



Biolinguistics and biological systems: a complex systems analysis of language

Ryan M. Nefdt¹

Received: 27 March 2022 / Accepted: 19 February 2023 / Published online: 6 March 2023
© The Author(s) 2023

Abstract

In their recent book, Ladyman and Wiesner (What is a complex system?, Yale University Press, 2020) delineate the bounds of the exciting interdisciplinary field of complexity science. In this work, they provide examples of generally accepted complex systems and common features which these possess to varying degrees. In this paper, I plan to extend their list to include the formal study of natural language, i.e. linguistics. In fact, I will argue that language exhibits many of the hallmarks of a complex system, specifically a complex biological system. Thus, my aim is to advocate *contra* the the ‘Minimalist Program’ (Chomsky, The minimalist program, MIT Press, Cambridge, 1995), which motivates simple underlying mechanisms (i.e. Merge) in their idealisations, that biolinguistics should embrace a ‘Maximalist Program’ in which multiple subfields contribute component explanations to an emerging whole.

Keywords Systems biology · Biolinguistics · Complex systems · Language · Emergence · Language evolution

*“nothing is gained by labeling the propensity for language as biological unless we can use this insight for new research directions – unless more specific correlates can be uncovered” -
Lenneberg 1964.*

Introduction

Biolinguistics is a term which broadly characterises a particular naturalistic approach to the study of language. Its precise methodological and ontological nature is the topic of the present article. In fact, my aim is both descriptive and normative. In so far as biolinguistics hopes to exemplify a strong analogy with biology, I argue it

✉ Ryan M. Nefdt
ryan.nefdt@uct.ac.za

¹ Department of Philosophy, University of the Cape Town, Cape Town, South Africa

needs to embrace a different path from what linguists have generally taken the field to be. In fact, Sect. 2 attempts to both disambiguate the different possible senses of the aforementioned term within the extant literature and also align the most plausible instantiation with its most prominent framework: the Minimalist Program (Chomsky 1995). In Sect. 3, I provide arguments against this conception of biolinguistics on the grounds that it fails to properly connect to the biological sciences. Finally, in Sect. 4, I present a novel interpretation of biolinguistics in terms of systems biology and thus an ontological picture of natural language as a complex biological system. The resulting programmatic suggestion, I call the *Maximalist Program*.

There have been a number of articles questioning the philosophical foundations of biolinguistics (Lappin et al. 2000; Behme 2015; Levine 2018). However, what sets the current work apart from these largely negative critiques is that I focus exclusively on the biological claims made by biolinguists and in addition I offer a novel path to underpinning the field within the biological sciences and complex systems analysis alike.

Three grades of biological involvement

There are a number of ways in which biology and linguistics can be combined or jointly pursued. Therefore, biolinguistics as a discipline encompasses a multitude of possible instantiations for scientific partnership. In the literature, it is not always clear which of these possibilities is being realised. The lack of clarity plagues the field. This is, of course, not entirely surprising or uncommon in any nascent scientific pursuit. Thus, in the following I will utilise the structure of Quine's (1976) infamous 'three grades of modal involvement' in order to differentiate between different ways of interpreting biolinguistics. The first grade carries the strongest biological commitment while the last incorporates the weakest. Each grade has its advantages and drawbacks and can include a number of extant frameworks.

As I see it, there are three possible and viable interpretations of biolinguistics to be found in the literature.¹ The options are as follows:

1. The inclusion of the study of language within the current biological sciences.
2. The formal study of language constrained by biological principles.
3. The extension of biology to include linguistics as a subdiscipline.

The first option makes the closest connection to biology in terms of methodology and ontology. I will provide examples of this approach while highlighting the problems with adopting it as a general biolinguistic strategy. The second interpretation is chiefly advocated within the current epoch of generative grammar, namely the Minimalist Program (MP). But importantly, it is not the only possible instantiation of this general approach. I will provide an alternative in Sect. 4. The last grade is the most superficial. It involves business as usual for linguists. Some

¹ The important word here is 'viable' since as Martins and Boeckx (2016) point out, there are many more senses to be found than this, most of which are nonstarters.

reinterpretation of the technology might be in order but overall the methodological agenda remains constant.

First grade: the neurobiology of language

Neurobiology is the broad and interdisciplinary study of the organisation of the nervous system and brain function of vertebrates like us. The nervous system itself consists of the brain, spine, various neural circuits and nerves that are spread throughout the body. The field encompasses methods from physics, physiology, molecular biology, genetics, anatomy *inter alia*. Although a recent science, the accomplishments of the field are already numerous.

Modern neurobiologists know basically how information is encoded in our nervous systems, can place constraints on the costs of this encoding, have a fairly complete (and surprisingly sparse) map of what connects to what in the human brain, and can even explain biophysical/molecular constraints on the form and costs of learning. (Glimcher 2014: 63)

The above claim overstates the current state-of-the-art to a certain extent and, with a strong reading of ‘know’, is controversial. There are still many open questions, including basic issues such as how words and concepts are stored in memory (Poeppel and Iscardi 2022). Nevertheless, the specific focus of this scientific purview and these techniques on the representation, production and acquisition of language is called neurolinguistics. Specifically, neurolinguistics is the study of how language is represented and mapped in the brain. Thus, the first way in which we could understand biolinguistics is as the neurobiological study of language.

What does such a study entail for the discipline, ontologically and methodologically? Essentially, the basic ontology would be shared by its parent discipline: neurobiology. Thus, our scientific explanations would involve glial cells and neurons (and their synaptic connections). While the former mostly serve support and regulatory roles, the latter form the crux of the computational functions of the brain.

The general methods also follow from neurobiology in incorporating brain imaging technology such as MRI, fMRI, computed tomography used in event-related potential research in which direct brain responses are measured with specific stimuli. In neurolinguistics, research which attempts to locate the brain regions activated during the comprehension or production of words would qualify. For instance, Friederici et al. (2017) identify the Broca’s area (BA 44) among others in syntactic processing (as well as the dorsal pathway). “[W]ith respect to Broca’s area, the activation of BA 44 as a function of syntax has been confirmed in many studies across different languages” (Friederici et al. 2017: 714). More specifically, they link these areas to the ‘Merge’ postulate of later generative grammar. Merge is an operation which takes two syntactic objects and composes a labelled (unordered) set containing these objects, iteratively. Merge is meant to capture the allegedly universal property of the hierarchical structure of syntax, i.e. sentences are composed of embedded phrases which can themselves be embedded,

represented by a tree-like structure. However, Frank et al. (2012) challenge this latter claim by reviewing computational and experimental studies that indicate that hierarchical syntactic structure, as postulated by constituency-based grammars, are not necessarily represented during on-line language processing, and that, instead, “sequential structure” (flat representations of grammatical or semantic dependencies between words) is often enough to explain human or model performance.

This points to the first difficulty with interpreting biolinguistics in terms of neurolinguistics, namely it is unclear what role linguistic theory plays or ought to play. One can study the brain regions activated during speech production or sign comprehension without consulting linguistic theory, generative or otherwise. Even optimistic studies aimed at realising theoretical linguistic architectures do not clearly correspond to linguistic practice at a more precise level. Berwick et al. (2013) claim that brain imaging studies reveal the basic linguistic structure of a primary computational core of autonomous syntax with ‘externalisation’ in phonological and semantic output:

At the neural level, core computations may be differentiable from a sensory-motor interface and a conceptual system [...] In this context, two different dorsal located pathways have been identified, one involving Brodmann area (BA) 44 and the posterior superior temporal cortex (pSTC) that supports core syntactic computations and one involving the premotor cortex (PMC) and the STC that subserves the sensory-motor interface. (93)

Besides the well-known issues of correlation versus causation in neuroscience, a more specific worry is that the posits of linguistic theory are more rarefied than just the division of syntax, semantics, phonology and their interfaces. Words might be identifiable in event-related potentials at some level, but hierarchical phrases, island effects, covert operators and other structures are much further from realisation. Yet these constructions are the mainstay of linguistic theory and they are completely invisible to neurobiology at this stage of the science. It seems that we are still quite far from the realisation of the objects of linguistic theory in the glial and neural cells or synapses of the brain tissue. The ontology is distinct and reduction some ways off.

In addition, methodologically the fields are not clearly related. Linguists generally construct representations of linguistic structures, such as sentences or clauses, into tree diagrams and rules or constraints based on the varying degrees of grammatical well-formedness. Formal semanticists, for example, are interested in compositional meaning in which larger units are built from smaller ones and they model this property by means of functional application and type theory (Heim and Kratzer 1998). fMRI and CT studies do not form part of their training or tools.²

Thus, research in neurolinguistics neither needs to include nor presuppose work in theoretical linguistics. Baggio (2020) explicitly considers the role linguistic theory can and does play in generating new knowledge the neuroscience of language or what he calls ‘unidirectional epistemic transfer’. After surveying a number of candidate options such as constructions, parameters, syntactic and semantic composition, he concludes that despite convergence between linguistics

² See Baggio (2018) for an integration of neuroscience and formal semantics.

and neuroscience being considered a desirable goal, “achieving it has proved exceedingly hard. Because the concepts, methods, and background assumptions of linguistics and neuroscience are so different, their results, although fundamentally about the same object of inquiry—i.e., language as a human mental capacity—are not immediately integratable” (Baggio 2020: 298). Thus, he advocates for methodological pluralism. A sentiment shared by Christiansen and Kirby (2003) when it comes to language evolution where they claim that the expectation that linguists contribute to the field is neither generally met nor necessarily cause for concern. Furthermore, the neural structures of animal signalling systems would qualify as neurolinguistic and say nothing of generative grammars or Merge. This is the so-called comparative approach in neurobiology.³ Some scholars at the intersection of these fields specifically favour the association with neuroscience over that of generative linguistics. As Embrick and Poeppel (2005: 2) point out:

The idea that language can be approached in these terms is stressed in some recent work under the heading of Biolinguistics [see e.g. Chomsky (2002)]. While we are sympathetic to many of the (mostly programmatic) suggestions in Chomsky’s work, in practice much of the work that falls under that particular heading differs markedly in focus from the programme that we advance here.⁴

Of course, direct brain realisation might not be the only conduit to the neurobiological interpretation of biolinguistics. As previously mentioned, the field is broad and interdisciplinary. One of its subfields is genetics. Research on specific genes involved in or responsible for language development such as FOXP2 gene have gained some traction in recent decades (Lai et al. 2001; Konopka et al. 2009). Here some features or more specific properties identified by linguists have proven useful. In fact, both genetics and evolutionary developmental biology (evo-devo) have played a role in the research into the so-called ‘language gene’.

Briefly, the FOXP2 gene is responsible for making a protein called forkhead box P2 which is a transcription factor. This basically means that it controls the activity of other (possibly hundreds of) genes. It was first associated with linguistic relevance in a particular family (the KE family). Researchers noticed that around half of the members of this family suffered from a very rare autosomal dominant speech and language disorder, called developmental verbal dyspraxia, “which was shown to be due to a heterozygous point mutation in FOXP2, inherited by all the affected, but none of the unaffected individuals” (Scharff and Petri 2011: 2128). The symptoms of this disorder range from the delayed onset of speech and stuttering in affected individuals to dyspraxia. Of specific linguistic significance, it was noticed that affected KE family members had serious difficulties with plural formation and tense. This morphosyntactic effect touches on the issue of agreement in syntax

³ Chomskyans are generally dismissive of this approach. They argue that the core computational system or the narrow faculty of language (FLN) is species-specific (and domain specific) such that “this is a challenge for comparative research, as there may be “not much to compare”” (Friederici, 2017: 717).

⁴ The ‘programme’ to which they refer is an attempt to unify what they call computational-representational theories of language in cognitive science with neurobiology.

(why **They sees a whale* is ungrammatical in Standard English). The hypothesis was that affected members of the KE family were cognitively indistinguishable from nonaffected members aside from subtle grammatical effects. Gopnik (1990), who first reported the discovery, intimates that the defect lies in an inability to follow general grammatical rules (such as add ‘s’ to plurals) as evinced by a failure to generate nonsense plurals with which average speakers of English see no difficulty. In fact, Pinker (1994), in his popular science book, goes as far as to suggest that this provides evidence for FOXP2 as ‘grammar gene’ of some sort. This was an exciting possibility, namely that specific genes could be linked directly with the processing of specific grammatical functions or rules, i.e. with posits of linguistic theory. Furthermore, theoretical claims such as the existence of innate linguistic structures were thought to be bolstered by this sort of negative evidence in which impaired individuals seem to lack some aspects of the Universal Grammar (UG) present in their kin. UG is thus confirmed as a genetic endowment.

Unfortunately, since the initial hype, the precise relevance of the gene for the development of language has been further complicated. Firstly, as Sampson (2005) notes, drawing on subsequent research and replication studies (e.g. Vargha-Khadem et al. 1995):

The press sensationalized this as discovery of ‘a gene for language’, but that was a misunderstanding. The FOXP2 mutation gives those who bear it low general intelligence relative to their unaffected kin; and among other things, it damages their ability to execute simple sequences of actions. Language is heavily dependent on the ability to execute complex sequences of actions rapidly and accurately, so there is little wonder that the affected KE family members have a wide range of problems with language. (125)⁵

Secondly, more recent comparative evo-devo studies show evidence for ‘deep homologies’ in the FoxP2 gene across species with no linguistic abilities akin to human language. The gene is expressed in our closest primate relatives, other mammals and even songbirds. Nevertheless, the comparative approach could shed light on why our expression of the Fox gene might have resulted or contributed to the emergence of natural language in humans. Much of this research, however, is related to speech production or vocalisation, what generative linguists call ‘externalisation’. This aspect of language is not considered to be its primary essence. So there is little hope of genetics offering a clear bridge to theoretical linguistics from biology. As Martins and Boeckx (2016: 7) state “[n]umerous practitioners in biology know that this gene-centric view is far too simplistic. There is no direct route from a linguistic entity...and a gene or genes.”

In the neurobiology of language, biology has a strong foothold but linguistic theory does not share an equal or even required position in the broader research enterprise. Although neurolinguistics is a thriving field of inquiry, biolinguistics cannot be identified with it on pain of losing the essence of linguistic explanation.

⁵ The gene has indeed been linked to motor function but it also seems to have specific linguistic relevance. Sampson’s claim ignores that the impairments resulting from this genetic defect largely affects speech and language production.

