

TOWARDS UNDERSTANDING THE ORIGINS OF LANGUAGE

Svetlana T. Davidova

Abstract

Recent interest in understanding the evolution of language has delivered remarkable depth and detail of knowledge, empirical and theoretical. Significant advances are made in understanding the processes by which modern language has evolved from earlier and simpler communication systems, i.e. protolanguage, or pre-language, defined by the predominant use of content words in their bare forms and labeled as grammarless.

That said, very little is known of the evolution from the communicative grunts of primates to protolanguage. It is fair to say that this period of the evolution of language has attracted much less interest, given the lack of factual support which in the past has driven scholars to resort to conjectures and speculations as a substitute.

The goal of the present article is to minimize this gap by making the best use of known findings and offering alternative analyses of these. The article argues that by evaluating traces of ancient forms in modern languages, i.e. “language fossils”, the formation of language systems *de novo* in modern times, the formation of modern semasiographic systems, one can detect general patterns and extrapolate valuable insights into language origins. In addition, by examining bio-cognitive features, participants in some language-relevant functions and shared with pre-human pre-linguistic species, one can project how a new communicative behaviour could have emerged from pre-existing bio-cognitive properties, i.e. without a Language Faculty.

Keywords: language origins, protolanguage, self-organization, signs, iconicity, Language Capacity, semasiographic systems

1. Introduction

Modern language is defined in terms of highly abstract grammatical machinery for organizing a rich lexicon in publicizing private human thoughts. Protolanguage is assumed to be an intermediate state of linguistic communication, preceding modern language and limited in both meaning and structure in its role of publicizing the private life of the human mind. The origins of language mark a distinct stage in evolution of life forms prompted by the desire or need to consciously and voluntarily share thoughts and experiences.

Language in its modern form is attributed to human species. Protolanguage as an intermediate stage of linguistic communication is attributed to pre-human species, e.g. Neanderthals, Denisovans, etc. This means that the origins of language, i.e. some form of communication, presumably a mixture of non-linguistic signs, e.g. vocalizations and gesticulations, and some form of protolanguage, should be attributed to various pre-human ancestral species.

Because scholars differ on their definitions of language, so do their visions on language beginnings.

1.1. What is language which has originated

In modern linguistics language is defined as Universal Grammar (Chomsky, N. 1972, 1980,

1986, 1995, and elsewhere) and the beginning of language is understood as the imposition of the grammar algorithm onto protolanguage (Bickerton D.1984, 1990 and elsewhere)

Alternatively, language is defined as a system of symbols organized for the propose of sharing propositional thoughts.

Regardless, language is understood as complex grammatical machinery deployed for the efficient externalization of the rich and complex human semiosis, i.e. highly complex meanings dressed in highly abstract grammar. Thus, human language is identified with modern languages spoken by modern civilized societies.

Protolanguage is identified as an earlier intermediate stage in the evolution from animal vocalizations to modern language. It is defined in terms of meaning-based short utterances composed mostly by lexical words with concrete meanings and often uncertain grammatical categorization, organized by the principles of information around the opposition old vs. new. (Bickerton, 1984, 1990 and elsewhere).

That said, some modern and fully functioning languages demonstrate striking resemblance to the hypothetical protolanguage, e.g. Piraha, Riau Indonesian, some languages of modern hunter-gatherers(Cysouw, Comrie 2013; Everett, D. 2005 and elsewhere; Gil, D. 2009).

And various linguistics (Jackendoff, Wittenberg 2014) have argued that there is no qualitative difference between protolanguage and modern language, suggesting that protolanguage-like modern languages fulfil all the major functions of language, i.e to codify and communicate infinite human thoughts by finite linguistic means, albeit with minim use of the grammatical arsenal hypothesized to be contained by UG. Thus, protolanguage is actually language.

And hunter-gatherers are biologically and cognitively normal language users. The difference between them and speakers of English and French is the material and social complexity of modern civilization and most notably, formal education, as environment and a backdrop for language use.

In this sense the origins of language is not protolanguage, but some even simpler communication system, intermediate between animal vocalizations and protolanguage.

But what could be even more rudimentary and cruder expression of human concepts?

The present article argues that language has begin with the reorganization of the multimodal animal communication and the formation of a novel subsystem with the first primitive spoken words used as a minor component of animal multimodal utterances.

1.2 .Language begins with the invention of the word: the notion of “ minimum complexity”

As per Pinker and Bloom's seminal article (1990) language has begun as a cultural invention, i.e. as a communicative novelty, prompting subsequent bio-cognitive adaptation of the human body in a Language Faculty.

In this sense the Complexity theory, initially designed to understand the spontaneous emergence of novel entities with unique internal organization in physical matter (Prigogine, Sengers, 1993; S. Kauffman 1995) and subsequently adapted for understanding complex design in a broad range of contexts (Heylighen1999) can provide some guidelines into understanding the beginnings of language as a communicative innovation. It states that complexity increases from an initial point of minimum complexity, which marks the point of irreversibility (Prigogine and Stengers,1993, p.301) So, complexity does not arise from simplicity, but with

some minimal level of complexity, some minimum of internal organization. Significantly, complexity does not start at random levels, it starts with middle-sized formations (Prigogine, Stengers, *ibid.*)

In language such minimum internal organization is the a primitive form of word, or protoword, identified as emergent pairs of a single concept and a stream of speech, by simultaneous particulation and association in semiosis and vocalizations.

To note, although in modern linguistic thought word is not well defined, in modern languages a word is usually identified as a codified bundle of semantic, grammatical and phonological features, thus, it is conceived as part of the structure of language. And although the generative approach attributes a marginal role for words in UG, alternative generative paradigms, e.g. the Parallel Architecture among others, places the word at the centre of language as interface of the three levels of structure, making it the cornerstone of the language system.

Significantly, there is linguistic communication with minimum grammar but there is no, and I am sure, never has been, a language without words. They are the first to emerge when a new language is formed, and the last to disappear when a language becomes extinct. They are the first to emerge in the developing mind and the last to remain in the damaged one.

In sum, it is logical to anticipate that in the context of multimodal non-linguistic communication the word emerged in its most primitive form and initiated the beginning of language.

2. Language as a system of signs

Communication is based on signs and understanding the origin of language requires some basic understanding of signs in nature and culture.

To remind, in the theory of signs by C.S. Pierce (see Stanford Encyclopedia of Philosophy) he distinguishes three categories of signs: icons, indices and symbols.

In icons the relationship between the meaning and the form is physical resemblance. Portraits are icons. Very few words are icons, e.g. swirl, splash, etc. which are vocal imitations of natural sounds.

Indices are patterns of correlation and co-occurrence of two sensory experiences, one of which points at the other, or indicates the presence of the other and becomes the form and the other, the meaning of an index. The lit window indicates that someone is in the house, or the smell of skunk indicates the near presence of the animal.

Both icons and indices are motivated signs in that the connection between the two sides, meaning and form, is motivated, either by physical resemblance or by correlation, spatial and/or temporal.

Symbols are different in that the connection between meaning and its representation is accidental and unmotivated. The flag of a country is a symbol of a nation, a dollar bill is a symbol of certain value.

Indices are ubiquitous in nature, they are at the foundation of communication starting at the lowest organic levels, e.g. between the molecules inside the body. Indexical communication is limited by the present, it cannot be detached from the here and now.

Iconic signs are found both in nature and in culture. In the immune system the antibodies block the “invader” by creating a mirror image of it and blocking its access to the rest of the

organism. In culture photographs are the clearest example of iconic signs.

The communicative power of indices and icons is limited as it is dependent on perception.

Social sciences explain symbols as products of cultural evolution resulting in simplification of iconic signs by truncation and/or abbreviation of the iconic form, prompted by the need to simplify the message due to the unbalance between the increased volume of information and the natural limitations of the perceptual capacities, on which iconic signs are based. As a result of this simplification the natural resemblance between the two sides of the sign, the form and the meaning, is lost and the relation becomes obscured, i.e. a symbol is formed.

Language makes use of all three types of signs.

2.1. Words are unique signs.

Most words are symbols. But the arbitrary association of an idea, i.e. a concept and its material representation, which is the definition of a symbol, is not all that words are.

Culture is based on symbols, from traffic signs to mathematical formulas, to films. But the cultural symbols are different from linguistic symbols. Cultural symbol systems fall into two categories, one is codes, i.e. systems composed of discrete symbols with unambiguous and stable meanings each with a well defined and unique place in the system. They are interpretable only within that system. Traffic signs and mathematical formulas fall into this category. Codes do not need grounding in time and space as they are context-independent, that is, they receive the same unambiguous interpretation in all contexts they appear.

Others, e.g. films and music, although may have discrete form, (music, for example) have holistic meanings whose interpretation depends on the context. The meaning of a painting and a music score is “in the eye of the beholder”. They belong to the so called inferential communication where the same symbol receives different interpretations by different interpreters and contexts, e.g. the same music score is often used in a film and in a commercial advertising and is interpreted differently in these very different contexts.

Words are different as they are unique in that they are both symbols and indices, i.e. they have a stable meaning - form associations as a part of the language system where the meaning is a concept, recognizable by every user. At the same time these associations are flexible as a word points to a different referent in each communicative act, i.e. words have sense and reference. Thus, the meaning of words is general and specific at the same time. Words are hybrids, a combination of two traits.

Hurford (2007) defines the beginning of language as the formation of the dual role of the linguistic sign: the combination of the role of denotation (intension, stable representation of a class of objects, a code,) and reference (extension, representation of a particular object in a particular situation). The inferential use of linguistic code in individual communicative circumstances is the central characteristic of the word. This hybrid nature of the word is unique in communication. The uniqueness of human language as both a code and inferential system begins with the word.

To note, languages differ in the types of grammars but all languages have words. Words are a linguistic universal.

3. Speculations on the beginnings of language

3.1. The debate on combination vs. decomposition of pre-linguistic communication

Modern linguistic paradigms identify language as a system of discrete units, phonemes, lexemes, morphemes, combinable in a multilevel hierarchically organized sentences. In this context the beginning of language is understood as the beginning of compositionality by the formation of discrete meanings and discrete vocalizations or manual signs(M. Tallerman, 2005, among many others).

