EasyChair Preprint

№ 2123

Inquiries about the Evolution of Merge

Dieter Hillert

January 4, 2020

Inquiries about the Evolution of Merge

Dieter Hillert

We discuss the Merge concept of the generative paradigm in context of the evolving language capacity in the human lineage. Merge is shared to some extent across cognitive domains and might have evolved for purposes other than language. Not only limitations, but also fruitful extensions of the generative Merge concept are discussed. A framework will be introduced, the RC Model, which emphasizes an integrative approach towards the phylogenesis of the Modern Language Capacity by considering linguistic or neurobiological evidence or fossil evidence. We show that external syntactic Merge meets the criterion of evolvability and may result from different precursor stages of the modern language capacity. We conclude that syntactic Merge as found in modern languages is a byproduct of sociocultural accumulations while phonological and semantic-conceptual properties may point to our genetic disposition for language presumably shared with H. erectus.

Keywords: *Broca's area, H. erectus, Syntax, Language Evolution, Merge*

1. Introduction

One of the most fascinating cross-disciplinary research programs is related to the fundamental question how the language-ready brain evolved after the final split from genus Pan about 6 million years ago (mya). Only snippets of indirect evidence based mostly on fossil records and genetic material are available to model the components leading to the language-ready brain for fully-fledged modern languages. Here, we use the term Modern Language Capacity (MLC) to differentiate between the capacity for modern fully-fledged languages and capacities for possible precursor stages of modern languages such as the Protolanguage Capacity (PLC). The evidence discussed below shows that language as part of the computational mind is the result of reciprocal biological and cognitive exaptations/adaptations. From an evolutionary viewpoint, the analysis of intrinsic computations is essential for understanding the emergence of the MLC. Intrinsic computations are considered as unconscious, possibly genetically triggered inherent processes of the cognitive mind only indirectly accessible by introspection and empirical research. A specific cytoarchitecture grounded in the human-specific genotype is the foundation for this intrinsic cognitive capacity to acquire intuitively fluent language skills during early

---

I am grateful to Michael Arbib and Masumi Wakita for their valuable comments and suggestions.

childhood in a social-interactive context that any kind of intentions, emotional states or thoughts can be expressed and exchanged. Human and nonhuman primates share to some extent cognitive components such as communication, social behavior, vocalization, gestures, categorization, or memory that might have evolved for purposes other than language (e.g., Arbib, 2005; Pinker & Jackendoff, 2005; Seyfarth & Cheney, 2014; Fitch, 2017).

It is also quite plausible that defining, categorizing, and notating nonverbal information may have significantly contributed to the enhancement of cognitive capacities and respective memory systems in the nonverbal domain. As such language might be the trigger for cognitive enhancement of various skillsets. For example, medieval monks and nuns invented in Western Europe music notations by using first dashes and dots (*lat.* notae) to illustrate melodic up- and downward movements. Later, church musicians put notes on a grid of lines and spaces (staff) to show the melodic distance. Over the centuries these signs developed into notational symbols essential for specifying rhythm and pitch precisely. An analogy can be found in writing notations such as the transformation from pictographs to kanji characters. At the same time, we find fundamental language-independent cognitive properties shared by different domains and grounded in the human genotype. For example, when listening to a melody or a sentence, we intuitively use phrases to segment the stream of sounds; or we use prototypes to organize word meanings, tones (middle C) or chords (root position), or we rely on preferred cognitive strategies which are perceptually grounded and thematically encoded, such as Agent-first and Focus-last (Bever, 1970; Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976; J. Fodor, Bever & Garrett, 1974; Jackendoff & Wittenberg, 2014); or we use certain patterns such as binary forms in language (e.g., adjective + noun: noun phrase), in music (e.g., AABB), or in visuospatial design (e.g., meeting place + tower: church). Here, we define the MLC not to be language-specific at the cognitive or biological level, but as a language-related manifestation of an innate cognitive capacity to be considered *ad hoc* as species-specific. This innate capacity may partly consist of universal principles or computations shared across domains, and modality- and domain-specific extensions such as the MLC are expressions of this underlying capacity. It is, however, open to debate, which aspects are the outcome of cultural accumulations and which are genetically anchored.

The generative tradition solely focuses in a narrow sense on the basic syntactic operation Merge (Berwick & Chomsky, 2016; Chomsky, 2017). Merge in its basic operation combines two syntactic objects α and β to form the new syntactic object γ including the set $\{\alpha, \beta\}$. Hereby, lexical features are checked to exclude ungrammatical combinations (Chomsky, 2000). For example, the word *eat* with the syntactic features “verb” [V] and nominal uninterpretable [uN] permits the set formation {eat, avocado} labelled with the head “eat” to form a verb phrase (VP). Accordingly, the new syntactic object γ (head) can be again combined for further set formations: $\{. . . \{\gamma, \{\alpha, \beta\}\} . . .\}$, that is, Merge can recursively apply to its own output in a binary fashion to generate a sentence. Theoretically, this internal Merge (IM) operation can be iterated indefinitely. While external Merge (EM) seems to correlate with argument structure, IM with discourse information such as topic or new and old information or scope

(Chomsky, 2005). Merge is considered as a unique property of the "narrow language faculty" of modern humans that generates complex hierarchical syntactic structures and interfaces with the conceptual-intentional (CI) and sensory-motor (SM) system (Hauser, Chomsky & Fitch, 2002).

A more recent debate on IM (syntactic recursion) questions the empirical validity of this statement (e.g., Everett, 1987; Nevins, Pesetsky & Rodrigues, 2009; Everett, 2009; 2013; Futrell, Stearns, Everett, Piantadosi & Gibson, 2016). According to Everett (2013), his fieldwork of the indigenous Amazonas language Pirahã shows that sentences of this language have upper bounds, that sentential long-distance dependencies, syntactic recursion (no phrase within phrases) and expandable possessives do not exist, and there is no clear evidence of PS-rules. There is, however, no disagreement about the fact that all languages are recursively unbounded at the discourse level.

In general, many polysynthetic languages with a rich verb morphology tend to avoid subordinate structures. For example, in Bininj Gun-wok (Australian Aboriginal language) a single embedding occurs at the morphological level (a verb can embed another verb) and can be described by a finite-state grammar; or Riau Indonesian does not make use of function words, syntactic categories and recursion, it has virtually no morphology, and the word order is based on semantics (Gil, 2014). As discussed below, other examples of present-day languages show that "linear grammar" is not an uncommon property. IM is based on a linguistic constituency analysis which focuses on English or English-like languages. Many languages, however, make only limited use of it or lack it altogether (Evans & Levinson, 2009). Only theoretically, natural languages can make use of a recursive grammar algorithm that guaranties an infinite output. Two center-embeddings (e.g., *Mike who Susan who Paul met greeted ate an avocado*) cannot be at once understood and even the meaning expressed by one center-embeddings is avoided in spoken language. Although tail-recursions (e.g., *Paul met Susan who greeted Mike who ate an avocado*) are much easier to understand, tracing back the information has its limits (Christiansen & Chater, 2015). The property "recursion" is therefore not a defining or biologically anchored feature of the MLC. The recursive capacity appears also not to be restricted to the language domain as it can be also found in music (Lerdahl & Jackendoff, 2006; Katz & Pesetsky, 2011), vision (Pinker & Jackendoff, 2005), action (Koechlin & Jubault, 2006), or pragmatics (Levinson, 2013).

The recursive syntactic capacity (not its usage) seems, however, to be limited to our species. Fitch and Hauser's (2004) experiment with cotton-top tamarins (*Saguinus oedipus*) show that these monkeys were not able to process the phrase-structure grammar (A^nB^n) but spontaneously succeeded to process the finite-state grammar (AB^n). Apparently contrary findings with European starlings or baboons turned out to be misinterpreted and were the outcome massive conditional training rather than of spontaneous cognition (Gentner, Fenn, Margoliash & Nusbaum, 2006; Rey, Perruchet & Fagot, 2012). In using similar artificial grammar materials, human fMRI studies were also criticized as participants did not learn the phrase-structure grammar but applied a word class counting strategy or grammaticality was confounded with repetition structure of letter strings (De Vries, Padraic, Knecht & Zwitserlood, 2008; Rohrmeier, Fu &

Dienes, 2012).

One important extension of Chomsky’s (1995) Minimalist Program (MP) is to use the concept of (external and internal) Merge for dimensions which interface with nonlinguistic systems or which are completely independent or on which language is “parasitic”. These dimensions include, for example, figurative language (Jackendoff, 1997; Hillert, 2014), music (Koelsch, Rohrmeier, Torrecuso & Jentschke, 2013; Fitch & Martins, 2014); arithmetic (Leslie, Gelman & Gallistel, 2008; Monti & Osherson, 2012; Dehaene, Meyniel, Wacongne, Wang & Pallier, 2015; morality (Hauser & Watamull, 2017); or action (Wakita, 2014; 2016; Fitch & Martins, 2014). Another important extension of the MP is to ask how Merge might have evolved in the hominin lineage (Berwick, Okanoya, Beckers & Bolhuis, 2011; Janik, 2013).

Syntactic modelling in the generative tradition started in the 1950s (e.g., Chomsky, 1953; 1957) and triggered since then experimental research in the cognitive and brain sciences. In terms of the evolution of the syntactic capacity not only principal design issues take center-stage but also questions about the ontology of linguistic syntax (Chomsky, 2017). From an evolutionary perspective, some basic syntactic structures recruit broad cognition such as Merge and Unification (Jackendoff, 2011b) or they are shared to some degree by different domains such as phrasing or long-distance dependencies in language and music (Rohrmeier, 2011) or they are specific to language such as the coordinate structure constraint (Keshev & Meltzer-Asscher, 2017). If the basic property of language, recursion *per se*, is not unique to language, then this property is not informative about the distinctiveness of language (Jackendoff, 2011a). In evolution, for example, action principles may have been exapted from vision principles and principles in language from action-vision principles. Accordingly, it is extremely critical that syntax in the generative tradition (or in other linguistic camps) is theory-immanent detached from neurologically or cognitively relevant properties (e.g., Arbib, 2019). Modelling the MLC requires research on how language might have evolved in the hominin lineage by focusing on those neurally implemented cognitive components humans and nonhuman primates share or do not share and which may have played a key role in the emergence of the MLC (Christiansen & Kirby, 2003). Formal-generative principles are indeed an excellent tool to model cognitive capacities across different domains but only if those formalisms are empirically grounded in terms of human cognition.

