Outlining the parameters for a linguistic nativist position

Mark de Vos

Department of English Language and Linguistics, Rhodes University, South Africa
E-mail: m.devos@ru.ac.za

Abstract

I will outline the so-called "nativist position" as it relates to language.[1] While this debate has a long history, the data on which we can now draw is much richer and more varied than that available when the issues were first articulated and debated in the '60s and '70s. Similarly, the research context is radically different: behaviourism is no longer as dominant as it once was; Linguistics and Cognitive Science are much more theoretically mature and diverse, sporting a rich array of subdisciplines and perspectives. Along the way, some arguments have proved richer and more sustainable than others and indeed, in my opinion, the nativist position has shifted over the years. In some sense, it might be fair to ask whether the field has outgrown the debate and whether it is still as polarising as it once was. This article attempts to outline a nativist position by looking at the fundamental research questions that define it. Nativism is one possible coherent way of navigating through these questions, though other routes may also be possible. Consequently, there may be new spaces for rapprochement between different linguistic disciplines, which are often concealed by our scientific discourses.

Keywords: Innatism, nativism, Humboldt, Plato, biolinguistics, Universal Grammar.

1.   Introduction

I will start by outlining a number of research problems that have underpinned a fundamental research thrust in Linguistics. What may be thought of as Humboldt's Problem (Chomsky 1965, 1986) (section 3) considers what constitutes knowledge of language. This question can be answered in two ways: (i) what do we, as linguists and cognitive scientists, know to empirically constitute language (section 3.1) and (ii) what theoretical tools, systems and structures are required to explain those facts (section 3.2). Whether these systems are considered to be 'knowledge of language' in an intrinsic or extrinsic sense depends on whether one considers them to be real or not. In section 4, I consider Plato's Problem

(Chomsky 1965, 1986), namely where this knowledge has its origins: in discourse, in experience or in innate structures. I will argue that nativism is the least extreme available option. I will end by looking briefly at a few arguments for a particular type of nativism, which argues that linguistic structures are the result of language-specific computation, rather than general, cognitive systems.

There are a number of constraints on the article. Given the breadth of the debate and the decades over which it has been articulated, it will not be possible to give a comprehensive review of the debate or the literature in linguistics and ancillary fields. Interested readers are referred to Boeckx (2015), Botha (1981, 1989), Sampson (2001), Newmeyer (2005; 2014a,b,c), Harris (1993), Seuren (2009), Holmberg (2010), Piattelli-Palmarini (2000), among others. In addition, readers will find illuminating the spirited interchange between Lapin, Levine, Johnson, Roberts, Holmberg, Piattelli-Palmarini, Reuland and Uriagereka in successive editions of *Natural Language and Linguistic Theory* (volumes 18 of 2000 and 19 of 2001), as well as the interchange between Newmeyer (2004, 2006), Roberts and Holmberg (2005), and Holmberg (2010).

The expansion of cognitive science to such varied disciplines as genetics, computer modelling, psychology, philosophy and a plethora of linguistic frameworks is both a testament to the vitality of the field but is also impossible to cover comprehensively. Because of this, I have adopted two strategies. Firstly, I have explicitly chosen to frame the issues through a number of fundamental research questions and to provide an overview of the types of scientific, epistemological choices that can be taken with regard to them. In other words, rather than trying to define different specific positions and frameworks, I am trying to position them in a broader, ontological framework. Secondly, for the sake of specificity, I have chosen to situate myself as a generative linguist of a Minimalist type and I consistently show how this type of linguist might choose to navigate these research questions. For the non-specialist reader, I will explain what this means as it becomes relevant.

2.   Five basic linguistic questions

Defining the object of scientific study is always an ideological struggle and advances in theoretical depth often come at the expense of empirical coverage. For example, the Ptolemaic universe included empirical coverage of the astrological interaction between stellar systems, mental states and human intentionality (Feyerabend 1993). The Copernican solar system was only feasible at the expense of removing these from the object of study. Similarly, the objects of linguistic study are contestable. Nevertheless, the following questions help identify major research areas for linguistics (Jenkins 2000, Chomsky 1986, Boeckx 2015).

a.   What constitutes knowledge of language?   **Humboldt's problem** (Chomsky 1986)
b.   How is this knowledge acquired?   **Plato's problem**
c.   How is knowledge of language put to use to achieve a functional end?
                          Descartes' problem[2]

To these, Jenkins adds two questions that have become relevant thanks to advances in neuroscience and in evolutionary biology, but which are natural extensions to the original

---

[2] I will not be able to deal with Descartes' problem in this article, as it is a question that does not play a role in defining an innatist position.

questions. In my view, these last two questions are only possible if one adopts answers −
particularly to Plato's problem − that are innatist and realist in some sense, namely that at
least some aspects of knowledge of language are biologically specified and are therefore
biological loci for natural selection. Together these five questions constitute a 'biolinguistic'
position that holds language to be the result of cognitive computation of real, biological
structures rooted in neurochemistry and genetics.[3]

    a.   What are the relevant brain mechanisms underpinning this knowledge?
    b.   How does this knowledge evolve in the species? (Jenkins 2000: 11)

3.   Humboldt's problem: What is language?

Humboldt's problem strikes at the heart of what the object of linguistic study is to be. In
principle there are at least two ways of approaching a scientific object: (i) by describing its
formal composition and then demonstrating how these result in function or (ii) by describing
its functionality and then relating these to formal make-up. Both may, in principle, be fruitful
but the issue may be discipline or problem-specific. For instance, in palaeontology, the
classification of Palaeolithic tools is often done according to their probable function. The
function can then be used to account for their diversity of forms. Similarly, in studies of the
evolution of the eye (Dawkins 2004), given that the functions of eyes are limited to vision, a
function that is dependent on the fixed physical properties of light, it is possible to infer the
set of possible biological structures that can fulfil this function. Accordingly, eyes have
evolved separately at least 16 times and use a range of different biological systems (Dawkins
2004). In contrast, in anatomy, an object such as 'the human hand' is first analysed in terms
of its formal structure (skeletal structure, opposable thumb and index finger, five digits for
grasping, high numbers of nerve endings on fingertips, etc.) and then the infinite range of
functions of the hand can be inferred from its structure (e.g. grasping, twisting, hitting,
picking, caressing, holding, stirring, a receptacle for water, making love, etc.).

In my opinion, the broadly formalist and the broadly functionalist positions are closer than they
would appear. Both commit scientists to detailed analysis of form: "elucidation of their
functional motivation does not excuse us from actually describing grammatical constructions in
full and explicit detail" (Langacker 2012: 7). Both are scientifically useful, in certain contexts,
and the choice of one over the other is really a pragmatic choice about the complexity of the
problems under study. However, it seems to me that for some scientific endeavours like the
studies of the human hand and language, to infer finite structures and mechanisms from an
infinite number of functions is more difficult than to describe the finite structures and
mechanisms and then to explain how they could be used in an infinite number of creative ways.

3.1.   What constitutes language?

Armed with this insight, it is germane to ask what the formal/structural properties of language
are that make it unique or which differentiate it from other cognitive, behavioural or
communicative abilities. Most cognitive scientists would agree that it must include, firstly, a
massive lexicon consisting of an open-class of lexemes which is infinitely extendable (Hauser
1996) and a finite set of closed-class functional items and features existing in a steady state in

---

[3] Jenkins (2000) traces the origin of this term to a conference in 1974 at which both Chomsky and Luria were
present. See also notes in Chomsky (2005).

individuals (e.g. mature speakers do not acquire new passive morphology, tenses, determiners, complementisers or case systems etc.). Secondly, there must be a derivational component that uses lexical items to create larger structures of some sort. Thirdly, the resulting structures are characterized by recursion/hierarchy, discrete infinity, domains, etc.