Others point at holistic elements in modern language and argue for holistic communication as a distinct stage in the continuous process of language evolution and a precursor to compositionality . Wray (2002, (p. 113 -) entertains the idea of hypothetical holistic protolanguage consisting of formulaic expressions where a stream of speech, or a vocal phrase, is mapped onto to a holistic meaning containing a proposition. Thus, there are no words and morphemes and strings of discrete sounds, i.e. speech, are mapped onto whole propositions, producing formulaic expressions, suggesting that phonology is the first combinatorial system to emerge.

Formulaic expressions have the function to manipulate physically and emotionally the behaviour of group members and are also instrumental to social bonding.

Wray's formulaic expressions resemble the words in polysynthetic languages seen from the perspective of a learner who has not yet analyzed them into component parts and perceives them as an indivisible whole.

Wray hypothesis has an advantage in demonstrating the continuity from the natural communication of modern primates in their function of manipulating the minds of conspecifics by dyadic communicative acts.

“...the holistic cries and gestures of our pre-human ancestors were transformed , over a long period of time, into a phonetically expressed set of holistic message strings with a manipulative function such as greeting, warning, commanding, threatening, requesting...” (A. Wray, 2002, p.115).

A hypothetical holistic protolanguage is attributed to Erectus. It is understood as a precursor to the emergence of proto-words and a lexical protolanguage as hypothesized by D. Bickerton (1990 and elsewhere), marked by the coexistence of the two systems where the older holistic messages continued to be used for socializing, while the compositional system filled the need for exchange of referential information as the proto-words, predecessors to the modern nouns and verbs, were , most likely, labels for concrete objects and actions.

Such mixture of holistic and compositional elements is found in modern rudimentary forms of linguistic communication, e.g. the linguistic communication of trained apes, early child language , Basic Variety, the communication of people with damaged brains, suggesting ancient roots.

“... the interaction of these two systems (generative and holistic) in the absence of hierarchical grammar, is a convincing match for existing grammarless systems, such as those of Kanzi, very young children, aphasic people, and foreign tourists.” (Wray, 2002, p.116)

Holistic elements persist in modern languages, as they have proven successful in their functions as fast, efficient and unambiguous expressions of the community's emotional life as formulas of social bonding. (A. Wray, 2002, p. 117-)

Wray's model of holistic protolanguage is criticized by M. Tallerman(2005) who reminds

that the reason phonemes exist is to differentiate word meanings, thus phonology implies a lexicon of word-sized units. Others dispute the possibility of a community coming up with an agreement on the mapping of a string of syllables and a proposition, while others see no difficulty in that, pointing at children's learning of songs. For detailed discussion of the various theories of protolanguage see T. Fitch (2009).

3.2. Holistic vocalizations, a precursor to speech: musical protolanguage

Various authors, as early as Darwin and later Jespersen, have expressed the idea of song-like holistic vocalizations as a precursor to speech, further developed in a theory which argues for primacy of music and singing as a distinct stage in language evolution (see T. Fitch 2009 for a review). In Jespersen a melody was a sign for a particular event, activity or an entity , a person or an object, as a first stage, later transformed by a process of analysis resulting in the formation of combinable units of form and meaning, a precursor to language as a combinatorial system.

Evidence for the ancient roots of learnable patterned sound-making are found in species as distant from us as birds in birdsong. Of course the term “ musical protolanguage “ does not refer to the modern understanding of music, it means that the meanings of the messages were unanalyzed, holistic, thus there was no direct correspondence between parts of the signal and part of the meaning. Thus, there was no duality of patterning.

Some scholars suspect that the natural propensity for music emerged initially from the mother-infant bonding, suggesting that “ motherese” has its evolutionary roots in musical protolanguage. It is also suggested that the universal nature of music and dance, present in all cultures, speaks to its ancient evolutionary roots when its primary function was markers of group identity. Most recently S. Mithen (2005) has developed his version of “ musical protolanguage” hypothesis with his Hmmm (holistic, manipulative, multimodal, musical) theory.

3.3. Prosody, a link between animal communication and language

Prosody and intonation are holistic components of modern language. It is the layer which remains holistic as part of predominantly discrete system.

Prosody and intonation bare some resemblance with animal vocalizations which are also holistic. On this basis some argue that the capacity for speech has evolved from the prehuman capacity for holistic vocalizations (P. Liebermann, 2007 and elsewhere) relying on findings that lower animals are innately predisposed to use vocalizations to communicate emotions . In higher animals there is a level of control over the vocal organs. The transition from uncontrolled expression of emotions to use of intonation for holistic expression of propositions is the easiest and most natural way to begin using already existing capacities in a slightly different way at a stage when the highly refined control of the articulatory organs , needed for modern type pronunciation has not yet evolved. Further, there is a one-word phase in early child language development, when one word, accompanied by intonation is used to communicate propositional meaning, suggesting that this could be another intermediate stage in a long and protracted process.

Intonation plays a major role in rudimentary forms of communication and to lesser extent in

spontaneous speech to communicate questions, grammatically shaped as statements, in imperatives shaped as statements, etc., intonation and stress is used routinely to resolve ambiguity, effectively reducing or eliminating the need for complex syntax.

In addition, in a study of a newly emerged sign language Al Sayid Sign language M.Aronoff et al. (2008) has found that prosody (stress, intonation) plays essential part in conveying syntactic structure.

“not only do prosodic signals serve to separate clauses into intonational phrases, they can also link them to one another to form complex propositions...complexity may be encoded in the prosody...prosody plays crucial role in marking constituents and dependencies from the very beginning“ (Aronoff, et al. 2008).

The persistent use intonation in modern linguistic communication is another window into earlier stages in the evolution of language as a remnant of its beginnings.

4. Traces of ancient linguistic forms in modern language

4.1.Language “fossils”

There is a small and heterogeneous group of linguistic forms in modern languages which deviate significantly from the standard definition of a linguistic form. Jackendoff (2002) terms them 'language fossils' and argues that they are relics, bits and pieces from earlier stages of pre-linguistic vocal communication and as such might provide clues for earlier stages of language evolution. These include :

- a. Holistic utterances with phonology, but no syntax, mostly exclamations of the sort 'Damn' 'ha-ha-ha', or 'Wow', which display more similarities to instinctive vocalizations (reactions to pain, etc.), than to linguistic entities as they are not consciously communicative. They exist as individual utterances, not integrated into the language system.
- b. Holistic utterances used intentionally for communicative purposes, although limited to specific situations and better defined as communicative noises, e.g. Shtt'. Their meaning is interpretable from the context, most of them have no distinct syllable structure nor grammatical features.
- c. Holistic utterances of the sort 'Hello', 'Good bye', 'Yes', 'No', etc. with phonological structure and holistic meanings, tied to specific communicative situations.

If these forms are indeed remnants of ancient pre-linguistic forms, the fact that they continue to exist incorporated in the modern linguistic system would suggest that they have retained their original functions and that these functions are relevant for modern humans as they were for ancients.

4.2.At the beginning there was iconicity: iconicity and the formation of the earliest language forms

Many linguists assume that words have originated as symbols since the beginning. Others disagree hypothesizing some kind of natural connection, iconic and/or indexical, between the

meaning and form of the first words, facilitating communication at time when the formation of words as systematic pairings of meaning and form were in early stages (R. Berling, 2005). Bickerton hypothesizes that to assure adequate reception of message by placing limits on the possible interpretations in situation of displaced reference the first words would have had to be iconic signals, that is, imitations of natural sounds. (Bickerton, 2009). Similar arguments are made by Clark, B. 2011).

Iconicity is a broad term which stands for actual or perceived closed similarity between the form and meaning, of a sign, as opposed to arbitrariness. Iconicity in vocalizations is based on a natural connection between the meaning and the sound.

Iconicity alleviates challenges in comprehension by providing transparency in the connection of form and meaning. It also demonstrates that language reflects human experience with the world. It has shown to be effective in small lexicons and is attested in language varieties with limited communicative functions, e.g. pidgins and early child language.

Early stages of ASL and the Al-SAYIID sign language (Aronoff et al. 2008) demonstrate high degree of iconicity, later lost and replaced by arbitrary signs, themselves elaborations of the original iconic signs.

In each case iconicity helps avoid ambiguity and illustrates how language is used to communicate one's experience with the world by providing the link between the gradient nature of our perceptual experiences with the world and the linguistic forms, signalling that linguistic forms are grounded in human experience. Both these needs, for disambiguation and the externalization of bodily experience through linguistic communication, are present at the time of language origination, suggesting the likelihood that iconicity has been a prominent feature in the earliest stages of language genesis.

Importantly, iconicity is a matter of degree and different languages use iconicity to different degrees.

Iconic relations work successfully in small vocabularies, where the meanings of words are not compositional as the iconic relationship exists only between a holistic, undecomposable meaning and a phonological form (Hurford (2012). Vocabulary enlargement brings the need for introducing arbitrariness as a way of introducing complexity in language. This is because "... arbitrariness ...allows for a maximum discrimination between entries in the lexicon. " (Pemiss, Thompson, Vignioeco, 2010, p. 17).

As the functions of primitive iconic communication increased, the frequency of use and the need for speedy communication must have increased. Faster and , as a result, sloppier pronunciation lead to truncation of the iconic signs, making the initial natural connection unrecognizable. As a result the natural connection between signifiant and signifie is lost resulting in the substitution of icons by symbols.

A similar tendency of "cutting corners" by using only parts of linguistic expressions as substitutes for the whole is alive and well in modern linguistic communication given the demand for fast speed of speech production, e.g. TV studio, phys-ed teacher, a co-op program, etc. The enlargement of the lexicon played a crucial role in the subsequent evolution of language as it brought structural changes in the system , e.g. with an enlarged lexicon compositionality and arbitrariness assumed a leading role.

These findings add support to the argument for language origin as largely a cultural, not biological process.

4.3. Iconicity, a general property of language

Although the idea of arbitrary nature of language both in the lexicon and of grammar dominates modern linguistic theories, the topic of iconicity has recently attracted the attention of scholars and a number of studies have been published on the topic.

Earlier studies by J. Hyman (1993) and more recent findings (Permis, Thompson, Vigliocco, 2010; Perniss, Vigliocco, 2014) argue that iconicity is far from insignificant little quirk, it is pervasive and universal.

Relying on a broad-based analysis of a wide variety of languages (mostly non-Indo-European, in both modalities), e.g. languages of sub-Saharan Africa, aboriginal languages of Australia, indigenous languages of South America, languages of south-east Asia, Japanese, Korean, etc. the authors determine that iconicity is one of defining properties of language, on par with arbitrariness. This is why the authors suggest that it must be reflected in the theory of language. To quote: "...any viable theory of language must include iconicity in addition to arbitrariness as a guiding principle..." (Perniss, Thompson, Vigliocco, 2010 p. 14).