The title of the present research has been questioned as Merge would not “meet the criterion of evolvability” (Chomsky, Gallego & Ott, 2018) and accordingly it has been speculated that Merge arose from a "cognitive innovation" triggered by a single-mutation in an individual subject ultimately spreading to a group. The key argument is that Merge cannot be decomposed into smaller operations such as “half-Merge” or “proto-Merge” as it is an all-or-nothing operation, either the operation exists or not (Berwick & Chomsky, 2016; 2019; see also R. G. Klein, 2008; Bolhuis, Tattersall, Chomsky & Berwick, 2014). This stance indeed defends the single-mutation account declaring that a small genetic alteration caused cortical rewiring and empowered Merge as a new powerful cognitive operation to generate hierarchical structures between syntactic objects in terms of a context-free grammar. The newly suggested and backdated timeline of this event 100-200 kya in the Middle Paleolithic refers to the emergence of anatomical modern humans along with the sudden appearance of undisputed technology and symbolic artifacts. Although this evolutionary scenario might be possible, it is a relatively isolated view in considering the multifaceted data available. We discuss below a series of findings regarding cortical growth, complex social behavior, backdating of symbolic activity and the emergence of different Homo species that point to a different scenario about how the MLC might have evolved.

Merge will be considered here as a cognitive operation not specific to language and subject to evolvability as we discuss further below. External and IM (recursion), for example, may have evolved in different steps, whereas IM depends on the emergence of EM (Martins & Boeckx, 2019). If it is argued that external and IM emerged at the same time in a whole step (no half-Merge), the debate implies the mind-body problem (Boeckx, 2017; Martins, 2019): At which level(s) of description did Merge actually emerge? So far, the computational level of Merge or any kind of mental computation cannot be directly mapped onto the physical world, ranging from biophysical properties of synapses to the function of neural circuits and systems. In fact, it is unclear how discrete mental computations are neurally implemented. It is important therefore to search for the neural building blocks that enable the formation of structural hierarchies. The biophysical properties inform us whether Merge is a language-independent, incremental process that evolved in different evolutionary steps. It is an empirical question whether Merge meets the criterion of evolvability. Because the IM operation replicates or copies the operation of EM, it may have been not subject to preceding steps. In the attempt to define language in neurobiological terms, however, it turns out that IM *per se* is not significance because of its limited use in many natural languages.

As we discuss below, however, it is highly plausible that EM evolved from intermediate steps. Here, the hierarchy of grammatical complexity is useful, from 1-, 2-word and string grammar to simple and recursive PS-grammar (Jackendoff & Wittenberg, 2014). Other intermediate steps are also possible such as that Merge built first a symmetric phrase without a head as in exocentric structures and without extended lexical constrains. The only constrains might have been that two distinctive concepts must be merged with basic lexical features (e.g. the object of *eat* must be [+food]). At the same time, it is plausible that the asymmetry found in Merge is not the result of a syntactic process. For example, Merge operates also at the morphological or phonological level respectively in form of affixations or syllabic structures. The asymmetry generated by Merge might have been therefore exapted from the consonant-vowel-consonant (CVC) cluster asymmetry [C [V C]] due to articulatory and acoustic reasons (Carstairs-McCarthy, 1999; Jackendoff, 1999). Presumably, also exaptations took place during the course of dynamic adaptions to mental computations placing Merge outside of the language domain (e.g., Fujita, 2009; 2017; Fitch, 2011; Jackendoff, 2011a).

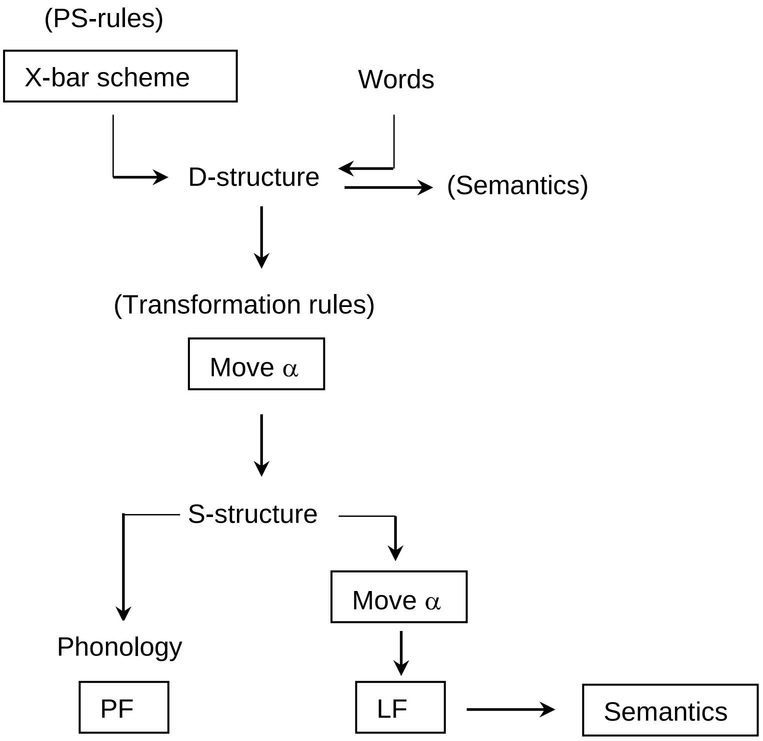
On theoretical grounds Merge is decomposable and may thus point to the gradual emergence of Merge rather than to a single-mutant account (e.g., Boeckx, 2013; Progovac, 2015; Corballis, 2017; De Boer, Thompson, Ravignani & Boeckx

2019). A variety of neurological data and fossil records, moreover, support a model that Merge results from an evolutionary process, which may have started already 2.4 mya with the appearance of H. erectus (Hillert, 2014; 2015; 2019; Martins, Maties & Boeckx, 2018). As we discuss below, some indirect evidence supports the view that the left dorsal white-mater fiber tract which connects temporoparietal regions with Broca's region evolved gradually in the hominin lineage. It is therefore possible that (pre-) H. erectus did not have the capacity for asymmetric and IM as the CVC cluster asymmetry has not been developed. Before returning to the extension of the MP in the context of the evolution of the MLC, let us first briefly review the generative tradition, on which Merge is based.

2. The Generative Tradition

Linguistic accounts in the mentalistic tradition share the argument that human cognition is equipped with a finite computational system that generates an infinite number of meaningful expressions. Generative models argue that these recursive linguistic computations cannot be solely inductively derived from linguistic experience, but that these computations must be based on an intrinsically rich innate structure.

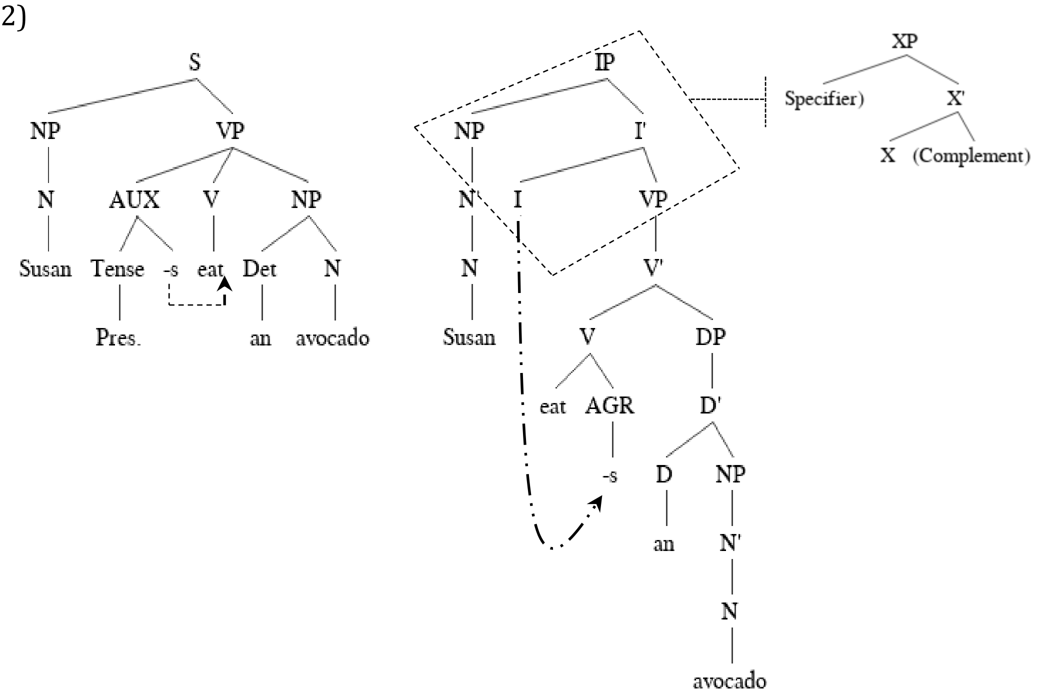
(1)



In contrast to the present view, a language-specific universal grammar (UG), as part of the human genotype, would determine these basic syntactic properties and would operate modular or domain-specific as various other human “mental organs” under different principles. Important pre-Minimalist milestones in the generative tradition are the Standard theory that introduced the term “transformational grammar” and the Government and Binding (GB) theory (Chomsky, 1965; 1981). In (1), components which apply only to the Standard theory are shown in brackets. Here, PS- (phrase structure) rules as well as atomic lexical items with their properties (e.g., subcategorization frames) generate an underlying syntactic structure of a sentence, the D- (deep) structure. The D-structure, which corresponds to a semantic representation, will be mapped onto the phonologically interpretable S- (surface) structure of the sentence by different transformation rules. This two-layer design has indeed the advantage that the number of PS-rules are significantly reduced. (1) shows GB components in boxes, the X-bar theory replaces the standard PS-rules and consists of a pre-determined three-layer scheme for each phrase of a lexical (e.g., noun, N; verb, V) or grammatic category (e.g., inflection, I).