This does not mean that linguistics and biology cannot fruitfully collaborate. But it does indicate a clear separation in methodology and present ontology.

Second grade: linguistics biologically constrained

The second grade of biological involvement aims for a balance between linguistic theory and biological constraints. In other words, it aims to restrain or constrain linguistic theory by means of principles derived from biology. The Minimalist Program (MP) (Chomsky 1995) is one prominent instantiation of this biolinguistic possibility. A full treatment of the framework would require explanations of concepts of optimality, interfaces, virtual conceptual necessity, Merge and so on. We'll briefly review the concepts necessary for understanding the basic idea behind the programme but our focus will be on the aspects the second grade of involvement.⁶

Firstly, MP is distinguished from linguistic theories of the past such as the Standard Theory, Extended Standard Theory and Government and Binding (GB) (Chomsky 1981). It was not meant as a theory itself but like biolinguistics, it is an approach or programme, i.e. there are many ways to practice linguistics in a Minimalist fashion. In an attempt to meet two essentially biological constraints, innateness and the biodiversity of languages, Minimalism takes on board what is known as the Principles and Parameters (P&P) model to different degrees.⁷ This model specifies a finite set of universal linguistic principles and derives parametric settings based on external experience of conventions found in one's linguistic community. In other words, UG can be characterised by universal principles and the variation found across the world's languages can be accounted for by the parameters imposed by the primary linguistic data (PLD) of the environment.

Minimalism marks a philosophical as well as architectural departure from earlier versions of generative grammar. In some ways it does attempt to exemplify aspects of option (2) more so than any other generative framework which came before. The first such attempt is the idea that prior linguistic theory was bloated in its explanations. For example, GB consists of a number of subtheories or modules such as Case theory, Theta theory, Binding theory, Bounding Theory, Control Theory, and Government Theory. Minimalism sheds most of this structure in the spirit of asking not 'how much' theory is required to explain language but 'how little', hence the minimalism. However, economy is not yet biology. That comes with Chomsky's further requirement on the ambitions of linguistic theory, a condition he calls 'beyond explanatory adequacy' (Chomsky 2004). Those familiar with early versions of the goals of generative grammar will recall the nested adequacy conditions of the *Aspects* model (Chomsky 1965). There the main goals were *descriptive adequacy* (basically capturing the grammatical intuitions of speakers or linguistic data) and

⁶ The focus in this section is on my analysis of the second grade of involvement from the biolinguistic literature. However, there has been promising non-generative work on the impact of cognitive biases on both typological distributional patterns among languages and individual learning. See Culbertson et al. (2013) for one such proposal.

⁷ As noted by an anonymous referee, the degrees range from viewing parameters as *bona fide* theoretical objects to eliminating them from Minimalism altogether.

explanatory adequacy or accounting for how a child acquires language. What MP adds is a layer beyond this by involving evolutionary considerations. As Johnson puts it:

Evolutionarily speaking, it is hard to explain the appearance of highly detailed, highly language-specific mental mechanisms. Conversely, it would be much easier to explain language's evolution in humans if it were composed of just a few very simple mechanisms (Johnson 2015: 175).

Minimalism attempts to do exactly this. It posits very simple mechanisms such as the aforementioned set-theoretic operation of Merge which outputs one syntactic object from the composition of two separate such objects. This mechanism is part of what is known as 'Bare Phrase Structure', a bottom up reversal of the top-down X-bar theory of early frameworks. Basically, unlike earlier versions of GG where you had deep structures from which you derive or transform surface structures like questions or passives, MP says all you have is a set of lexical items and a bare tree structure ('bare phrase structure'). Then there's a procedure for taking an item from the set to create a subtree. You then merge that product with another item and project a head creating another tree and so on until you derive sentences ready for pronunciation/interpretation, i.e. to be sent off to the interfaces. "The system linking these interfaces is the *minimal* system that satisfies "legibility" constraints or conditions imposed by both the [perceptual/articulatory] system and the [conceptual/intentional] system" (Ludlow 2011: 36). In fact, MP aspires to an ideal in which the system is perfectly designed not as a result of messy evolutionary biological processes but rather driven by 'virtual conceptual necessity'. This ideal acts like a benchmark to which reality can be compared. This is the so-called 'Strong Minimalist Thesis' (SMT) which states that language is the optimal matching between sound and meaning as per the demands of the phonetic and semantic interfaces. Moreover, "the SMT holds that the merge function, along with a general cognitive requirement for efficient computation and minimal search for agreement and labelling operations, suffices to account for much of human language syntax" (Friederici et al. 2017: 714). Merge answers the question of innateness anew by linking it to the evolution question and specifying a simple mechanism at the core of the genetic inheritance of language users like us.

So what does a biolinguistic analysis look like on this view? Well, it's best to compare both the architecture of the language faculty under MP to that of its immediate predecessor, GB. In MP, there is basically a lexicon that feeds into a computational system upon which merge operates to create syntactic objects. This is an internal, isolated system aspects of which are externalised via interfaces to our sensory-motor systems and conceptual-intentional systems respectively (see figure below).⁸ In other words, the Lexicon feeds into a computational procedure which creates formal structural mental objects. These objects are then translated to be read

⁸ There is also a distinction between overt and covert syntax related to the interfaces. As Müller (2018: 128) explains "[o]vert syntax stands for syntactic operations that usually have a visible effect. After overt syntax the syntactic object is sent off to the interfaces and some transformations may take place after this Spell-Out point". However, these transformations are hidden from PF (unpronounced) so they form part of the covert component related to the semantic interface. We'll return to some biological problems with the postulation of covert material in Sect. 4.

(and used) by the systems of externalisation (like phonological output) and semantics (Fig. 1).

In Government and Binding theory, there are by contrast four levels of representation. The underlying or D-structure is mapped onto S-structure by the all encompassing *move- α* operation (move anything anywhere), a version of which also facilitates the relationship between S-structure, or a representation of surface word order, with the interface responsible for semantics: Logical Form (LF). The D-structure representation itself is fed by the Lexicon, the phrase-structure rules (X-bar theory), and the Θ -theory (which contains constraints on the X-bar representation). The case filter makes sure that every noun phrase has a case at S-structure, even in languages such as English or Afrikaans where inflection has mostly overtly disappeared. The systems acts like a conveyor belt along a complex assembly line, where every structure passes through each point whether or not the final product reflects it on the surface form. We'll return to some details of GB in Section 4.2 but for now the point is that the various constraints and features of the language faculty under this paradigm were motivated by theory internal linguistic considerations not the human brain or evolutionary concerns. The movement from this model to MP was meant to correct this trajectory of linguistic theory in favour of strong biological constraints by means of purging it of much of its structure (Fig. 2).

The Minimalist architecture, unlike the GB one, is meant to be motivated by biological considerations such as simplicity, efficiency and evolution. In other words, what kind of linguistic system is likely to have emerged within a short period of time. A highly complex, highly modular, multilevel language specific model like GB? Or a simple computational system with interfaces like Minimalism?

With these considerations (with some details left out), we have a viable interpretation of biolinguistics not as identical to MP but as exemplified by the latter's constraints on any theory of language worthy of the programme. Simplicity, economy, and evolutionary design are the alleged driving factors which should constrain any particular grammatical representation and linguistic architecture. Thus, the three specific biological constraints which the second grade of biological involvement (under Minimalism) imposes on linguistic theory are: (1) an explanation of the role of innateness, (2) an explanation of the immense diversity of world languages, and (3) a plausible evolutionary account of the emergence of natural language in our species. I believe, and will show, that these and other constraints can be met by a very different biological picture of the science, one based on complexity and biological systems. However, before we get to that, there is one more grade of involvement that needs to be explored.

Third grade: linguistics is biology

The last grade of biological involvement simply states that biolinguistics and generative linguistics are one and the same. This possibility is contrary to the scientific progression of the field suggested in the previous section. Indeed some practitioners do claim that biolinguistic concerns were essentially 'there from the start'. Boeckx and Grohmann (2007) call this the 'weak sense' of biolinguistics or

‘business as usual’ while the ‘strong sense’ genuinely attempts interdisciplinary influence and confluence. Although I agree with their perspective, I hope to sketch a possibly stronger interpretation of the ‘business as usual’ model.

The idea is that instead of including the study of language within biology (first grade) or constraining it by means of biological principles (second grade), we merely extend biology to include linguistics as is as a proper subset. This might seem like a distinction without a difference. But there is a subtle one to be gleaned. Whereas the methodological implications of grades one and two might require either finding strong analogies (or even homologies) between linguistic grammars and biological entities and structures or shedding biologically implausible structure, this view claims that linguistic grammars and posits were in some sense biological all along. The claim is captured by what Martins and Boeckx (2016: 5) call the “linguistics is biology at a suitable level of abstraction mantra”. In distinguishing the generative grammar perspective from views which characterise knowledge of language as an ability, Chomsky (2000: 50) writes:

This view contrasts with the conception of a language as a generative procedure that assigns structural descriptions to linguistic expressions, knowledge of language being the internal representation of such a procedure in the brain (in the mind, as we may say when speaking about the brain at a certain level of abstraction).

The key phrase is “certain level of abstraction”.⁹ In other words, linguistics is biology at a certain level of abstraction. But what does this mean exactly? It is unclear from the literature. Perhaps that biology includes linguistics after all, just perhaps as one of its more inchoate subdisciplines. This is one way of appreciating what is meant by ‘language is a biological object’. And furthermore, biological objects can be studied from varying points of view. Linguistic theory offers us one such perspective.¹⁰

One problem with this grade of involvement is that it belies one of the main scientific aims of biolinguistics, namely methodological convergence. Smith (2000: vii) highlights this scenario when he reflects on Chomsky’s biolinguistic view as “[h]uman language is therefore a psychological, ultimately a “biological object,” and should be analysed using the methodology of the natural sciences”. Similarly, Fitch writes “I consider it self-evident that the appropriate models for biolinguistics come from the natural sciences, such as physics” (2009: 291). But linguistics under the third grade does not advocate incorporating the same methodology as psychology (see Soames 1984), biology, or the natural sciences. Arguably it is psycholinguistics

⁹ Anderson and Lightfoot (2002: 17) similarly state “the object of study, *I*-language, is an aspect of the structure of the mind/brain, the language organ, defined at some level of abstraction, which must be studied as such and whose study can lead to further insights about the architecture of cognition.” This statement includes the organ terminology and the levels of abstraction mantra.

¹⁰ Interestingly, at some point in the heyday of the early generative revolution, biologists even considered how other systems might be like generative grammars. See Jerne (1985) for a generative grammar of the immune system, and also Marr (1982) for a like between generative grammar and vision at the computational level.

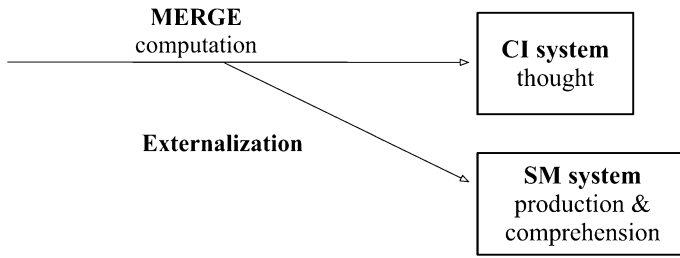


Fig. 1 Basic MP architecture (from Wiltschko 2022)

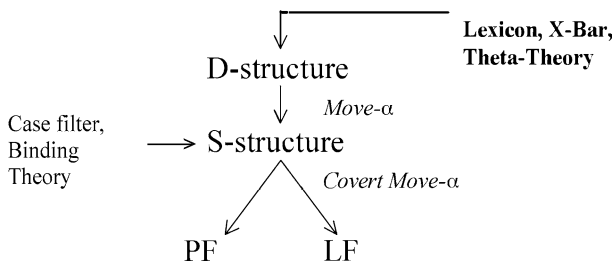


Fig. 2 GB architecture

that incorporates tools and techniques from psychology into the study of language and, as we have seen, neurolinguistics that does so for biology. My account in Section 4.3 will aim to connect linguistics to physics and other complexity sciences. Furthermore, both Poeppel and Embick (2005) and Mondal (2020) argue that there are categorical and ontological mismatches between the explanatory levels of linguistic theory and those of neurobiology and cognitive science respectively.¹¹ If this is true, then the ‘level of abstraction mantra’ is at risk on falsehood.

Moreover, Chomsky (2000) himself has heavily criticised what he considers the methodologically dualist approach of traditional philosophers of language who incorporate a priori reasoning into their theories of language and mind. He contrasts this practice with ‘methodological naturalism’ in which the tools of the natural sciences are used to investigate these domains. But true methodological naturalism would surely also militate against the ‘business as usual’ approach since it fails to connect with contemporary biology, even in the form of constraints the likes of which we saw in the previous section. In fact, the formalist and computationalist methodology of generative linguistics has also been a point of disconnection between the field and its cognitive scientific cousins (Jackendoff 2002; Sinha 2010).

¹¹ Poeppel and Embick introduce the *Ontological Incommensurability Problem* to capture the incompatibility between the posits of linguistic theory and those of neuroscience.

An important caveat: at no point can a view in second and third grades be incompatible with first grade accounts of the neurobiology of language. Given that there is still much to learn about the neural mechanisms and structures behind language, this isn't a strong constraint as it stands. In fact, it resembles what Ladyman and Ross (2007) call the 'primacy of physics constraint' on the special sciences. In other words, linguistics might represent a different level of description of phenomena but it cannot present one which is incompatible with the facts that emerge from neurobiology (nor physics for that matter). I personally believe this is strong enough to capture methodological naturalism (without methodological reductionism).