Iconic mappings are found both in lexicon and grammar. In the lexicon in onomatopoeia, the spoken form of a word is an imitation of the natural sound of the class of object it denotes. Monaghan et al. (2014,) term this transparent connection "absolute iconicity, an example of which is "miau" as the sign for the sound a cat makes, distinct from relative iconicity, i.e. partial similarity with the natural sound the class of named objects makes., exemplified in a tendency to associate the sounds of small animals with high vowels, for example birds "chirp", while the sound of large animals are associated with low vowels, e.g. the lion "roar". In grammar there is iconicity of sequence, where the sequence of verb forms (and the sentences formed around them) mirrors the sequence of events they denote. That is, spatial proximity of linguistic forms reflect the proximity of meanings these forms denote. This means that words as signs reflect the human experience with the world. Iconicity is found in all grammatical categories: nouns, verbs, adjectives, adverbs.

4.3.1. Synesthesia (sound symbolism), a property of early language

Synesthesia (as defined by scientificamerican.com) is "...anomalous blending of the senses in which a stimulation in one modality simultaneously produces sensation in a different modality". In language synesthesia is a type of iconic relation between meaning and form of a word where iconic (natural) relation is viewed between a sound and the meaning it represents. For example, in many languages the meaning of SMALLNESS is represented by words with the high front vowel /i/, and BIGGNESS, HEAVINESS is represented by words with low back vowel /a/. Moreover, arbitrariness and iconicity are not binary notions, in sign languages as many linguistic forms display both in various degrees. That is, "...the degree of iconicity or arbitrariness ascribed to individual signs is best understood as a continuum" (in Permis, Thompson, Vigliocco, 2010 p. 5).

Iconicity is a universal phenomenon in language indicating sensitivity of human mind to it which informs about the internal organization of the brain and the way it processes language. A

natural predisposition for mapping acoustic and visual properties of objects is shown to be universal. For example, the meaning of upward motion is represented in spoken language by higher pitch. Iconic effects are shown to influence language learning in children and adults. Both children and adults learn faster iconic words compared to purely arbitrary ones in both modalities.

The assertion that sound symbolism is a universal property of language challenges the view of language as an abstract system unrelated to experience. It shows that language is intertwined with both sensory experience and with other aspects of cognition. In addition, it challenges the view that different aspects of language, phonology, syntax, semantics, are processed independently and demonstrates that linguistic theories do not reflect the way language is processed by the human brain.

Interestingly, cases of synesthesia and sound symbolism are not as rare as previously thought. Hurford (2012) shows that in many unrelated languages the sound symbolism is part of the language system. For example, in English the consonant cluster /sn / at the beginning of words, e.g. sneak, snigger, etc. is associated with the meaning of unpleasantness. (Hurford 2012, p. 128). This correlation of meaning and form is used in the creation of new words.

One would assume that conventionalization of sound symbolism has played a role in the very early stages of language evolution by facilitating word learning by making it predictable as similarly sounding words are likely to have similar meanings.

Moreover, in many languages the conventionalized sound symbolism is extended to grammar : there is a correlation between the phonetic form and the grammatical category. In Dutch, French, Japanese, English, phonological cues help distinguish content from function words and nouns from verbs, thereby facilitating language learning. Moreover, a large number of nouns of high frequency of use have back vowels, and high frequency verbs have front vowels.

Thus, iconic mappings facilitate not only the learning of lexicon, but also of grammar, effectively eliminating the need to postulate a grammar organ or Universal Grammar.

Significantly, processing of iconicity in language involves the formation of neuronal connections across multiple, especially proximate cortical areas and multiple modalities.

Ramachandran and Hubbard (2001) argue for neural basis of synesthesia as simultaneous activation of various sensory and motor maps in the brain, properties of human cognition thought to precede the emergence of language as one of its pre-adaptations, subsequently co-opted for use in service of language-related functions.

4.3.2. The limits of iconicity in language

Iconicity may have been an advantage in earliest stages in language evolution.

In evolutionary linguistics modern words are hypothesized to have evolved from imitating emotional animal and primate vocalizations. Others point at iconic vocalizations imitating natural sounds, or onomatopoeia, e.g. “ splash”, “ swoop” i.e the name reflects the sound the referent makes, which limits the vocabulary to words of sound-making entities.

Iconicity has been a much more prominent feature in the initial stages in the evolution of language systems, spoken, signed, written (Perniss, Vigliocco, 2014). It is attributed a key role as a bridge between the indexicality of animal communication systems, which relies on the spatial and temporal proximity of a sign and a referent, and the displaced reference as a unique property of modern language . The replacement of indexes with icons replaces

reference to an object, i.e. “ functional reference”, with reference to a concepts or “ conceptual reference” where the concept is represented by an image .

The iconic representation of reality has deficiencies as icons are necessarily limited and fixed in number and they refer only to existing entities in reality. In this sense in a changing environment they represent outdated information which makes it for survival. Moreover, it puts learners in disadvantage as learning information which represents earlier environments is useless for survival in new conditions.

The formation of symbols has replaced to a significant degree the role of iconicity in modern language and this transformation was motivated by key advantages of symbols, i.e displaced reference and discreteness in meaning and form, resulting in duality of patterning .

That said, despite its inherent limitations in encoding linguistic meanings, iconicity has retained a role in modern language as it contributes to ease of learning and fast and accurate understanding of the message.

On the other hand, if we take motivated signs to mean semantic transparency , then most of the newly invented lexical items are motivated signs. A closer look would reveal that most of the new lexical inventory are assembled of reused and recombined phonological and/ or semantic material as a demonstration of transparency in the meaning-form pairings, while genuinely novel words are rare. In this sense one could say that in all languages for the overwhelming majority of the words in the lexicon the relationship of meaning and form is motivated. This preference for transparency is explicable by memory constraints imposed on language by the human brain as acquisition / learning of any language includes the memorization of thousands of words which presents a great challenge for memory.

5. Public labels for private concepts : the formation of the lexical word

Humans are the only species which share their private mental life with conspecifics in precise and clear ways. They achieve that through attaching public labels to private concepts, which has two important implications: one, to share thoughts and in this way to influence other minds and two, to organize one's thoughts in a more clearly defined ways.

Language is the main vehicle for publicizing human thoughts and experiences though linguistic signs.

5.1. Common ground, the foundation of signs

The ability to conceptualize reality in repeated patterns is nothing new as many non-human mammal, not only ape species, have this ability. Humans, as many other social species, are creatures of habit, they represent reality in repeated patterns, thus, they memorize patterns not only of natural occurrences but of behaviour of animals and conspecifics.

In a closed group of interacting individuals the members are united by common daily experiences as interactions with one another and with the material reality. Eventually common experiences become routinized and common patterns of behaviour emerge and become recognizable as part of group identity, i.e. group members are expected to behave in certain predictable ways.

In cognitively advanced species shared behavioural routines furnish common

conceptualization of reality, which in humans becomes the foundation for the formation of signs, based on agreement on common meanings as behavioural common ground, the foundation of culture.

Common ground is the knowledge and understanding of the surrounding world an individual shares with the members of the group, including knowledge of oneself, which is gained through interactions of the individual with the environment, including one's own body. In communicatively advanced species the fact that it is shared means that it does not need to be explicitly communicated as it is implied. The understanding of the self as a helpful group member is a pre-requisite for the formation of norms of behaviour and social conventions, such as language. It presupposes shared knowledge and the pre-existence of some kind of primitive culture, i.e. patterns of behaviours and practices not directly shaped by survival necessities and based on shared meanings, likely to be assembled on a small scale, in a small group of people who interact in a daily basis, live in close proximity, very likely, family members. The formation of common ground relies on theory of mind, or the cognitive ability to understand that all individuals (conspecifics) have minds and have representation of others' minds, i.e. I know that you know.

Croft (2017) distinguishes various types of common ground knowledge:

- * natural common ground, based on the recognition of other humans as conspecifics with common human anatomy, physiology, cognitive capacities
- * perceptual common ground, based on the cognitive capacity for joint attention and categorization
- * communal common ground, based on joint practices among community members.
- * actional common ground: the use of conventions, such as language, for coordination of joint activities.

Crucially, the formation of common ground and the ability to participate in joint behaviours is based on the innate propensity for socialization, inherited from ancestral species, but elevated significantly in humans.

It is plausible to suspect that the formation of common ground is one of the prerequisites and a stepping stone in the origination of language, as it leads to the formation of common meanings, an essential component of the linguistic sign.

Human minds, and, presumably other, non-human language using species, evolved beyond the ape reciprocal altruism as a result of selection pressure to participate in group activities as survival necessity, initially in a small group of genetically related individuals. Later the advanced capacities for socialization were extended beyond the family circle with the formation of larger and more complex social structures, e.g. tribes, kingdoms etc.

The uniquely human propensity for cooperation has resulted in the formation of common ground, which is one of the sine qua non of any type of communication. Behavioural common ground becomes a prerequisite for the formation of the lexicon, and later, of semantic structure and, subsequently, the formation of syntax in language.

The uniquely human propensity for cooperation has resulted in the formation of common ground, which is one of the sine qua non of any type of communication. Moreover, behavioural common ground becomes a prerequisite for the formation of the lexicon, and later, of semantic structure and, subsequently, the formation of syntax in language.

It is natural to be doubtful in the efforts to trace the beginnings of language, preliminary dated at the very beginning of humanity and shared with non-human species. It is especially difficult

to trace the origin of the first lexicon given the fast pace of change in lexicons of modern languages. That said, understanding the origin of the most ancient linguistic forms is not an impossible task and for that one does not need a time machine to travel into the distant past as one can extrapolate knowledge of processes of formation of new language systems and other sign systems from scratch in modern times and the formation of new words in modern languages.

5.2. Word formation in modern times, a window into the beginning of language

The formation of linguistic norms as conventions is hypothesized by J Aitchison (2000) to have started out from a chaotic situation where a concept was labeled differently by each speaker. At this stage communication is impossible, there is only an attempt to communicate which with time results in communicators converging on single sound/meaning pair, established as a norm and is passed to the next generation. Because direct evidence for the formation of the first sociolects is unavailable scholars resort to indirect sources of information as a substitute. One alternative is the use of artificial intelligence where robots mimic the behaviour of human individuals to some degree and interact and communicate in similar fashion.