Another property of language is the fact that there are clear and obvious similarities in structural organization between unrelated languages. Generative theory has been applied to hundreds of languages since its inception and theoretical concepts are transferable. For example, the work of Julien (2002) shows relative orders of tense, aspect and mood are consistent with the Mirror Principle (Baker 1985) across 530 languages. Furthermore, the fact that Greenbergian implicational 'universals' (Greenberg 1966; Greenberg, Osgood and Jenkins 1963) exist even as tendencies suggests deeper correspondences between human languages (Scalise, Magni and Bisetto 2008).

Another concrete example of an unexpected – but grammatically quite specific – similarity across unrelated languages concerns agreeing prepositions. Consider the following data cited in Baker (2010: 6).

(1) a.   Kambale a-ka-kanay-a   na-(*bo) abasyakulu
           Kambale 3S-PRES-speak-FV with-2   2.old.people
           'Kambale is speaking with the old people.'

     b.   Aba-syakulu b-o   Kambale a-ka-kanay-a   na-bo
           2-old.people 2-FOC Kambale 1s-PRES-speak-FV with-2
           'It's old people that Kambale is speaking with'   (Kinande)

The Kinande data replicate what was originally observed in Kilega by Kinyalolo (1991) and is found in several Bantu languages. Example (1a) shows that in Kinande, a DP *abasyakulu* ("old people") occurs as the complement of a preposition-like category *na* ("with"). Importantly, the preposition cannot agree with the DP in this construction. However, if the DP is fronted, stranding the preposition, then agreement is obligatory (1b). Thus, extraction of the prepositional complement is licensed by agreement. The same generalization is also visible in Fijian as shown by the following data, also cited in Baker (2010: 6).

(2) a.   Eimami saa qaaqaa   a   'ai-Boumaa [i-na   drano]
           we   ASP victorious ART native-place about-ART lake
           We, the natives of Boumaa, were victorious concerning the lake.'

     b.   A   drano 'eimami saa qaaqaa   ['i-na]   a   'ai-Boumaa.
           ART lake   we   ASP victorious about+3.SG ART native-place
           The lake, we the Boumaa people were victorious concerning it.'   (Fijian)

In (2a), the DP *drano* ("lake") occurs as the complement of a preposition-like category *i-na* ("about"). When the DP is fronted, stranding the preposition (2b), the morphological form of the preposition changes reflecting agreement (Baker 2010 citing Dixon 1988). Kayne (1994) notes a comment from Kenneth Hale that, in SVO languages, prepositions generally do not exhibit agreement – i.e. unless the prepositional complement has moved through the specifier of P. Therefore, we can take this to be a fairly general effect across languages.

(3) a. Ek   stem *voor   /vir die kandidaat
         I.1SG vote   for.AGR/for the candidate
         '(i) I am voting for the candidate'
         '(ii) I am voting in front of/before the candidate'

     b. Watter kandidaat moet ek   **voor**   /*vir stem
         which candidate must I.1 SG for.AGR /*for vote
         'Which candidate must I vote for?'   (Oosthuizen 2000: 72)

Finally, (3) shows the same pattern in Afrikaans. In (3a), the preposition *vir* ("for") takes a DP complement and corresponds with reading (i). It should be noted that the preposition *voor* ("front/before") is actually grammatical in this sentence but it means something quite different, namely reading (ii). It is important here that *voor* is not consistent with reading (i) in (3a). When the DP is extracted in (3b) and reading (i) is kept constant, the preposition undergoes a morphological change to *voor* ("for"), which is argued by De Vos (2013) to be agreement. Extraction is illicit without this morphological change on the preposition. In other words, it is not the case that languages are infinitely variable – commonalities exist that are oddly specific and these must surely count as a property of human language to be explained.

Another type of odd congruence is the discoveries of 20th century generative linguists concerning islands. Ever since Ross (1967), it has been known that languages exhibit island effects, a subtype of the domain-boundedness property of languages. Islands are logically possible but ineffable constructions. For example, (4) demonstrates a superiority effect: it is grammatical to extract a subject in a multiple-WH construction but not an object, which would "move across" the subject WH-item.

     (4) a. Who saw what?
         b. *What did who see ~~what~~?   [WH island]

Example (5) demonstrates another movement island. In (5a), the DP *a story about who* is in its unmoved, in situ position; it is possible to extract the WH-item from this constituent. Example (5b) shows a similar construction but the DP *a story about who* has moved to subject position as part of passivization; it is not possible to extract the WH-item from this moved position. Thus, extraction is illicit when the constituent has already moved.

     (5) a. Who was there [DP a story about ~~who~~] in the paper
         b. *Who was [DP a story about ~~who~~] printed ~~[a story about who]~~ in the paper?
                                                 [movement island]

Example (6a) illustrates so-called "Across-the-Board" extraction of a WH-item from both conjuncts of a coordinate structure, while (6b) shows that a WH-item cannot be extracted from only a single conjunct of a coordinate structure. For the sake of contrast, the same extraction is licit when the coordinator is replaced with an approximately semantic equivalent (6c). The crucial difference between the two types of construction is that (6b) is structurally coordinative while (6c) is structurally subordinative.

(6) a.   Who did John like ~~who~~ and Mary despise ~~who~~?
     b.   *Who did John like ~~who~~ and Mary despised Tumi?   [coordination island]
     c.   Who did John like ~~who~~ while Mary despised Tumi?

It appears that the existence of islands and domain-bounded phenomena are linguistic universals and therefore an important defining property of language. I am not claiming that all languages exhibit all island effects in the same ways; with the possible exception of the coordinate structure constraint, islands play out differently in different languages. I would also like to thank an anonymous reviewer for pointing me to the work of Hofmeister, Casanto and Sag (to appear), and references cited therein, which attempt to explain certain island effects in terms of processing. Islands and domain-bounded phenomena are major empirical discoveries that are not anticipated in earlier literature or frameworks. These types of facts were not part of earlier linguistic traditions because Generative Linguistics not only asks what empirical phenomena are present, but also which empirical phenomena are absent (i.e. gaps in the data). The methodological step that allows these data to be identified is the identification of ungrammaticality/ill-formedness and is central to a lot of generative work.[4] It is important to note that (i) ungrammaticality is a generativist position, not solely a nativist one, and (ii) it is possible to do generative linguistics without access to ungrammaticality evidence (e.g. historical linguistics where one generally only has access to corpora of positive evidence); the important distinction here is that all generativists accept the possibility of admitting such evidence. All of these characteristics constitute part of the growing body of knowledge that constitutes our knowledge of language.

3.2.   What equipment do we need to represent/analyse these structures?

The next question that naturally arises is, given these properties of language, what theoretical equipment is needed to represent and/or analyse them? At this point, there is another theoretical disagreement that may arise between two camps: those who are committed to representation – let us call them the "generativists" – and those who are not – let us call them the "poststructuralists".

Non-representational frameworks eschew representations of language in itself. This is not to say that they do not use models or representations of the architecture or computational system, but rather that they do not produce mental representations of the object of study itself. Skinnerian behaviourism was a framework of this type, which considered the mind (and therefore mental representations) to be beyond empirical study. Rather, it focused on inputs and behavioural outputs and the various types of conditioning that mediated the two. In a more contemporary context, connectionism bears many similar hallmarks. Connectionist models are mappings between inputs and outputs mediated by neural networks with various weightings. The representation in such a model, if it exists at all, is highly abstract and not easily evaluable by humans. It is next to impossible for humans to look at the end state of a neural network and to infer knowledge from it. Correspondingly, if one reads the literature in the field, it is clear that the focus is on the models and the mappings and the computational complexities thereof. Other non-representational frameworks include many discourse analytic and sociolinguistic frameworks. For these, the aim of linguistic study is not to identify,

---

[4] Note that negative grammaticality judgements are not "negative evidence" in the sense of absence of a measurable effect, which is governed by the dictum "absence of evidence is not evidence of absence". Ungrammatical sentences are not simply "gaps" in the data.

describe, explain or predict language structure, but rather to uncover the complexes of social meanings that emerge from text as a whole. Therefore, they emphasise E-language rather than I-language (Chomsky 1986).

