol is to form words. Once words are available, initial 'algebraic' operations can take place far more efficiently between words themselves, rather than between individual concepts: 'red-wood', 'water-hole', etc.; in fact, it is totally unclear what is meant for combinatorial or other operations to take place on concepts themselves. Given the symbolic resource offered by words, it is plausible to suppose that the resource would have had significant effect on the growth of mental powers as Darwin indicated: more sound, more meaning; more meaning, more sound. The shift to words endowed the species with radically enhanced mental powers.

Having said that the power of speech grew with continued use, Darwin never explained where the system of speech came from except suggesting that strong vocal cords were already available from singing. So, Darwin assumed that singing (= music) was already available prior to words. Even a rudimentary form of singing requires putting together individual units of sound in some order, thus requiring another combinatorial operation.

Further, it is implausible that units of music—tones—directly go into the making of speech except in a very limited way in rare cases. Therefore, the units of speech—phonemes—must have been picked up from some other resource, and put together to form signs that entered into formation of words: *ba-na-na*, *ri-pe*; this requires another combinatorial operation. As combinatorial operations are beginning to come out of our ears, let us suppose that eventually these operations also cover the formation of

inflected words such as *un-ripe*, *banana-s*.<sup>6</sup> Given an existing stock of morphemes, a finite stock of inflections enables further flexibility in human mental powers. For example, for English speakers, a single inflection *-s* for pluralisation classifies all countable objects into individuals and collections, while the lexical store remains virtually unchanged (Chomsky et al. 2002). Once words were available, Darwin simply assumed that they helped in forming long trains of thought in an ‘algebraic’ fashion. Darwin never told us what that critical algebra was.

As we know, it was Noam Chomsky’s work which finally clarified how the required algebra worked in the human case to generate unbounded structures from finite resources. As far as we know, the specific algebra implemented as the ‘basic property’ of language was made available all at once as a saltation, since there is not even a partial analogue of it elsewhere in the organic world. As Tattersall (2019) put it, archaeological evidence also suggests that the ‘language as we recognize it today originated suddenly, at some definable point in the human past’.

However, in the light of what we saw, it is unclear when that ‘definable point’ happened. It might not have been the case, as Berwick and Chomsky suggest, that the saltation occurred only after ‘pretty much’ like the human SM and CI systems were

---

<sup>6</sup> An intriguing issue is whether different combinatorial operations are at work or whether a single powerful operation accomplishes all combinatorial operations across various domains. Parsimony suggests that it is the latter. Since we need the powerful operation Merge for combining words in an unbounded manner in any case for generating the structures at the interfaces, could it be that Merge itself executes both bounded and unbounded operations. The discussion of this issue is beyond the scope of the paper, but see Uriagareka (2008) for a similar suggestion from a different direction in terms of the Chomsky hierarchy.

already in place. Given that the evolution of human language uniquely needed a variety of combinatorial operations to first achieve word-formation itself, it is perhaps parsimonious to think of the event of saltation as occurring prior to the establishment of SM and CI systems. A fuller discussion of the issue is beyond the scope of this paper (see Mukherji Forthcoming, Chapter 5). The present historical point is that, since Darwin made no significant proposal about generativity beyond vague metaphors like speech ‘acted on’ mental powers and ‘long trains of thoughts’ were carried out, my interest in Darwin’s proposals about word-formation is dissociated from the gradual versus saltation debate on the origin of language.

### **Is SAH tenable?**

Despite a variety of infirmities as noted, Darwin’s narrative on language evolution contained at least two attractive ideas. First, given the uniqueness of human language in almost every aspect, he attempted to restrict the narrative mostly to what was plausibly available in the proximal hominid line itself to minimize assumptions about borrowing evolutionary material from nonhuman sources. Thus, the power of speech was explained from the prior availability of elaborate singing and minimal assumptions about phonemic units in nonhuman species. Further, he made minimal assumptions about borrowing mental powers from the apes to let it grow with the rapid growth in brain size in the hominid line.

Second, once these two ‘powers’ were independently available, he imagined them to be acting on each other to reach the exponential richness of human language. According to this narrative, word-formation was the central achievement of the first part of the narrative, while word-combination leading eventually up to unbounded structures was the accomplishment of the second part.

In my view, Darwin's narrative, if valid, casts doubt on the validity of SAH. Words, viewed as structures of sound-meaning correlation, are central to Darwin's and Tattersall's narratives. The basic problem with SAH is that, if the sound system is ancillary and it evolved *after* the origin of language, then the entire burden of the rich and complex structure of human thought is to be borne by units of human thought itself. Given that the units of human thought were not available in the apes, it remains a mystery in SAH how thought evolved at all. It is true that Darwin also needed a minimal stock of mental units so that they could grow with brain size etc.; so, the first availability of human thought is a mystery for everyone. However, in Darwin's account, the *further* growth of the thought systems happened due to the presence of sounds *via* words, as we saw. This account is not available to the proponents of SAH.