Experiments by L.Steels (1995) and others with robots, designed to resemble human interaction and communicative behaviour demonstrate the emergence of shared vocabulary. Similar simulations demonstrate how individual agents through coordinated interactions converge onto a unified symbolic system. Here "...individual agents select an action from a set of possible actions, seeking to align their selections with those of other agents. (Dale Barr, 2004, p.938).

M.Dowman's (2007) robots came up with a communication system which consisted of a mixture of holophrastic words, single concept words and various mixtures of the two, which co-existed. Thus, the holophrastic words and atomic words represent two extremes of a continuum. The words of this form of language also showed different degrees of generality and specificity of meanings. Given the fact that modern language is also a mixture of linguistic meanings of various degrees of generality and forms of various degrees of combinatoriality along a continuum, in addition to forms representing the two extremes of that continuum, I find this picture quite realistic.

Jill Morford (2002, p.338) looks for answers for language origins in the abnormal ontogenetic experience of individuals deprived from normal access to spoken language, e.g. homesigners (deaf children of hearing parents who have not been exposed to sign systems in early childhood and invent their own gestures to communicate, i.e. homesigns). In this case a simple communication system is created de novo in real time without any input where emergence of conventions of form-meaning pairings, thus, formation of a few word-like symbols are negotiated between a pair of communicators.

Thus, unique cases of formation of language-like systems from scratch by modern humans, witnessed in modern circumstances have shown to mirror to close approximation the original circumstances in which language emerged.

Crucially, the process of language emergence in vastly different contexts seems to display common patterns which makes the search of the beginnings of language not so impossible task.

5.3. Semasiographic systems, a potential window into the origins of language

The reconstruction of the conditions of language origins will always be a challenge. The lack of evidence for language origins can be further partially compensated by studying the processes of formation in cultural sign systems, other than language, e.g. art, writing, musical notation, maps, mathematical symbols, etc. which are also uniquely human behaviours. The common denominator which unites them in a group is that they are conventions for encoding meaning via signs. These systems are certainly different from language in that they are restricted systems, which encode very specific meanings. That said, their relevance here is in the fact that they also share a number of similarities with language. This suggests that by the history of sign systems other than language can be viewed as another window into the genesis of language.

In Croft, (2017) the knowledge of how sign systems other than language are formed is examined and parallels with the history of language formation are suggested. In each case the cooperation is voluntary and is based on common ground, or shared knowledge, and although each participant has its own agenda, all manage to coordinate their individual actions by which a system of social conventions as arbitrary rules of behaviour emerges.

Semasiographic systems display the following features:

- * they encode information for the purpose of sharing it , thus, they are communication systems.
- * they are systematic, the same meaning is consistently encoded in the same sign
- * they are combinatorial systems
- * start as simple and increase in complexity over time
- * they are applied biology, (Koster, 2007) thus, they use the available biological and cognitive resources of the human organism each in a different but equally creative ways

The formation of semasiographic systems displays a clear evolutionary path:

- a. begin with a narrow semiotic envelope, i.e. limited number of meanings in limited contexts, e.g. numbers were initially used to count tangible, concrete objects, e.g. life stock, etc. and only later their use was extended to other countable entities. The same trajectory is found in writing: it was initially used to record only certain types of information which made them unusable in broader contexts.
- b. reference to present before reference to past and future: initially the signs in writing systems, were used to refer to factual experience of the present, and only later came to refer to past and future, as in memories and plans.
- c. from icons to symbols: signs in these systems evolved over time from more iconic to increasingly more abstract(symbolic) .
- d. initial multimodality: a mixture of signs were combined in writing and in music annotation.
- e. initial reliance on contextual interpretation : at first the meanings in sign systems were only vague representations of concepts with significant reliance on contextual interpretation.

The formation of the semasiographic systems is a very long process, e.g. it has taken thousands of years from the first representations of visual art , 50,000 ya (years ago) to the invention of writing, some 6,000 ya.

That said, none of the above discussed systems display duality of patterning, which is a unique characteristic of language. This difference is attributed to the fact that discussed systems are restricted in their functions and types of information they encode, unlike language with its ability to encode a significant part of human semiosis and a very broad communicative function.

Despite notable differences there are similarities in the general pattern in the process of formation of sign systems and language from the very beginning to their further elaboration. These are as follows:

- a. both evolve incrementally
- b. from simpler versions
- c. originate as depending on context for disambiguation and evolve context-independence
- d. originate with very restricted semiotic envelope and incrementally broaden it
- e. originate as motivated (icons and indexes) and evolve into arbitrary signs
- f. from multimodality to single modality

The here identified parallels affords the opportunity to extrapolate and use this information as another credible source contributing to our understanding of the genesis of language.

6. Language origins: making a new product with old machinery

6.1. The continuity vs. discontinuity argument

The fact that at present humans are the only language users has prompted the argument for human exceptionality in the living world. Depending on one's definition of language, scholars differ in their views on the question of continuity, cognitive and communicative, vs. discontinuity of humans as language users from the rest of life forms.

The defenders of the discontinuity position are scholars who identify language origins with the onset of modern grammar and attribute this to Universal Grammar (Lenneberg, 1967; Chomsky, 1972, 1980, 1986, 1988, 1995, 2000 and elsewhere), explicable with transformations in the human genome, most notably with the human version of the FOXP2 gene and its role in the Broca's area as location of the UG processor.

Others, mainly archeologists, evoke behavioural discontinuity as a demonstration for cognitive discontinuity between subhuman species and homo sapiens (W. Noble, I. Davidson, 1996).

But if one argues that language begins with the formation of the most rudimentary language forms, combined into primitive utterances, attributed to some known to us pre-human species, as I have in earlier segments, one must seriously consider the continuity argument.

And while human exceptionality is defended mostly by linguists, experts in life sciences usually reject the notion of human exceptionalism and are interested in the question what components of the ape organism could have served as precursors to language (Fitch, 2009, Liebermann, 2007 and elsewhere, among many others). They point out the genetic similarity between us and extinct homo species and assume behavioural, including linguistic continuity. Especially relevant is the genetic continuity in the evolution of FOXP2 gene, also known as “ grammar gene”. Genetic comparison of humans, Neanderthals and primate species, which constitute the common ancestor of both Neanderthals and homo sapiens reveal that “ ...the

Neanderthals carried a FOXP2 protein that was identical to that of present-day humans in the only two positions that differ between human and chimpanzee. ...this establishes that these changes were present in the common ancestor of modern humans and Neanderthals....Whatever the two amino acid substitutions might have for human language ability, it was present not only in modern humans, but in late Neanderthals.”(Krause, et al. 2007).

Importantly, arguments for exceptionality of us as species are obviously anthropologically biased as the very concept of speciation implies some unique traits defining each species. Thus, “Uniqueness is not unique” as per Matt Ridley in *The agile gene*.

In what follows I will briefly remind of findings which demonstrate continuity of humans with the rest of life forms and especially with primate and pre-human ancestors.

6.1.1.Continuity in brain architecture and functions

We share most of our genes, the basic body plan and functions, including essential brain functions, with the rest of life forms.

The neocortex is a layer of neuronal cells which envelops the inner layers of brain tissue (Fitch 2009). Mammalian species have evolved a thicker neocortex where the cells are not pre-specified for a specific function and can easily be grown in volume, making them highly flexible and adaptable for various functions, from specific sensory functions to abstract thought, under the influence of internal and/or external factors. Thus, the neocortex is one explanation for the adaptability of mammalian species. It is thicker in volume in apes and homo species, suggesting that quantitative change in neuronal cells can result in qualitative difference in cognitive functions. The overgrown cortex in humans provides one possible explanation for the extreme adaptability of the human species, while at the same time providing an example of physiological and cognitive continuity between us and the rest of the living world.

Moreover, continuity is found in brain functions despite differences in brain size as brain organization and functions in small-brained and in big-brained species is roughly the same. “...the basis for transforming patterns of experience to patterns of neural behaviour -, i.e. learning and memory – and it appears that even very simple nervous systems are built around essentially the same basic learning principles as the most complex ones.” and “...little evidence that basic learning abilities differ significantly across species of vastly different brain size” (T.Deacon, 1997, p.163).

In addition, many non-human species have rich mental lives. Animals, including primates have two types of memory, semantic and procedural. Semantic memory, or “ declarative memory ” is static knowledge of the way the world is. It contains concepts and categories acquired innately or through experience with the world through sensory experiences as well as generalizations from these. Procedural/episodic memory is the ability to recall stored episodes of individual experiences by creatively combining concepts from the semantic memory. Humans have inherited the same two types of memory, but they typically are understood as directly linked to the use of language, the contents of declarative memory being expressed linguistically in declarative sentences and the contents of the procedural memory imperative sentences.

Nevertheless, studies summarized in Hurford (2007) show that both stable concepts and their

use in humans are independent of language, implying continuity.

6.1.2. Continuity in cognition

Studies of animal cognition consistently show strong continuity between animal and human cognitive abilities (Fitch, 2009).

Various animal species, monkeys, birds, dogs, dolphins, etc. , thus, not just primates, display the following abilities :

- * form categories by generalizing from experience
- * remember past events and use these memories in plans for the future
- * use learned information in creative ways in solving novel problems
- * basic arithmetic abilities for representing small number of objects
- * form cognitive representations from multiple sensory sources
- * tool use, thus, hierarchical organization of mental representations and planning
- * social intelligence : in interaction with conspecifics understand dominance hierarchies, kin relations
- * follow gaze, precursor to a theory of mind and pragmatic inference
- * as a result of training many species are capable of learning numerous words and use them referentially in understanding a signal (dog, parrots)

I will sum up with a quote by Fitch (2009 p. 148) “ animals have surprisingly rich mental lives and surprisingly limited abilities to express them in signals.”

6.1.3. Discontinuity in communication

The natural communication of non-human species is very different from modern language, which has prompted the argument, popular some linguists of generative persuasion, that language has evolved de novo, with no communicative precursors.

Scholars focus on the following characteristics (Fitch (2009 p. 181-)

- * animal signals are innate, i.e. unintentional and each species, including humans, has a specific set of signals, although some species are capable of some voluntary control. e.g. some species of monkeys are capable of deceptive alarm calls,
- * animal vocal signs are arbitrary, thus, form is independent of meaning,
- * most vocalizations have simple structure, although some bird songs display similarities to syllable structure ,
- * meanings are holistic, tied to specific set of emotions and referential information, (warning calls for predators or food sources) ,

On the other hand, animals are capable of making pragmatic inferences , thus, the cognitive capacity for pragmatic interpretation is definitely not specific to language and to humans and predates the evolution of language .