The single movement rule “Move α” allows anything to move anywhere and replaces the transformation rules. In addition, phonological (PF) and logical form (LF) respectively interface with phonology and semantics. PS-rules in the Standard theory follow the pattern X → Y Z; GB uses, as mentioned before, the X-bar scheme as a primitive, which sets three layers for each lexical and grammatical category, including an inflection phrase (IP). Complements and specifiers are optional, in addition to adjuncts, binary and unary branching is permitted.

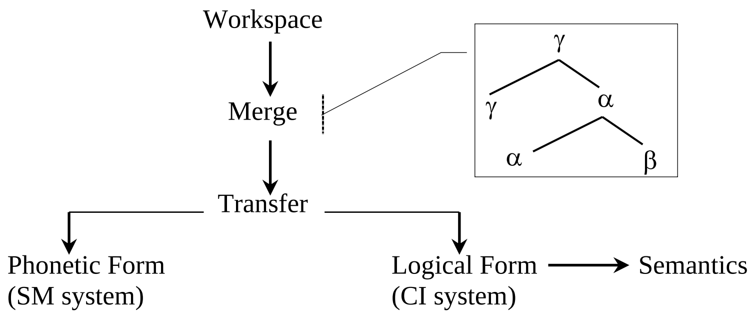
(2)



The sentence *Susan eats an avocado* would be analyzed in the Standard theory by the PS-rules: S → NP VP; VP → (AUX) V NP; NP → (Det) N; round brackets indicate optional categories, and S stands for Sentence, NP for noun phrase, VP for verb phrase, Aux for auxiliary and Det (or D in GB) for determiner. Syntactic features of a lexical item would be inserted at the S-structure; here the grammatical features [Present] [+3rd person singular] in Aux and the verb suffix *-s* will be realized by "affix hopping", which is an obligatory transformation. In GB, the agreement feature Agr has been lowered to V (see 2). Consistent with the Standard theory, syntactic restructuring by means of Move α is required for most sentences (e.g., Agent deletion, extrapositions, *wh*-fronting, negation, passive, raising, reflexives, number agreement).

Many extensions and modifications of this generative paradigm were developed in theoretical linguistics. All generative accounts have in common that it is the syntactic component in a narrow sense which is the engine of the human language system (Hauser, Chomsky & Fitch, 2002). Within and outside of this generative approach, however, concepts are further developed mostly according to economy principles. Theories and sub-theories change routinely. What stays the same are the linguistic phenomena across natural languages. The internal linguistic debate continues with the MP that tries to reduce the complexity of syntactic computations to basic cognitive operations (Chomsky, 1995). The stated UG is a 6-tuple (Collins & Stabler, 2016): <Pho, Syn, Sem, Select, Merge, Transfer>. The first three sets of features constitute a lexical item, the last three universal operations. Pho contains phonological features, segments and sequential restrictions, Syn syntactic categories such as N, V, Adj; subcategorization features and principles that each clause requires an NP or DP in Subject position (Extended Projection Principle, EPP); unvalued features [uF] valued by the operation Agree. Sem includes features such as [Event] and thematic roles such as Agent, Recipient, Experiencer. Select, Merge, and Transfer (aka Spell-Out) are universal operations. The operation Select refers in the early MP version to the selection of atomic lexical items from a lexical array (numeration), which are used for syntactic derivations. The numeration includes set of pairs, the lexical item, and an index about how often it is selected.

(3)



The main syntactic operation Merge, which works bottom-up in contrast to the top-down or left-to-right PS-rules in pre-Minimalist accounts, performs pairwise derivations. One of them is the head and determines the label of the resulting unit, the phrase. This form of derivation implies (in contrast to GB for example) that all structures generate binary branching (Kayne, 1984; Larson, 1988). EM implicates that a word merges with another syntactic object, IM that an already merged lexical or phrasal set can recursively re-merged to generate binary branching of an arbitrary size (see Uriagereka, 2002). Transfer maps the merged syntactic object to phonetic and semantic entities interpretable at the respective interfaces to the CI or SM system.

Alternatives were proposed for many components and operations such as that tokenization can be used instead of numeration to avoid co-indexing among lexical items. For example, a lexical array may include the tokens <avocado$_{i}$, avocado$_{ii}$> instead of the numeration <avocado, 2> in the sentence *The avocado next to this avocado* and syntactic objects of the lexicon are taken directly from a workspace (WS) for syntactic derivations without using the Select operation (Collins & Stabler, 2016; Chomsky, Gallego & Ott, 2018). Also, parallel WS make IM redundant (Jayaseelan, 2017). Binary Merge operations with the verb as a single head are shown below for the sentence *Susan ate an avocado* (see Fujita, 2017; T = tense).

(4) a. (an) (avocado) > {an, avocado}
b. (ate {an, avocado}) > {ate {an, avocado}}
c. (Susan {ate {an, avocado}}) > {Susan {ate {an, avocado}}}
d. (TENSE {Susan {ate {an, avocado}}}) > {TENSE {Susan {ate {an, avocado}}}}
e. (Susan {T {Susan {ate {an, avocado}}}}) > {Susan$_{i}$ {T {__$_{i}$ {ate {an, avocado}}}}}

Merge as discussed before is a binary set operation which can be applied to its own output to create hierarchical structures, possibly to free up buffer space for new input items. The basic critique against generative approaches as mentioned before holds also for the MP. The stipulated functions and operations are not only detached from and immune to empirical verifications, they also do not relate to any evolving precursor stages. Here, we focus on Merge and research how this property, in its full extent, might have evolved in modern humans.

3. The Neural Substrate of Merge

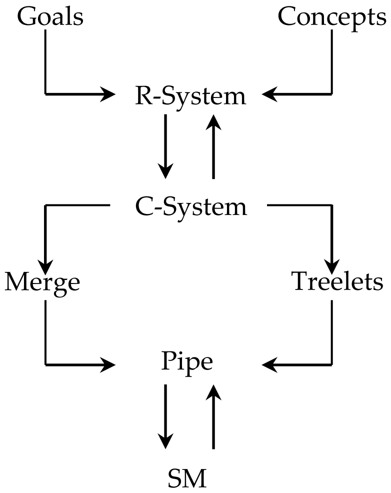
We may all agree that syntax is the result of biological and cognitive properties, but what syntax is and how specifically linguistic syntax might have emerged in the hominin lineage is highly controversial (e.g., Dor & Jablonka, 2000; Christiansen & Kirby, 2003; Corballis, 2013; Dunbar, 2016). In considering a broad range of empirical data, we argue that syntax is a modality- and domain-independent capacity of the human mind which follows innate nonlanguage-specific universal principles. Thus, syntactic operations as exemplified in the previous section may have their roots in nonlinguistic intrinsic operations at the conceptualization level and may be the result of a phylogenetic process that lasted more than 2 million years as we discuss below.

3.1. The RC Model

In nonhuman primates, syntax *per se* has a minor function for expressing meanings in a broad sense such as emotions and goals/intentions. Thus, launching a phylogenetically relevant design of the language system may focus on the intrinsic computations. Here, we call this output relevant sequencing operation Pipe. Intrinsic computations are generative in production and comprehension. In the former case results are sequential executions via Pipe, in the latter case the results are internal representations. We adopt implicitly a design of the generative approach as addressed above but expand it to nonlinguistic domains.

The "Representational System" (R-System) comprises knowledge in form of conceptualizations intrinsically generated by the "Computational System" (C-System). This distinction is based on the account that cognition involves neural computations over neural representations (Bechtel, 2008). The R-System is linked to meta-cognitive planning or strategies such as Goals, and includes Concepts including for example memories about factual, episodic, or sensorimotor (SM) information. In production, the C-System can generate in the language domain complex syntactic relations among lexical items (Merge), but also implements or prepares these generated structures for Pipe. Merge may be an optional function as the C-system may access fixed representations of PS-rules such as syntactic subtrees (treelets). In context of the present framework and in contrast to the MP, a Merge approach will be favored that stipulates not only bottom-up but also top-down processes (as most non-derivational theories do) and *n*-ary branching to account for symmetric structures in coordination or compounds.

(6)



The intrinsically generated information will be expressed via articulation, gestures, signing, or in any specific form of sequential motor activities as required for dancing or playing a music instrument. The term generative is reserved in the present context for intrinsic computations. What made us unique among extant primates, is the cognitive capacity to generate different types of computations. Conceptual information will be converted to hierarchical clusters and long-distance dependencies to manage restricted working memory (WM) capacities, and information will be sequenced to adjust to the constraints of our SM system. Merge itself is generative in sentence production but also in comprehension as rehearsal operations contribute to language comprehension.

The RC Model in (6) shows the general cognitive architecture proposed here, but interface components are necessary to account for domain- and modality-specific properties. In the case of spoken language, for example, articulatory motor innervations for output, phonological rehearsal operations or prosody play a key role. The framework of this model is in the spirit of Jackendoff's (2009) Parallel Architecture. Both approaches consider what we know about neural functions, linguistic and non-linguistic cognitive capacities and about interactions among these different cognitive capacities. In contrast to the Parallel Architecture, the universal cognitive components of the RC Model are biological grounded as they *underly* all cognitive domains. To illustrate and exemplify the interactions between single domains and universal components and/or among single domains, a refinement of this model is required to be addressed in a different context.