Notice, before we proceed, that the preceding problem with SAH makes no appeal to the phenomenon of communication at all. Hence, Darwin's account does not require a 'social' narrative of emergence of language proposed by Michael Tomasello (2003) and many others. Darwin proposed that sounds were needed to enhance the power of thought itself. In that sense, sound-meaning correlations were internally required in the language system for the system to take off as a generative procedure. Externalization of thought was not a necessary feature of the device at all.

Nonetheless, once words were available, then assuming other human facilities such as demonstration, imitation and the like, some form of communication was also available as an ancillary capability. In this perspective, communication was a bonus, it was not a part of either the basic structure or the primary function of language. This proposal fits well with the widely-attested phenomenon that human languages are not very well designed for communication.

In this connection, we recall that we never quite made sense of the idea of what it means for a ‘pure’ concept to be there, how do they ‘look’ like, what it means to combine them. According to the suggested reading of Darwin’s cryptic proposals, words were needed precisely to mark what Descartes called ‘hidden thoughts’; the marked objects of thought both acquired identity and facilitated combinatorial operations. Perhaps, we may be able to strengthen this idea to argue that, in fact, there were no concepts before there were words; human concepts *emerged* with words. No wonder nonhuman animals do not have concepts like us because they do not have words (Mukherji 2019).

## References

- Asano, R. and C. Boeckx. 2015. Syntax in language and music: What is the right level of comparison? *Front. Psychol.* 6:942. doi:10.3389/fpsyg.2015.00942
- Berwick, R. 2011. All you need is Merge. In A. M. Di Sciullo (Ed.) *Biolinguistic Investigations*. Oxford: Oxford University Press, 461-491.
- Berwick, R. C., A. Friedrici, N. Chomsky, and J. J. Bolhuis. 2013. Evolution, Brain and the Nature of Language. *Trends in Cognitive Science*, 17.2.
- Berwick, R. C. and N. Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, Ma.: MIT Press.
- Brooks, A. S., J. E. Yellen, R. Potts, A. K. Behrensmeyer, A. L. Deino, D. E. Leslie, S. H. Ambrose, J. R. Ferguson, F. d’Errico, A. M. Zipkin, S. Whittaker, J. Post, E. G. Veatch, K. Foecke, J. B. Clark. Long-distance stone transport and pigment use in the earliest Middle Stone Age. *Science* 360, 90–94, 6 April.
- Carroll, S. 2005. *Endless Forms Most Beautiful*. New York: Norton.
- Chomsky, N. 1993. A Minimalist Programme for Linguistic Theory. *MIT Working Papers In Linguistics*, reprinted as Chapter 3 in Chomsky (1995a).

Chomsky, N. 1994. Bare Phrase Structure. *MIT Working Papers In Linguistics*, reprinted in G. Webelhuth (Ed.) *Government and Binding Theory and The Minimalist program*. Oxford: Basil Blackwell.

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

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry*, 36.1, 1-22.

Chomsky, N. 2012. *The Science of Language: Interviews with James McGilvray*. Cambridge: Cambridge University Press.

Chomsky, N., A. Belletti, L. Rizzi. 2001. *An Interview on Minimalism*. University of Siena, Nov 8-9, 1999.

Darwin, C. 1871. *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.

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

Flake, G. 1998. *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation*. Cambridge: MIT Press.

Fujita, K. 2014. Recursive merge and human language evolution. In T. Roeper and M. Speas (Eds.) *Recursion: Complexity in Cognition*. Heidelberg: Springer.

Gallistel, C. 1998. Symbolic processes in the brain; the case of insect navigation. In D. Scarborough and S. Sternberg (Eds.) *An Invitation to Cognitive Science: Methods, Models, and Conceptual Issues*, Volume 4. Cambridge: MIT Press, 1-51.

Hale, K. 1966. Kinship reflections in syntax: some Australian languages. *Word*, 22.1.

Hauser, M. 1996. *Evolution of Communication*. Cambridge, Ma.: MIT Press.

Hauser, M. 2001. What's so special about speech? In E. Dupoux (Ed.), *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler*. Cambridge: The MIT Press, 417-434.