Moreover, an act of communication involves at least two parties, a sender and a receiver, and communicative acts are dyadic, i.e. between a sender and a receiver, and triadic, i.e. between a sender, receiver and a referent. Animal communicative acts are 1. dyadic, i.e., non-referential, 2. they are social acts, which implies some form of social organization, a sense of membership in a group. Dyadic acts are non-referential, e.g. commands and greetings involve a sender and a receiver and do not refer to a third entity, thus, they do not represent propositions and do not

have truth values. As such they do not represent propositional thought. Dyadic communication is rampant in many animal species, not just in apes .

Primate communication is largely illocutionary, i.e. performative(Hurford, J. 2007).

Thus, animal signals are limited in range, in the information they encode and they are predominantly innately fixed, not learned. They are not different from human innate signals, e.g. laughter, smile, cries, facial expressions, etc. which usually accompany and complement linguistic signals.

Despite obvious significant discontinuity in meaning, form and function, elements of non-human communication persist in human communication and are argued to have been stepping stones in the evolution of language. Dyadic communication predates triadic communication ontogenetically and is argued to have predated also phylogenetically (Hurford, J. 2007) Illocutionary speech acts are central to language use even today, as language is used largely for “ doing things”(A. Wray, 2003, p. 126) and any referential statement can have an interpretation as a demand or request , suggesting that the ancient manipulative function of earlier forms of language is persistent although expressed in modern linguistic forms.

6. 2. Continuity in vocalizations, potential precursors for speech

6 .2.1.Vocalizations in birds and mammals

It is common knowledge that many species express their physiological and cognitive states by vocal signals. The comparative method has produced evidence for general characteristics of vocal production. All “...vocalizations are produced by a source, which converts air into sound, modified by a filter, which filters or “ sculpts “ this signal with a set of formant frequencies. “ (Fitch, 2009, p.306).

The articulatory organs, e.g. tongue, larynx, jaws, lips, velum, pharynx, are participant in vocal production in all sound-producing animals. They are dynamic biological structures which constantly alter their configurations.

Most animal calls are innate, thus, species-specific, innately predetermined, which means uncontrollable, and immutable, and so are human non-linguistic vocalizations expressing uncontrollable emotional reactions, suggesting continuity.

On the other hand, the fact that pre-human vocalizations are involuntary and under the control of the limbic system, while human speech organs are intentional and under the control of the pre- frontal cortex, pointing at discontinuity. This transition from involuntary to intended signals is believed to be one of the uniquely human adaptations for language.

Nevertheless, studies of non-human species , summarized by Hurford (2012) show that the voluntary control of vocalizations is not a uniquely human trait. For example, Gamba et al. state that “ ...facts suggest that the neocortical area homologues to the human speech area takes part in the generation and control of monkey vocalizations ...” (Hurford, ibid. p. 106). What is unique in humans is “ fine control of articulation” (Hurford, ibid. p. 107).

Significantly, great apes have voluntary control also over their facial movements.(Hurford, 2012, p. 107).

The dropping of the larynx , previously thought to be a uniquely human adaptation for speech, is found to be common to many mammal species, including dogs, primates and hominid

species, and as per Fitch (2009, p. 321) in red deer and other species has evolved for body size exaggeration, i.e. for self-defence. In humans lowering of the larynx happens in males during puberty and is part of natural growth, with no relevance to speech.

This also means that the evolutionary trajectory of the vocal organs is about 200 million year old.

Importantly, discreteness and combinatoriality in the perception and production of the sound signal, essential for the formation of phonemes and phonological systems, thought to be unique feature of language, is found in the songs of some birds as hierarchically organized combinations of discrete sounds, i.e. demonstrating phonological syntax (A.Hilliard, S.White, 2009), thus, displaying some similarities to speech. That said, bird songs display combinatoriality, but not compositionality as the sound combinations they produce have holistic meanings i.e. individual notes are not analogue to morphemes or words. In short, birdsongs lack double articulation, essential feature of human language.

In addition, vocal learning is found in birds, some marine animals and even bats and elephants (A.Hilliard, S.White, 2009; J. Petri, C. Schraff, 2011, p. 2125). The most relevant to language learning is the case of vervet monkeys, whose calls are mostly innate, although there is some learning involved as young vervet monkeys must learn to use their vocalizations in the appropriate context (A.Hilliard, S.White, 2009).

Vocal learning suggests a critical period, found in song birds. (A.Hilliard, S.White, 2009). Macaques (T. Fitch et al. 2016) and bats have demonstrated a complex system of vocalizations and echo-detection. Bats learn new vocalization throughout their lives, their vocalizations have dialects as markers of group identity, a pre-requisite for sign-based communication and their communicative interactions resemble human dialogues in turn-taking, a pre-requisite to participation in a dialogue (S. Vernes, 2017).

Nevertheless, despite obvious similarities, human physiology is specifically adapted for facilitating speech: the human speech organs are very agile: the speech organs are capable of rapidly altering their position due to efficient motor control which results in rapid change in formant frequencies. And although other animals are also able to alter rapidly the positions of their vocal organs, and Neanderthals show some capacity for speech, (Liebermann Ph. 2007 and elsewhere), the human vocal tract has a unique L shape in part because the tongue root in humans is permanently positioned down into the pharynx. This allows humans the unique ability to produce the so called “quantal vowels” /i, a, u/, which are language universals. And only in humans these abilities are exploited to the fullest for communication purposes.

Similarity in vocalizations suggests similarity in perception. Sound perception and production are asymmetrical in all non-human species. Production is much more constrained and species-specific. On the other hand, as Fitch (2009) shows, the basic biological machinery of the auditory system is shared not only in all mammals, but even more generally in vertebrates.

Categorical perception of speech sounds, i.e. the ability to form categories of speech sounds and perceive them differently, e.g. categorization of stops, initially thought to be uniquely human, is within the abilities of many animal species. In Fitch, (ibid, p. 326) the perceptual abilities for discriminating consonants is identical in humans and chinchillas. This makes sense as a general tendency in evolution, given that the accurate perception and identification of a wide range of sounds plays an important role in survival and also supports an argument that sound production, including speech production, has co-evolved with sound perception.

In sum, “...speech perception is based on perceptual processing mechanisms largely shared

with other animals.”(T. Fitch, 2009,. p. 327)
And uniqueness grows from similarity.

6 .3. The ancestors, apes and homo species

As paleontological evidence shows , all species of great apes, hominids and humans, evolved in Africa .The exact specification and labeling of the different species in the Homo branch is a matter of debates. The information presented as follows reflects the broad consensus based on the studies of Richard Leakey (1994), a recognized classic in the field.

From 7mya to 2 mya following the divergence from the great apes, the hominid evolution is marked by various stages:

- *Australopithecus, a bipedal chimp, or the walking ape, evolved about 4-2 mya. A representative of these species in the paleontological record is the skeleton known as Lucy, a bipedal, small-brained ape with brain size of about 400cc (cubic centimetres). Bipedalism as change in locomotion is explained as adaptation to change in climate and resulting formation of the savanna or grass land in east Africa.

Bipedal locomotion is not unique to these species, it is well established in bird species and dinosaurs, although unusual in mammals. It brought a cascade of physiological and cognitive alterations in the Australopithecus species, the most significant of which is the evolution of big brains and facilitated a range of new behaviours e.g. faster running after prey, freed hands for carrying food, tool use etc. Australopithecus used tools at the level of modern chimpanzee.

- * a branch of Australopithecus evolved into the genus Homo, species with larger bodies and brains. The defining characteristic of the Homo species is big brains as a ratio to the body size. The earliest representative known as Homo Habilis lived in Africa about 1,9 mya (million years ago). Large brains coincide with the appearance of primitive tools known as the Oldowan culture.

- * * the evolution of Homo Erectus from a species sister to Habilis, which diverged from a common ancestor, represents a major key stage in the evolution of the Homo branch. The Erectus species were the first to leave Africa about 2 mya and dispersed in Asia and beyond. The name Ergaster usually refers to the African version of these species and Erectus refers to the Asian version, represented in the paleontological record by the skeleton of the “ Turcana boy” with larger body, brain size of 750- 1200 cc, very close to that of a modern human , i.e.1000-2000 cc as per Fitch (2009). Homo Erectus is credited with producing major innovations, e.g. developing more complex tools, known as the Acheulian culture, which remained in use for about 1 million years, the invention of fire about 400,000 ya which requires a significant degree of cooperation. Scholars attribute the beginning of some form of primitive linguistic communication to Erectus species.

- * Homo Heidelbergensis is a Homo species with traits intermediate between Erectus and modern human, or post-Erectus hominid, the common ancestor of Neanderthals and modern human , also known as Homo Rodesiensis. The branch of these species which emigrated out of Africa evolved into Neanderthals and the branch which remained in Africa evolved into Homo sapiens .

- * Although Neanderthals evolved as a divergent branch of Homo, they display very significant bio-cognitive and behaviour similarities with us e.g. large brain, comparable to human size, large body with strong built, well adapted to cold climate, used fire, hunted large

game, demonstrated care for elders and of burials, developed the so called Mousterian culture, some attribute to them simple art and symbolism. They lived in Asia and Europe for about 500,000 years.

6.3.1.Behavioural and physiological continuity of apes and humans

Primate brains show left hemisphere bias in perception and production of conspecifics' calls, in attention to visually represented meaningful symbols, ability also detected in various species of mammals, birds and amphibians (Hurford, J.,2012, p. 119-).

Thus, for one, the ape brain is ready and able to learn word-like signs, and two, these abilities are concentrated in the left hemisphere, as are in humans.

Chimpanzees share with humans the following characteristics : large body, large brain, long life, unusually low reproductive rate.i.e one child at a time, long childhood, thus, long period of child dependence on parental care, rich social life as they live in large social groups, with primitive division of labour, e.g. male individuals hunt. They also hunt for meat, kill conspecifics, self-medicate using plants, use caves for shelter, practice tool making and use, pair binding and parental care for infants (Fitch, 2009, p. 238-)

This means that the last common ancestor of humans and chimps also had these characteristics Great apes have shown superior cognitive abilities compared to non-ape species.