In contrast, representationalists believe that there is an intrinsic good to having a representation of underlying structure that is amenable to human reason, by which the relations between constituents can be made explicit and can be evaluated and tested.

Chomsky's (1965) claim that the knowledge of language of an ideal speaker-hearer can be represented by a phrase structure grammar, should be read as a commitment to representationalism of this type:

> The competence attained by the normal speaker-hearer is represented by a transformational generative grammar which determines the representation of each sentence on the levels of phrase structure and transformational structure (inter alia).
>
> (Chomsky 1975: 45, in Jenkins 2000)

This position can be critiqued in terms of its specific technical requirements. Thus, whether or not linguistic knowledge can indeed be represented by a context-free vs context-sensitive grammar or whether linguistic knowledge includes rewrite rules or constraints remain interesting theoretical questions, which distinguish different representational frameworks. Furthermore, the methodological decision to rely on the idealization of a speaker-hearer can certainly be revisited in the context of big data and corpus studies. However, critiques of this sort are beside the point and do not touch on the scientific ideology behind the statement. It is perfectly possible to do generative analyses using corpus data (as all historical linguistic research does). Rather, the postulation of a mental system known by an ideal speaker-hearer (which was obviously an idealization from the start) is rather a commitment to a systemic mental representation, which underpins the messiness of performance. The acceptance of a competence-performance distinction (reflected in the I/E-language binary) is one that distinguishes many representational frameworks. I would argue that accepting the distinction is inevitable in a representational system; after all, if language/text is represented, then it seems to follow that the representation cannot be the "surface" text in itself, but must be an abstraction of it. Conversely, if one does not accept a competence-performance distinction, then there cannot be an abstract representation behind the text and one is committed to the text in and of itself.

At this point, another theoretical split appears around the question of whether linguistic representations are real or whether they are simply convenient ways of characterising the data. By "real" I intend that the representations have an ontological existence in themselves and are present in the mind or brain and are therefore available for computation. The other tradition goes back to Ernst Mach, James Maxwell, Gregor Mendel, Pierre Duhem and Jean-Baptiste Fourier, among others, who argued that their contributions were merely algebraic abstractions that did not necessarily have an independent existence but which captured the data accurately nonetheless (Besson 2014). To this end, for example, Maxwell promulgated his equations, which explain the properties of electromagnetism although at that time there was no way of explaining how they linked to reality. It was only with the advent of quantum theory that the equations could be explained realistically (Goldman 2015). It is pointed out by

Goldman (2015) that, although it is logically coherent to maintain a non-realist perspective, scientific discourse is replete with realist assumptions and that it becomes increasingly difficult to maintain a non-realist position in the light of the subtle and accurate predictions that constitute the success of modern scientific theorising. Similarly, in linguistics, it is perfectly possible to postulate structures without necessarily being committed to their reality. In fact, almost any structural theory can be interpreted in this way. Chomsky has consistently been committed to what he would label the more stringent hypothesis, namely that linguistic representations have independent existence and are instantiated in the mind/brain. However, it seems to me that the distinction is largely paradigm-internal insofar as scientists working with either set of assumptions can productively work together and contribute toward a common knowledge project.[5]

Within representational frameworks, there are differing positions on the nature of the representations themselves. For example, within generative frameworks, especially in the '80s and '90s, there was robust debate on whether representations were derived dynamically in accordance with various principles (the so-called "derivationalist position") or whether the principles acted as constraints on representations (the so-called "representationalist position").[6]

While the exact technical content of this representational system differs from framework to framework (e.g. Merge, Move, Agree, Parameters, LFG L-structures, Categorial Grammar types, etc.), it is possible to identify common assumptions that characterise many, if not all, frameworks of this type. Firstly, allied to the guiding principle of representationalism is the principle of symbolic abstraction, which is used to derive the ontology of categories over which representations are computed. These categories are usually represented by features of various types, e.g. nouns, verbs, selectional features, EPP feature on T; Dative case on a preposition; animacy feature on a pronoun; EPP on P, etc. Another way of representing abstraction is the notion that language involves creation of sets. Sets are used as a modelling language for representations in semantics, categorial grammars, HPSG, LFG, Principles & Parameters, and Minimalism (Montague 1974, Epstein and Hornstein 1999, Uriagereka 1999, Chomsky 2001b). From these can be inferred various relations defined over sets including hierarchy, constituency, quantification, c-command, asymmetry, antisymmetry, minimality, extension conditions, etc. as well as set-building or set-modifying operations such as Merge, Move and Agree.

Linked to the concern with symbolic abstraction is the principle of digitisation, which is meant to capture the discreteness in "discrete infinity". Digitization is the ability to impose a discrete schema onto the continuity of experience. Digitization is not necessarily inherent in representation (although it is very closely linked), but rather encodes empirical facts about language: in many ways, the linguistic system appears to be digital. For example, the following two phrases capture the same truth-conditional content although (7a) is a finite clause headed by a verb, while (7b) is a noun-phrase headed by a noun.

---

[5] Where the commitment to realism becomes important is when one attempts to expand the domain of enquiry beyond Plato's, Humboldt's and Descartes' questions to the biolinguistic questions. These latter two questions are only amenable to study if one takes the realist position, i.e. that linguistic mechanisms are "real" and therefore available as loci of natural selection.

[6] The terminology is confusing as I use the term "representationalism" in a broader sense in this article.

(7)   a.   Scipio destroyed Carthage
      b.   Scipio's destruction of Carthage

Important here is the way the arguments are introduced into each of these. In (7a) the direct object "Carthage" is assigned accusative case by the verb (specifically, "little *v*"), while in (7b) the semantic object is assigned case by a semantically empty preposition. This is because it is a digital property of nouns that they do not assign case, while it is a property of transitive verbs that they assign accusative/objective case. Moreover, grammaticality changes sharply if one inserts a case marker or removes it, as illustrated in (8).

(8)   a.   *Scipio   destroyed **of** Carthage
      b.   *Scipio's destruction   Carthage

Thus, the structural properties of nouns and verbs are different and appear to operate in a binary fashion and are not derived from semantics or function. In other words, it appears as though the mind imposes a digital mapping on the continuity of experience.

In summary, there is an epistemological choice to be made about whether or not it is the task of linguistic science to lay bare the structures that produce the empirical data of the world. If one chooses a non-representational route to grammar, then the focus tends to be on the architecture of apparatus that produces the data and that data become a means to the end of understanding the machine.[7] If one chooses a representational route, then the focus is on the internal, computational operations within the mechanism that generate representations of the data in itself.

4.   Plato's problem: Where does this apparatus come from?

Once scientists have committed to representations and systems and categories to drive them, it becomes necessary to ask from whence it all comes. There are multiple answers to these questions, each implying different paradigms. Depending on one's answers to the previous questions, one may take a number of positions here.

4.1.   The social, discursive construction of knowledge of language

One type of answer is framed in terms of society: the structures and categories on which language depends are not a property of biological individuals but are discursively co-constructed in society in the same way that culture is. This constructivist orientation is predominant around discourse studies, e.g., inter alia, Critical Discourse Analysis (Wodak 2007, Fairclough 2001), Speech Act Theory (Butler 2010, Searle 1969), Rapport Management (Spencer-Oatey 2000), etc.), literary criticism, etc. In these fields, it is natural to consider language as a functional resource to achieve discourse ends, where language is infinitely variable, infinitely changeable and dependent on shared conventions established by culture within discursive communities. These approaches are strongly non-representational in their epistemological commitments and consequently cannot seriously grapple with the main question at hand, namely where does the representational equipment come from. Thus, while these approaches describe how language is used in society to achieve communicative aims,

---

[7] I only mention non-representational approaches to grammar here. However, there are many non-representational approaches of a sociolinguistic or discourse-analytic nature that do not focus on grammar.

they are silent on why language has the structures it does and not others, why language is uniquely human and how it evolved in the species. This is precisely because the technical details of language appear to be biologically and not societally endowed. Consequently, there has been no explanation for discrete infinity, the existence of islands, movement, negative polarity items, agreement or for the existence or non-existence in particular languages of specific constructions within this framework. The language-in-society framework is therefore largely irrelevant here.