Linguists, on this grade, often attempt to adjust the traditional terminology in favour of more biologically aligned terms. Universal grammar is described as a species-specific genetic endowment present in all non-impaired human beings at birth and responsible for linguistic development (see Chomsky 1986, 1995, 2007; Pesetsky 1999). However, the precise demarcation of UG has never been settled (see Dabrowska 2015). In fact, with the advent of Minimalism (canonically associated with biolinguistics), UG was minimised to include only basic operations and constraints on the grammar. Anderson and Lightfoot (2000), who argue for a biological analogy of language as a distinct 'organ' based on poverty of stimulus and modularity arguments, make the link between genetics and UG when they state:

[L]anguage emerges through an interaction between our genetic inheritance and the linguistic environment to which we happen to be exposed [...] At the center is the biological notion of a language organ, a grammar. (703)

By 'grammar' they mean the standard tool in linguistics involving finite rule systems which generate structural descriptions of possible human linguistic expressions. These grammars are infinite in output based on the recursive rule representations used in linguistic models. But such claims, linking formal models to genetics, are far from the programmes suggested in the first two grades above. They suggest that linguists should keep conducting research as before but with a new set of terms to describe their work when pressed. Hornstein et al. (2005: 3) add that "generative grammarians have postulated that children come biologically equipped with an innate dedicated capacity to acquire language - they are born with a language faculty". Statements like these are commonplace in linguistics textbooks where talk of 'organs', 'genotypes', and Mendelian genetics are often found. However, not many of these concepts figure in the linguistic analysis of WH-movement, anaphora, covert operators, and structural constraints on trees, under this grade of involvement at least. There is a strong sense that the biology is 'conservative' in the nominalist sense of Field (1980). In other words, the claims of linguistic theory on this grade of biological involvement can be made with or without the biological analogy.

However, there is a line of objection which might support the first grade. The thought is that biology itself is heterogeneous. Since it involves no methodological boundary on the kinds of investigations it licenses, linguistics, psychology, ethology etc. can all look very different but still count as biological, perhaps in the mere fact that they target a biological domain of inquiry. In the case of linguistics, the domain is the human mind or some subsystem thereof. A shift in terminology, such as 'life sciences', might then obviate much of the controversy surrounding

the idea of calling these disciplines biological.¹² I'm certainly sympathetic to these sorts of worries. The boundaries between sciences should not be fixed by fiat. In fact, the view I will put forward in section 4.3 might be considered to be making a similar move. Nevertheless, the grades of involvement, as I have defined them, are ordered in terms of methodological overlap. And without this component, this kind of argument can easily overgenerate to allow for almost every discipline from sociology to ethics to count as biology.¹³

As an approach that respects both linguistic theory and aspects of the biological sciences, the second grade of involvement is still the best candidate for biolinguistics. Despite this, I hope to show in the next section, that biolinguistics *qua* minimalism fares less favourably upon closer inspection.

Biological problems with biolinguistics

In this section, I focus my attention on incongruities between biolinguistics (under minimalism) and the biological sciences. The aim is to show that MP is not the only instantiation of the second grade of involvement, and given the complexities of language evolution specifically, other approaches might fare better with relation to explanatory depth.

Chomsky's controversial gambits

One of the most profound claims of modern linguistic theory has its roots in the early recursion and proof-theoretic leanings of the Standard Theory (Chomsky 1956, 1957; Lobina 2017).¹⁴ The idea is that natural language is in some important sense infinite. Where recursion theory comes in is as a means of capturing how a

¹² Of course, many non-generativists do not hold that linguistics has as its target a psychological or biological system and still hope to maintain a majority of the current methodology. See Itkonen (1978), Kac (1992), Devitt (2006).

¹³ Again, my interlocutor could insist that this is indeed a benign possibility, highlighting the arbitrariness of such distinctions in the first place. But the instrumentalist in me holds on to the idea that some disciplinary boundaries prove themselves useful, even if they aren't in the nature's joint carving business. I thank an anonymous referee for pushing me on this challenging line of thought.

¹⁴ The scientific underpinnings of generative linguistics go further than this to what is often referred to as 'the Galilean Method' which allows for the rejection of recalcitrant data. For instance, Chomsky (1995: 26) admits that natural language is a complex biological system which emerged from messy evolutionary processes and accidental environmental conditions. But nonetheless maintains that a more "fruitful working hypothesis" is that the basic structure of language results from simplicity and elegance not common to complex organic systems. This is one of the ontological imports of the method which Allott et al. identify as "nature is mathematically organized, and that the underlying principles of nature are in some sense elegant or simple" (2021: 522). Thus, the data associated with language might be messy and complex but the Galilean method allows us to abstract away from those details to home in on the true nature of language given to us by Merge and other conceptual necessities. Boeckx (2006: 124) states of the connection between the Galilean Method and MP that "[s]ometimes, the mathematical results are too beautiful to be untrue, so that it seems justifiable to stick to the theory, while setting aside problematic or even conflicting data."

finite system like the human brain can generate an infinite output, what generative linguists call ‘Humboldt’s Problem’ in honour of the linguist Wilhelm von Humboldt (Boeckx 2015). Formal ‘generative’ grammars are devices for capturing this essential property of natural language. Recursive rules are incorporated into the grammars which allow for discretely infinite output. The infinitude claim takes on many forms. In some cases, the essential property is described as ‘discrete infinity’, in others it is ‘recursion’.¹⁵ In some ways, these properties are orthogonal. As Pullum and Scholz (2010) show, recursive structures do not entail infinite output.¹⁶ In other words, not all grammars with recursive rules allow for infinite output. And when discrete infinity is a feature of a formal grammar’s output, this doesn’t mean that the languages the latter models inherit this property. It could merely be a feature of the formal model and not the target system (Tiede and Stout 2010; Nefdt 2019). What is more important for our purposes is that this property, whatever label it takes, limits the application of analogies from biology, according to these linguists.

Some basic properties of language are unusual among biological systems, notably the property of discrete infinity. A working hypothesis in generative grammar has been that languages are based on simple principles that interact to form often intricate structures, and that the language faculty is nonredundant, in that particular phenomena are not “overdetermined” by principles of language. These too are unexpected features of complex biological systems, more like what one expects to find (for unexplained reasons) in the study of the inorganic world. (Chomsky 1995: 154)

Thus, biolinguistics starts with a very controversial assumption, namely that its actual target is biologically anomalous. This might indeed be the case, but I do not believe that the biological resources were exhausted prior to this determination. In fact, as I will show, the allegedly biological anomaly of natural language draws from a controversial claim about its emergence or evolution.¹⁷

To see how this is the case, let us consider the evolutionary claim of Hauser et al. (2002), developed further in Berwick and Chomsky (2016). The central Merge operation, which produces a single object from two separate syntactic objects and projects the head of one of them to the overarching structure, is said to be an evolutionary mutation responsible for the alleged rapid emergence of human language.

At some time in the very recent past, apparently sometime before 80,000 years ago, if we can judge from associated symbolic proxies, individuals in a

¹⁵ This is often referred to as ‘the Basic Property’ or “each language provides an unbounded array of hierarchically structured expressions that receive interpretations at two interfaces, sensorimotor for externalization and conceptual-intentional for mental processes” (Chomsky 2013: 647).

¹⁶ They do so by limiting the lexicon of a toy example to create space for inapplicable recursive rules. A recursive structure without the lexical items to apply *ad infinitum* does not produce infinite output.

¹⁷ It is important to note that I am not saying that human language (or competence) is not biologically unique. It might very well be the case that it is. Nor am I claiming that biological uniqueness is out of step with biology *simpliciter*. My argument is that biological uniqueness is a stronger claim that requires stronger evidence to prove. Unfortunately, the counter evidence is compelling, as we shall see.

small group of hominids in East Africa underwent a minor biological change that provided the operation Merge - an operation that takes human concepts as computational atoms and yields structured expressions that, systematically interpreted by the conceptual system, provide a rich language of thought. These processes might be computationally perfect, or close to it, hence the result of physical laws independent of humans. (Berwick and Chomsky 2016: 87)

There are a few components to this evolutionary thesis. The first is that Merge is supposed to be a single genetic mutation which emerged in an individual or a few individual hominid ancestors of ours. It was a *macro mutation* or a mutation that has a massive effect on the organism going forward, one that essentially rewired the human brain and gave rise to language. The reason for the need for a single macro mutation, according Berwick and Chomsky, is temporal. Language emerged around 100,000 years ago in our species.¹⁸ Thus, the usual resources of natural selection are unavailable to us since their processes tend to take much longer to effect change.

Another limiting assumption of this kind of proposal is the claim that language might have evolved exclusively for the purpose of internal thought and not communication. Different generativists disagree on the extent of this claim (with some following Chomsky in even denying any real significance to the idea that language evolved for any specific purpose). In the strong sense, any evidence from the neurobiology of speech production, animal vocalisations, symbolic processing in bees or other species is rendered ‘peripheral’ at best by this assumption. The internal computational system at the heart of this view is central and “there is no empirical evidence that any non-human species has such a system, suggesting that language is human-specific” (Friederici et al. 2017: 717). While they also claim that “communication is merely a possible function of the language faculty, and cannot be equated with it” (Friederici et al. 2017: 713). Of course, for Chomskyans, evolutionary biology holds limited analogies but it isn’t out of step entirely with a Darwinian picture since the mutation was selected for its benefits to thinking which is meant to also explain the rapidity of its spread.¹⁹

Furthermore, there are no half measures when it comes to the emergence of Merge. The emergence was purchased wholesale not piecemeal from nature. This, again, is in part motivated on the basis of the timeline assumption, and in part on the nature of discrete infinity. However, Martins and Boeckx (2019) use the theory-internal claims about Merge, i.e. that it is separated into ‘internal Merge’ applying to its own products and ‘external Merge’ applying to two distinct objects, to argue that Merge could have emerged in more than one step. They, therefore, separate the emergence of the process of Merge from the property of recursion.

Thus, the biological anomaly that is language emerged from a single mutation, in one instantaneous step, around 100,000 years ago in our recent ancestors and led to the property of recursion or the production of “an infinite array of hierarchical

¹⁸ The number in the above quote is 80,000 but it tends to vary between this number and 120,000 years.

¹⁹ I thank an anonymous reviewer for this clarification.

structured expressions” (Berwick and Chomsky 2016: 107). Essentially, minimalism reduces language evolution to the evolution of the computational system *via* its proxy Merge. Each assumption of this picture is prompted by the Minimalist aims of economy, simplicity and computational efficiency. But how does this view fare on the three desiderata mentioned at the end of Section 2.2?

As an explanation of what is innate about language, it returns the answer of Merge or the computational system. In terms of explaining the diversity of the world’s languages, it opts for an alternative route. In fact, the answers are related. The innate initial state of the language faculty is remarkably simple and similar across languages and persons. What differs are peripheral externalisation characteristics. These give the impression, or rather illusion, that language itself is diverse. But language in a minimalist sense recall, is just Merge. Perhaps a more charitable interpretation would have Berwick and Chomsky making use of the scientific versus manifest image distinction in the philosophy of science. *Prima facie*, it certainly seems that the emergence of a vast array of different languages is an evolutionary *explanandum*. But what the science tells us, in this case evolutionary biology, is that for language to have evolved so rapidly it needed to be an extremely simple macro mutation. “[T]he appearance of complexity and diversity in a scientific field quite often simply reflects a lack of deeper understanding, a very familiar phenomenon” (Berwick and Chomsky 2016: 93).

As a theory of language evolution, Berwick and Chomsky’s saltation account might be plausible given their assumptions. Of course, it really depends on what we mean by ‘plausible’. One popular problem in evolutionary theory is that “there is no end to plausible storytelling” (Lewontin 1998: 129). Nevertheless, even if this controversial claim is true, some accounts are more plausible than others. In other words, if you start with the empirical assumptions involving some formal notion of linguistic infinity or recursion, a timeline of emergence between 80 to 100 000 years, and a biologically unique or anomalous subject matter, the methodological route laid out by the Minimalist Program offers a viable option. However, Berwick and Chomsky eschew complexity in favour of a particular brand of simplicity, in the next section we will challenge this assumption and its resulting vision of language evolution showing that even with the assumptions they take onboard, an alternative is possible.²⁰

The complexity of language evolution

There are many points of contention in the above account. In this section, I want to highlight two recent objections both of which are related to the failure of minimalist accounts of language evolution to appreciate the complexity of language.²¹ The first

²⁰ For practically-minded scientists, even if one is compelled by their picture, until all the empirical facts are settled, one might develop alternatives for the sake of hedging one’s bets. I thank an anonymous reviewer for this possible angle.

²¹ There are more objections to more specific elements of linguistic theory from a biological perspective. Bickerton (2014a), for instance, mounts an argument against the use of theory-internal covert movement in MP. He suggests that it amounts to what Poeppel and Embick (2005) call a ‘granularity mismatch’ in that there is a mismatch between the levels and/or processes involved in syntactic analysis and those of

concerns the issue of the timeline, which plays a central role in saltation accounts such as Berwick and Chomsky's. The second challenges the evolutionary logic of their response to their own assumptions. In both cases, the issues point to a more complex target.

Let us begin with the first claim, namely that the paleontological record strongly suggests that human language evolved between 80,000 to 120,000 years ago within our hominid lineage. Dediu and Levinson (2013), Steedman (2017), and Everett (2017) all dispute this projection. For the former, evidence from various sources - including genetics, brain size, cultural artifacts and skeletal morphology - indicate that early *homo sapiens*, Denisovans, and Neanderthals had some form of language. This pushes the timeline back at least 400–500,000 years, “language as we know it must then have originated within the 1 million years between *H. erectus* and the common ancestor of Neandertals and us” (Dediu and Levinson 2013: 10). They do not, however, suggest that ‘full language’ was present prior to modern humans and allow for the possibility that syntax, speech and vocabulary size were significantly impoverished in this common ancestor. The cross-species prevalence of the FOXP2 gene as well as evidence that suggests its (or a variant's) possible presence in Neandertals (Krause et al. 2007) also serves to challenge the uniqueness and rapidity claims of saltation views (which assumed the gene was unique to humans). Everett (2017), on the other hand, goes further to suggest that the epoch which produced full language was that of *homo erectus*. He rejects the notion of protolanguage on the basis of claims about the culture, cranial capacity and vocal capabilities of this early hominid. Everett's account stretches the timeline for the emergence of language back to around 1.9 million years ago. Both the tools of natural selection and sexual selection are thus fully available to us.²² Steedman (2017) homes in on a piece of evidence which he considers more suggestive of the presence of language, namely the lengthening of the vocal tract with the *homo* genus which created a much wider array of possible sounds than any other primate vocalisations. He suggests that “[t] his evolutionary adaptation has been so rapid and extreme as to leave adult humans alone among animals in not being able to swallow and breathe at the same time, a change that would otherwise seem to be maladaptive as it can cause them to die prematurely by choking on food” (Steedman 2017: 581). Taking fossil evidence into consideration, this vocal tract lengthening (and larynx lowering) started at least 2 million years ago.²³

The important point, for our purposes, is that rejecting the strictures of the timeline assumed by Chomsky and Berwick opens us up to the standard resources of

Footnote 21 (continued)

neurobiology. Unlike overt movement in which the same syntactic object lives at two locations but is pronounced at only one of them, Bickerton claims that covert movement is unlikely to find equal or any empirical footing.