They display extended social learning, i.e. learning from conspecifics, pass on animal traditions through imitation, e.g. the potato-washing monkeys, can learn by imitation and are capable of self-recognition in a mirror, have the concept of discreteness, i.e. perceive discrete objects and events and thus, have the concept of same and different, understand abstract relationships, understand complex events as ordered actions and know when to participate in them, have some concept of group membership and can organize a war, can learn a limited vocabulary (M. Donald, 1993)

These cognitive capacities in non-human animals are very similar, and probably identical, to human's, demonstrating clear continuity of cognitive functions in human evolution.

Moreover, pertinent to our understanding of origins of speech and language is an ape capacity (albeit limited), to alter the innate calls by learning and control innate emotional calls and manual gestures (Hurford 2012, p. 119-)

Significantly, great apes and even monkeys anticipate calls to have a communicative function, suggesting a primitive form of Gracian principles of signal interpretation. They demonstrate awareness of social structure as they tailor their calls to reflect the social rankings of the receivers. Apes also produce call combinations with cumulative meanings different from the meanings of the individual calls in isolation and, in addition, demonstrate ability to interpret calls in context-dependent way. Importantly, recursive conceptual structure has been demonstrated in primates. Primate communication, similar to human conversations, is multimodal, consisting of vocalizations, gesticulations, facial expressions, body posture (K. Zuberbuhler, 2015).

Some monkey species combine individual calls to create new complex meanings, suggesting some form of capacity for syntax (Hilliard, White, 2009)

Importantly, primates' thoughts are said to be proposition-like, suggesting that they intuitively understand that a thought is composed of parts, which stand in a relation to one another. And although primate propositions are primitive, e.g. with limited number of arguments, compared

to humans', given limitations in episodic memory capacities, this clearly suggests cognitive continuity.

That said, ape thoughts remain private, while humans have invented a way to make their thoughts public by linguistic and other communicative technologies, a major point of discontinuity.

6.3.2. The linguistic abilities of modern trained apes

Modern primates have been raised in human families and exposed to the same environment human children are raised, where they display remarkable capacities for learning, remembering and creatively using a large number of symbols, e.g. some have learned over a hundred symbols. Similar experiments with similar results have been conducted with an African Grey Parrot and dolphins. Thus, the ability to understand and use signs and learn large vocabularies is not specific to humans, not even apes.

Chimpanzees raised in human environment learned to communicate by gestures at a level of complexity comparable to this of young children and display the same initial stages of language learning as human children, e.g. using language in dyadic performative acts, using combinations of signs for concrete objects (S. Savage-Rumbaugh 1986 and elsewhere).

Moreover, chimps have demonstrated that they do not just memorize signs, they form categories and have symbolic representations of these, thus, they use symbols and simple combinations of them in a meaningful way, in some cases spontaneously.

Kanzi, the famous bonobo, "...have demonstrated a well-developed ability to comprehend all types...of sentences, including conditionals" (Fetzer, 2005, p.65).

Kanzi's linguistic achievements are described as follows: "It wasn't a complex language, not a language with syntax. It was more a culture language, complex set of behaviours that was the way chimps' lives were lived in the laboratory. It made one think of Homo Sapiens without sophisticated spoken language- intelligent, sensitive creatures, able to communicate and coordinate their behaviour in a collective subsistence effort" (J. Fetzer, p.67)

So, chimps and bonobos, living in labs and trained in human language (Savage-Rumbaugh 1986 and elsewhere) have displayed human-like behaviour. Thus, animals/primates in captivity develop and use, although not at the level of humans, almost all human capacities and display only quantitative differences with humans in handling the major aspects of what we call human uniqueness.

On a different note, conceptualization in any species is based on perception, i.e. concepts are embodied in all species and human concepts are reflections of human perceptual experiences with reality. And because human bodies are very similar to pre-human's, given the well known genetic, physiological and behavioural similarities, one must assume also cognitive similarity and from this, similarity in conceptualization of reality by pre-human species and sapiens, especially given the long lasting interaction among Sapiens, Neanderthal, Denisovan, and probably other similar species.

And as language is a way of making the concepts in the individual mind public, one would anticipate that communication systems of Neanderthals, Denisovans etc. and others, whether called languages or not, had similar meanings.

6.3.3. Primate social organization

Primates are social species who live in groups with complex social organization with its internal hierarchical structure. They learn from conspecifics, a sign of social intelligence(Cheney, Seyfarth, 2012, 2017;(Hurford, J. 2007, p. 191-) and pre-requisite for complex communication.

That said, their social structure is dictatorship- like. In despotic societies the social standing of individuals is fixated by the dictator's attitudes and does not allow for social mobility. This , on the other hand, does not encourage the development of negotiating skills in individuals, which, in turn, does not encourage the development of communicative complexity.

An egalitarian society, on the other hand, encourages the development of communicative skills and communicative complexity, as individuals do have prospects for improving their status by negotiating improvements in social standing.

Thus, primates have complex social lives determined by their place in the hierarchy of the social group. This is a learned behaviour, as in different social groups an individual of the same species can have a different social status. Their behaviours as group members is obviously very different from human's. Nevertheless, we have inherited from them 98% of the genome, including propensities for social skills.

6.4. The “ mirror neurons” , a precursor to protosign

To remind, “ mirror neurons” (G.Rizzolatti , M.Arbib, 1998) are a type of neurons which control motor movements and become excited both when the individual performs an action and it witnesses the action performed. It was found in brains of numerous species, from monkeys to humans, ergo, they can be traced to the last common ancestor of us and monkeys. The presence of mirror neurons is hypothesized to signal a capacity to interpret gestures as signs, thus, attach meaning to body movements, a pre-requisite for the formation of protosign and an ability to form iconic signs, learn by imitation, detected later in primate evolution, and ground symbols in physiology. It is suggested that the mirror neurons paved the way to a transition to vocal communication and speech and even the foundations of syntactic structure in motor routines or “ action grammar” (Arbib, 2004)

Arbib recently has argued for a gradual transition from a multimodal communication, dominated by gestural signs and marginal use of vocalizations, into dominant use of vocalizations and formation of speech, given the facility of speech to encode symbols.

Physiologically the shift is explained with the extension of conscious control from the monkey' s area F5 to the adjacent area of orofacial movements and sound making, resulting in replacement of gestural communication with some form of proto-speech which propels the development of spoken protolanguage (M. Arbib. 2007, p. 21-47).

For extended discussion and criticism see T. Fitch (2009 p. 462-).

Although the mirror system hypothesis is not without critics, it adds clarity in our understanding of the initial stages of language genesis by suggesting cognitive, physiological and neurological continuity.

7 . Language as invention

At present the human body has specific adaptations for the use of language. That said, given

the evolutionary principles of adaptation to pre-existing environmental features, it is natural to assume that at the very beginning language was formed by pre-linguistic bio-cognitive hardware. That is, language begun as an invention , a communicative novelty without a Language Capacity, i.e. a new behaviour made possible with available bio-cognitive machinery, adapted for use in language from other non-linguistic functions. This makes the usage-based/complexity argument for language as a communicative technology highly relevant in understanding language origins. Koster has argued that the language system is one of many forms of culture, e.g. traditions, music, science, etc. and as such arises without any pre-conceived biological blueprint (J. Koster 2009) formed as the unintended result of repeated, intentional communicative interactions of all participants, thus, the community is the collective author in the invention of language. The implication is that biological organs are flexible and can be recruited for new functions on demand i.e. “ The set of functions that can be fulfilled by a given structure is potentially as infinite as the set of possible environments “ (Koster, *ibid.* p. 9).

Koster's argument has deficiency explaining the formation of modern language given the proven presence of a Language Faculty. That said, the argument is much better suited for explaining the earliest stages of language as the initial environmental pre-conditions prompting the evolution of a Language Faculty.

In sum, language origins are reasonably explained with the “ ceaseless, unpredictable creativity” of the bodies and minds of ancient homo and human species.

7.1.Producing speech without speech capacities

Cultural innovations are made by capacities evolved for essential biological functions: writing and painting is performed by the hands and arms, the feet operate bicycles, the whole body is the instrument in dancing.

On the assumption that language has originated as a new function by reusing available capacities, one would expect the first attempts to articulate speech to have been made with pre-existing capacities evolved for essential functions of biological survival.

The frame-content theory of speech evolution (Davis, MacNeilage 2004; MacNeilage , 1998) argues that lip smacking motions in primates were co-opted for the formation of the syllable and its most basic structure as CV. In addition, a plausible argument is made by Ralph-Axel Muller (1996) that the syllable is an instance of a “ general motor alphabet”, a unit of body motion, which has assumed a new role in speech production. The emergence of the syllable as a linguistic unit is traced back to the “ rhythmic jaw oscillation” and the rhythmic nature of the movements of the body-extremities. Muller also suggests that this “general motor alphabet “, not specific for any group of muscles and comparable across species, when applied to the human vocal tract, yields the set of phonemes found in human languages.” (Muller, *ibid.*). Vowels are argued to originate from neonate “comfort noises” (Muller, *ibid.*).

Thus, the phoneme and the syllable as linguistic units have originated from general tendencies and principles of biology and physiology, used in pre-linguistic life forms for purely biological functions which have received a new function of producing speech . Needless to say that early speech was slow, clumsy and inarticulate, i.e. inefficient in both production and perception.

7.1.1.The emergence of language as self-organization in primate multimodal communication

Multimodal communication is defined as a composite integrated signal containing simultaneous emission of signals from multiple modalities, e.g. vocalizations, gesturing, body posture, facial expression.

Recent studies of primate communication have the potential to shed light on the initial stages of the beginning of language as a vocal component of a multimodal communicative complex which includes non-linguistic vocalizations, intentional and instinctive, manual gesticulations, facial expressions, etc..

Slocombe, Waller and Liebal (2011) show that modern primates' communication is multimodal. For example, chimps in captivity communicate with signals emitted simultaneously from a variety of modalities. They recognize and produce visual, tactile, vocal signs in their communication with caregivers. "... primate communication is inherently multimodal, at both at behavioural and neuronal level, indicating that unimodal research tells only part of the complex story." (Slocombe, Waller, Liebal 2011).

Waller et al. (2013) show that often one modality provides multi-sensory stimulus, e.g. a gesture provides visual, tactile, auditory, etc information, thus, a signal is usually composite, a combination of various components which act in concert to deliver the complete message. In this sense studying each component in isolation ignores the role of the rest. For example, "When a slap is paired with a playface it leads to play. Isolating the slap from the playface will not, therefore, help us understand the signal better." (Waller et al. 2013 p. 540).