4.2.   Strong empiricism

Another school of thought draws on the empiricist tradition of Locke, Hume, and later Skinnerian behaviourism, connectionism and "corpus-driven" (Tognini-Bonelli 2001, Biber 2009) linguistics. This school holds that knowledge of language is acquired entirely through one's experience of the world, i.e. through an interaction with data.[8] Consequently, Mind is seen as a "black box" or treated in homogenising terms such as "general cognition". It is important to see these claims in historical context. During the '50s and '60s, cognitive science was in its infancy. Neuroscience was practically non-existent and the only means of looking inside the brain was with X-rays. The first attempt to scan a human organ (the brain) with ultrasound had only taken place in 1942 and, until the invention of the CT scan in 1972, the ability to look into a functioning human brain was extremely limited. Moreover, until the discovery of DNA (1953), the manner in which biological structure could be transmitted was not understood. In addition, the basis for mental computation and learning was not in place until Farley and Clark implemented a Hebbian neural network in 1954 (Farley and Clark 1954). In this particular historical context, it made sense to adopt a strong empiricist position because the alternatives were too imprecise. Against this backdrop, when Chomsky (1965) and Lenneberg (1967) made claims about the mental and biological substrates of language, it was revolutionary. Perhaps for the first time, linguists were able to glimpse into the workings of cognition through the lens of language when other means of doing so were extremely limited. Since then, much more information has become available within linguistics.

> There has been an explosion of knowledge […] from studies of languages and dialects that now number in the thousands. […] Evidence bearing on these questions is now available from numerous areas; e.g., syntax, semantics, morphology, phonology, articulatory and acoustic phonetics, language acquisition, language change, specific language impairment, language perception, sign-language, neurology of language, language-isolated children, creole language, split-brain studies, linguistic savants, and electrical activity of the brain, among others.
>
> (Jenkins 2000: ix)

However, it is my contention that the evidence from linguistics is no longer the most important or the most prominent in the debate on the nature of Mind; linguistic evidence is just one of many types of evidence now available. Recently, there has been a resurgence in interest in nature–nurture debates. Prominent among these are Jesse Prinz's (2012) *Beyond*

---

[8] Unlike the innatist position, this is not a straw man: empiricists really do make claims about the absoluteness of experience, e.g. Karmiloff-Smith Plunkett, Johnson, Elman and Bates (1998) explicitly (and persuasively) point out that knowledge cannot be inherited because neural nets develop in experience-dependent ways. Those that do not – i.e. those that hold that some systems are acquired while others are innate – are really holding a nativist position.

Human Nature and Steven Pinker's (2003) *The Blank Slate*. In each of these, discussions of innatism in language are only one of several that include a range of cognitive sciences. Since the early '60s, the range of evidence and information available to cognitive scientists has expanded dramatically. Linguistic evidence for nativism is now only one of many; insights about cognition come from medicine, computer science, psychology, neurology, genetics, evolutionary biology and many others including linguistics. The brain can be studied with CT and (f)MRI imaging, it can be modelled with Bayesian networks among many others, it can be studied through experimental behavioural studies, and its functioning can be affected by chemicals and drugs. Given these advances, there is no reason to continue to place the Mind and brain off limits and, consequently, there is no reason to necessarily adopt an a priori empiricist stance.

What these advances tell us is that the biology of the physical body is specialized and modular: lungs, hearts, brains, pancreases and livers all perform different functions with different physical and biochemical processes, each specialized to its own task. The difference between physical organs and mental ones is that we are better able to characterise the chemical and functional transitions from a multicellular blastula to a functioning organism.[9] We should ask ourselves why the psychology of the Mind or the biology of the brain should a priori be any different. During the heyday of behaviourism, the notion of the mind was anathema. Consequently, there could never be anything specific within it; to the extent that cognition was theorisable within behaviourist frameworks, it necessarily had to be of a general nature. However, in modern cognitive science, particularly neuroscience, the debate about modularity is alive and well (Calabretta 2007, Frankish 2011). It is not the case that the better ability to see and infer the brain's workings through f/MRI, CT/CAT, DOI/DOT and EROS imaging has put the question of modularity to bed. The visual cortex is particularly well studied and it is known that there are more than 30 specialized modules involved in visual perception in primates (Van Essen, cited in Ramachandran 2011), at least some of which are physically localisable as well as being mentally modular. There are distinct mental modules for facial perception and visual body perception: evidence from akinetopsia, achromotopsia, and visual agnosia point to a module for motion processing (Zihl, Von Cramon, Mai and Schmid 1991); experiments on selectively blinding kittens show there may be modules for horizontal and vertical edge detection (Hubel and Wiesel 1962); the mind has specialized modules for recognising faces and within those modules, specialized components for identification of mouths and eyes (Arcurio, Gold and James 2012). Even at the level of individual neurons there is specificity: there are specialized neurons for mirroring action

---

[9] It might be countered that physical organs are not modular in the sense that they all derive from stem cells based on cellular experience (e.g. chemical gradients) and general principles of cellular development. This is true of course, just as it is true for the development of neural architecture; neurons that are locally stimulated develop more connections in response to those experiences, which may lead to specialization of subsystems of neurons. The argument goes that since the specific properties of any particular organ (e.g. the heart) can be derived from more basic tissue via general principles, the heart is then not modular. This is, I think, a straw man insofar as nobody denies this extremely general notion of non-modularity: clearly the language faculty is embedded in neuronal development. Moreover, it misses the point that the principles governing the proper functioning of the heart (e.g. systolic and diastolic phases, the coordination of the atria, ventricles, integrity of the pericardium, etc.) are not the same principles governing the functioning of the liver (e.g. glycogenesis, gluconeogenesis, amino acid synthesis), i.e. that these organs are functionally modular irrespective of their ontogeny. As applied to language, this translates into there being a distinct difference in the claim that synchronic language processes necessarily require only "general cognitive principles" and the claim that the ontological development of the language faculty is consistent with the principle that "what fires together wires together".

(mirror neurons; Rizzolatti and Craighero 2004), neurons associated with actions that fire on visual stimulus (cannonical neurons; Grèzes, Armony, Rowe and Passingham 2003) and many other types of neurons with different functions and biochemistry. In other words, when we look carefully at what we might consider "general cognition", it devolves into fractal-like specificity. There is nothing "general" about "general cognition".

4.3.   Nativism

Another type of answer is to claim that language is entirely innate. An example of this is James IV of Scotland who allegedly locked up children minded only by a deaf and dumb caretaker and waited to see what language would emerge. The experiment was considered successful when the children uttered syllables akin to Hebrew (Pinker 1994). The point here is that this is a straw man because generative innatists have never been innatists in this narrow and indefensible sense.[10] "From the staunchest Chomskyan nativist to the most domain-general empiricist, all scientists concur that development involves contributions from both genes and environment" (Karmiloff-Smith et al. 1998: 588). It has always been patently clear that generative innatists considered that certain components of language were innate and biological, while others were gained from experience. Therefore, it should be borne in mind that what is called "nativism" or "innatism" is actually an intermediate position, where the real question is what and how much is innate vs acquired.