3.2. *Neurolinguistic Evidence*

The proposed cognitive architecture for human language processing is based on neurolinguistic evidence. *Figure 1* refers to Rauschecker & Scott's (2009) dual auditory processing scheme of the human brain. In the case of forward mapping, speech will be decoded in the antero-ventral stream, including Broca's area with the pars opercularis (pOp) and pars triangularis (pTr) of the left inferior frontal gyrus and ventral premotor cortex (vPreMC), and will be transformed in a motor-articulatory representation. Brodmann's cytoarchitectonic area (BA) 44 correspondences approximately to pOp, BA 45 to pTr and BA 8 to vPreMC. In turn, transfer occurs to the inferior parietal lobe (IPL) and the posterior superior temporal gyrus (pSTG) as an efference copy. Inverse mapping uses the postero-dorsal stream and attention-related changes in IPL control the preparation of motor programs in Broca's area and PreMC. The IPL seems to be the hub, where forward mapped information will be matched with feedback from sensory areas. Both streams continuously orchestrate in-real time information during speech production and perception. For the present purpose, we consider here the ventral and dorsal pathways as functions connecting different cortical regions, although they correspond to white-matter fiber tracts as mentioned below.

Based on lesion, electrophysiological and neuroimaging studies, here we add a syntactic and semantic component to address the neural substrate of higher-ordered structural formations such as Merge. The standard view is that specifically Broca's area plays an key role in syntactic processing of complex sentences (e.g., Caramazza & Zurif, 1976; Grodzinsky, 2000; Friederici, 2011; Fridriksson, den Ouden, Hillis, Hickok, Rorden, Basilakos, Yourganov & Bonilha, 2018; Iwabuchi, Nakajima & Makuuchi, 2019). Several findings speak, however, against this neuroanatomical model. The deficits found in agrammatic Broca's aphasic patients only indicate that Broca's area may be involved in the comprehension process of complex sentences. Numerous studies show that Broca's aphasic patients are still able to judge the well-formedness of a sentence (e.g., Linebarger, Schwartz, & Saffran, 1983) and that lesions to Broca's area are not necessarily associated with sentence comprehension deficits (e.g., Dronkers, Wilkins, Van Valin, Redfern & Jaeger., 2004; Thothathiri, Kimberg & Schwartz, 2012). In fact, Broca's aphasic patients typically suffer also from lesions in subcortical structures (D'Esposito & Alexander, 1995). Lesions to Broca's area alone do not cause the deficits in question rather than the neural circuits associated with this area and neighbored regions. This broader circuit includes the frontal operculum (FO), the anterior middle frontal gyrus (BA 46; mid-anterior to BA 45 of Broca's area), the pars orbitalis (pO; BA 47; inferior anterior to BA 45), the supplementary motor area (BA 6), and subcortically the basal ganglia (BG) and thalamus. We use here the term "Broca's region" to refer to this broad area (see also Ardila, Bernal & Rosselli, 2016).

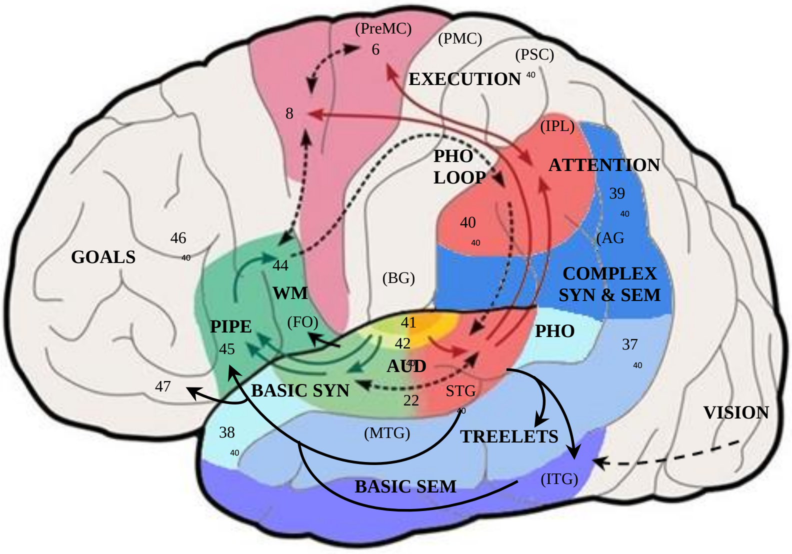


Figure 1. Dual auditory processing model extended by the syntactic and semantic component (see text for details; figure adapted and modified based on Rauschecker & Scott, 2009; © CC BY license) numbers: Brodmann's cytoarchitectonic areas; cortical regions: AG, angular gyrus; BG, basal ganglia; FO, frontal operculum; IPL, inferior parietal lobe; PMC, primary motor cortex; PreMC, premotor cortex; PSC, primary sensory cortex; (S/M/I)TG, superior/middle/inferior temporal gyrus; functions: AUD, auditory; PHO, phonology; SEM, semantics; SYN, syntax; WM, working memory).

Others again argue that the effects in Broca's area are the result of WM capacities and/or of integrative control functions (e.g., Hillert, 2000; Kaan & Swaab, 2002; Rogalsky & Hickok, 2011; Novick, Trueswell & Thompson-Schill, 2010). These data support a neuroanatomical model that considers the posterior temporoparietal region as the primary region for more complex syntactic-semantic and phonological processing in production and comprehension rather than Broca's area or region (e.g., Pillay, Binder, Humphries, Gross & Book, 2017; Matchin & Hickok, 2019). This interpretation is also consistent with the findings that Broca's area will be engaged not only in the language domain, but also in context of actions, music, or calculations. Depending on the specific task, such as execution, listening or imaging, the cortical activation does not only involve Broca's area/region to various degrees but also the connected cortical network similar to the language domain (Fadiga, Craighero & D'Ausilio, 2009). A more recent fMRI study reported two distinct subregions for language-specific (sentences > nonwords) and domain-general tasks in at least 90% of individual subjects overlapping the border of BAs 44 and 45 in Broca's area (Fedorenko, Duncan & Kanwisher, 2012). The question is how to define the term "language-specific." If we find identical or overlapping computations for phrasing in sentences and music, the language-selective region becomes less specific or more general. For example, violations of sequence learning elicit the same event-related potential (ERP) brainwave responses (P600) as ungrammatical sentences (Christiansen, Conway & Onnis, 2012) or musical sequences (Patel, Gibson, Ratner, Besson & Holcomb, 1998), and fMRI studies show that Broca's area will be activated when sequences are violated (e.g., Petersson, Folia & Hagoort, 2012).

The recruitment of particular sections of Broca's region results from the neural connectivity with temporoparietal regions, a functional task-dependent circuit which corresponds to dorsal and ventral white-mater fiber tracts. The dorsal pathway, including the arcuate fasciculus, connects pSTG and IPL with the Broca's area and generates complex linguistic information. In addition, a less direct pathway is possible via the superior longitudinal fasciculus involving the PreMC. Lemma information can be specific to a single word or generic and includes abstract information about syntactic category, subcategorization, theta-roles and grammatical properties (e.g., Kempen & Huijbers, 1983; Levelt, 1993). In addition, lemmas may include PS-structures in form of treelets which are fragments of sentential trees (e.g., J. D. Fodor, 1998; Sakas & J. D. Fodor, 2012; Jurafsky, 1996; Vosse & Kempen, 2000). In contrast to a word's lexeme information, which includes phonological and morphological properties, treelets are small templates of syntactic nodes typically underspecified in some respects of sentential tree structure. Local phrasal structures, Agent-first strategy, treelets are transferred ventral via the uncinate fasciculus from the STG and/or middle temporal gyrus (MTG) to the FO and Broca's area. Similar basic lexical semantic information reaches ventral BAs 47 and 45 of Broca's region via the extreme capsule fiber system (Van der Lely & Pinker, 2014; DeWitt & Rauschecker, 2012).

The function of Broca's area may be primarily related to WM operations (e.g., Smith, Jonides, Marshuetz & Koeppe, 1998) which keep and monitor syntactic and morphological information for forward processing in terms of Merge in the posterior temporoparietal region. As part of the C-System, pSTG maps sequential phonological information onto the posterior part of the MTG and the inferior parietal lobe (IPL) in the proximity of the angular gyrus (AG), where complex semantic as well as syntactic operations are generated in terms of Merge. In turn, these syntactic structures are mapped onto the conceptual structures of the R-System which is cortically distributed. The C-System may access representations of treelets and/or generate Merge to compute syntactic information, and this information is piped into sequences in pTr of Broca's area for output in the SM system. Independent of the type of linguistic or of domain-specific information in general, top-down driven attentive or more conscious computations seem to take place in the posterior temporoparietal region and computed via the dorsal pathways, while automatic access to basic information, local or fixed, seems to computed via the ventral pathways.

4. The Evolution of Merge

In considering neurolinguistic data but also fossil records, we find evidence for the assumption that Merge meets the criterion of evolvability. Furthermore, on theoretical grounds Merge is decomposable and may thus point to the phylogeny of Merge rather than to a single-mutant account (e.g., McBrearty & Brooks, 2000; Boeckx, 2013; Tallerman, 2014; Progovac, 2015; De Boer, Thompson, Ravignani & Boeckx 2019). Another interesting approach is the Motor Control Origin of Merge stating the linguistic syntax evolved from a motor control capacity, first unrelated to language (Fujita, 2014). A related but different approach is the action-perception framework of the Mirror System Hypothesis, which involves different steps from imitations and pantomimes to protosigns and the "language-ready brain" (Arbib, 2012). Macaque's mirror system for grasping is associated with the cortical area F5, homologous to the human Broca's area. Since Merge seems to be a function of posterior temporoparietal regions rather than being restricted to Broca's areas or closely connected regions, it seems not to be directly related to the mirror system of the inferior frontal areas. Here, we assume that the function of Merge evolved in different, possibly parallel steps since the appearance of genus Homo. This evolving process may have started already 2.4 mya with the appearance of H. erectus (Hillert, 2014; 2015; 2019).