²² However, the ‘stretch’ is often based on controversial assumptions about the culture and vocal capabilities of *erectus*. His view also delves into the internecine debates on recursion, Pirahã, and universal grammar.

²³ Various empirical studies have challenged the claim that linguistic sounds were due to lengthening of the vocal tract. See Fitch et al. (2016), Fitch and Reby (2001), Boë et al. (2017).

evolutionary biological. Specifically, it is the strict timeline, or what Martins and Boeckx (2019) call the ‘Great Leap Forward’, which was supposed to force us toward the biological uniqueness of the emergence (and subsequent nature) of natural languages. Without this assumption, language can be treated like any other biological phenomenon in need of evolutionary explanation. Progovac (2015), for instance, proposes a gradualist, incremental approach to the evolution of syntax and language more generally.²⁴ Her account starts with MP but then quickly moves beyond it. She, unlike Everett, embraces the possibility of proto-grammar by identifying (fossilised) flat paratactic binary compounds found across languages (like *rattle-snake*, *cry-baby*) as the foundation for later hierarchical more complex syntax. Her account fits the adjusted timeline and more importantly, unlike Berwick and Chomsky, relies on language variation as a source of evolutionary insight. Furthermore, she envisages testable neuroimaging hypotheses:

When linguistic reconstructions can identify ancestral proto-structures, and distinguish them from more recent structures, neuroscience can test if these distinctions are correlated with a different degree and distribution of brain activation, and genetics can shed light on the role of some specific genes in making necessary connections in the brain possible. (Progovac 2016: 10)

On a gradualist, incrementalist approach, we are not compelled towards the reduction of language to syntax and syntax to Merge on simplicity grounds. Language, and syntax, is thus more complex than the minimalist assumptions would have it. Moreover, Progovac suggests that postulates of syntactic theory, such as Subjacency,²⁵ are explained by this approach. Not only does this serve a second grade of biological involvement level but given the generation of possible testable theses, it could move us closer to a first grade level (or perhaps grade 1.5).

The second objection, I want to briefly consider, derives complexity from Berwick and Chomsky’s own assumptions. De Boer et al. (2020) take onboard, for the sake of argument, all of the assumptions of their particular minimalist saltation theory and show that it is incorrect from a probabilistic perspective. “Specifically, we formalize the hypothesis that fixation of multiple interacting mutations is less probable than fixation of a macro mutation in this time window, and show that this hypothesis is wrong” (de Boer et al. 2020: 452). They use extreme value theory to determine the a priori probability of a mutation occurring and diffusion analysis to plot the probability of it leading to fixation in a population (modelled on the evidence of likely human population sizes around 140,000 years ago). The result of their study is that it is more likely, even within the limited time period, that smaller biological changes contributed to the emergence of language gradually rather than

²⁴ See Pinker and Bloom (1990) and Pinker and Jackendoff (2005) for earlier arguments for this kind of picture as opposed to the saltation account.

²⁵ The literature on Subjacency is vast and the idea has been refined and updated a number of times. But basically the ‘subjacency condition’ was proposed in Chomsky (1973) to account for multiple island effects, where an ‘island’ is a constituent that blocks items from moving outside of it (like a relative clause), within one structural principle. More specifically, it is a condition on movement to ensure that all movement is local such that an element cannot move over more than one cyclic node like a noun phrase.

the rapid saltation scenario involving a single macro mutant. Furthermore, their view does not rule out communication as a selective advantage, nor the possible ‘smaller biological changes’ involving phonology, gestures, or pragmatic elements (or a combination of all of these).

The precise details are beyond the present scope but in the next sections we will embrace both the methodological pluralism and the possibility of multiple interacting elements at various levels resulting in the emergence of language. It seems that even with the Chomskyan gambits in hand, complexity is a more likely outcome (and initial state) than simplicity. And what’s more “the alternative scenario for gradual evolution of linguistic ability proposes that the evolution of language happened in a way that is far less exceptional by biological standards” (de Boer et al. 2020: 453). As Martins and Boeckx (2019: 5) note:

The evolution of something as complex as human language deserves integration of results and insights from different corners of the research landscape, namely the fields of neurobiology, genetics, cognitive science, comparative biology, archaeology, psychology, and linguistics.

This is all to show that the minimalist moves are not benign, biologically speaking. Accepting the timeline relegates many of the resources of comparative biology and natural selection to the periphery. Endorsing the single mutation logic of Merge and recursion as metonymous for language, attenuates biodiversity of languages and accepts a parochial concept partly divorced from other biological phenomena.

Berwick and Chomsky believe that unlike the laws of physics, biology is more like case study. In what follows, I will argue that this only seems like the case on the assumption of individual biology but patterns and generalisations emerge when the purview is shifted from the individual to biological systems.

The maximalist program

Much like the Minimalist Program, the Maximalist Program or MP+ I will advocate is not a theory but an approach or a strategy. In Section 4.1, I will outline what a complex system is with reference to the ten standard features outlined in Ladyman and Wiesner (2020). In section 4.2, I will briefly mention three prominent examples of this general approach each of which exemplifies a few of these features before proffering my own account which aims to incorporate additional features. Finally, I explain why my view is biolinguistic in Section 4.3.

Before we get into some of the details, I want to discuss exactly where this approach is pitched with relation to MP or Minimalism. As I stated above, both frameworks are meant to be above the level of individual theory. This allows for the possibility of some compatibilism between the approaches. For instance, as Christiansen and Kirby (2003) emphasise “there is general consensus that to understand language evolution, we need a good understanding of what language is” (301). However, they go on to state that this issue is precisely the point at which the field is divided. Thus, if different theorists have different views on what language is, a

public system of conventions, an internal computational system, a formal mathematical object, then they will naturally evoke different methods for investigating it. On one reading of the framework to come, MP+ could be a union of all of these components. Thus, it could subsume Minimalism. The latter might target an internal psychological system while some of other theories target speech behaviour and so on in the division of labour. I think some compatibilism is most certainly possible. However, genuine divergence between the programmes remain. Minimalism asks us to abstract away from interaction effects and isolate simple basic structures via processes like Merge. Maximalism, homes in on the interactions between subsystems and finds the methodological (sometimes statistical) tools to tame these hybrid targets. Minimalisms asks us to start idealising at the individual biological or organism level while Maximalism tells us that genuine discovery with relation to language evolution is only possible at the systems level. There's an interesting analogy with Feminist social science and the concept of intersectionality here. Crenshaw (1989) famously argued that identifying the experience and dimension of discrimination related to Black women cannot be adduced by combining the values of Black and woman additively. Attempting a minimalist isolation of these components would lose sight of a genuine scientific target and reduce the possibility of redress in this particular case, it is argued. Of course, nothing hinges on the veracity intersectionality and it remains methodologically controversial (see Gasdaglis and Madva 2020). Systems biology and 4E approaches to cognition make similar cases for a loss of target at the individual level, as we shall see. Thus, MP+ marks a modelling strategy, or type of idealisation, which rivals MP in so far as MP makes claims that minimal structures can be studied independently of their environment or interactions with other systems.

What is a complex system?

This section's titular question is related to the methodological question of what complexity science is. The answer to the latter is multifarious and pluralist. Complex systems are studied from various angles, with a bent towards computational and probabilistic methods. Language too can be studied from various angles. Yet the dominant position for generations has been that the science of language cannot be a 'science of everything' and the scientific demarcatable aspects of such a study is an *I-language* (Chomsky 1965, 2000). An *I-language* is an *internal, individual, intensional*²⁶ mental/brain state of an individual language cogniser (at the appropriate 'level of abstraction'). With this focus, social *externalia* hold little currency among generatively-minded linguists. Such pursuits are best left to philosophers of language, sociolinguistics, ethnographers or abandoned entirely, it is often argued. As we saw with MP, grammars are considered to require the postulation of simple

²⁶ *Internal* means internal to the language user or intracranial, *individual* pertains to the fact that the grammar does not consider relational facts outside of the individual speaker, and lastly *intensional* marks the functional or generative procedure for getting at the expressions of a language as opposed to the external set containing all such expressions.

mechanisms responsible for grammatical complexity. There is, however, another way to do things.

The science of language could embrace complexity without becoming a theory of everything and without eschewing scientific idealisation. Chomsky himself hints at this option before dismissing it lines later.

[L]anguage is a biological system, and biological systems typically are “messy”, intricate, the result of evolutionary “tinkering”, and shaped by accidental circumstances and by physical conditions that hold of complex systems with varied functions and elements. (Chomsky 1995: 29)

A few lines later he reaffirms the need to be minimalist (with a nod to the competence-performance distinction) as a ‘working hypothesis’ of the basic structure of language based on simplicity and elegance. Language use might indeed be more complex in the above sense, he concedes,²⁷ pragmatics, and other linguistically relevant cognitive systems while the latter only includes narrow syntax. Many linguists accept that FLB might meet the criteria of complex systems but Chomskyans insist that FLN does not. I thank an anonymous reviewer for the useful clarification. Before I argue for complexity science based linguistics more holistically, a few properties of complex systems need to be considered. I will closely follow Ladyman and Wiesner (2020) recent characterisation, although there are other excellent introductions, for the sake of clarity and due to their ecumenical approach (which draws from a range of other work).²⁸

One way of thinking about complex systems involves emergent phenomena. Indeed emergence is an important perhaps inextricable aspect of complexity science. Individual honey bees exhibit simple random behaviour in isolation but when they act in unison they display highly complex collective behaviour including advanced symbolic communication (the famous ‘waggle dances’), the ability to deliberate on hive creation, temperature regulation, and swarming. Eusocial insects like bees and ants often display the adaptive behaviour characteristic of multicellular organisms for survival of the colony above that of any individual (see Hölldobler and Wilson 2008). This is one expression of how a complex system can emerge from simple components or as Ladyman and Weisner put it ‘complexity can come from simplicity’, which arguably MP also espouses to a certain extent. In fact they highlight a number of other ‘truisms of complexity science’ including that coordinated behaviour doesn’t require centralised control, complex systems are often modelled as networks and information processing systems, the field is interdisciplinary, inorganic systems can produce order and so on. More directly to the point of this section is the list they provide of the standard features of complex systems.

²⁷ It should be noted that Chomsky is assuming something like the distinction between the faculty of language broadly (FLB) and narrowly construed (FLN) (later proffered by Hauser et al. (2002)). The former might include some aspects of semantics.

²⁸ See Morin (2008), Mitchell (2011), Thurner et al. (2018) for other prominent accounts.

1. Numerosity.
2. Disorder and Diversity.
3. Feedback.
4. Non-equilibrium.
5. Order and Self-Organisation.
6. Nonlinearity.
7. Robustness.
8. Nested Structure and Modularity.
9. History and Memory.
10. Adaptive Behaviour. (Ladyman and Wiesner 2020: 10)

The list above does not constitute a set of necessary and sufficient conditions. Rather some systems exemplify some of these features and not others. For instance, they argue that adaptive behaviour is a hallmark of living systems (although artificial neural networks do exhibit a facsimile of this capability). Different complex systems also display these features to different degrees. The universe involves more numerosity in terms of its elements and its interactions than other systems do, while many systems such as the climate and economies strive toward equilibrium (but don't necessarily achieve it), non-equilibrium physical systems such as chemical reactions can be captured by stochastic characterisation. Complex systems tend to be open, dynamic and not at equilibrium (Kauffman 1995). Systems like the human brain also exhibit feedback (and reinforcement) *via* millions of neuronal and synaptic connections, and hierarchical nested structure in its functions and organisation. Evolution tends towards the establishment of robust structures for the sake of stability without which it would be nearly impossible to attempt to describe a complex organic system. However, the diversity of targets and methods is part of the reason complexity science has taken so long to establish itself as distinct discipline. We cannot, of course, discuss every feature and its instantiation in particular complex systems but with these features in mind, and the examples provided, we can describe what a complex system can generally be.

Complexity science studies how real systems behave. The models of the traditional sciences often treat systems as closed. Real complex systems interact with an environment and have histories. Complexity is not a single phenomenon but the features of complex systems identified [above] are common to many systems. If it is right that the hallmark of complex systems is emergence and that there are different kinds of emergent features of complex systems, then instead of defining kinds of emergent features of complex systems, it is possible to identify different varieties of complex systems according to what emergent features they exemplify. (Ladyman and Wiesner 2020: 126).

This is the key to understanding the Maximalist approach to language sciences. We'll be guided by this latter insight going forward. By focusing on which features of complex systems are present in or gives rise to linguistic phenomena and how one might measure these to produce or support theoretical claims, we can work within

the overarching framework. In other words, instead of viewing language as an isolated biologically unique outlier characterised entirely by simple mechanisms, we should embrace the complexity and attempt to show how language emerges from the interaction of many parts. Thus, our working definition of language will be something of the following sort:

Language is a complex system in which robust structures emerge from the dynamic interaction of multiple interconnected parts.

The evolution of language can similarly be approached from multiple angles and not reduced to the emergence of any single factor as de Boer et al. (2020) suggested (Section 3.2). And more importantly, there are new ways in which to connect the study of language to biology. There are many possible arguments and evidence to explore starting from different features of complex systems exhibited by language and each specific hypothesis is a member of the Maximalist Program or MP+. In the next section, I will briefly show how different features have led to different complex systems analyses of natural language before offering my own sketch of a theory in the final section.²⁹ As Ladyman and Wiesner emphasise, not all complex systems exhibit all ten features listed above.

Examples of linguistics as complexity science

There are three extant linguistic accounts that treat natural language as a complex system I think worthy of discussion here.³⁰ Each focuses on one or two different features of complex systems as the core linguistic *explananda*.

The first is perhaps the most controversial entrant into the space, namely GB of generative grammar (Chomsky 1981). In this framework, the linguistic system is divided into two classes of subsystems, those pertaining to the rule system and those pertaining to the principles. In the former, we find the lexicon, syntax and both interfaces (Phonetic Form and Logical Form). In the latter, bounding, government, Θ , binding, Case, and control theories.