Within the nativist camp there are diverse answers to the question of exactly where innate linguistic knowledge comes from. Karmiloff-Smith et al. (1998: 588) distinguish between architectural innateness ("pre-specified learning mechanisms") and representational innateness ("pre-specified language content"). There is no doubt that biological architectures can be genetically specified. Thus the liver, heart and brain anatomy all develop in response to genetically-coded building instructions. When Chomsky (2007) refers to the language faculty as an organ, he is presumably invoking architectural innateness; Universal Grammar (UG) as a means of acquiring language should be read in terms of architectural innateness.[11] There are two problems here: the first is that there appears to be, what Jenkins (2000) calls "methodological dualism", i.e. a double standard applied to cognitive sciences. While we have no trouble accepting genetically pre-specified physical structures such as hearts and livers, or pre-specified complex behaviour such as bee dances, there seems to be a reticence among many cognitive scientists to allow cognitive structures to be developed similarly.
The second problem is that we do not yet know enough about the language faculty to specify exactly how it is encoded genetically; more precisely, although we understand a lot about language as a mental object, we do not yet understand how particular cellular configurations translate into what we know about Mind. In a real sense, we are in a similar position to

---

[10] Even Descartes, it should be noted, did not discount experience. The point of the parable about the demon revolves around an acknowledgement that experience plays a role in how we see the world; Descartes' argument was that experience could not be known to be infallible and so his sceptical method was aimed at distinguishing what was known from experience vs what could be known intrinsically.

[11] A reviewer pointed out that Chomsky (2001) assumes an innate set of features and that this might be considered representational innateness. I do not think it is as clear as that and an argument can be made that the types of features Chomsky (2001) talks about could be part of the pre-specified derivational architecture. For instance, these features are not "semantic" in the sense that their semantics play no role in the computation; the computation is blind to whether features are for Case, Tense, Agreement, etc., and operates by blindly matching interpretable/uninterpretable pairs. Seen from this perspective, a feature need be no less architectural than Merge. At the very least, I think it is incumbent on linguists to be clear about what type of innateness they refer to.

Gregor Mendel who was able to show that genetic characteristics were inherited and to develop coherent mathematical models of inheritance but who was unable to specify the precise biological mechanism. Similarly, Charles Darwin and Alfred Wallace were able to formulate the theory of evolution through descent, even though they were unable to identify the genetic mechanism. It is worth pointing out that Mendel's and Darwin's inability to specify the inheritance mechanism in no way invalidated their results or made them less scientific – even though the exact means by which heritable traits were transmitted biologically were only outlined over a hundred years later in 1953. This is a very different scientific standard to that applied to Language – compare this to the hyperbolic claim of Karmiloff-Smith et al. (1998: 589) that "to say anything scientifically useful [about innatism] one must be explicit about the pathway from genome to phenotypic outcome".

The questions of the genetic encoding of language remain for the biolinguistic programme. In Chomskyan frameworks, UG is a module of Mind – ultimately, of course, reducing to the cellular level – and is taken as being those aspects of language that are innate/inherited. UG is more than a general propensity to acquire language; at its broadest, UG is a genetic endowment, which limits/defines the set of learnable languages in all their deep, grammatical complexity (Chomsky 2007).

The need for UG to shape the learning process has traditionally been argued for on the basis of the Poverty of the Stimulus argument. This argument claims that children are exposed to highly variable, low-quality linguistic inputs where it is not possible to differentiate well-formed from ill-formed input. Sometimes, this has been combined with the claim that the amount of data in itself is insufficient. Over the years, these claims have been disputed by showing that caregivers do sometimes correct children, that the pace of acquisition is extremely variable, etc. (Reali and Christiansen 2005, Pullum and Scholz 2002) However, it seems to me that there is an additional reason for the postulation of UG: the "richness" of the stimulus. An insight of neural networks, as applied to applications like speech recognition, is how intensely rich the primary data actually are. Audible sounds within a child's (or a neural net's) context include child-directed language and non-child-directed language, but also, environmental sounds, traffic, telephones, television jingles, music, different accents and voice pitch, and all of this occurs simultaneously: the stimulus is extremely rich. In the cacophony of quotidian life, it is hard to know what to pay attention to and what to ignore and to factor out the irrelevant information. And yet, anecdotal experience tells us that a parrot (i.e. an empiricist learner) acquires many random sounds, including the ringing of the telephone, while a child brought up in the same environment acquires language. From this practical perspective, a genetic predisposition to sieve the data of experience in predetermined ways would be extremely useful. The same point can be made mathematically: for any finite data set, there are an infinite number of possible analyses consistent with it (cf. the Duhem-Quine Problem). Specifically, Gold (1967) shows that a regular language is not learnable from text given infinite time.[12] In other words, for language to be learnable from data there needs to be an innate means of restricting the range of possible languages.

This restriction has traditionally been the role of UG. We could characterise UG as the initial state of a grammar before experience is applied. This can be represented schematically as:

---

[12] See http://groups.lis.illinois.edu/amag/langev/paper/gold67limit.html (Accessed 3 April 2015). A similar argument is presented by Carnie (2002: 15), where he shows that for any finite data set, there is more than one possible algorithm to generate it.

Learned Language (Final State) is a product of experience applied to an initial state. Freidin (2012) points out that the initial state is by definition innate. In the words of Lightfoot (2002: 506), "[l]anguage acquisition takes place by an interaction of UG, the PLD, and nothing else."

    (9)   Initial State + Experience = Final State
    (10) UG + Primary Linguistics Data = I-language (Freidin 2012: 30)

Taking into account Chomsky's (2007) "third factor" arguments, this might be modified to something like (11) where additional architectural properties of and constraints on the cognitive system may play a role. Importantly, this is only possible because the properties of the Initial State are, in my opinion, fairly restricted.

    (11) Initial State$^{\text{Language Faculty}}$ + Experience + Cognitive properties = Final State

One question that arises is the characterization of this initial state and how it might be translated into cognitive functions. "We need a list of computations that linguistic theorists deem indispensable to solve their particular problem" (Fitch 2009: 298, in Boeckx 2015). Boeckx (2015: 7) argues that:

> Until we have reached this level of 'computational primitivism', until we have done this translational research, we won't be in a robust position to tell the neurophysiologist, or the comparative psychologist, or the geneticist, what to look for.

Within this narrative, Humboldt's question has already been answered: knowledge of language (i.e. the initial state) is constituted by Merge operating over syntactic features:

> […] biolinguistics […] is an attempt to go beyond Humboldt's problem. That is to say, work in biolinguistics […] regards Humboldt's problem as "essentially solved". Biolinguists do so, for example, by assuming that "Merge+interfaces" is the answer to that problem (Hauser et al. 2002). Put another way, they consider the traditional focus of theoretical linguistics to be a thing of the past, and ask how the results of that literature can be used to answer the four remaining questions.
>
>   (Boeckx 2015: 8-9)

Appealing to *Deus ex machina* sits uncomfortably with scientists – although it has not prevented them from invoking it, e.g. Newton's gravitational action at a distance, Pauli's suggestion of the neutrino, Einstein's criticism of the EPR paradox in quantum theory as spooky action at a distance, etc. However, in these cases, the additional assumption could be provided with specific content: in Newton's case, gravitation could be measured; in Pauli's case, the precise properties of neutrinos could be predicted; in Einstein's case, his prediction was proven correct (albeit counter-intuitively) by quantum entanglement.

Similarly, postulating well-defined components of UG with specific content allows hypotheses to be formulated and tested. The testing of hypotheses about what theoretical tools are required to analyse language are the "bread and butter" of theoretical linguists. An example is the

various definitions of command: (i) according to Kayne (1994), c-command can look out of a left branch; (ii) Barbiers (1995: 29) argues for c-command that allows a constituent to look out of a specifier in a right branch; and (iii) M-command allows a constituent to look "upwards" in a tree, etc. These are all incredibly well-defined, nuanced and precise characterisations, which give specific content to the notion of 'UG'. This is not to say that our knowledge of UG is necessarily definitive or complete, or that these characterizations are uncontested: this is not how science works. But at least some linguistic characterizations are specific enough to be useful in other cognitive sciences. One such characterization of UG is thus "Merge + interfaces". In my own work, I develop this perspective: I make specific claims about the content of UG, namely that it consists of the ability to construct a lexicon, syntactic features and SetMerge; principles of knowledge representation known as "Normal Forms" (De Vos 2008, 2014a) and possibly a linearization component (De Vos 2013; De Vos 2014b,c). Within this framework, the unique structure of human language is a result of these simple components of UG interacting. Whether or not one adopts my own particular, radically Minimalist framework or any other, I would contend that UG can be made specific and contentful enough to at least start to provide the basis on which other cognitive scientists can construct testable hypotheses in their disciplines.