4.1. The Biological Capacity of Merge

In considering Bräuer's (2008) de-inflated model of the speciation taxonomy, we divide here between two different Homo genera in the hominin lineage: The genus "H. sapiens s.l." (sensu lato) represents a polytypic species with an anagenetic relationship and includes anatomical modern humans (oldest fossils dated ca. 315 k years old at Jebel Irhoud in Marocco; Hublin, Ben-Ncer, Bailey, Freidline, Neubauer, Skinner, Bergmann, Le Cabec, Benazzi, Harvati & Gunz, 2017). Denisovans, who diverged from Neanderthals about 430 kya, and (pre-) Neanderthals / archaic H. sapiens (presumably including H. heidelbergensis) split from late H. erectus ca. 700-800 kya. The genus "H. erectus s.l." comprises in addition to H. habilis and H. ergaster (early H. erectus: 2.4 mya) late H. erectus, who went extinct 50-100 kya. During the epoch of H. erectus s.l., the cranial capacity increased from 600 cc to 1,000 cc, reaching almost the size of modern humans with a range of ca. 1,200-1,400 cc. The *C*-variant of the *SRGAP2* copy-code mechanism responsible for cortical growth and increase of neural connectivity (Charrier, Joshi, Coutinho-Budd, Kim, Lambert, de Marchena, Jin, Vanderhaeghen, Ghosh, Takayuki & Polleux, 2012; Fossati, Pizzarelli, Schmidt, Kupferman, Stroebel, Polleux, & Charrier, 2016) emerged 2.4 mya, which coincides with the appearance of H. erectus. At the behavioral level, the refinement of tool traditions (from Oldowan to Acheulean) indicates that cytoarchitecture and cognitive capacities specifically co-evolved during the H. erectus s.l. period of about 2 million years. This interpretation is also supported by the findings that H. erectus travelled long distances and used vessels to reach Java (Dubois, 1894). The significant increase of mobility requires complex social and strategic planning, which in turn presumably put selective pressure on the improvement of communicative skills. The increase of cognitive capacities may have therefore also included the cognitive operations associated with the language circuits, probably in form of a PLC.

The idea that language is not a unique trait of modern humans but that a PLC was already in place and used by extinct Homo species is endorsed not only by morpho-syntactically less complex languages or by the typology of certain language stages but also by linear structures re-surfing in fully-fledged modern languages (e.g., Bickerton, 1990; 1981; Jackendoff, 1999; Jackendoff, 2002; Progovac & Locke, 2009; Progovac, 2010; Progovac, 2012; Tallerman, 2014; 2017; but see Nóbrega & Miyagawa, 2015). As addressed before, languages that are based on a linear grammar use semantic strategies such as Agent-first and Focus-last to determine word order and typically lack inflectional morphology such as Tense or Case. This applies to pidgin or creole languages, emerging sign languages, certain stages in first and second language acquisition, symptoms of grammatical impairments in aphasia or grammar acquisition by feral children.

We already mentioned the Malayan dialect Riau Indonesian which served in its history as a lingua franca. It has virtually no syntactic categories (nouns vs. verbs) or inflections, and the word order is based on semantic principles (Gil, 2005; 2014). Also, signers of a newly emerging language apply a perceptual syntactic strategy such as deaf children who create home signs to communicate with their hearing parents. Similar to spoken language, children go through two gestural stages and their developed home-sign system is more complex than the gestures used to support speech.

The famous case of the Nicaraguan Sign Language shows also a two-stage process: The first generation used signs to refer to objects they needed to talk about and stringed them together in two-word phrases (Agent-verb). The next generation of deaf children elaborated on these structures (Senghas, Kita & Ozyurek, 2004). Further examples are the emerging sign language Al-Sayyid Bedouin Sign Language (Sandler, Meir, Padden & Aronoff, 2005) and the isolated village sign language Central Taurus Sign Language (Arbib, 2012; Caselli, Ergin, Jackendoff & Cohen-Goldberg, 2014). Again, adults who learn a second language without explicit instructions show across all examined pairs of first and second language a basic linguistic competence, called The Basic Variety (W. Klein & Purdue, 1997; Jackendoff, 1999) that is similar to the hypothesized protolinguistic stage.

Second language speakers tend to acquire in the first stage words without inflections and a word order which is based on semantic-pragmatic principles rather than on syntactic principles such as subclauses. Most interesting, the default semantic strategy Agent-first, which often applies together with the principle Focus-last, is efficient to interpret strings of the order V-N-N (e.g., *hit girl boy* will be always interpreted as *The girl hit the boy* and not as *The boy hit the girl*). Focus-last represents often the result or significance caused by the Agent. Semantically based parsing of word strings is, however, often irrelevant because typically pragmatics informs amount the intended meaning (e.g., *drink milk Bob* and *drink Bob milk* means always the same: *Bob drinks milk*). Again, first language acquisition involves after the holophrastic stage always a two-word and telegraphic stage in which inflections and function words are rarely used. Further evidence can be found in individuals experiencing neurological deficits. In the case of agrammatism, which is a cardinal symptom of Broca's aphasia, patients fall back on an Agent-first strategy (e.g., comprehension disorders of non-canonical sentences structures such as "Theme-first" as in reversible passives: *Mike was chased by Paul*) and drop most inflections and function words (e.g., Schwartz, Saffran & Marin, 1980). Another example are feral children who have difficulties acquiring the grammatical competence of native speakers. Genie, a well-known victim of severe child abuse, was not exposed to language until the age of 13 years (Curtiss, 1977). She quickly acquired words after Genie was discovered, but her grammar remained far behind despite of many years of intensive training.

Linear structures, which are sometimes called "linguistic fossils", re-surface also in modern languages. In addition to Agent-first and Focus-last strategies, Jackendoff (1999) also addresses grouping (e.g., modifiers are typically adjacent to what they modify), compounds (e.g., N+N are concatenated and the meaning of the new bigger N results from pragmatics as found in many idioms), and adverbial expressions (e.g., adverbs can be relatively freely placed). These other examples point to distinct stages in the history and/or evolution of language. They are the building blocks for Merge, for higher order headed units and PS-rules. For example, Agent-first strategy applies not anymore to a single word but to a phrase, and phrasal syntax generates hierarchical structures and long-distance dependencies. Inflections which mark for example Case, Tense, and Agreement allow a more flexible word order although used to various degrees across languages. The variety of typology shows that fully-fledged modern languages use more cognitive resources than creoles which in turn use more resources than pidgins or new sign languages. The linguistic analysis, moreover, supports the assumption that the syntactic property Merge evolved from more semantically based linear structures. Members of the genus Homo sapiens s.l. are equipped with the MLC to develop fully-fledged modern languages, but depending on the conditions of cultural accumulations, these resources are used to various degrees as described above. It is an open question whether Merge is the product of cultural accumulations based on the PLC. In this scenario, the biological capacity for protolanguages and modern languages would be identical.

The phylogenetic and ontogenetic development of white-matter fiber tracts connectivity indirectly supports the idea of a PLC. We know that in particular the dorsal route is less well developed in our extant closest relatives, common chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) and does not reach the homologous area of Broca's area or the temporoparietal regions (Rilling, Glasser, Preuss, Ma, Zhao, Hu & Behrens, 2008). In humans and compared to chimpanzees, the left-sided Broca's area expanded more (BA 44: 6.6x; BA 45: 6x) in relation to the right-sided Broca's area (BA 44: 4.1x; BA 45: 5x), the whole brain (3.6x), the frontal cortex (4.6x) and other cortical areas (e.g., the left primary visual area V1: 1.8x). Also, Broca's area is more strongly connected to the temporoparietal regions (Schenker, Hopkins, Spocter, Garrison, Stimpson, Erwin, Hof & Sherwood, 2014). The left Broca's area underwent a specific expansion in evolution since the split from genus Pan. This is consistent with the results that the dorsal pathway does not connect to Broca's area in human newborns. At this early developmental stage, the ventral pathway connecting the inferior frontal gyrus with the temporal lobe is already fully developed (Perani, Saccuman, Scifo, Anwander, Spada, Baldoli, Poloniato, Lohmann & Friederici 2011). The dorsal pathway is apparently a phylogenetically more recent structure and may have played a key role in the evolution of complex sentence processing.

In considering the above-mentioned speciation model, it might be possible that the dorsal pathway of early H. erectus was not fully developed as compared to H. sapiens s.l. Again, Merge computations might have gradually evolved without fully integrating sensory-motor processes (PN, Arbib). In this case, H. erectus s.l. might have used hierarchical structures in terms of Merge in nonverbal actions but not in verbal actions.

4.2. From Calls to Merge

Our approach focuses here on the PLC, which may constitute on cultural and/or biological grounds a direct precursor stage of the MLC in H. sapiens s.l. It is outside of the present scope to discuss the full range of the evolving cognitive components associated with modern languages. We consider here, however, relevant cornerstones of the full phylogenetic spectrum to specify a possible pre-Merge stage in (late) H. erectus. In this vein, the monkey call capacity has been often taken as a starting point for the evolution of human cognitive capacities. The monkey call capacity goes beyond the generation of single calls as they can combine two calls, but not more. The frequency of repeated two-call sequences seems not to be relevant.

For example, female Diana monkeys (*Cercopithecus diana*) produce from a set of calls two-call combinations (*LA, HA, RA*). These combinations appear not to be compounds creating a new meaning based on the combined individual calls (Zuberbühler, Cheney & Seyfarth, 1999). Campell's monkeys (*Cercopithecus campbelli*) combine the calls *krak* (ground alert) and *hok* (in-air alert) with the call -*oo*, which itself is not produced in isolation but only combined in a fixed order *krak-oo* or *hok-oo* and not with the non-predator call *boom*. The predator- or alert-specific calls *krak* and *hok* seem to change to a non-predator disturbance meaning such as [falling tree] or [neighboring group] (Ouattara, Lemasson & Zuberbühler, 2009). These new meanings are not compositional but idiomatic. Also, male putty-nosed monkeys (*Cercopithecus nictitans*) use *pyows* or *hacks* sequences in isolation to warn respectively about a loitering leopard or eagles from overhead but combine them to *pyow-hack* sequences, with small numbers of *pyows* and *hacks*. Again, Arnold & Zuberbühler, (2012) observed that these sequences are not compositions of the individual meanings but apparently a new meaning leading to group travelling. They also suspected that the number of *pyows* in *pyow-hack* sequence indicates the planned travel distance (Schlenker, Chemla, Arnold & Zuberbuḧler, 2016).