²⁹ What's interesting is that many of the historical developments of modern linguistics, such as influence of cybernetics, dynamical systems theory etc. were pursued under the umbrella of the classical cognitive revolution of the mid-twentieth century. Linguistics was a core member of that approach to the study of mind (Miller 2003). Its early methods, such as formal language theory and automata theory, are directly related to fields such as cellular automata theory which laid the foundations for complexity science. Yet although Ladyman and Wiesner add the brain to their list of complex systems, they leave out the mind and language. They include them only as an emergent phenomena of the brain but this is too parochial and internalist. No other general introductions or advanced discussions include language explicitly. And yet language exhibits many of the fundamental features of complex systems. Besides multiple interacting parts, in sound, form and meaning, feedback is characteristic of the practice of communication and information transfer as well as decentralised organisation and modularity. Nested structure is the hallmark of linguistic analysis with hierarchical trees forming the primary methodology. More below.

³⁰ There are probably many more such accounts. Hopper (1987) emergent grammar is one such paradigm. Clark (1996), for instance, treats language as an emergent coordination system like a dance. Lexical Functional Grammar is another theoretical approach in which modularity and the interplay between different systems plays a major role (Bresnan et al. 2016).

The system that is emerging is highly modular, in the sense that the full complexity of observed phenomena is traced to the intersection of partially independent subtheories, each with its own abstract structure. (Chomsky 1981: 135)

One aspect of this analysis is that language is broken up into many different component parts each with their own constraints and mechanisms. Although the theory still maintains that the faculty of language is autonomous from other cognitive systems, it does stratify the concept of UG to include the idea that a large portion of grammar is common to all languages. This becomes a move towards more complexity and modularity. The previously mentioned principles is the part of UG which acts like well-formedness conditions or constraints on the representations of each level of the grammar (D-structure, S-structure, PF and LF).

Of course, GB remains an internalist and isolationist approach to grammar along the lines of the competence model discussed above. Some streamlining of the many rules of transformational grammar takes place (*via* X-bar theory), focus on learnability is increased, and a number of subtheories are introduced. In this way, the features of complex systems analysis it incorporates are (1) numerosity given the number of interacting parts, (5) order and self-organisation since language is said to emerge from independent modules, and (8) nested structure and modularity for obvious reasons. What is lacking is the interaction with non-linguistic aspects of the environment and fellow language users (through dialogue data or corpus research), the biological analogy as well as interdisciplinary methods. In other words, language is not an ‘open’ system in the terminology of complexity science. Nevertheless, it is the closest that generative grammar comes to being a complexity science.

The next candidate more explicitly embraces the idea of language as a complex system. Kretzschmar (2015) homes in on the features of (1) numerosity, (3) feedback, and (6) non-linearity or what he calls the “A-curve” in the corpus data he evaluates. He states that “no linguist can afford to ignore the fact that human language is a complex system” and that furthermore “[a]ll approaches to human language must begin with speech, and all speech is embedded in the complex system” (Kretzschmar 2015: 2). Generative linguists or Minimalists wouldn’t agree with either statement but certainly not the first part of the latter, namely that approaches to human language must begin with speech, even if they might grant that speech is a complex system. This is especially the case given that generative grammar has often relied on what they call ‘negative data’ or mistakes that language learners do *not* make, elements unlikely to be present in corpora. Furthermore, as Pullum (2007) notes, despite their merits, corpora often do not contain rare but possible constructions which can inform linguistic theory. One specific complex feature which Kretzschmar shows to be omnipresent in his corpus data is an emergent nonlinearity characteristic of market economies, namely the Pareto or 80/20 principle in which 80% of wealth is concentrated within 20% of the populace based on Zipf’s Law.

Perhaps the most striking evidence for speech as a complex system is the nonlinear distribution of the variants for any given linguistic feature. Linguists will recognize Zipf’s Law, a frequency ranking of words in texts that always

finds that rank is roughly inversely proportional to frequency. (Kretzschmar 2015: 24)

He claims that the Pareto principle shows up all over the data at various levels and that “we in language studies can and should make good practical use of the 80/20 Rule on a conceptual basis” (2015: 85). He reflects on how to ‘make good practical use’ of the principle by comparing the size of certain compendia of English Grammar (as evidence of the wrong kind of complexity), he suggests that the appreciation of the 80/20 Rule should improve grammars by insisting that generative linguists focus on infrequent constructions given that they “only study just the top-ranked variants” (Kretzschmar 2015: 99).³¹ What’s useful about Kretzschmar’s work is that he applies complexity science to one physical output of human language, namely speech. He discovers emergent patterns and principles, like his non-linear A-curve, present in a particular subsystem.

Neither GB nor Kretzschmar’s work really engages with the biological aspects of linguistics. This is unsurprising since neither prescribe to biolinguistics explicitly. A complex systems analysis of language which does aspire to biolinguistics is the work of Simon Kirby and his colleagues on computational evolutionary theory. He considers the approach “a new way of thinking about the role of cultural transmission in an explanatory biolinguistics” (Kirby 2013: 460). Kirby focuses on the idea that language is an adaptive system or feature (10).

The evolutionary approach to this challenge [explaining why language has the structural features it does and not others] is one that attempts to explain universal properties of language as arising from the complex adaptive systems that underpin it. (Kirby 2013: 460)

Kirby too embraces complex systems analysis and designs his models so to capture the essence of numerosity (1), i.e. not only the role of individual elements in emergent structure but also their interactions. ‘Iterated learning models’ in computational language evolution research aim to explain how complex syntactic structure, such as discrete infinity, is generated by creating highly simplistic models involving generational simulations of populations with no language to begin with (see Brighton and Kirby 2001). This exemplifies both the truism ‘complexity can come from simplicity’ and the feature of role of history and memory (9). ‘History’ and ‘memory’ are distinguished in complexity science by the latter’s effect on behaviour in adaptive systems. Hence, ants pheromone trails can be thought of as external memories. Iterated learning models exhibit a kind of system memory. Importantly, Kirby and his colleagues see themselves as in some ways starting from

³¹ This claim strikes me as false. Generative grammar has always considered less frequent, sometimes even previously untokened, phrases and constructions. The debates concerning center-embedding and recursive phenomena were crucially about *possibilia* and not frequency (van der Hulst 2010). In fact, if there is criticism in this vein, it probably goes in the opposite direction. The mischaracterisation of generative linguistics as isolating frequent constructions and then making inductive generalisations based on these can be found throughout the book. But it is not my purpose to mount a critique of Kretzschmar’s account, which is on the whole a huge step in the right direction.

a very different perspective to the Galilean method of Chomskyans (see footnote 14). Specifically, they claim that “rather than abstract away details about population structure or patterns of interaction, computational modellers will typically retain these complexities” (Kirby 2013: 461). They choose to abstract away from other features in their models. Thus, the biolinguistic perspective is stretched to include population level dynamics (similar to Kretzschmar) but with the focus on language evolution *via* the emergence of phenomena like innate signalling and the role of iterated learning.

The specifics of these accounts are beyond the present scope but they do serve as a proof of concept and examples of the MP+ at work. They specifically highlight Ladyman and Weisner’s features (1), (3), (5), (6), (8), (9), and (10). If this were my only aim in the paper, I could stop here. But I also endeavour to provide a sound interpretation of biolinguistics which can attract those present practitioners already inclined towards the second grade of biological involvement. In the last section of the paper, I provide a sketch of a view I call *systems biolinguistics*.

Systems biolinguistics

With the exception of Kirby’s work, most of the extant offerings of linguistics as a complexity science, or what I’ve been calling MP+, do not take the biological analogy seriously. Hence, they do not offer the biolinguist a way to ground or constrain their field in terms of the biological sciences. I will outline the beginnings of such an approach here while focusing on systems biology, what I will call *systems biolinguistics*. My strategy is to show that many of the *sui generis* concepts of biolinguistics (and MP) can be reinterpreted within this framework (and complexity science more generally) to yield more scientific, less isolationist, and more measurable results. In what follows, my chief goal will be to show that this novel perspective can offer three main advantages over other theories within the second grade of biological involvement: (1) significant theoretical incorporation/integration, (2) better naturalisation of concepts in linguistic theory as *per* the goal of biolinguistics, (3) a specific route to methodological pluralism.

One criticism of MP was that it resorted to a strong uniqueness claim about language, severing it from case studies in other biological sciences. Uniqueness leads to isolation and cognitive modularity. If language is an outlier in the biological world, then it cannot be easily integrated with other systems of which we might know more. Thus, knowledge transfer is hindered. MP+ rejects this assumption and views aspects of language such as phonetic distribution (Kretzschmar), symbolic signalling (Kirby), and semantic significance as emergent phenomena within a complex network of interacting internal and environmental factors. The first question to confront is how to apply a complex systems analysis to language via biology? The novel answer I provide is that this possibilities should be relocated within an understanding of systems biology.

Systems biology is a holistic approach to the life sciences. It is an extremely collaborative interdisciplinary field which includes biology, computer science,

physics, engineering, and mathematics. Whereas the nexus of traditional biology might have been individual organisms, cells, plant life etc. systems biology abstracts away from these to home in on their complex interactions with the environment. There are a number of specific sub-disciplines of this larger field, such as metagenomics or the study of diverse microbial communities.

Like many theoretical offshoots, systems biology started with critical reflection on the limitations of both standard microbiology with its focus on microbes such as viruses and bacteria and mainstream biology with its focus on individual macroorganisms such as plants and animals. For instance, classical concepts such as multicellularity are ill-defined on the entity-based accounts since they fail to capture the multi-cellular nature of symbiotic organisms like lichens which exhibit interdependent existence. Cellular cooperation, competition, communication, and certain developmental processes require a broader perspective than the object-oriented accounts can provide. Some have put forward the claim the microbial communities can be considered multicellular organisms themselves (O'Malley and Dupré 2007).

Dupré and O'Malley (2007) survey the literature on metagenomics or environmental genomics which “consists of the genome-based analysis of entire communities of complexly interacting organisms in diverse ecological contexts” (835). In this field, microorganisms are not placed in isolated artificial settings but rather assumed to be essentially coupled with their environments and interactions with other organisms. A proper investigation of biodiversity seems to require the analysis of metagenomes or large amounts of DNA collection within the environment. One additional reason for this shift is that evolution seems to require a larger perspective of this kind. As they state:

Conceptually, metagenomics implies that the communal gene pool is evolutionarily important and that genetic material can fruitfully be thought of as the community resources for a superorganism or metaorganism, rather than the exclusive property of individual organisms. (Dupré and O'Malley 2007: 838)

On this view, one might consider human bodies to be complex symbiotic systems composed partly of human cells, viruses, the bacteria hosted by prokaryotes and so on. But this perspective is also too limited. Systems biology assumes that there is no non-arbitrary distinction to be had between an individual organism and its environmental conditions. No clear ‘self’ versus ‘other’ is discernible. The immune system is a clear case where the human host and the prokaryote communities form one complex system which benefits the organisms (Kitano and Oda 2006). Dupré and O'Malley use these considerations and more to suggest an ontological shift is necessary and/or present in biology, one that moves from entities or organisms to processes and systems as the basic ontological categories. There is no useful concept of a static genome-organism correspondence as “[g]enomes, cells, and ecosystems are in constant interactive flux: subtly different in every iteration, but similar enough to

constitute a distinctive process” (Dupré and O’Malley, Dupré and O’Malley (2007): 841).³²

Systems biology conceives of biological entities at the systemic level, not only as individual components, but interacting systems, processes and their emergent properties. In this way, linguists such as Clark (1996) are correct that language is like a dance in which coordination between partners plays a major role. What they leave out is the interaction between the dancers and the dance hall, the other dancers at the party and human microbiota who call us home while we sweat and salsa.

In order to accommodate the analysis of big data, the complex inter-organism interactions and their environments, statistical and network approaches have become prominent. Thus, biological systems are usually represented as dynamic networks which form complex sets of binary interactions or relations between different entities and their contexts. Graph theory has been a very useful tool in the representation of biological networks. The vertices represent different biological entities such as proteins and genes in biological networks, and edges convey information about the links or interactions between the nodes. The links can be weighted or assigned quantitative values to encode various properties of interest, either topological or otherwise. More complex networks of networks can model the interaction between systems themselves (Gao et al. 2014).³³ It is this aspect of systems biology that I think makes it especially applicable to biolinguistics as per the ‘truism’ that “complex systems are often modelled as networks or information processing systems” (Ladyman and Weisner 2020: 9).

These networks can take the form of *trees* or *forests*. See Fig. 3 below for different kinds of networks used on plant systems biology. Some networks model correlations across multiple conditions (a), while others (b) model sets of molecular interactions (or ‘interactomes’), (c) shows hierarchical regulatory networks of genes with another way of modelling this shown in (d) by means of graphs that resemble finite-state automata.

The idea in complexity science and systems biology is that these (graph-theoretic) tools are not merely instruments but tell us something ontologically important about organic life and reality respectively. Silberstein (2022: 600), for instance, claims that “reality is more like multiscale complex networks or structured graphs of extrinsic dispositions”. He insists that this view is commonplace in network neuroscience. Deacon (2008) argues that life itself is a third-order emergent property characterised by self-organisation and processes which involve some form of history or memory. This fits with Kirby’s biolinguistic view of the evolution of language.

Footnote 32 (continued)

³² The advent of Big Data has had a profound influence on the life sciences as it has on computational linguistics. The data explosion that originated in biological research compelled the development of systemic approaches to data analysis and departures from the single gene/protein frameworks of the past.

³³ For instance, Goldberg (2013) also emphasises the “network” aspect of the approach in which “[p]hrasal constructions, words, and partially filled words (aka morphemes) are related in a network in which nodes are related by inheritance links” (1). Network and inheritance hierarchies are also utilised in formal frameworks such as Word Grammar, Head-Driven Phrase Structure Grammar, and Sign-based Construction Grammar.