Hauser, M., C. Yang, R. C. Berwick, I. Tattersall, M. J. Ryan, J. Watumull, N. Chomsky and R. C. Lewontin. 2014. The mystery of language evolution. *Front. Psychol.*, 07 May, doi: 10.3389/fpsyg.2014.00401



- Hinzen, W. 2007. *An Essay on Names and Truth*. Oxford: Oxford University Press.
- Hinzen, W. and M. Sheehan. 2013. *The Philosophy of Universal Grammar*. Oxford: Oxford University Press.
- Jenkins, L. 2000. *Biolinguistics: Exploring the Biology of Language*. Cambridge: Cambridge University Press.
- Jones, D. 2010. Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, 33, 367–416. doi:10.1017/S0140525X10000890
- Joordens, J, F. d’Errico, F. P. Wesselingh, and others. 2015. *Homo erectus* at Trinil on Java used shells for tool production and engraving. *Nature*, 518, 228–231.
- Katz, J. and D. Pesetsky. 2009. The identity thesis for language and music. Ms. MIT. <https://ling.auf.net/lingbuzz/000959>.
- McGonigle, B., M. Chalmers, and A. Dickinson. 2003. Concurrent Disjoint and Reciprocal Classification by Cebus Apella in Seriation Tasks: Evidence for Hierarchical Organization. *Animal Cognition*, 6, 185–197.
- Miyagawa, S., S. Ojima, R. C. Berwick and K. Okan. 2014. The Integration Hypothesis of Human Language Evolution and the Nature of Contemporary Languages. *Front. Psychol.* 5:564. doi:10.3389/fpsyg.2014.00564.
- Miyagawa, S. 2017. Integration hypothesis: A parallel model of language development in evolution. In S. Watanabe, M. Hofman and T. Shimizu (Eds.) *Evolution of the Brain, Cognition, and Emotion in Vertebrates*. Springer, 225-247.
- Miyagawa, S., C. Lesure and V. A. Nóbrega. 2018. Cross-Modality Information Transfer: A Hypothesis about the Relationship among Prehistoric Cave Paintings, Symbolic Thinking, and the Emergence of Language. *Front. Psychol.*, 20 February 2018. doi.org/10.3389/fpsyg.2018.00115.
- Moro, A. 2016. *Impossible Languages*. Cambridge: The MIT Press.
- Mukherji, N. 2009. Doctrinal Dualism. History of Science, Philosophy and Culture in Indian Civilization (General Editor: D. P. Chattopadhyaya), Volume xii Part 5: *Materialism and Immaterialism in India and the West: Varying Vistas* (Ed. P. Ghose).
- Mukherji, N. 2010. *The Primacy of Grammar*. Cambridge: The MIT Press.

- Mukherji, N. 2019. Sound of thoughts. Invited Lecture at international conference on Linguistic Explanation, CEP-2019, University of Heidelberg, 14 February.
- Mukherji, N. Forthcoming. *Mind in the Mirror of Language: Humans, Mentality, Merge*. London: Bloomsbury Publishers.
- Povinelli, D. J. 2000. *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford: Oxford University Press.
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, Ma.: MIT Press.
- Ramus, F., M. Hauser, C. Miller, D. Morris and J. Mehler 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349-351.
- Schino, G., B. Tiddi, and P. di Sorrento. 2006. Simultaneous Classification by Rank and Kinship in Japanese Macaques. *Animal Behaviour* 71: 1069-1074.
- Seyfarth, R. M., Cheney, D. L. and T. L. Bergman. 2005. Primate Social Cognition and the Origins of Language. *Trends in Cognitive Sciences* 9: 265-266.
- Stewart, I. 1995. *Nature's Numbers: Discovering Order and Pattern in the Universe*. London: Weidenfeld and Nicholson.
- Stewart, I. 2001. *What Shape is a Snowflake?* London: Weidenfeld and Nicolson.
- Stout, D. 2011. Stone toolmaking and the evolution of human culture and cognition. *Philosophical Transactions of Royal Society of London B Biol Sci.*, Apr 12, 366 (1567), 1050–1059.
- Tattersall, I. 2012. *Masters of the planet : the search for our human origins*. New York: Palgrave Macmillan.
- Tattersall, I. 2016. A Tentative Framework for the Acquisition of Language and Modern Human Cognition. *Journal of Anthropological Sciences*, 94, 157-66. doi: 10.4436/JASS.94030.
- Tattersall, I. 2019. The Minimalist Program and the Origin of Language: A View From Paleoanthropology. *Frontiers in Psychology*, 02 April. doi.org/10.3389/fpsyg.2019.00677
- Tomasello, M. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Uomini, N. T., and G. F. Meyer. 2013. Shared brain lateralization patterns in language and Acheulean stone tool production: a functional transcranial doppler ultrasound study. *PLoS One* 8:e72693. doi: 10.1371/journal.pone. 0072693.

Uriagereka, J. 2008. *Syntactic Anchors on Semantic Structuring*. Cambridge: Cambridge University Press.