In sum, monkeys can generate single calls and concatenate them to two-call sequences. This is a basic combinatory operation and further recursive operations are not possible. This is not surprising as monkeys can only keep a stimulus trace for a moment in time or it will be overwritten by a subsequent stimulus (Scott, Mishkin & Yin, 2012; Rizzi, 2016). Monkey's auditory short-term storage capacity is not only limited to one unit for a moment in time, but they also lack the capacity to analyze the temporal structure of discrete sounds (Wakita, 2019). Overall, the analysis of monkey calls shows that they do not represent a direct precursor stage for a language capacity. Within a timeframe of about 6 million years, however, memory resources must have evolved along with the semantic and phonological capacity and enhanced later by a morphosyntactic capacity to communicate and reflect about experiences stored in memory.

How could have Merge evolved from simple monkey calls? In drafting here our ideas, we first assume as outlined before that cortical rewriting in terms of volume and connectivity increased in the hominin lineage in a non-saltate, possible punctuated gradual form. Behavioral changes may have contributed to the emergence of basic computational capacities realized in different cognitive domains. The roots of language may be therefore at the conceptual level. At some point in the evolutionary path, our extinct relatives started to gesture and/or vocalize these conceptualizations, whereas concept development and expressiveness mutually benefited from each other. Naming allows to categorize and share experiences and is the origin of awareness. These conceptual computations refer thus to different representational levels, including semantic and phonological representations. The next step implies the ability to create distinct entities beyond the level of immediate sensory experiences, to name those entities (words) and to concatenate them to generate sequences. Articulating two to three words in a row are the building blocks of grammar and perceptual strategies such as Agent-first and pragmatics may have been sufficient to comprehend the intended meaning. From an evolutionary viewpoint, the use of EM may be a more recent event. Symmetric Merge emerged presumably first along with basic lexical constrains before asymmetric Merge in which typically the first-mentioned entity modifies the head. In considering the evolving view presented here IM (recursion) is less complex than EM on which it is based. The development of fully-fledged languages, including complex morphology, PS-rules and dependency relations is the outcome of cultural developments due to population growth and communicative needs.

In the case of the MLC, Merge describes the structural relationship among and between words (w) and phrases (p). Thus, words are merged with words (ww), but also words are merged with phrases (wp) and phrases with phrases (pp); for example, without considering morphological processes: [the avocado]$_{ww}$; [eat [the avocado]$_{wp}$; [[the chimp] [eats [the avocado]]]$_{pp}$. As mentioned before, at present it is difficult to predict the language capacity for a certain period in the hominin lineage. Because of its relatively modern behavior, we assume here that the PLC of late H. erectus s.l. included not only Merge of "ww" but possibly also "wp". It is still, however, unclear whether H. erectus s.l. or H. sapiens s.l. made use of Merge operations in language at all to increase the complexity of phrases. H. erectus s.l. might have used solely short linear structures, because of a limited WM capacity and/or sociocultural developments as compared to modern humans; for example, for the phrase *I saw a big fish*, a flat linear structure such as [saw] [big] [fish] might have been used instead of the Merge operations {saw {big {fish}}}, whereas ambiguity is solved by pragmatics.

Merge is certainly the result of the attempt to generate longer and more complex phrases. For the combination of up to 3-4 words, Merge has been probably not used or created. Specifically, written language, which is a more recent cultural achievement (first literary writings dating ca. 2,400 BC), may have triggered the process to reflect on spoken language and enabled a more complex use of morphosyntactic structures. A plausible assumption is that archaic H. sapiens much like H. erectus relied as well on linear sequences without forming PS-rules. Cultural development of the language trait may have resulted in modern times (around 50 kya) to more complex structures and to the generation of PS-rules, that is to Merge operations at the word and phrasal level. At this point we are meeting again the generative paradigm. It is possible that IM operations were created in some languages to accommodate the increasing complexity of sociocultural activities associated with behavioral modernity.

5. Conclusions

Here, we discussed evolving steps of Merge and the MLC in general by focusing on the syntactic component of the generative-linguistic paradigm. By confronting syntactic Merge with empirical findings, it became clear that an extension of this concept is essential to understand not only the associated cognitive and neural computations in modern humans but also its phylogenesis in relation to the speciation process in the hominin lineage. The proposed RC Model considers Merge as an intrinsic computation *underlying* various cognitive domains. Its basic cognitive function is to generate more complex structures at the phrasal level. Merge operations seemed to be neurally implemented within the temporo-posterior region, whereas the key role of Broca's area is related to WM functions to update temporarily stored units to pipe sequenced structures for speech and/or sign output. Again, fossilized phrasal structures were discussed which re-surface in modern languages and are present in emerging and ancestral languages. They point to a precursor stage of the MLC indirectly supported by genetic, neuroanatomical, and behavioral-cultural stages, respectively associated with H. sapiens s.l. and H. erectus s.l.

We conclude that the evidence favors the hypothesis of a PLC in H. erectus s.l. While we cannot rule out that the PLC included more EM operations (e.g., asymmetry or recursion) as well as treelets processes, linearized flat structures may have been the typical language patterns, and it may have been a pragmatically efficient form of communication to cope with the ecological and sociocultural conditions during the Pleistocene epoch. The increase of morphosyntactic complexity may be therefore primarily the result of cultural accumulations and refinements while phonological and semantic-conceptual properties may be regulated by genetic constraints. In considering the evidence currently available, we conclude that H. erectus s.l. was already language-ready much like modern humans are.

References

Arbib, Michael A. 2005. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28 (2), 105-124.

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

Arbib, Michael A. 2019. The aboutness of language: From Berwick and Chomsky to comparative neuroprimatology. *Chicago Linguistics Society (CLS).*

Ardila, Alfredo, Bernal, Byron & Monica Rosselli. 2016. How localized are language brain areas? A review of Brodmann areas involvement in oral language. *Archive Clinical Neuropsychology* 31(1), 112-122.

Arnold, Kate & Klaus Zuberbühler. 2012. Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language* 129 (3), 303-309.

Bechtel, William 2008. *Mental Mechanisms.* New York: Taylor and Francis.

Berwick, Robert C., Okanoya, Kazuo., Beckers, Gabriel J. L. & Johann J. Bolhuis. 2011. Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences* 15, 113–121.

Berwick, Robert C. & Noam Chomsky. 2016. *Why Only Us.* Cambridge: MIT Press.

Berwick, Robert & Noam Chomsky. 2019. The siege of Paris. *Inference: International Review of Science* 4 (3).

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and Language Development*, 279–362. New York: Wiley & Sons.

Bickerton, Derek. 1981. *Roots of Language*. Ann Arbor: Karoma Publishers.

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

Boeckx, Cedric. 2017. Not only us. *Inference: International Review of Science* 3 (1).

Boeckx, Cedric. 2013. Biolinguistics: Forays into human cognitive biology. *Journal of Anthropological Sciences* 91, 1–28.

Bolhuis, Johann J., Tattersall, Ian, Chomsky, Noam & Robert C. Berwick. 2014. How could language have evolved? *PLOS Biology* 12 (8), e1001934.

Bräuer, G Günter. 2008. The origin of modern anatomy: By speciation or intraspecific evolution? *Evolutionary Anthropology* 17, 22–37.

Caramazza, Alfonso & Edgar B. Zurif. 1976. Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language* 3 (4), 572-582.

Caselli, Naomi K., Rabia Ergin, Ray Jackendoff, R. & Ariel M. Cohen-Goldberg. 2014. The emergence of phonological structure in Central Taurus Sign Language. *Conference proceedings, From Sound to Gesture, Padua*.

Carstairs-McCarthy, Andrew. 1999. *The Origins of Complex Language*. Oxford: Oxford University Press.

Charrier, Cécile, Joshi, Kaumudi, Coutinho-Budd, Jaeda, Kim, Ji-Eun, Lambert, Nelle, de Marchena, Jacqueline, Jin, Wei-Lin, Vanderhaeghen, Pierre Ghosh, Anirvan, Takayuki, Sassah & Franck Polleux. 2012. Inhibition of SRGAP2 function by its human-specific paralogs induces neoteny during spine maturation. *Cell* 149, 923–935.

Chomsky, Noam. 1953. Systems of syntactic analysis. *Journal of Symbolic Logic* 18 (3), 242- 256.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1965) *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Mouton de Gruyter.

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

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

Chomsky, Noam. 2004. Beyond explanatory adequacy. In A. Belletti (ed.), *Structures and Beyond*, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + Recursion = Language,* 1-29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36, 1–22.

Chomsky, Noam. 2017. The Galilean challenge: Architecture and evolution of language. *Journal of Physics: Conference Series 880*.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Howard Lasnik, R. Martin, D. Michaels & Juan Uriagereka (eds.), *Step by Step: Minimalist Essays in Honor of Howard Lasnik*, 89-155. Cambridge: MIT Press.

Chomsky, Noam, Gallego, Ángel J. & Dennis Ott. (2019. Generative Grammar and the faculty of language: Insights, questions & challenges. *Catalan Journal of Linguistics, lingbuzz/003507*.

Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.

Christiansen, Morton H. & Kirby, Simon. 2003. Language evolution: Consensus and controversies. *Trends in Cognitive Sciences* 7 (7), 300-307.

Christiansen Morton H., Conway, Christopher M. & Luca Onnis. 2012. Similar neural correlates for language and sequential learning: evidence from event-related brain potentials. *Language and Cognitive Processes* 27, 231–256.

Christiansen, Morton H. & Chater, Nick. 2015. The language faculty that wasn't: a usage-based account of natural language recursion. *Frontiers in Psychology* 6, 1182.

Collins, Cris & Stabler, Edward. 2016. A formalization of minimalist syntax. *Syntax* 19 (1), 43-78.