4.3.1. Domain-general vs domain-specific nativism

There is considerable disagreement about the nature of this innate restriction centring around the generality vs domain-specificity of the mechanism. It is perhaps this that is the pivot of theoretical disagreement in the linguistic world. "The ravine between theories lies in how rich and how domain-specific the innately specified component is claimed to be […]" (Karmiloff-Smith et al. 1998: 588).

Whatever one's position on the exact content of UG, there may also be genetic endowments that are not uniquely human, which may nevertheless form part of the language faculty (FLB; Hauser, Chomsky and Fitch 2002). These might include, for instance, (i) a genetic cognitive predisposition to experience the world in particular ways, such as being able to distinguish dynamic events vs static entities which are then grammaticalized as verbs and nouns respectively; (ii) old and new information; (iii) Figure vs Ground, and (iv) analogical reasoning, etc.

Chomskyan linguistic frameworks explicitly adopt the working hypothesis that the language faculty also includes domain-specific computation (i.e. the modularity hypothesis).[13] It is worth distinguishing two types of modularity within Minimalist-type approaches. The first is of a technical nature and describes the interactions between different components of the Y/T model of language, e.g. the PF interface may not introduce symbols that require computation at LF; similarly, LF may not introduce items that require PF computation into the representation, etc. I will not be dealing with this conception of modularity here. The second, related, notion of 'modularity' is the claim that the operations applied within language are not available elsewhere in the cognitive system, i.e. as far as we know, there are empirical properties of

---

[13] This is not the same as claiming that the language faculty is unique to humans. It is quite possible that highly specific cognitive systems can be used in the human species to compute language but that these could be adaptations of specific cognitive systems in a precursor, but which are not "general cognitive systems". Consequently, whether these domain-specific systems are part of FLN or FLB (Hauser et al. 2002) is not relevant here, I think.

language that appear to be unique to language and appear not to be true of other behaviour; similarly, the types of computation that occur in language appear not to have obvious correlations in other mental systems, and therefore these computations are domain-specific.

When seen against the backdrop of a cognitive science historically dominated by behaviourism, it is evident that the modularity hypothesis was an important means of differentiating generative linguistics from behaviourism and to distinguish mental from external behaviour. It is notable that generative theory has adjusted its position on modularity over the years from a very rich UG in the Principles & Parameters/Government & Binding approaches to a relatively lean one in early Minimalism (Chomsky 1995), an even leaner version in later Minimalism (Hauser et al. 2002), and finally to making allowance for "third factor" (broad, cognitive, UG-external) influences (Chomsky 2007). For example, the Principles & Parameters framework balances universal principles with language-specific settings for parameters (Lightfoot 1993, Holmberg and Platzack 1995, Holmberg 2010, Holmberg and Roberts 2014). This means that the set of principles and the set of language-specific, learned parameter settings are potentially very large. Minimalist syntax places all language-specific features and settings as lexical items, and universal principles are radically reduced to Merge, Move and Agree and then, in Hauser et al. (2002) to just recursion (presumably instantiated by Merge) although this move is resisted by some who prefer a richer UG (Chomsky 2001a, Pinker and Jackendoff 2005).[14] Principles are thus restricted to these three. Parametric settings may still be a very large set but there are constraints placed on the nature of parameters: they must be expressible by features on lexical items (i.e. parameters may not be expressed as properties of constructions).

Furthermore, the insight behind the Minimalist Program also affects modularity in subtle ways. The Strong Minimalist Thesis contends that the properties of narrow syntax are determined by the bare output conditions of the PF- and LF interfaces – and nothing else (Chomsky 2000 *et seq.*). In other words, the representational output of the syntactic component must be visible to – and evaluable by – the conceptual-intentional interface and the sensory-motor interface: even if syntactic computation may be domain-specific, its outputs cannot be. In addition, there have been a number of studies attempting to show that linguistic-type principles are involved in musicality, numeracy, vision (Gregory 1997) and even one's moral sense (Roedder and Harman 2015). Given that the modularity hypothesis has adjusted in these ways, it is possible that this hypothesis is no longer as central a tenet as it once was. If it was conclusively demonstrated that all linguistic computation was, say, exapted from the visual cortex, then there might not be any immediate consequences for Minimalist theory: the analyses would stand, the data and insights would stand, the Strong Minimalist Hypothesis would also most likely stand and might even receive support from the finding.[15, 16]

---

[14] Chomsky (2001a) incorporates an innately pre-specified feature inventory. De Vos (2014) argues that reducing UG to Merge still requires a substantial set of operations and features upon which Merge relies and that, consequently, it is only a superficial reduction in complexity.

[15] However, the question would still remain why visual neurons had been exapted in this way in humans but not in any other seeing animal – an interesting question for biolinguistics.

[16] To assess the plausibility of this claim that a retreat from modularity would not necessarily impact the practice of linguistic analysis, it might be useful to consider that we have case studies where it has already happened. Consider the distinction between FLN and FLB introduced by Hauser, Chomsky and Fitch (2002). This paper radically and suddenly made the claim that recursion might be the only component of syntax that might not have precursors elsewhere. This meant that "narrow syntax" suddenly became much narrower (Faculty of Language Narrow; FLN) while the broader language system (Faculty of Language Broad; FLB) was, in principle, open to

However, I think that there are several good empirical and methodological reasons to maintain the modularity working hypothesis. Firstly, modularity should be kept because the data demand it. Language is qualitatively different to any other communicative behaviour that we are aware of (see section 3 above). While a fair amount of research has attempted to demonstrate recursion and context-unboundedness in other animal behaviour, as far as I know, there is no other cognitive behaviour that displays anything like island constraints, or like the conditions on extraction from PPs mentioned above. Similarly, Principle A/B, ATB, and A/A-bar effects, among others, seem specific to language. If language only uses "general cognitive principles" then surely these types of effects would be commonplace in the animal kingdom. But island effects appear to be limited to language and language appears to be restricted to humans. In the context of these facts, it makes sense to posit modularity as a working hypothesis.

Secondly, modularity should be kept because it is good science: it challenges scientific assumptions as the literature on the debate amply attests. If the hypothesis explicitly stands to be refuted then this deepens our understanding of cognition and how language is embedded in the brain. Despite decades of research, the modularity hypothesis has still not been persuasively disproven, in my view. This is not because it is inherently undisprovable. If it could be shown that another cognitive process had the same characteristics as language, then the modularity of that process would be disproved, e.g. if island effects occurred during visual processing; if Principle A/B effects could be found in the maze-solving abilities of mice; if quantifier-raising effects were exhibited in discrimination tasks among pigeons; if problem-solving could be shown to involve functional "layers" reminiscent of the distinction between CP, TP and vP; if an A/A-bar distinction arose in kinship systems in bonobos, etc. Any of these would be sufficient to show that the types of computation that are fundamental to language are also available in other types of cognitive processes. Decades of close and nuanced analysis of diverse languages have shown these effects to be present in language. The burden of proof lies on those who would claim otherwise, and abandoning the modularity hypothesis as a methodological principle of conservativeness would mean that cognitive scientists are no longer held to this rigorous standard of proof.[17]

Importantly, the ways in which innate and learned learning architectures may be evolutionarily related are also becoming better understood. Frankish (2011) points out that specialized subsystems are characteristic of evolutionary systems and that specialized cognitive systems are probably more efficient than general cognitive systems. Jacobs and Jordan (1992) show that an undifferentiated learning network can develop internal specialisation. However, this finding must be interpreted in the light of the findings of

---

other cognitive processes. This was a retreat from modularity and has subsequently been confirmed by Chomsky's (2007) suggestion that linguists look to broader, cognitive "third factors". Despite this departure from modularity and a radical opening up of the field, minimalist syntax continues in a robust fashion.