In short, "communicative complexity may be less about how each single modality is used, and more about signal integration." (Waller et al. 2013, p.540). In this sense for Waller and colleagues the common practice of studying the primate communication by breaking it down into distinct unimodal components, each studied in isolation, is mistaken, given that facial expressions as well as vocal signals in subhuman species are used simultaneously and are under voluntary control. This isolationist method is argued to be inadequate as it fails to capture an important side of communication, its inherent multimodality.

Traces of these remote beginnings of language are observed as modern human communication is also multimodal, as humans routinely combine language with non-linguistic sound-making and other non-verbal cues, e.g. facial expressions, body posture etc., to complement the linguistic message.

In this context one would estimate that from that initial point language evolved into as the major component in this multimodal complex which has partially displaced the role of the rest of components, attributing them only a supportive role. That is, language has emerged as the result of self-organization in multimodal communication as formation of social conventions in Heylighen's sense (Heylighen 2013)

7.1.2.Multimodality in early child language development

The argument that ontogeny represents a condensed version of phylogeny is well known in the

literature and ,although it has its supporters and critics, it provides a possibility to infer information, permanently lost and non recoverable, pertinent to the evolution of language, from observation of development. In this sense the multimodal nature of the first communicative engagements at early age could shed light at initial stages in language evolution.

In the first weeks and months after birth the infant displays affinity for socializing and participating in communicative behaviour by visible body actions, e.g. smile in response to a smile, imitate facial expressions, demonstrate turn-taking in non-verbal sound making and playful interactions ,facial expression, gesticulations, body posture, direct attention by sound making, looking, pointing. The youngster's first experimentation with language begins with babbling , followed by one-word utterances where children 's communicative interactions are predominantly by body movements with language having a limited and supporting role, which broadens with the next stages of language development with two-word utterances and beyond. As development progresses, although the child's linguistic abilities become more sophisticated, his/her linguistic output continues to be complemented by non-linguistic body signals. During the development of syntactic competence, as children often omit arguments or elements of constructions, potentially compromising the overall understanding of the message, the lack of full linguistic competence is mediated by body actions, e.g. pointing, etc. Linguistic behaviour is constantly synchronized with communicative use of other modalities as a component of a multimodal unified system (Morgenstern A.,2014). And although the use of body gesticulations as compensation for linguistic immaturity gradually decreases as the child's linguistic skills develop in full , to paraphrase Morgenstern A.(2014), body signals continue to illustrate, specify, reinforce, modify the meanings of linguistic production (ibid. p. 18) In sum, the participation of multimodal non-linguistic signals in early linguistic behaviour in young children suggests that early language also emerged as a component of a multimodal communicative ecology.

7.1.3.Speech as emergent

The crucial stage in the origin of speech is the ability to produce and perceive discrete sounds and sound combinations.

A general principle of particulation , known as “ particulate principle of self-diversifying systems” (Studderd-Kennedy, M. 1998) as part of the general process of self-organization in nature by which holistic entities are broken down into discrete primitives, further combinable into hierarchical units , e.g. chemical compounds, the genetic code etc. is evoked to explain the formation of speech.

In recent years various experiments have been performed which illustrate the emergence of speech where the specifics of the human organism are mimicked by units of artificial intelligence. The formation of reusable units is shown as crystallized points along a sound continuum, formed under the pressures for ease of articulation and auditory distinctiveness. (Oudeyer, P., 2006; de Boer 1998). The process develops as follows :

1.Particulation, the formation of discrete units as clusters along a continuum.

“... from the continuous space of possible gestures , speech carves out basic building blocks which it reuses systematically. The phonemic and gestural continuum becomes discretized. “ (Oudeyer,P. 2006, p.24)

2. Formation of phonetic gestures, units of action.

3. Formation of stable combinations of parallel coordinated gestures, phonemes, which are reused according to patterns of vocalizations. The articulation of a phoneme is a parallel articulation of several phonetic gestures. Gesture combinations which produce narrow constriction or full closure of the air passage produce consonants, vowels are produced with minimal or no constriction.

4. Formation of reusable phonotactic patterns, i.e. syllables, displaying universal tendencies in syllable structure.

In the experiments the robots interact and as a result converge on a limited number among the many possible combinations of gestures or phonemes, and phonotactic patterns. The fact that different languages have similar but not identical choices of phonemes and syllables shows that they are the result of self-organization processes and not innate phonological universals. The experiments in artificial intelligence describe a hypothetical state of affairs as they are based on formalisms which only approximate the functions of the human organism. That said, the results are in unison with the linguistic reality: the majority of languages use about 30 phonological segments, about 22 consonants and 5 vowels and 95% of languages use only 3 places of articulation.

Two constraints which bootstrap and shape the phonological systems, both are of biological nature, are pointed at:

1. Articulatory-acoustic interdependence. The sound patterns are formed through dynamic interactions between the organs of perception on one hand, and the organs of sound production on the other. Patterns which are the most frequently heard are the patterns most frequently pronounced, creating patterns of joined neuronal activity. In this way speech patterns emerge as a result of repeated interactions.

2. Energy cost. Achieving more with less effort is one of the main principles of in any activity. In speech the power of this constraint is exemplified in syllable structure as CV is universally preferred (to VCV, CCVVC etc.) as it requires less muscular energy from the articulatory organs, even though the less preferred options contain the same phonemes. (Oudeyer P. 2005) In this sense, one would agree with Liebermann (2007 and elsewhere) that the quantal vowels /i,a,u/ which require more articulatory flexibility than other vowels, would be beyond the articulatory abilities of the ancient communicators and excluded from the phonemic arsenal at the very beginning of language. Given the argument for speech emergent in infancy with underdeveloped speech capacities, one would assume similarities with the speech abilities of the ancient language speakers and extrapolate that the first sounds babies master, e.g. /p, t, k, m, n, a, h/. would provide a window into the phonemic arsenal ancient words were composed.

7.2 .Self-organization in the conceptual space

The principles of self-organization which explain the formation of the phonological system can also explain the formation of meanings of words.

Similar experiments with artificial agents were designed to illuminate the formation of vocabulary (L. Steels, 1995).

Studies of early language attainment by youngsters could provide a road map to understanding the origins of word meanings at the earliest stages of language evolution. Emergentist perspectives of language development understand word learning in infancy as process of

emergence of novel neuronal pathways in the young brain with no pre-established innately preconfigured neuronal assemblages (B.MacWinney 2005). Here word learning is a cognitive process of coordinated self-organization among three “ local maps” of neurons during which patterns of associations are formed and reinforced, i.e. auditory map, a concept map and articulatory map . This means that the conceptual space, and the sound space are configured through experience without a Language Faculty.

The assumption that language learning in human infants is made in the absence of a Language Faculty is challenged as propensities for word learning are detected at the very first experiences with language (P. Bloom, 2000)). In this sense it is pertinent to talk not of absent, but of underdeveloped Language Faculty. Nevertheless, it is reasonable to suggest similarities between the two and extrapolate from accessible knowledge of language attainment, and form plausible speculations on the beginning of language. Thus, findings and theoretical ramifications of language learning by human infants in modern times could provide insights into language origins.

Koster (2009) points out that , word meanings are stored as “ information clusters” produced by self-organization both at the level of the idiolect and the sociolect and vary from person to person, depending on people's experiences, thus, they are not universal. They are stored in association with a form, e.g. sound combinations. One could assume a similar process of self-organization in the minds of the ancient language speakers.

Importantly, “ information clusters” are said to contain vague information about some part of the environment, underspecified potential meanings, which become specified in the context. They are stored in the brain only in association to something external, something in the environment, which points at possible contexts of use . Thus, linguistic entities cannot be detached from experience, they must point at potential applicability in communication.

To note, in order for communication to be possible, there must be a considerable overlap in the stored information of the speakers because they share the same natural environment, have very similar physiology and function in the same culture and thus, have very similar experiences.

A similar convergence of combination of information storage and potential for referential use is likely to have happened in the minds of ancient speakers.

Thus, some theoretical perspectives outlining the evolution of modern linguistic forms are better suited for understanding the initial stages in language origins.

8. What can you say without a Language Capacity

The formation of primitive language forms as novel behaviour presumably implicates the coordination of a number of pre-existing traits, e.g. cooperation among various cognitive capacities, between cognitive and physiological capacities and activities. In this sense the emergentist perspective, designed to explain the formation of modern language, can provide guidelines for understanding the origins of language from coordinated synergistic activities at multiple levels:

a. synergy among organs and systems in the individual organism:

* synergy among the articulatory organs for the purpose of speech production

* synergy among cognitive capacities :

capacity for reference (to represent a class of objects through signs (as special case symbolic

thought, symbolic representation)

capacity to form categories (things and actions are universal categories)

capacity for mind-reading, or theory of mind

capacity for self-monitoring, or metacognition, (Studard-Kennedy, Knight, Hurford, 1998).

consciousness (awareness that one's person and mind differ from others)

intentionality (stimulus-free initiation)

capacity to learn , extended memory

imagination, planning or displacement, i.e. capacity to refer to distant referents, distant from here and now

capacity for socialization i.e, need for the company of conspecifics

* cognitive and behavioural synergies among the members of a group:

synergy in behaviours resulting in common cultural practices,

synergy in conceptualization resulting in the so called “ common ground", i.e. standardized views of reality reflected in myths, folklore, etc. which makes possible the formation of a vocabulary.

Importantly, some of these capacities are uniquely found in human species, e.g. consciousness, imagination and planning, others are demonstrated to some lesser degree in pre-linguistic species, e.g. intentionality, socialization, learning and use of signs, and yet others underly broad continuity with distant branches on the tree of life, e.g. capacity to learn, capacity to form categories, to communicate by vocalizations, etc. suggesting that, although modern human bodies and brains have evolved specific adaptations making modern language possible, in the absence of such adaptations some form of language-like sign system is possible by reliance on creative repurposing of available pre-linguistic properties.

Summary and conclusions

The search for the origins of language is challenging and rife with conjectures and speculations as credible evidence of processes and events pertaining to our remote past is hardly available.

Nevertheless, the present could be a good indication of the past.

Given the close similarity of form and function between language and other sign systems, the formation of these, observable in modern times, one would anticipate, share common patterns and principles of formation. For example, formation de novo of new languages and other cultural sign systems could inform of language origins by extrapolating common principles and processes underlying the formation of sign systems.