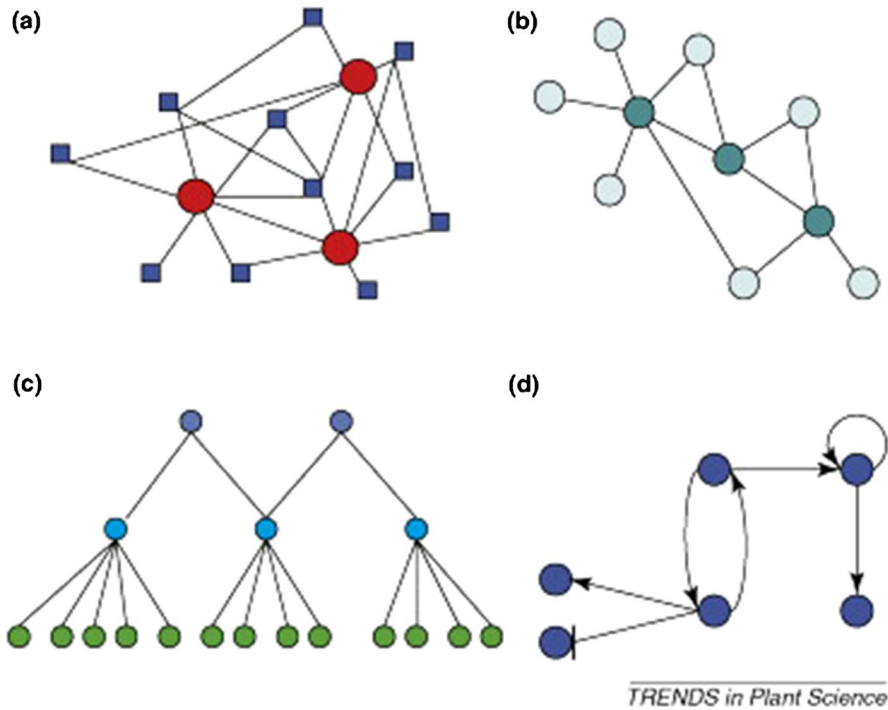


Fig. 3 Plant system networks (from Yuan et al. 2008: 166)

Thus, language would also count as case of third-order emergence characteristic of living organisms which “inevitably exhibits a developmental and/or evolutionary character” (Deacon 2008: 137). For him, the robustness of these emergent structures and patterns sustained over time involves a kind of ‘self-similarity maintenance’ which “[i]n the jargon of complexity theory, such patterns are called ‘attractors’, as though they exerted a ‘pull’ toward this form” (Deacon, 2008: 120). In linguistics, the attractors could be universal forms or so called ‘statistical universals’ that connect the world’s languages. For example, languages in which verbs precede objects (SVO, VSO) tend to have prepositions while languages in which verbs follow objects (SOV, OVS) usually have postpositions.³⁴ Deacon’s example of choice is snow crystal formation in which external environmental factors can shape individual snowflakes whose general form is compelled by the crystal lattice structure.

Deacon’s picture of complexity and emergence involves three nested kinds of emergent phenomena arranged into a hierarchy of increasing topological complexity. Third-order emergent processes (‘teleodynamics’), where he locates life and mind,

³⁴ Greenberg (1963) catalogues a number of these conditional patternings.

require second-order emergent processes ('morphodynamics') or chemical processes as necessary conditions, while at the base are self-amplifying (non-equilibrium) first-order emergent processes ('thermodynamics') to create their necessary conditions, basically the laws of physics. What's interesting for us is that both Deacon and Silberstein's accounts of complex systems allow for law-like patterns emerging not just at the level of physics and chemistry but also biology and cognition. However, in both cases a systems purview is required to appreciate these patternings.

It is well-known that formal linguistics since the mid-twentieth century embraced very similar network and graph-theoretic analyses of language. Early formal language theory emphasised the importance of the nested hierarchy of formal languages which characterise the rules and complexity of human language. The type of rules a generative grammar possesses maps its output to a given class of formal languages. Regular grammars express the regular languages. Context free grammars produce the regular languages and the context free languages. In context free languages, we find patterns like $a^n b^n$ (ab, aabb, aaabbb...) but not more complex (and harder to parse) patterns like $a^m b^n c^m d^n$ (aaabbbccdd).³⁵ The formalisms of formal language theory (which can be represented as both graphs and automata as in the plant systems) are not just supposed to be tools but reflect the actual structure and complexity of language.³⁶

If the formal language hierarchy represents the relationship between different complex configurations (i.e. languages) at the systems level, the individual tree diagrams represent the individual construction level where most linguists ply their trade. Figure 4 shows the nested Chomsky Hierarchy with each corresponding accepting class of automata. Homing in on any ring of the Hierarchy, such as the context free ring, produces similar kinds of graph-theoretic structures used across systems biology, shown in Fig. 5. Technically, grammars produce strings and languages are composed of strings. This is called 'weak generative capacity' in the literature. However, each string generated by the grammar is also associated with a 'structural description', a tree or graph. This is called 'strong generative capacity' (Chomsky 1963). In fact, Fig. 5 is a hierarchical tree diagram which represents the context free rules (like $S \rightarrow NP, Aux, VP$) similar to the hierarchical gene regulatory networks in plant systems biology.³⁷

Of course, graphs and networks are common mathematical tools across disciplines. As we have seen, most linguists believe that hierarchical constituent structure is the essence of language. For them, language is in a sense graph-theoretic. Whether or not we hold this strong syntactic view, network structure clearly plays an important role in every linguistic discipline from phonology to pragmatics.³⁸ Moreover, the connections between different linguistic systems are often modelled as mappings

³⁵ It is widely assumed that natural language syntax lies somewhere between context free and context sensitive languages, in the so-called mildly-context sensitive class.

³⁶ See Hopcroft and Ullman (1979) for a still classic introduction to the general field.

³⁷ The corresponding automaton would be a pushdown automaton which incorporates an element of memory or the 'stack'.

³⁸ Nefdt (2021), for instance, argues for an ontic structural realist position in linguistics in which the chief vehicle of the scientific continuity is structure.

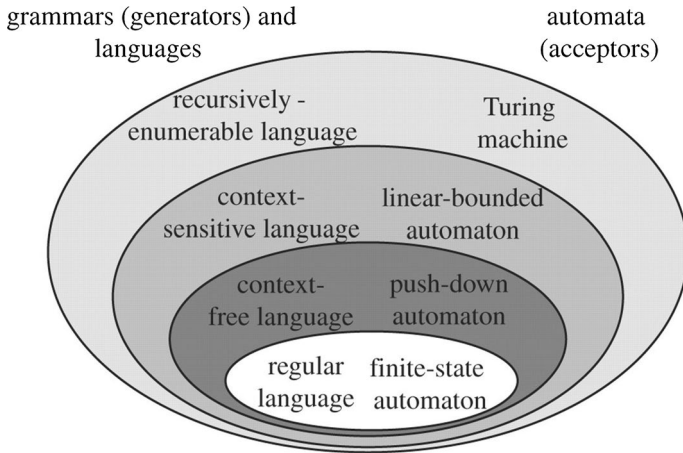
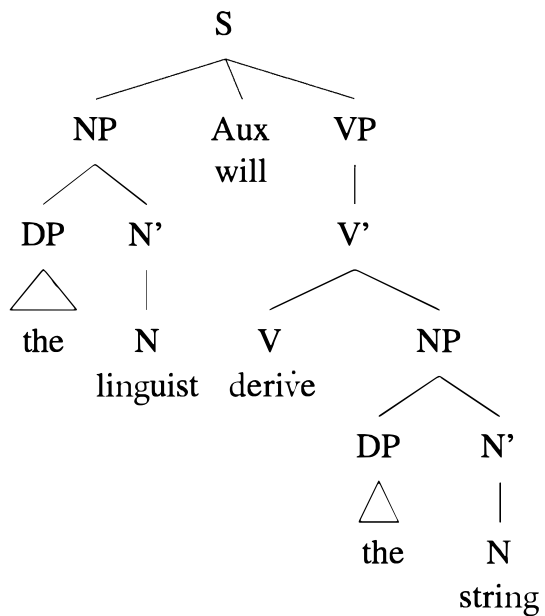


Fig. 4 The Chomsky Hierarchy (From Fitch and Freiderici (2012: 1936))

Fig. 5 Tree for *The linguist will derive the string*



or *structural morphisms*. Jackendoff (2002) parallel architecture (PA) is one prominent example of multiple generative systems with interface principles linking them to and across one another (in opposition to the syntax-centred approaches of classical and minimalist generative grammar). The mappings are rarely complete (or rather injective), allowing for semantic structure without a syntactic component, and phonological structure without semantic interpretation etc.

The key insight to systems biolinguistics is to ascend to the level of grammars which characterise more than just individuals by adding more systematic information from other aspects of language and the social environment. What the biological systems (and complexity science) perspective brings with it is a clear way to integrate information from different systems as networks of networks. The tendency among biolinguists under MP has been to simplify trees and isolate the syntactic information from the phonetic, semantic and pragmatic. This doesn't mean that they are not connected but merely that they are explanatorily autonomous. But there are deep, ontologically important, interactions between these elements that can be modelled as networks of networks. One clear example which aims to capture the compositional connections between syntax and semantics is Shieber and Schabes (1991) framework of synchronous grammars. If we allow ourselves for a moment to take a grammar to be a network of some sort (since it can be represented as a tree or graph structure), then formalisms which map one or more grammars onto each other are networks of networks.³⁹ The resulting complex analysis is rather structural but this is in keeping with both systems biology (French 2011) and complexity science.

The general idea behind synchronous grammars is to create nested information structures with syntax and semantic information encoded as couples. Specifically, take a pair of trees, one representing the syntax and other the semantics of a particular sentence. Some nodes in the trees form links. These links then conjoin the nodes such that operations on the tree pairs occur on both sides of the link.⁴⁰ So if you move one part or constituent of the syntactic tree you move its semantic couple. The insight is that single operations (such as adjunction or substitution) can happen on pairs of trees and not just segments of individual trees. In principle, there is nothing stopping us from incorporating contextual (or environmental) parameters, phonological markers and even neurological regions creating quadruples or further tuples of trees and tree segments. The important aspect is finding the links between systems which become the units of our analysis over and above isolated fragments such as syntactic constituents. These are the nodes of our networks of networks. For instance, what might count as grammatical is in part based on community standards and conventions and these can vary between dialects of the same language in distinct regions. Grammars are not (only) inside the head!

What I am advocating is similar to a practice in cognitive neuroscience in which researchers construct multiple distinct graphs and look for invariant structure across them. As Sporns (2014: 653) notes “studies of brain networks using a variety of parcellations [...] have converged on a set of fundamental attributes of human brain organization that are largely consistent with those found in nonhuman primates”. These studies have uncovered empirically significant features such as robust ‘hubs’ in particular brain regions where a hub is a node which has the most number of

³⁹ In computational linguistics and machine translation, *transducers* compare two or more languages.

⁴⁰ This marks a departure from frameworks such as LFG which split the syntax into constituent structure (C-structure) and functional structure (F-structure) reserving semantic correspondence for the latter only. Other aspects of LFG, however, are highly conducive to a systems biolinguistic or MP+ approach more generally. See Börjars (2020).

edges attached to it. More specifically, there are two general types of brain networks. The first type, anatomical or structural networks, are identified by means of Magnetic Resonance Imaging (MRI) techniques such as diffusion tensor imaging (DTI) in which the diffusion of water molecules are used to study neural tracts and the white matter organisation of the brain. The second kind are functional networks which are not completely reducible to anatomical connections and thus not completely amenable to the latter MRI (and fMRI) techniques. They are composed of “patterns of statistical dependence among neural elements” or functional connections (Sporns 2013: 248). In order to study these networks, ‘parcellation packages’ are created which are basically graph-theoretic segmentations of brain regions into regimented borders and clusters according to activation patterns and the like. In order to determine the ‘real’ network, convergence of packages is needed such that invariant structure is revealed. Hubs are useful markers in this process as they are thus nodes with maximum convergence. As Yan and Hricko (2017) put it: “the brain networks that cognitive neuroscientists seek to investigate are presumed to exist independently of the choice of parcellation scheme - a real network must be parcellation-independent (4)”. This process can be modelled by something like a node in a synchronous tree that allows for the most links or connections across grammars. Of course, a connection as strong as that is not needed. The underlying idea is that there is a ‘common argument pattern’ (in the sense of Kitcher (1989)) between certain kinds of modelling practices in linguistics and cognitive neuroscience.

In terms of theoretical integration or the first stated advantage of this approach, notice that this picture can retain the computationalism of generative grammar without endorsing its individualism. The most prominent example of a network of networks is one of Ladyman and Weisner’s cases of a complex system, namely the Internet. Consider a local area network or LAN. These can be configured in a number of ways, but ring and mesh networks seem most appropriate as models of linguistic communities since either each computer is connected to neighbouring computers to form closed circuits (ring) or each computer is connected to every other computer in a distributed fashion (mesh). In order to communicate or exchange information certain protocols need to be observed between senders and receivers. In evolving systems, these interactions can shape future structures and create robustness.

Hutchins (1995) applies a very similar idea to ship navigation on board a small aircraft carrier. Navigation is, in a sense, an emergent computational phenomenon which draws from the hierarchical and socially distributed connections of individual officers without a central controller. Again, one of Ladyman and Wiesner’s truisms of complexity science is that ‘coordinated behaviour does not require an overall controller’. What Hutchins develops is a cognitive social computational model which abstracts away from individual cogniser’s internal states but still incorporates environmental conditions constitutively. In language, the individual cognitive states are important (as the CPUs are in computer networks) but they do not determine the language. The language emerges when a number of these states are connected in the right kinds of ways within a particular environment toward shared and varied tasks. Evolution plays a central role in what kinds of networks evolve for which purposes and how certain structures are stabilised over time. But many distinct components could have evolved simultaneously as De Boer et al. (2020) argue for

language evolution (Section 3.2). In fact, Seyfarth and Cheney (2014) specifically integrate formal language theory, social cognition, neurobiology, and comparative evolutionary biology into a single framework. They argue that many of the discrete combinatorics characteristic of human language can be found in simpler forms within nonhuman primate social cognition. They focus on features of the complex social groupings of baboons and argue that “human and nonhuman primates exhibit many homologous brain mechanisms that have evolved to serve similar social functions” (Seyfarth and Cheney 2014: 5). Again, they show that social cognition offers a system-level purview from which to appreciate the connections of social structure and language evolution involving “discrete, combinatorial, rule-governed, and open-end systems of communication in which a finite number of signals can yield a nearly unlimited number of meanings” (Seyfarth and Cheney 2014: 7).⁴¹

The idea of situated social cognition invites analogies with the 4E approaches to cognition, which have dominated the cognitive scientific landscape recently. Both systems biolinguistics and 4E approaches start with the criticism of the individualistic computationalist approach to language and cognition respectively. *Prima facie*, the move to the 4E approaches to cognitive science resembles the move to systems from individual organisms in biology. Most of the 4E approaches take environmental factors to be constitutive of cognition and advocate integrating social sciences into the cognitive sciences. The idea is that cognition (and ‘mind’ itself) is *embodied, extended, embedded* or situated, and *enacted* in the environment and not located squarely within the skull of the cogniser (Varela et al. 1991). The last three components emphasise the sometimes active (in the case of *enacted*) role the environment plays in mental phenomena. Take the concept of extended cognition for a moment (Clark and Chalmers 1998). This framework allows for ‘cognitive coupling’ in which an external device can be connected with internal processes for the completion of a task such as a calculator making certain calculations possible. Similarly, Google translate (or even a dictionary) can be said to operate in tandem with a language user to linguistically interact with her environment, thereby extending the language.