[17] The degree to which these linguistic phenomena are specific to language is evidenced by the fact that it is hard to even imagine what they might look like in non-linguistic domains. Perhaps this is a significant portion of the problem: there is a collective failure of the scientific imagination in being able to think of how an object of analysis in one domain can be effectively transferred to another with very different scientific objects. There are interesting findings that deserve deeper investigation to see if they can be developed into proofs of this sort. For example, if it could be shown that apes use abstract features in analogical reasoning tasks (something apparently unique to apes) then the feature-driven properties of language could be shown to be non-modular. Ultimately, features are part of the discreteness characteristic of language. Similarly, Roedder and Harman (2015) identify island phenomena in moral behaviour. If these effects could be validated, replicated and developed in greater detail then they may be very persuasive arguments against modularity.

Calabretta (2007), who develops a theory of the evolution of modularity itself. She argues that such learning in undifferentiated (non-modular) networks is subject to "neural interference" and "genetic interference" which can be reduced if the learning network is inherited through sexual reproduction: "entrusting the task of finding the neural network architecture to evolution and that of finding the network connection weights to learning is a way to completely avoid the problem of genetic interference" (Calabretta 2007: 403). If these types of results hold up to scrutiny, they strongly suggest that modularity is compatible with modern understanding of both language and the brain, and a new consensus is possible.

5.   Conclusion: The content of UG

In this paper, I have, starting with fundamental research questions, outlined a set of epistemological choices that cognitive scientists need to make. The aim of this is twofold. Firstly, it provides a descriptive framework by which we can evaluate different theories and positions the better to compare and critique them. Secondly, it explains the nativist position against its alternatives.

  a.   General research programme:
       (i)   Does the framework accept the three main research questions?

  b.   Humboldt's problem: the nature of representations:
       (i)   Is the framework representational?
       (ii)   Are the representations real or abstractions over data?

  c.   Plato's problem: the origin of knowledge:
       (i)   Is knowledge from society, from experience, from internal specification, from a (nativist) combination of internal and external specification?
       (ii)   If a nativist position is chosen: How domain-specific are the computations?

Although this paper – which was based on a debate at a conference – started around questions of nativism vs empiricism, I have argued that this is only one of many epistemological choices made by linguists. The differences between the theoretical clusterings characterized as "generative" or "functionalist" do not necessarily hinge on formalism vs functionalism, but rather on their commitments to a range of issues such as modularity, representation and empiricism/nativism which, in turn, can be seen as choices prompted by Humboldt's and Plato's problems. I have tried to show how many of these same choices are present in other fields of scientific enquiry. I argue that generative nativism is one particular way of navigating these choices and that, in fact, generative linguists have made considerable progress toward resolving and moving beyond them.

References

Arcurio, L., J. Gold and T. James. 2012. The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia* 50: 2454–2459.

Baker, M. 1985. The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3): 373–415.

Baker, M. 2010. On parameters of agreement in Austronesian languages. In M. Raphael, E. Potsdam and L. Travis (Eds.) *Austronesian and Theoretical Linguistics*. Amsterdam: John Benjamins. pp. 345–374.

Barbiers, S. 1995. *The syntax of interpretation*. HIL Disserations 14. The Hague: Holland Academic Press.

Besson, U. 2014. Teaching about thermal phenomena and thermodynamics: The contribution of the history and philosophy of science. In M. Matthew (Ed.) *International handbook of research in history, philosophy and science teaching*. London: University of Toronto Press. pp. 245–283.

Biber, D. 2009. Corpus-based and corpus-driven analyses of language variation and use. In B. Heine and H. Narrog (Eds.) *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press. pp. 159–192.

Boeckx, C. 2015. Beyond Humboldt's problem: Reflections on biolinguistics and its relation to generative grammar. *Language Sciences* 50: 127–132.

Botha, R. 1981. *The conduct of linguistic inquiry: A systematic introduction to the methodology of generative grammar*. The Hague: Mouton De Gruyter.

Botha, R. 1989. *Challenging Chomsky: The generative garden game*. Oxford and New York: Blackwell.

Butler, J. 2010. Performative agency. *Journal of Cultural Economy* 3(2): 147–161.

Calabretta, R. 2007. Genetic interference reduces the evolvability of modular and non-modular visual neural networks. *Philosophical Transactions of the Royal Society* 362: 403–410.

Carnie, A. 2002. *Syntax: A generative introduction*. Malden MA: Blackwell.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

Chomsky, N. 1986. *Barriers*. Cambridge MA: MIT Press.

Chomsky, N. 1995. *The Minimalist program*. Cambridge MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (Eds.) *Step by Step*. Cambridge MA: MIT Press. Reprinted. pp. 89–155.

Chomsky, N. 2001a. Beyond explanatory adequacy. *MIT Occasional Papers in Linguistics* 20. Cambridge MA: MIT Press.

Chomsky, N. 2001b. Derivation by phase. In M. Kenstowicz (Ed.) *Ken Hale: A life in language*. Cambridge MA: MIT Press. pp. 1–52.

Chomsky, N. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1–22.

Chomsky, N. 2007. Approaching UG from below. In U. Sauerland and H.-M. Gärtner (Eds.) *Interfaces + Recursion = Language? Chomsky's Minimalism and the view from semantics.* Berlin: Mouton de Gruyter. pp. 1–29.

Dawkins, R. 2004. *The ancestor's tale: A pilgrimage to the dawn of life (Kindle edition).* Houghton-Mifflin.

De Vos, M. 2008. Deriving narrow syntax from principles of lexical organization. *Lingua* 118: 1864–1899.

De Vos, M. 2013. Afrikaans mixed adposition orders as a PF-linearization effect. In T. Biberauer and M. Sheehan (Eds.) *Theoretical approaches to disharmonic word orders.* Oxford: Oxford University Press. pp. 333–358.

De Vos, M. 2014a. The evolutionary origins of syntax: Optimization of the mental lexicon yields syntax for free. *Lingua* 150: 25–44.

De Vos, M. 2014b. Head movement is an artefact of optimal solutions to linearization paradoxes. *Stellenbosch Papers in Linguistics PLUS* 44: 23–48.

De Vos, M. 2014c. The spec-head vs head-spec asymmetry: Post-syntactic linearization and the embedded-topicalization paradox. *Stellenbosch Papers in Linguistics PLUS* 43: 81–117.

Dixon, R.M.W. 1988. *A grammar of Boumaa Fijian.* Chicago: University of Chicago Press.

Epstein, S. and N. Hornstein. 1999. *Working Minimalism.* Cambridge MA: MIT Press.

Fairclough, N. 2001. *Language and power.* London: Longman.

Farley, B. and W. Clark. 1954. Simulation for self-organizing systems by digital computer. *IRE Transactions on Information Theory* 4(4): 76–84.

Feyerabend, P. 1993. *Against method* (3rd ed.). London, New York: Verso.

Frankish, K. 2011. Cognitive capacities, mental modules, and neural regions. *Philosophy, Psychiatry, & Psychology* 18: 279–282.

Freidin, R. 2012. *Syntax: Basic concepts and applications.* Cambridge: Cambridge University Press.

Gold, M. 1967. Language identification in the limit. *Information and Control* 10(5): 447–474.

Goldman, S. 2015. *Science wars: What scientists know and how they know it.* Available online: http://www.thegreatcourses.com/courses/science-wars-what-scientists-know-and-how-they-know-it.html (Accessed June 2015).