Language learning in infancy by underdeveloped human brains can shed light on language formation with limited bio-cognitive resources.

Modern languages contain some clues in traces of ancient forms preserved in modern language are windows into the otherwise unknowable past of language origins.

Theoretical ramifications of language emergence in the absence of a Language Faculty, although aim to explain modern language formation, are better positioned for tracing language origins.

In addition, principles and processes underlying broader nature, e.g. self-organization and emergence could explain the formation of the earliest linguistic forms in the minds and

communities of ancient language speakers.

There is a lot yet to discover about the beginnings of language, but there is a lot we can learn by looking closely at what we already know, where knowledge of the past is hidden in our present. We just have to notice and use it.

References

Aitchison, J. (2000) *The seeds of speech: language origin and evolution*, Cambridge University Press

Arbib, M. (2004) From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics, in *Behavioural and Brain sciences* , 28: 105-167

Arbib M.(2007) The mirror system hypothesis, how did protolanguage evolve, in Tallerman,M. ed. *Language origins, Perspectives on Evolution*,Oxford University Press, p. 21-47,).

Aronoff, M., Meir, I., Padden, C., Sandler, W.(2008) The roots of linguistic communication in a new language, *Interaction studies*, 9:1,p, 133-153, doi: 10.1075/is.9.1.10aro

Barr, D. (2004) Establishing conventional communication systems. Is common innate knowledge necessary? , in *Cognitive Science*, 28, p.937-962

Berling, R. (2005)*The talking ape, how language evolved*, Oxford University Press

Bickerton, D. (1984) The Language Bioprogram Hypothesis, in *Behaviour and Brain Sciences*, vol.7, p. 173 -221

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

Bickerton, D. (2009) *Adam's tongue: how humans made language and how language made humans*, Hill and Wang, Publishing, New York

Bloom, P. (2000) *How children learn the meanings of words*, MIT Press

Chomsky, N. (1972) *Language and Mind*, Cambridge University Press

Chomsky, N. (1980) *Rules and Representations*, Columbia University Press

Chomsky, N. (1986) *Knowledge of Language, Its nature, origin and use*, Greenwood Publishing

Chomsky, N. (1988) *Language and the Problem of Knowledge*, MITPress

Chomsky, N. (1995) *The Minimalist Program*, MITPress

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

Clark, B, (2011) Scavenging, the stag hunt and the evolution of language, *J.of Linguistics*, 47, p.447-480

Croft, W.(2017) Evolutionary complexity of social cognition, semasiographic systems and language, in Mufwene, S., Coupe, C., Pellegrino, F. eds. Complexity in language, developmental and evolutionary perspectives, Cambridge University Press, chap. 5

Cysouw and B. Comrie (2013) Some observations on the typological features of hunter-gatherer languages , in Bickel et all. Eds., Language typology and historical contingency, in honour of Johanna Nickols, John Benjamins p. 383-394, doi.org/10.1057/tsl.104.17cys

Davis,B., MacNeilage, P.(2004) The frame-content theory of speech evolution: from lip smacks to syllables, Primatologie, vol 6, p. 305-328

Deacon, T. (1997) The symbolic species, Norton

de Boer, B.,(1998) Emergence of sound systems through self-organization,
www.ai.rug.nl/~bart/naic98.pdf

Donald, M. (1991) Origins of the modern mind: three stages in the evolution of culture and cognition, Harvard University Press

Dowman, M. (2007) Protolanguages that are semi-holophrastic, European Conference on Artificial Life: Advances in artificial life, p.435-444

Everett, D. (2005) Cultural constraints on grammar and cognition in Piraha: another look at the design features of human language, Current anthropology, 46(4), 621-634

Fetzer, J. (2005) The evolution of intelligence: are humans the only animals with minds, Open court, Chicago

Fitch, T. (2009) The evolution of language, Cambridge University Press

Fitch T.et all. (2016) Monkey vocal tracts are speech-ready, Science Advances,vol.2, n.12, 2016, DOI: 10.1126/sciadv.1600723

Gil, D. (2009) How much grammar does it take to sail a boat? In G. Sampson, P. Trudgill, eds. Language complexity as an evolving variable, Oxford Univ. Press

Heilighen,F. (2013) Self-organization in communicating groups, the emergence of coordination, shared references and collective intelligence , DOI: 10.1007/978-3-642-32817-6_10

Hilliard A.,White S.(2009) Possible precursors of syntactic components in other species, in Bickerton D., Szathmary E. eds. Biological Foundations and Origin of Syntax, MIT Press, p.161-185

- Hurford J. (2007) The origins of meaning, Oxford University Press
- Hurford, J. (2012) The origins of grammar, Oxford University Press
- Hyman, J. (1983) Iconic and economic motivation, Language, vol 59, No.4, p. 781-819
- Jackendoff, R. (2002) Foundations of Language, Oxford University Press
- Jackendoff, R., Wittenberg, E., (2014) What can you say without syntax, a hierarchy of grammatical complexity, in Newmeyer, F., Preston, L. eds. Measuring Linguistic Complexity, Oxford University Press, chap.4
- Koster, J. (2009) Ceaseless, unpredictable creativity, Language as technology, in Biolinguistics, vol.3 No.1, p. 061-092
www.let.rug.nl/koster/paper/language%20as%20technology.pdf
- Krause et al. (2007) The derived FOXP2 variant of modern humans was shared with Neanderthals, Current Biology, p.1908-1912
DOI: 10.1016/j.cub.2007.10.008
- Leaky, R. (1994) The origin of humankind, Basic books
- Lenneberg, E. (1967) The Biological Foundations of Language, Wiley and Sons
- Lieberman, Ph. (2007) The Evolution of Human Speech, its Anatomical and Neural Bases, Current Anthropology, vol 48, No.1. Feb., p. 39-66
- MacNeilage, P. (1998) The Frame-Content Theory of Evolution of Speech Production, BBS, 21, p. 499-546
- MacWinney, B. (2005) Language emergence, five timeframes and three illustrations, in Burmeister, Piske, Rohde, eds. An Integrated View of Language Development, Papers in honour of Henning Rhode, p.17-42, Trier: Wissenschaftliche Verlag
- Mitten, S. (2005) The singing Neanderthals: the origins of music, language, mind and body, Harvard University Press
- Monaghan, P. et al. (2014) How arbitrary is language, Philosophical transactions, The Royal Society Publ. Doi:10.1098/rstb.2013.0299
- Morford, J. (2002) Why does exposure to language matter? in Givón, T., Malle, B. eds, The evolution of language out of pre-language, John Benjamins Publishing, p. 329-343
-

Morgenstern ,A., (2014) Children's multimodal language development, in Facke,C. ed. Manual of language acquisition, De Gruyter, p. 124-142

Mueller R. A.(1996) Innateness, autonomy, universality? Neurobiological approaches to language, BBS , ,Vol. 19 ,issue 4, p. 611-675

Noble, W. Davidson ,I., (1996) Human Evolution, Language and Mind, a psychological and archeological inquiry, Cambridge University Press

Oudeyer, P.Y. (2005) Self-organization of combinatoriality and phonotactics in vocalization systems, Connection Science , vol 17,no. 3-4, 2005,Taylor and Francis p.325-341

Oudeyer, P.Y. (2006) Self-organization in the evolution of speech , Oxford University Press

Perniss, P., Thompson,R., Vigliocco, G.(2010) Iconicity as a general property of language , evidence from spoken and signed languages in Frontiers of psychology, 1:227,

Perniss, P.Vigliocco, G.(2014)The bridge of iconicity: from world of experience to experience with language, Philosophical transactions of Royal society B, Royal society publ.

Doi:10.1098/rstb.2013.0300

Petri, J., Schraff, C.,(2011) Evo-devo, deep homology and FOXP2, implications for the evolution of speech and language, Philosophical Transactions of Royal Society B, p. 1-17; doi: 10.1098/rstb.2011.0001

Pinker, S. , P. Bloom (1990) Natural language and natural selection. BBS 13(4), p. 707-784

Prigogine, I., Stengers, E.(1993) Order out of chaos, man's new dialogue with nature, Bantam books

Ramachandran, V., Hubbard, E.(2001) Synaesthesia, a window into perception, thought and language, Journal of consciousness studies, 8,No.12, p. 3-34

Ridley, M. (2004) The agile gene: How nature turns on nurture, Harper Perennial

Rizzolatti, G. Arbib, M. (1998) Language within our grasp, Trends in neuroscience, 21(5)p.188-194

Sauvage-Rumbough, S. (1986) Ape Language : from conditioned response to symbol, Columbia University Press

Seyfarth, R., Cheney,D. (2012) Primate social cognition as a precursor to language, in Tallerman, Gibson eds. Oxford Handbook of Language Evolution, Oxford Univ. Press, DOI: 10193/oxfordhb/97801119.13.0004

Seyfarth, R., Cheney, D.,(2017) Precursors to language, social cognition and pragmatic inference in primates, in *Psychon Bull rev*, 24(1), p. 79-84
DOI: 10.3758/s13423-016-1059-9

Slocombe, K., Waller, B., Liebal , K. (2011) The language void: the need for multimodality in primate communication research, *Animal behaviour*, 81 (2011) 919-924

Steels, L. (1995) A self-organizing spatial vocabulary, *Artificial life*, 2(3), p.319-332, MIT Press

Studdert-Kennedy, M. (1998) The particulate origins of language generativity: from syllable to gesture , in Studdert-Kennedy, Knight, C., Hurford, J. eds. *Approaches to the evolution of language*, p.

Tallerman, M. (2005) *Language origins, perspectives on evolution*, Oxford University Press

Vernes, S. (2017) What bats have to say about speech and language, *Psychonomic bulletin and review*, vol. 24, issue 1, p. 111-117

Waller, B., Burrows, A., Liebal, K., Slocombe, K., (2013) How can a multimodal approach to primate communication help us understand the origins of communication, *Evolutionary psychology*, 11(3) 538-549

Wray ,A.(2003)*The transition to language*, Oxford University Press

Zuberbuhler, K. (2015) Linguistic capacities of non-human animals, *Wire's Cognitive Science*, vol.6.issue 3, Wiley Online, p.313-321

* *Svetlana T.Davidova* is a linguist, unaffiliated researcher based in Toronto, Canada
address for correspondence: svetlana.t.davidova@gmail.com