In terms of the second advantage over rival approaches, complexity science has the tools to *naturalise* a number of notions in MP and biolinguistics more generally. The concept of naturalisation here tracks the extent of biological involvement it contains. I’ll consider two such possibilities here. The first is the idea of an I-language or steady state of the language faculty. This term is meant to capture the idea of a mature state achieved by a language learner after the PLD has set various parametric settings of the innate UG capacity (Chomsky 1986). Unlike the alleged externalised or socio-political concept of a language like English or kiSwahili spoken in a particular community, I-languages are supposed to be more scientifically tractable. However, a common criticism of this picture is that it produces a static view of

⁴¹ I am not saying that extending computationalism beyond individual cognition and modular systems isn’t a fraught task. For instance, the property of cognitive impenetrability posited by Pylyshyn (1999) seems to insulate certain computations from external influences or confounding factors. An analogous case in social cognition is tricky to identify.

language and ignores various dynamic aspects of the system. This is because that steady state or I-language is identified with a narrow concept of syntactic or the computational component of the faculty of language (Hauser et al. 2002). Where complexity science can assist is by reinterpreting this steady state of a language learner as a dynamic equilibrium where “a system is said to be in ‘dynamic equilibrium’ or ‘steady state’ if some aspect of its behaviour or state does not change significantly over time” (Ladyman and Wiesner, 2020: 72). In biological systems, this state is related to the concept of *homeostasis*. Homeostasis is in turn related to feedback from the environment (e.g. linguistic interlocutors in your community) and robustness of structure [features (3) and (7) in the list above]. Notice that the proposal here is not merely about nomenclature. Homeostasis is intimately linked to the environment. It is not a completely isolated internal system or UG only reflecting some sort of activation by external stimulus. Mature language is then not an internal component of a human mind or brain but a complex steady state attained by intricate calibration with the linguistic (and non-linguistic) environment, i.e. individual networks are fine-tuned or updated by connections to other networks like in the LAN case.⁴² The upshot of this shift in interpretation is that, unlike the previous view, dynamic equilibrium is measurable. We have tools from biology, chemistry and physics to use as templates. In addition, it tracks linguistic maturity better than a static view. Consider the concept in chemistry. Dynamic equilibrium happens when the rate of the forward reaction is equal to the rate of the reverse reaction. It can look like nothing is changing but processes are happening continuously. It’s a steady state but also a moving target. In language, our environment places learning constraints on us which require us to quickly achieve a state in which we can communicate effectively (we might also be helped by innate catalysts). There is a ‘critical period’ in which our internal machinery is particularly attuned to environmental stimuli. But mastery of language is an ongoing process. Static or mechanical equilibrium, by contrast, occurs when the reaction has stopped completely. Sometimes generative linguists seem to imply this idea when they speak of a mature state of the language faculty being ‘set’ or ‘achieved’, but this is misleading.

This brings me to the second concept in need of naturalisation in terms of systems biology, namely the infamous idea of a linguistic community. Generative linguists have long argued that the linguistic community has no significant theoretical or scientific role to play in a theory of language. It is too amorphous and thus not conducive to formal characterisation. The idea of an external environment of speakers linguistically interacting in sometimes imperfect ways was considered a ‘theory of everything’ (Chomsky 2000) and as such a scientific nonstarter. Conventions,

⁴² For those worried about acquisition and poverty of stimulus issues here, I refer them to the growing literature on neural networks and pattern recognition (Youguo et al. 2007; Mattson 2014). These are learning models completely based on empiricist mechanisms and sometimes minimal inductive biases programmed in and represent yet another way in which complexity science can illuminate linguistic phenomena and processes such as language acquisition. See Yang and Piantadosi (2022) for a recent empiricist model of language learning which uses data from 74 languages. For studies that promote the role of reinforcement and environmental interaction in acquisition, see Ladd et al. (2015) and Sturdy and Nicoladis (2017).

regularities, and patterns among speakers within such a community, although favoured by some philosophers (Lewis 1975; Millikan 1984) have thus not received due theoretical investigation within the philosophy of linguistics. With these elements, the social aspect of language has been banished to the realms of sociolinguistics and anthropology. But systems biology offers us a means of reintegrating many of these elements within theoretical biolinguistics. We can start by asking what a system is on this view?

Importantly for our purposes, there are two concepts of ‘system’ in systems biology. They differ in terms of ontological commitment. As O’Malley and Dupré (2005: 1271) state:

The first account is given by scientists who find it useful for various reasons (including access to funding) to refer to the interconnected phenomena that they study as ‘systems’. The second definition comes from scientists who insist that systems principles are imperative to the successful development of systems biology. We could call the first group ‘pragmatic systems biologists’ and the second ‘systems theoretic biologists’.

The pragmatic approach dominates in the field. However, some systems biologists insist that such an approach offers little philosophical insight. Taking systems to be some collection or conglomeration of parts misses aspects of interconnection, emergent structures and symbiosis. The alternative, one I endorse here, is that “[s]ystems are taken to constitute a fundamental ontological category” (O’Malley and Dupré 2005: 1271). In our case, the linguistic community is a complex semiotic system and language is an emergent phenomenon therein. The system involves language users, learners, gestures, external linguistic resources (books, computers etc.), non-linguistic animals, and the external environment. If biolinguists are skeptical about the latter’s inclusion, it has actually been well-documented in dialectometry for years that geographical location affects language variance in systematic ways. This is not to endorse anything as strong as the Sapir-Whorf hypothesis which states that language, cognition, and location are linked deterministically (see Reines and Prinz 2009). Omar and Alotaibi (2017) conducted a study to show that geographical distance can influence the use and frequency of intensifiers (*really*, *very*, *extremely* and so on) across populations of the same language (Arabic) speakers based on location (in Egypt and Saudi Arabia) (see also Huisman et al. 2019; Reed 2020). Thus, the linguistic community is even more broad than many philosophers have taken it to be. There seems to be an underexplored link between the concepts of linguistic diversity and other types of conditions which influence biodiversity in plants and animals.

Again, there are various tools, some from neglected fields like dialectometry and cognitive anthropology and others from complexity science such as network analysis and Shannon information theory, which can aid us in understanding the complex dynamics that give rise to linguistic structure. Besides Kirby’s work on signalling systems, Skyrms (2010) adds elements of deception and the introduction of new symbols thereby connecting semantics to information theory. Mapping the interconnected aspects of language, communication and the environment offers a much more promising analogy with the emergence and structure of genes and genetics than does

the claims of organ-hood along the lines of a more individualist ontology. A methodological cornucopia unfolds.

Returning to the issue of how biological constraints might play a role in biolinguistics we can see that simplicity and optimality conditions such as those discussed within MP are not enough to shape the field into a more biological direction, even from an evolutionary perspective. Language evolution must take culture and general cognition into consideration. One prominent example of such an approach is Bickerton (2014b) who aims to connect MP to cultural evolution and primatology. According to him, each component only tells one part of an interconnected story of how complex language evolved in human populations. His story involves the property of discreteness (symbolic representation) witnessed in bee and ant colonies transposed to a particular primate, *homo sapiens*, triggering brain reconfiguration due in part to the construction of a new niche imposed by a change in the hunting environment of our ancestors. Culture then shaped the linguistic diversity we find across the world. Bickerton's work remains highly speculative in parts but as we have seen, in Section 3.2, many biologists and biolinguistics have objected to the single mutate theory of MP precisely on complexity grounds. For instance, the possibility of niche construction theory playing a role in language emergence and variation is empirically approached by Blasi et al. (2019) who assess the impact the transition from prehistoric forager societies to more industrialised agricultural societies had on our spoken language by means of paleodental data. Under MP, this evidence is peripheral at best, under MP+ it's much more central because it tells us how the environment might have exerted a force on our linguistic development.⁴³

The last advantage already indicated by the myriad possible theoretical convergences of systems biolinguistics is the methodological pluralism this perspective forces into linguistics. What were considered rival theoretical and formal frameworks such as Lexical Functional Grammar, Head-driven Phrase Structure Grammar, Dependency Grammar, Construction Grammar, Probabilistic Linguistics and more semantic approaches like Dynamic Syntax all have a place within MP+. Synchronous grammars, sociolinguistics, pragmatics, social cognition, and neurobiology are especially important for systems biolinguistics, as I have described it, more specifically. But the possibilities extend beyond traditional avenues of connection. By adding natural language to the established list of complex systems examples such as brains, economies, climates, eusocial insects, the Internet and the universe itself, we open ourselves up to analogies and models drawn from these well-studied phenomena no longer relegating the study of language to the realm of the biologically unique.

In terms of the complex systems features in use in systems biolinguistics, this view would aim to incorporate (1) numerosity, (2) feedback, (7) robustness, (8) nested structure and modularity, (9) history and memory, and (10) adaptive behaviour into the study of language. We have mostly seen snapshots of (1), (7), (8), and (10) here. Of course, future work would precisify these aims but for now the chief

⁴³ I thank an anonymous reviewer for suggesting this connection.

goal is to present an argument for a Maximalist approach to language sciences as a means of capturing the true essence of a viable biolinguistics.

Conclusion

In this article, I have had a number of related goals, primary among them has been to provide a sound scientific and biological basis for biolinguistics. I developed and argued for a Maximalist Program in contrast to the Minimalism of contemporary biolinguistics. MP+ is a complexity science and my specific take on biolinguistics involves a shift to systems biology. I showed that there are already accounts which might fit into it before offering a sketch of my own systems biolinguistic approach.

Acknowledgements I am grateful to audiences at the University of Durham's CHES seminar, the Kleiner Lecture Series at the University of Georgia, and the Philosophy Colloquium at the University of Miami for excellent feedback and comments on the content of this article. I would also like to thank Emmanuel Chemla and three anonymous referees for their thorough and comprehensive reviews. Any remaining errors are due to my pure stubbornness.

Funding Open access funding provided by University of Cape Town. This research was funded by the National Research Foundation of South Africa and the Vice-Chancellor's 2030 Future Leaders Fellowship at the University of Cape Town.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allott N, Lohndal T, Rey G (2021) Chomsky's "Galilean" explanatory style. In: Allott N, Lohndal T, Rey G (eds) *A companion to Chomsky*. Wiley, Inc, pp 517–528
- Anderson SR, Lightfoot DW (2000) The human language faculty as an organ. *Ann Rev Physiol* 62(1):697–722
- Baggio G (2018) *Meaning in the brain*. MIT Press
- Baggio G (2020) Epistemic transfer between linguistics and neuroscience: problems and prospects. In: Nefdt R, Klippi C, Karstens B (eds) *The philosophy and science of language*. Palgrave Macmillan, Cham, pp 275–308
- Behme C (2015) Is the ontology of biolinguistics coherent? *Lang Sci* 47:32–42

- Berwick R, Friederici A, Chomsky N, Bolhuis J (2013) Evolution, brain, and the nature of language. *Trends Cognit Sci* 17(2):89–98
- Berwick R, Chomsky N (2016) *Why only us: language and evolution*. The MIT Press, Cambridge
- Bickerton D (2014) Some problems for biolinguistics. *Biolinguistics* 8:73–96
- Bickerton D (2014) *More than nature needs: language, mind, and evolution*. Harvard University Press
- Blasi D, Moran S, Moisik S, Widmer P, Dediu D, Bickel B (2019) Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363(6432):eaav3218
- Boë LJ, Berthommier F, Legou T, Captier G, Kemp C, Sawallis TR, Fagot J (2017) Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS ONE* 12(1):e0169321
- Boeckx C (2006) *Linguistic minimalism: origins, concepts, methods, and aims*. Oxford University Press, Oxford
- Boeckx C (2015) Beyond Humboldt's problem: reflections on biolinguistics and its relation to generative grammar. *Lang Sci* 50:127–132
- Boeckx C, Grohmann K (2007) The biolinguistics manifesto. *Biolinguistics* 1(1):1–8
- Börjars K (2020) Lexical-functional grammar: an overview. *Ann Rev Linguist* 6(1):155–172
- Bresnan J, Asudeh A, Toivonen I, Wechsler S (2016) *Lexical-functional syntax*. Wiley, Blackwell
- Brighton H, Kirby S (2001) The survival of the smallest: stability conditions for the cultural evolution of compositional language. In: Kelemen J, Sosik P (eds) *Advances in artificial life*. Springer, Heidelberg
- Chomsky N (1956) Three models for the description of language. *IRE Trans Inf Theory* IT 2:113–123
- Chomsky N (1957) *Syntactic structures*. Mouton Press, The Hague
- Chomsky N (1963) Formal properties of grammars. In: Duncan LR, Bush RR, Galanter E (eds) *Handbook of mathematical psychology*. II. Wiley, pp 323–418
- Chomsky N (1965) *Aspects of a theory of syntax*. MIT Press, Cambridge
- Chomsky N (1973) Conditions on transformations. In: Anderson S, Kiparsky P (eds) *A Festschrift for Morris Halle*. Holt, Rinehart & Winston, New York, pp 232–286
- Chomsky N (1981) *Lectures on government and binding*. Foris, Dordrecht
- Chomsky N (1986) *Knowledge of language: its nature, origin, and use*. Praeger, New York
- Chomsky N (1995) *The minimalist program*. MIT Press, Cambridge
- Chomsky N (2000) *New horizons in the study of language and mind*. Cambridge University Press, Cambridge
- Chomsky N (2002) *On nature and language*. Cambridge University Press, Cambridge
- Chomsky N (2004) Beyond explanatory adequacy. In: Belletti A (ed) *Structures and beyond: the cartography of syntactic structures*. Oxford University Press, Oxford, pp 104–131
- Chomsky N (2007) Approaching UG from below. In: Sauerland U, Gärtner H-M (eds) *Interfaces+recursion=language? Chomsky's minimalism and the view from semantics*. Mouton de Gruyter, Berlin, pp 1–30
- Chomsky N (2013) Lecture 1: what is language? *J Philos* 110(12):645–662
- Christiansen M, Kirby S (2003) Language evolution: consensus and controversies. *Trends Cognit Sci* 7(7):300–307
- Clark H (1996) *Using language*. Cambridge University Press
- Clark A, Chalmers D (1998) The extended mind. *Analysis* 58(1):7–19
- Crenshaw K (1989) Demarginalizing the intersection of race and sex: a black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *Univ Chic Legal Forum* 140:139–167
- Culbertson J, Smolensky P, Wilson C (2013) Cognitive biases, linguistic universals, and constraint-based grammar learning. *Topics in Cognit Sci* 5:392–424
- Dabrowska E (2015) What exactly is Universal Grammar, and has anyone seen it? *Front Psychol*. <https://doi.org/10.3389/fpsyg.2015.00852>
- Deacon T (2008) Emergence: the hole at the wheel's hub 1. In: Clayton P, Davies P (eds) *The re-emergence of emergence: the emergentist hypothesis from science to religion*. Oxford University Press, Oxford
- De Boer B, Thompson B, Ravignani A, Boeckx C (2020) Evolutionary dynamics do not motivate a single-mutant theory of human language. *Sci Rep* 10(1):1–9
- Dediu D, Levinson S (2013) On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Front Psychol* 4:397
- Devitt M (2006) *Ignorance of language*. Oxford University Press, Oxford