Greenberg, J. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (Ed.) *Universals of language*. Cambridge MA: MIT Press. pp. 73–113.

Greenberg, J., C. Osgood and J. Jenkins. 1963. Memorandum concerning language universals. In J. Greenberg (Ed.) *Universals of language*. Cambridge MA: MIT Publishing. pp. xv–xxvii.

Gregory, R. 1997. Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society* 352: 1121–1128.

Grèzes, J., J.L. Armony, J. Rowe and R.E. Passingham. 2003. Activations related to 'mirror' and 'canonical' neuron in the human brain: an fMRI study. *NeuroImage* 18(4): 928–937.

Harris, R. 1993. *The linguistics wars*. Oxford: Oxford University Press.

Hauser, M. 1996. *The evolution of communication*. Cambridge MA: MIT Press.

Hauser, M., N. Chomsky, and T. Fitch. 2002. The faculty of language: What is it, who has it and how did it evolve? *Science* 298: 1569–1579.

Hofmeister, P., L. Casanto and I. Sag (To appear). Islands in the grammar? Standards of evidence. In M. Sprouse and N. Hornstein (Eds.) *Experimental syntax and the islands debate*. Cambridge: Cambridge Scholars Publishing. pp. 1–22.

Holmberg, A. 2000. Scandinavian stylistic fronting: How any category can become an expletive. *Linguistic Inquiry* 31(3): 445–483.

Holmberg, A. 2010. Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics* 36(1): 1–48.

Holmberg, A. and C. Platzack 1995. *The role of inflection in Scandinavian Syntax*. Oxford Univeristy Press.

Holmberg, A. and I. Roberts. 2014. Parameters and the three factors in language design. In M. Picallo (Ed.) *Linguistic variation in the minimalist framework*. Oxford: Oxford University Press. pp. 61–81.

Hubel, D. and T. Wiesel. 1962. Receptive fields binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* 160(1): 106–154.

Jacobs, R. and M. Jordan. 1992. Computational consequences of a bias toward short connections. *Journal of Cognitive Neuroscience* 4(4): 323–336.

Jenkins, L. 2000. *Biolinguistics: Exploring the biology of language*. Cambridge: Cambridge University Press.

Julien, M. 2002. *Syntactic heads and word formation: A study of verbal inflection*. New York: Oxford University Press.

Karmiloff-Smith, A., K. Plunkett, M. Johnson, J. Elman, and E. Bates. 1998. What does it mean to claim that something is 'innate'? Response to Clark, Harris, Lightfoot and Samuels. *Theoretical Linguistics* 13(4): 588–597.

Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge MA: MIT Press.

Kinyalolo, K. 1991. Syntactic Dependencies and the SPEC-Head Agreement Hypothesis in Kilega. PhD thesis, UCLA.

Langacker, R. 2012. Substrate, system, and expression: Aspects of the functional organization of English finite clauses in English. In M. Brdar, I. Raffaelli, and M. Žic Fuchs (Eds.) *Cognitive Linguistics between universality and variation*. Newcastle: Cambridge Scholars Publishing. pp. 3–52.

Lenneberg, E. 1967. *Biological foundations of language*. New York: John Wiley & Sons.

Lightfoot, D. 1993. Why UG needs a learning theory: Triggering verb movement. In C. Jones (Ed.) *Historical linguistics: Problems and perspectives*. Harlow: Longman. pp. 123–138.

Lightfoot, D. 2002. Grammatical approaches to syntactic change. In B. Joseph and R. Janda (Eds.) *The handbook of Historical Linguistics*. Malden MA: Blackwell. pp. 495–508.

Montague, R. 1974. *Formal philosophy: Selected papers of Richard Montague.* New Haven: Yale University Press.

Newmeyer, F. 2004. Againts a parameter-setting appraoch to typological variation. In P. Pica (Ed.) *Linguistic variation yearbook*. John Benjamins. pp. 181–234.

Newmeyer, F. 2005. *Possible and probable languages: A Generative perspective on linguistics*. Oxford: Oxford University Press.

Newmeyer, F. 2006. *A rejoinder to 'on the role of parameters in universal grammar: a reply to Newmeyer' by Ian Roberts and Anders Holmberg.* Available online: http://ling.auf.net/lingBuzz/000248 (Accessed 10 December 2015).

Newmeyer, F. 2014a. Getting the word out: The early generativists' multi-pronged efforts to diffuse their ideas. *Language* 90(1): 241–268.

Newmeyer, F. 2014b. Syntactic change: Between parameters and fuzzy grammar. In T. Åfarli and B. Maehlum (Eds.) *The sociolinguistics of grammar*. Amsterdam: John Benjamins. pp. 37–66.

Newmeyer, F. 2014c. Where do motivations compete? In B. MacWhinney, A. Malchukov and E. Moravcsik (Eds.) *Competing motivations in grammar and usage*. Oxford: Oxford University Press. pp. 299–314.

Oosthuizen, J. 2000. Prepositions left and right in Afrikaans. *Stellenbosch Papers in Linguistics* 33: 67–90.

Piattelli-Palmarini, M. 2000. The metric of open-mindedness. *Natural Language and Linguistic Theory* 18(4): 859–862.

Pinker, S. 1994. *The language instinct: The new science of language and mind*. London: Penguin.

Pinker, S. 2003. *The blank slate*. London, New York: Penguin.

Pinker, S. and R. Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95: 201–236.

Prinz, J. 2012. *Beyond human nature: How culture and experience shape our lives*. London, New York: Allen Lane.

Pullum, G. and B. Scholz. 2002. Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9–50.

Ramachandran, V. 2011. *The tell-tale brain*. London: William Heinemann.

Reali, F. and M. Christiansen 2005. Uncovering the richness of the stimulus: Structure dependence and indirect statistical evidence. *Cognitive Science* (29): 1007–1028.

Rizzolatti, G. and L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27: 169–192.

Roberts, I. and A. Holmberg. 2005. On the role of parameters in universal grammar: A reply to Newmeyer. In H. Broekhuis, N. Corber, R. Huibregts, U. Kleinhenz and J. Koster (Eds.) *Organizing grammar: a Festschrift for Henk van Riemsdijk*. Berlin: Mouton de Gruyter. pp. 538–553.

Roedder, E. and G. Harman. 2015. *Moral grammar*. Unpublished manuscript. Available online: http://www.princeton.edu/~harman/Papers/Moral_Grammar.pdf (Accessed 10 August 2015).

Ross, J. 1967. Constraints on Variables in Syntax. PhD thesis, Massachusetts Institute of Technology.

Sampson, G. 2001. *Empirical linguistics*. London: Continuum.

Scalise, S., E. Magni and A. Bisetto. 2008. *Universals of language today. Studies in natural language and linguistic theory*. Dordrecht: Springer.

Searle, J. 1969. *Speech acts*. Cambridge: Cambridge University Press.

Seuren, P. 2009. Concerning the roots of transformational generative grammar [review article]. *Historiographia Linguistica* 36: 97–115.

Spencer-Oatey, H. 2000. Rapport management: A framework for analysis. In H. Spencer-Oatey (Ed.) *Culturally speaking: Managing rapport through talk across cultures*. London: Continuum. pp. 11–46.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at work*. Amsterdam: John Benjamins.

Uriagereka, J. 1999. Multiple spell-out. In S. Epstein and N. Hornstein (Eds.) *Working Minimalism*. Cambridge MA: MIT Press. pp. 251–282.

Wodak, R. 2007. Pragmatics and critical discourse analysis: A cross-disciplinary enquiry. *Pragmatics and Cognition* 15(1): 203–225.

Zihl, J., D. Von Cramon, N. Mai and C. Schmid. 1991. Disturbance of movement vision after bilateral posterior brain damage. *Brain* 114(5): 2235–2252.