Corballis, Michael C. 2013. Mental time travel: a case for evolutionary continuity. *Trends in Cognitive Sciences* 17, 5–6.

Corballis, Michael C. 2017. *The Truth About Language: What It Is, and Where It Came From*. Chicago: University of Chicago Press.

Cowan, Nelson. 2001. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Science* 24(1), 87-114 (discussion, 114-85).

D'Esposito, Mark & Michael P. Alexander. 1995. Subcortical aphasia: Distinct profiles following left putaminal hemorrhage. *Neurology* 45, 38-41.

De Boer, Bart, Thompson, Bill, Ravignani, Andrea & Cedric Boeckx. 2019. Evolutionary dynamics do not motivate a single-mutant theory of human language. *BioRxiv.*

De Vries, Meinou H., Padraic, Monaghan, Knecht, Stefan & Pienie Zwitserlood. 2008. Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition* 107, 763–774.

DeWitt, Iain & Josef P. Rauschecker. 2012. Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Science U.S.A.* 109, E505–E514.

Dehaene, Stanislas, Meyniel, Florent, Wacongne, Catherine, Wang, Liping & Christophe Pallier. 2015. The Neural Representation of Sequences: From transition probabilities to algebraic patterns and linguistic trees. *Neuron* 88, 2–19.

Dor, Daniel & Jablonka, Eva. 2000. From cultural selection to genetic selection: A framework for the evolution of language. *Selection* 1, 1-3, 33-55.

Dronkers, Nina F., Wilkins, David P., Van Valin, Robert D. Jr., Redfern, Brenda B. & Jeri J. Jaeger. 2004. Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92 (1-2), 145-177.

Dubois, Eugène. 1894. *Pithecanthropus Erectus. Eine Menschenaehnliche Uebergangsform aus Java*. New York: G.E. Stechert.

Dunbar, Robin. 1996. *Grooming, Gossip and the Evolution of Language*. London, UK: Faber and Faber.

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

Everett, Daniel L. 1987. *A Lingua Pirahã e a Teoria da Sintaxe: Descrição, Perspectivas e Teoria*. Campinas, Brazil: Editoria da Unicamp.

Everett, Daniel L. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85, 405–442.

Everett, Daniel L. 2013. The shrinking Chomskyan corner: A final reply to Nevins, Pesetsky & Rodrigues, *lingbuzz/000994, volume 4.*

Fadiga, Luciano, Craighero, Laila & Alessandro D'Ausilio. 2009. Broca's area in language, action, and music. *Annals of the New York Academy of Science* 1169, 448-458.

Fitch, W. Tecumseh. 2011. The evolution of syntax: an exaptationist perspective. *Frontiers in Evolutionary Neuroscience* 3 (9).

Fitch, W. Tecumseh. 2017. Dendrophilia and the evolution of syntax. In L.-J. Boë, J. Fagot, P. Perrier & J.-L. Schwartz (eds.), *Origins of Human Language: Continuities and Discontinuities with Nonhuman Primates*, 305-328. Frankfurt am Main: Peter Lang Press.

Fitch, W. Tecumseh & Marc D. Hauser (2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 16, 303 (5656), 337-380.

Fitch, W. Tecumseh & M. Dias Martins. 2014. Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences* 1316, 87-104.

Fodor, Janet D. 1998. Unambiguous triggers. *Linguistic Inquiry* 29, 1-36.

Fodor, Janet D. & William G. Sakas. 2017. Learnability. In Ian Roberts (ed.), *The Oxford Handbook of Universal Grammar*. Oxford: Oxford University Press.

Fodor, Jerry A., Bever, Thomas G. & Merrill F. Garrett, M. 1974. *The psychology of Language*. New York: McGraw-Hill.

Fossati, Matteo, Pizzarelli, Rocco, Schmidt, Ewoud R., Kupferman, Justine V., Stroebel, David, Polleux, Franck & Cécile Charrier. 2016. SRGAP2 and its human-specific paralog co-regulate the development of excitatory and inhibitory synapses. *Neuron* 91 (2), 356–369.

Fridriksson, Julius, den Ouden, Dirk-Bart, Hillis Argye E., Hickok, Gregory, Rorden, Chris, Basilakos, Alexendra, Yourganov, Grigori & Leonardo Bonilha. 2018. Anatomy of aphasia revisited. *Brain* 141 (3), 848-862.

Friederici, Angela D., Bahlmann, Jörg, Heim, Stephan, Schubotz, Ricarda I. & Alfred Anwander. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America* 103 (7), 2458–2463.

Friederici, Angela D. 2011. The brain basis of language processing: From structure to function. *Physiology Review* 91, 1357-1392.

Fujita, Koji. 2009. A prospect for evolutionary adequacy: Merge and the evolution and development of human language. *Biolinguistics* 3, 128-153.

Fujita, Koji. 2014. Recursive merge and human language evolution, In Thomas Roeper & Margaret Speas (eds.), *Recursion: Complexity in Cognition*, 243-264. New York, NY: Springer.

Fujita, Koji. 2017. On the parallel evolution of syntax and Lexicon: A Merge-only view. *Journal of Neurolinguistics* 43 (B), 178-192.

Futrell Richard L. J., Stearns, Laura, Everett, Daniel L., Piantadosi, Stefen T. & Edward Gibson. 2016. A corpus investigation of syntactic embedding in Pirahã. *PLoS ONE* 11 (3): e0145289. doi.org/10.1371/ journal.pone.0145289

Gentner, Timothy, Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207.

Gil, David. 2005. Word order without syntactic categories: How Riau Indonesian does it. In Andrew Carnie, Heidi Harley & Sheila A. Dooley (eds.), *Verb First: On the Syntax of Verb-Initial Languages* (Linguistics today 73), 243–263. Amsterdam: John Benjamins.

Gil, David. 2014. Sign languages, creoles, and the development of predication. In

F. Newmeyer & L. Preston (eds.), *Measuring Grammatical Complexity*, 37-64. Oxford: Oxford University Press.

Grodizinsky, Yosef. 2000. The neurology of syntax: Language use without Broca's area. *Behavioral Brain Sciences* 23 (1), 1-21 (discussion 21-71).

Hauser, Marc D. & Watumull, Jeffrey. 2017. The universal generative faculty: The source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics* 43 (B), 78-94.

Hauser, Marc D., Chomsky, Noam & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569-1579.

Hillert, Dieter (2000) The grammar of agrammatism. *Behavioral and Brain Sciences* 23 (1), 36-37.

Hillert, Dieter 2014. *The Nature of Language. Evolution, Paradigms and Circuits*. New York, NY: Springer.

Hillert, Dieter 2019. We were not alone: The gradual evolution of the language capacity. *Workshop EvoLinguistics, University of Tokyo.*

Hillert, Dieter G. 2015) On the evolving biology of language. *Frontier Psychology* 6 (1796).

Hublin, Jean-Jacques, Ben-Ncer, Abdelouahed, Bailey, Sarah E., Freidline, Sarah E., Neubauer, Simon, Skinner, Matthew M., Bergmann, Inga, Le Cabec, Adeline, Benazzi, Stephano, Harvati, Katerina & Philipp Gunz, 2017. New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature* 546 (7657), 289-292.

Iwabuchi, Toshiki, Nakajima, Yasoichi & Michiru Makuuchi. 2019. Neural architecture of human language: Hierarchical structure building is independent from working memory. *Neuropsychologia* 123 (107137).

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge: MIT Press.

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

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jackendoff, Ray. 2009. The parallel architecture and its place in cognitive science, In Heine, Bernd & Heiko Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 593-668. Oxford: Oxford University Press.

Jackendoff, Ray. 2011a. What is the human language faculty? Two views. *Language* 87 (3), 586-624.

Jackendoff, Ray. 2011b. Alternative minimalist visions of language. In Robert D. Borsley & Kersti Börjars (eds.), *Nontransformational Syntax*, 268-296. Oxford: Wiley-Blackwell.

Jackendoff, Ray & Eva Wittenberg, E. 2014. What you can say without syntax: A hierarchy of grammatical complexity. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring Grammatical Complexity*, 65-82. Oxford: Oxford University Press.

Janik, Vincent M. 2013. Cognitive skills in bottlenose dolphin communication. *Trends in Cognitive Sciences* 17 (4), 157-159.

Jayaseelan, Karthik. 2017 Parallel work spaces in syntax and the inexistence of internal Merge. In Gautam Sengupta, Shruti Sircar, Madhavi G. Raman & Rahul Balusu (eds.), *Perspectives on the Architecture and Acquisition of Syntax*, 115-134. Singapore: Springer.

Johnson, David. & Lappin, Shalom. 1997) A critique of the Minimalist Program. *Linguistics and Philosophy* 20 (3), 273-333.

Jurafsky, Daniel. 1996. A probabilistic model of lexical and access and disambiguation. *Cognitive Science* 20, 137-194.

Kaan, Edith. & Tamara Y. Swaab. 2002. The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences* 6 (8), 350-356.

Katz, Jonah & David Pesetsky. 2011. The identity thesis for language and music. *lingbuzz/000959*

Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht, Holland: Foris Publications.

Kempen, Gerard & Pieter Huijbers. 1983. The Lexicalization process in sentence production and naming: Indirect election of words. *Cognition* 14 (2), 185-209.

Keshev, Maayan & Aya Meltzer-Asscher. 2017. Active dependency formation in islands: How grammatical resumption affects sentence processing. *Language* 93 (3), 549-568.

Klein, Richard G. 2008. Out of Africa and the evolution of human behavior. *Evolutionary Anthropology* 17, 267-281.

Klein, Wolfgang & Clive Perdue. 1997) The basic variety, or: Couldn't language be much simpler? *Second Language Research* 13, 301–347.

Koechlin, Etienne & Thomas Jubault. 2006. Broca's area and the hierarchical organization of human behavior. *Neuron* 50 (6), 963-974.