- Dupré J, O'Malley M (2007) Metagenomics and biological ontology. *Stud Hist Philos Biol Biomed Sci* 38(4):834–46
- Everett D (2017) *How language began: the story of humanity's greatest invention*. Norton, New York
- Field H (1980) *Science without numbers: a defense of nominalism*. Oxford University Press
- Fitch T (2009) Prolegomena to a future science of biolinguistics. *Biolinguistics* 3(4):283–320
- Fitch WT, Friederici AD (2012) Artificial grammar learning meets formal language theory: an overview. *Philos Trans R Soc Lond B Biol Sci* 19:367(1598):1933–1955 <https://doi.org/10.1098/rstb.2012.0103>
- Fitch T, Reby D (2001) The descended larynx is not uniquely human. *Proc R Soc Lond Ser B Biol Sci* 268(1477):1669–1675
- Fitch T, De Boer B, Mathur N, Ghazanfar A (2016) Monkey vocal tracts are speech-ready. *Sci Adv* 2(12):e1600723
- Frank S, Bod R, Christiansen M (2012) How hierarchical is language use? *Proc R Soc B Biol Sci* 279:4522–4531
- French S (2011) Shifting the structures in physics and biology: a prophylactic promiscuous realism. *Stud Hist Philos Biol Biomed Sci* 42:164–173
- Friederici A, Chomsky N, Berwick R, Moro A, Bolhuis J (2017) Language, mind and brain. *Nat Hum Behav* 1(10):713–722
- Gao J, Li D, Havlin S (2014) From a single network to a network of networks. *Natl Sci Rev* 1(3):346–356
- Gasdaglis K, Madva A (2020) Intersectionality as a regulative ideal. *Ergo* 6(44):1287–1330
- Glimcher P (2014) Introduction to neuroscience. In: Glimcher P, Fehr E (eds) *Neuroeconomics: decision making and the brain*. Elsevier Inc, pp 63–75
- Goldberg A (2013) Constructionist approaches. In: Hoffmann T, Trousdale G (eds) *The Oxford handbook of construction grammar*. Oxford University Press
- Gopnik M (1990) Feature-blind grammar and dysphasia. *Nature* 344:715
- Greenberg JH (1963) Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg JH (ed) *Universals of Language*. Cambridge, MA: MIT Press, pp. 40–70
- Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: what is it, who has it and how did it evolve? *Science* 298:1569–1579
- Heim I, Kratzer A (1998) *Semantics in generative grammar*. Wiley-Blackwell
- Hölldobler B, Wilson E (2008) *The superorganism*. Norton, New York
- Hopcroft J, Ullman J (1979) *Introduction to automata theory, language, and computation*. Addison-Wesley, Reading
- Hopper P (1987) Emergent grammar. *Berkeley Linguist Soc* 13:139–157
- Hornstein N, Nunes J, Grohmann K (2005) *Understanding minimalism*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge
- Huisman J, Majid A, van Hout R (2019) The geographical configuration of a language area influences linguistic diversity. *PLoS ONE* 14(6):e0217363
- Hutchins E (1995) *Cognition in the wild*. MIT Press, Cambridge
- Itkonen, E (1978) *Grammatical Theory and Metascience: A Critical Investigation Into the Methodological and Philosophical Foundations of "Autonomous" Linguistics*. John Benjamins
- Jackendoff R (2002) *Foundations of language: brain, meaning, grammar, evolution*. Oxford University Press, Oxford
- Jerne, NK (1985) *The Generative Grammar of the Immune System (Nobel Lecture)*. *Angew Chem Int Ed Engl* 24: 810–816 <https://doi.org/10.1002/anie.198508101>
- Johnson K (2015) Notational variants and invariance in linguistics. *Mind Lang* 30(2):162–186
- Kac M (1992) *Grammars and grammaticality*. John Benjamins Publishing, Amsterdam
- Kauffman S (1995) *At home in the universe: the search for the laws of self-organization and complexity*. Oxford University Press, Oxford
- Kitano H, Oda K (2006) Self-extending symbiosis: a mechanism for increasing robustness through evolution. *Biol Theory* 1:61–66
- Kitcher P (1989) Explanatory unification and the causal structure of the world. In: Kitcher P, Salmon W (eds) *Sci Explan*. University of Minnesota Press, Minneapolis
- Kirby S (2013) Language, culture and computation: an adaptive systems approach to biolinguistics. In: Boeckx C, Grohmann K (eds) *Cambridge handbook of biolinguistics*. Cambridge Handbooks in Language and Linguistics. Cambridge University Press

- Konopka G, Bomar J, Winden K, Coppola G, Jonsson Z, Gao F, Peng S, Preuss T, Wohlschlegel J, Geschwind D (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* 462(7270):213–217
- Krause J, Lalueza-Fox C, Orlando L, Enard W, Green R, Burbano H, Hublin J, Hänni J, Fortea J, Rasilla M, Bertranpetit J, Rosas A, Pääbo S (2007) The derived FOXP2 variant of modern humans was shared with Neanderthals. *Curr Biol* 17:53–60
- Kretzschmar W (2015) Language and complex systems. Cambridge University Press
- Martins P, Boeckx C (2016) What we talk about when we talk about biolinguistics. *Linguist Vanguard* 2(1):1–15
- Ladd RD, Roberts SG, Dediu D (2015) Correlational studies in typological and historical linguistics. *Annu Rev Linguist* 1:221–241. <https://doi.org/10.1146/annurev-linguist-030514-124819>
- Ladyman J, Ross D (2007) Everything must go: metaphysics naturalized. Oxford University Press, New York
- Ladyman J, Wiesner K (2020) What is a complex system? Yale University Press
- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001) A Forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413(6855):519–523
- Lappin S, Levine R, Johnson D (2000) The structure of unscientific revolutions. *Natl Lang Linguist Theory* 18:665–671
- Lenneberg E (1964) A biological perspective of language. In: Lenneberg EH (ed) *New directions in the study of language*. MIT Press, Cambridge
- Levine R (2018) ‘Biolinguistics’: some foundational problems. In: Behme C, Neef M (eds) *Essays on linguistic realism*. John Benjamins, Amsterdam, pp 21–60
- Lewis D (1975) Languages and language. In: Gunderson K (ed) *Minnesota studies in the philosophy of science*. University of Minnesota Press, pp 3–35
- Lewontin R (1998) The evolution of cognition: questions we will never answer. An invitation to cognitive science, vol 4. MIT Press, Cambridge, pp 107–132
- Lobina D (2017) Recursion: a computational investigation into the representation and processing of language. Oxford University Press, Oxford
- Ludlow P (2011) *Philosophy of generative grammar*. Oxford University Press, Oxford
- Martins PT, Boeckx C (2019) Language evolution and complexity considerations: The no half-Merge fallacy. *PLOS Biology* 17(11): e3000389
- Marr D (1982) *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W.H. Freeman.
- Mattson M (2014) Superior pattern processing is the essence of the evolved brain. *Front Neurosci* 8(265):1–17
- Miller G (2003) The cognitive revolution: a historical perspective. *Trends Cognit Sci* 7(3):141–144
- Millikan R (1984) *Language, thought, and other biological categories*. MIT Press, Cambridge
- Mitchell M (2011) *Complexity: a guided tour*. Oxford University Press, Oxford
- Mondal P (2020) *Language, biology and cognition*. Palgrave Mcmillan
- Morin E (2008) *On complexity*. Hampton Press
- Müller S (2018) *Grammatical theory*. Language Sciences Press, Berlin
- Nefdt R (2019) Infinity and the foundations of linguistics. *Synthese* 196(5):1671–1711
- Nefdt R (2021) Structural realism and generative linguistics. *Synthese* 199:3711–3737
- O’Malley M, Dupré J (2005) Fundamental issues in systems biology. *BioEssays* 27:1270–1276
- O’Malley M, Dupré J (2007) Size doesn’t matter: towards an inclusive philosophy of biology. *Biol Philos* 22:155–191
- Omar A, Alotaibi M (2017) Geographic location and linguistic diversity: the use of intensifiers in Egyptian and Saudi Arabic. *Int J Engl Linguist* 7(4):220–229
- Pesetsky D (1999) Linguistic universals and universal grammar. In: Wilson RA, Keil FC (eds) *The MIT encyclopedia of the cognitive sciences*. MIT Press, Cambridge
- Pinker S (1994) *The language instinct*. Harper Perennial Modern Classics, New York
- Pinker S, Bloom P (1990) Natural language and natural selection. *Behav Brain Sci* 13:707–726
- Pinker S, Jackendoff R (2005) The faculty of language: what’s special about it? *Cognition* 95:201–236
- Poeppl D, Embick D (2005) Defining the relation between linguistics and neuroscience. In: Cutler A (ed) *Twenty-first century psycholinguistics: four cornerstones*. Lawrence Erlbaum Associates, Mahwah, pp 103–118
- Poeppl D, Iscardi W (2022) We don’t know how the brain stores anything, let alone words. *Trends Cognit Sci* 26(12):1054–1055

- Progovac L (2015) *Evolutionary syntax*. Oxford University Press, Oxford
- Progovac L (2016) A gradualist scenario for language evolution: precise linguistic reconstruction of early human (and Neandertal) Grammars. *Front Psychol* 7(1714):1–14
- Pylyshyn Z (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav Brain Sci* 22(3):341–423
- Pullum G (2007) Ungrammaticality, rarity, and corpus use. *Corpus Linguist Linguist Theory* 3:33–47
- Pullum G, Scholz B (2010) Recursion and the infinitude claim. In: van der Hulst H (ed) *Recursion in human language*, vol 104. *Studies in Generative Grammar*. Mouton de Gruyter, Berlin, pp 113–138
- Quine WV (1976) *Ways of paradox and other essays*. Harvard University Press, Cambridge
- Reed P (2020) Place and language: links between speech, region, and connection to place. *Wiley Interdiscip Rev Cognit Sci* 11(3):e1524
- Reines M, Prinz J (2009) Reviving Whorf: the return of linguistic relativity. *Philos Compass* 4(6):1022–1032
- Sampson G (2005) *The ‘Language Instinct’ Debate*. Continuum Press
- Scharff C, Petri J (2011) Evo-devo, deep homology and FoxP2: implications for the evolutions of speech and language. *Philos Trans R Soc B* 366:2124–2140
- Seyfarth R, Cheney D (2014) The evolution of language from social cognition. *Curr Opin Neurobiol* 28:5–9
- Silberstein M (2022) Context is king: contextual emergence in network neuroscience, cognitive science, and psychology. In: Wuppuluri S, Stewart I (eds) *From electrons to elephants and elections*. Springer, Cham, pp 597–640
- Sinha Chris (2010) Cognitive linguistics, psychology, and cognitive science. In: Geeraerts D, Cuyckens H (eds) *The Oxford handbook of cognitive linguistics*. Oxford University Press, Oxford, pp 1–30
- Shieber S, Schabes Y (1991) *Synchronous tree-adjointing grammars*. Technical Reports (CIS)
- Skyrms B (2010) *Signals: evolution, learning and information*. Oxford University Press, Oxford
- Smith N (2000) Introduction. In: Chomsky N (ed) *New horizons for the study of mind and language*. Cambridge University Press, Cambridge
- Sporns O (2013) Structure and function of complex brain networks. *Dialogues Clin Neurosci* 15:247–262
- Sporns O (2014) Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci* 17:652–660
- Soames S (1984) Linguistics and psychology. *Linguist Philos* 7:155–179
- Steedman M (2017) The emergence of language. *Mind Lang* 32:579–590
- Sturdy C, Nicoladis E (2017) How much of language acquisition does operant conditioning explain? *Front Psychol* 8(1918):1–5
- Turner S, Klimek P, Hanel R (2018) *Introduction to the theory of complex systems*. Oxford University Press, Oxford
- Tiede H, Stout L (2010) Recursion, infinity and modeling. In: van der Hulst H (ed) *Recursion and human language*. Mouton de Gruyter, Berlin/New York, pp 147–158
- van der Hulst (2010) In: Harry (ed) *Recursion and Human Language*. DeGruyter Mouton.
- Varela F, Thompson E, Rosch E (1991) *The embodied mind: cognitive science and human experience*. MIT Press, Cambridge
- Vargha-Khadem F, Watkins K, Alcock K, Fletcher P, Passingham R (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proc Natl Acad Sci* 92:930–933
- Wiltshcko M (2022) Language is for thought and communication. *Glossa J Gen Linguist* 7(1):1–19
- Yan K, Hricko J (2017) Brain networks, structural realism, and local approaches to the scientific realism debate. *Stud Hist Philos Sci Part C Stud Hist Philos Biol Biomed Sci* 64:1–10
- Yang Y, Piantadosi S (2022) One model for the learning of language. *PNAS* 119(5):e2021865119
- Youguo P, Huailin S, Tiancai L (2007) The frame of cognitive pattern recognition. In: *Chinese control conference*, pp 694–696
- Yuan J, Galbraith D, Dai S, Griffin P, Stewart N (2008) Plant systems biology comes of age. *Trends Plant Sci* 13(4):165–171