Koelsch, Stefan, Rohrmeier, Martin, Torrecuso, Renzo & Sebastian Jentschke. 2013. Processing of hierarchical syntactic structure in music. *Proceedings of the Academy of Sciences of the United States of America* 110 (38), 15443–15448.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19, 335-391.

Lee, Seung Jin, Lee, Soo, Song, Ji Yeong, Kim, Ga Young & HyangHee Kim. 2015. White matter connectivity as a neurophysiological mechanism for auditory comprehension in the neurologically normal and impaired. *Communicative Disorders* 20 (1), 121-132.

Lerdahl, Fred & Ray S. Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge: MIT Press.

Lerdahl, Fred & Ray S. Jackendoff. 2006. The capacity for music: What's special about it? *Cognition* 100, 33–72.

Leslie, Alan M., Gelman, Rochel & Charles R. Gallistel. 2008. The generative basis of natural number concepts. *Trends in Cognitive Sciences* 12 (6), 213-218.

Levelt, Willem J. M. 1993. *Speaking: From Intention to Articulation*. Cambridge: MIT Press.

Levinson, Stephen C. 2013. Recursion in pragmatics. *Language* 89, 149-162.

Linebarger, Marcia S., Schwartz, Myrna F. & Eleanor M. Saffran. 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 13 (3), 361-392.

Martins, Pedro Tiago. 2019. Progress overlooked. *Inference: International Review of*

Science 4 (3).

Martins, Pedro Tiago, Maties, Marí & Cedric Boeckx. 2018. SRGAP2 and the Gradual Evolution of the Modern Human Language Faculty. *Journal of Language Evolution* 3 (1), 67-78.

Martins, Pedro Tiago & Cedric Boeckx. 2019. Language evolution and complexity considerations: The no half-Merge fallacy. *lingbuzz/004509*.

Matchin, William & Gregory Hickok. 2019. The cortical organization of syntax. *Cerebral Cortex* (in press).

McBrearty, Sally & Alison S. Brooks. 2000. The revolution that wasn't: a new interpretation of the origins of modern human behavior. *Journal of Human Evolution* 39, 453-563.

Monti, Martin M. & Daniel N. Osherson. 2012. Logic, Language and the Brain. *Brain Research* 1428, 33–42.

Neubert, Franz-Xaver, Mars, Rogier B., Thomas, Adam G., Jerome Sallet & Matthew F. S. Rushworth. 2014. Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron* 81 (3), 700-713.

Nevins, Andrew, Pesetsky, David & Cliene Rodrigues. 2009. Pirahã exceptionality: A reassessment. *Language* 85 (2), 355–404.

Nóbrega, Vitor A. & Shigeru Miyagawa. 2015. The precedence of syntax in the rapid emergence of human language in evolution as defined by the integration hypothesis. *Frontiers Psychology* 6, 271.

Novick, Jared M., Trueswell, John C. & Sharon L. Thompson-Schill. 2010. Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass* 4, 906–924.

Ouattara, Karim, Lemasson, Alban & Klaus Zuberbühler. 2009. Campbell's monkeys use affixation to alter call meaning. *PLoS ONE* 4, e7808. doi:10.1371/journal.pone.0007808

Patel, Aniruddh D., Gibson, Edward, Ratner, Jennifer, Besson, Mireille & Phillip J. Holcomb. 1998. Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience* 10, 717–733.

Perani, Daniela, Saccuman, Maria C., Scifo, Paola, Anwander, Alfred., Spada, Danilo, Baldoli, Christina, Poloniato, Antonella, Lohmann, Gabriele & Angela D. Friederici. 2011. Neural language networks at birth. *Proceedings of the National Academy of Sciences of the United States of America* 108 (38), 16056–16061.

Petersson, Karl-Magnus, Folia, Vasiliki & Peter Hagoort. 2012. What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language* 120, 83–95.

Pillay, Sarah B., Binder, Jeffrey R., Humphries, Colin J., Gross, William L. & Diane S. Book. 2017. Lesion localization of speech comprehension deficits in chronic aphasia. *Neurology* 88 (10), 970–975.

Pinker, Steven & Ray Jackendoff. 2005. The Faculty of Language: What's special about it? *Cognition* 95, 201–236.

Progovac, Ljiljana. 2010. Syntax: Its evolution and its representation in the brain. *Biolinguistics* 4, 234-254.

Progovac, Ljiljana. 2012. Compounds and commands in the evolution of human language. *Theoria et Historia Scientarum* 9, 49–70.

Progovac, Ljiljana. 2015. *Evolutionary Syntax*. Oxford: Oxford University Press.

Progovac, Ljiljana & Locke, John. 2009. The urge to merge: Insult and the evolution of syntax. *Biolinguistics* 3, 337–354.

Rauschecker, Josef P. & Sophie K. Scott. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience* 12 (6), 718–724.

Rey, Arnaud, Perruchet, Pierre & Joel Fagot. 2012. Centre-embedding structures are a by-product of associative learning and working memory constraints: Evidence from baboons (Papio Papio). *Cognition* 123, 180–184.

Rilling, James K., Glasser, Matthew F., Preuss, Todd M., Ma, Xiangyang, Zhao, Tiejun, Hu, Xiaoping & Timothy E. J. Behrens. 2008. The Evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience* 11, 426–428.

Rizzi, Luigi. 2016. Monkey morpho-syntax and Merge-based systems. *Theoretical Linguistics* 42 (1-2), 139-145.

Rogalsky, Corianne & Gregory Hickok. 2011. The role of Broca’s area in sentence comprehension. *Journal of Cognitive Neuroscience* 23, 1664-1680.

Rohrmeier, Martin, Fu, Qiufang & Zoltan Dienes. 2012. Implicit learning of recursive context-free grammars. *PLoS ONE* 7 (10): e45885. doi:10.1371/journal.pone.0045885

Rohrmeier, Martin. 2011. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music* 5, 35–53.

Rosch, Eleanor H., Mervis, Carolyn B., Gray, Wayne D., Johnson, David M. & Penny Boyes-Braem. 1976. Basic objects in natural categories. *Cognitive Psychology* 8 (3), 382–439.

Sakas, William G. & Janet D. Fodor. 2012. Disambiguating syntactic triggers. *Language Acquisition* 19, 83–143.

Sandler, Wendy, Meir, Irit, Padden, Carol & Mark Aronoff. 2005. The emergence of grammar in a new sign language. *PNAS* 102 (7), 2661–2665.

Schenker, Natalia M., Hopkins, William D., Spocter, Muhammad A., Garrison, Amy R., Stimpson, Cheryl D., Erwin, Joseph M., Hof, Patrick R. & Chet C. Sherwood. 2010. Broca's area homologue in chimpanzees (Pan troglodytes): Probabilistic mapping, asymmetry, and comparison to humans. *Cerebral Cortex* 20 (3), 730-742.

Schlenker, Phillipe, Chemla, Emmanuel, Arnold, Kate & Klaus Zuberbühler. 2016. Pyow-hack revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua* 171, 1-23.

Schwartz, Myrna F., Saffran, Eleanor M. & Oscar S. M. Marin. 1980. The word order problem in agrammatism: I. Comprehension. *Brain and Language* 10 (2), 249-262.

Scott, Brian H., Mishkin, Mortimer & Pingbo Yin. 2012. Monkeys have a limited form of short-term memory in audition. *Proceedings of the Academy of Sciences of the United States of America* 109 (30), 12237-12241.

Senghas, Ann, Kita, Solaro & Asli Ozyurek. 2004. Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305 (5691), 1779-1782.

Seuren, Pieter A. M. 2004. *Chomsky's Minimalism*. New York: Oxford University Press.

Seyfarth, Robert M. & Dorothy L. Cheney. 2014. The Evolution of language from social cognition. *Current Opinion in Neurobiology* 28, 5–9.

Smith, Edward E., Jonides, John, Marshuetz, Christy & Robert A. Koeppe. 1998. Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences of the United States of America* 95 (3), 876-882.

Tallerman, Maggie. 2014. No syntax saltation in language evolution. *Language Sciences* 46 (B), 207-219.

Tallerman, Maggie. 2017. Can the integration hypothesis account for language evolution? *Journal of Neurolinguistics* 43 (B), 254-262.

Thothathiri, Malathi, Kimberg, Daniel Y. & Myrna F. Schwartz. 2012. The neural basis of reversible sentence comprehension: Evidence from voxel-based lesion symptom mapping in aphasia. *Journal of Cognitive Neuroscience* 24, 212-222.

Tyler, Lorraine K. & William D. Marslen-Wilson. 2007. Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions Royal Society B: Biology Science* 363, 1037-10.

Uriagereka, Juan. 2002. Multiple spell-out. In J. Uriagereka (ed.), *Derivations: Exploring the Dynamics of Syntax*, 45-65, London: Routledge.

Van der Lely, Heather K. J. & Pinker, Steven. 2014. The biological basis of language: insight from developmental grammatical impairments. *Trends in Cognitive Sciences* 18 (11), 586-595.

Vosse, Theo & Gerard Kempen. 2000. Syntactic Structures Assembly in Human Parsing. A computational model based on competitive inhibition and a lexicalist grammar. *Cognition* 74, 105-143.

Wakita, Masumi. 2014. Broca's area processes the hierarchical organization of observed action. *Frontier Human Neuroscience* 7 (937).

Wakita, Masumi. 2016. Interaction between perceived action and music sequences in the left prefrontal area. *Frontiers in Human Neuroscience* 10, 656.

Wakita, Masumi. 2019. Auditory sequence perception in common marmosets (Callithrix jacchus). *Behavioural Processes* 162, 55–63.

Wilson, Stephen M., DeMarco, Andrew T., Henry, Maya L., Gesierich, Benno, Babiak, Miranda, Miller, Bruce L. & Maria Luisa Gorno-Tempini. 2016. Variable disruption of a syntactic processing network in primary progressive aphasia. *Brain* 139, 2994-3006.

Zuberbühler, Klaus., Cheney, Dorothy L. & Robert M. Seyfarth. 1999. Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology* 113 (1), 33-42.