biolinguistics

volume 8 2014

ISSN 1450-3417

Syntactic Structures as Descriptions of Sensorimotor Processes

Alistair Knott

In this paper I propose a hypothesis linking elements of a model of theoretical syntax with neural mechanisms in the domain of sensorimotor processing. The syntactic framework I adopt to express this linking hypothesis is Chomsky's Minimalism: I propose that the language-independent Logical Form (LF) of a sentence reporting a concrete episode in the world can be interpreted as a detailed description of the sensorimotor processes involved in apprehending that episode. The hypothesis is motivated by a detailed study of one particular episode, in which an agent grasps a target object. There are striking similarities between the LF structure of transitive sentences describing this episode and the structure of the sensorimotor processes through which it is apprehended by an observer. The neural interpretation of Minimalist LF structure allows it to incorporate insights from empiricist accounts of syntax, relating to sentence processing and to the learning of syntactic constructions.

Keywords: embodied cognition; generative syntax; neural models of language; reaching-to-grasp; sensorimotor processing

1 Two Strategies for Investigating the Neural Correlates of Syntax

At least since Chomsky's *Aspects* (Chomsky, 1965), it has been traditional for linguists to think of syntactic models as describing something in the domain of psychology, rather than just as characterising sets of well-formed sentences. In practice, theoretical linguists rarely try to express the detailed claims of their models in terms which would be intelligible to neuroscientists or experimental psychologists. However, if a syntactic theory really does describe the cognitive system, then ultimately we might expect the theoretical devices it employs (representing syntactic structures or operations) to correspond to brain mechanisms which are identifable in their own right. For instance, if a theory uses a particular device to represent how phrases are formed, or how long-distance syntactic dependencies are sanctioned, we might expect to be able to find a mechanism in the brain which 'implements' the device, and ultimately explains why it has a role to play in the overall theory. The mechanism should be independently identifiable using the techniques available to neuroscience. Neuroscience is maturing rapidly; there are now many well-established methods for investigating brain structures and processes at several levels of hierarchy. So eventually we might expect to be able to

I am grateful to the Marsden Fund of New Zealand for support of this work. I would like to thank the two anonymous reviewers for their helpful comments.

ISSN 1450–3417

Biolinguistics 8: 001-052, 2014
http://www.biolinguistics.eu

express a hypothesis linking terminology or devices from syntactic theory to identifiable brain mechanisms.

Poeppel and Embick (2005) provide a useful discussion of this expectation. They argue that it is not a foregone conclusion: that some syntactic theories are expressed at a level of detail which is unlikely ever to correspond to measurable observations in neuroscience. In order to move towards a linking hypothesis, they suggest that syntactic theories should be formulated in ways that highlight the general *classes of computation* which must be carried out by the syntactic system—and to seek theories in which the proposed types of computation are plausibly implementable by neural machinery. It is not clear what syntactic formalisms Poeppel and Embick intend to rule out through this proposal, but it is certainly helpful for a theory to be explicit about the general types of computation which the syntactic mechanism must be capable of, and to focus the search for a linking hypothesis on a search for neural mechanisms which can implement computations with the right properties.

There is another strategy for seeking brain mechanisms underlying syntactic devices, which is in some ways complementary to that proposed by Poeppel and Embick: namely to focus the search for a linking hypothesis on the semantic domains which have been most intensively studied within neuroscience. Promising areas of neuroscience to consider are those which study perceptual and motor mechanisms (which I shall refer to jointly as 'sensorimotor' mechanisms). Models of sensorimotor mechanisms are among the most detailed in neuroscience, because it is through these mechanisms that the brain interfaces with the world, and they are therefore the easiest to study empirically. If we want to look for neural mechanisms underlying a particular syntactic structure, it may be helpful to begin by studying examples of the structure featuring concepts with obvious links to sensorimotor mechanisms, rather than concepts taken from some arbitrary semantic domain. For instance, if we are interested in the neural mechanisms underlying predication, it might be helpful to look first at 'concrete' sentences like *The cat is white* rather than arbitrary predications like *The idea is popular* or *The company is solvent*. We know quite a lot about the neural mechanisms involved in perceiving concrete objects and properties, so if we begin with *The cat is white*, we have a point to start from on the neuroscience side.

The idea of focussing on concrete language is central to a recent programme of research in psychology and neuroscience, investigating the claim that language somehow supervenes on, or recruits, sensorimotor mechanisms (see e.g. Rizzolatti and Arbib, 1998; Glenberg and Kaschak, 2002; Zwaan and Taylor; 2006; Pulvermüller, 2010; Meteyard *et al.*, 2012). These claims form part of a broader 'embodied' conception of cognition which is currently quite influential in cognitive science (Barsalou, 2008). Of course a hypothesis about the neural mechanisms underlying syntax must eventually extend beyond concrete sentences to all sentences. If we look first at concrete sentences, we must remember that we are looking in the sensorimotor domain for *instantiations* of general neural mechanisms, which we expect to also find in other more abstract domains. But it may still be helpful to start by studying concrete sensorimotor domains.

2 A Project Which Combines the Two Strategies

In this paper, I will summarise a research project which pursues both of the strategies outlined above. The aim of the project is to suggest how the devices of a formal syntactic theory can be linked to neural mechanisms. In line with the second strategy just discussed, the project focusses on sentences in a concrete domain, which express propositions that could conceivably be apprehended through well-studied sensorimotor processes. In line with the first strategy, the syntactic formalism used to analyse these sentences is one which emphasises simple computations, for which there is some prospect of a neural implementation.

The project focusses on a single example episode: one in which a man reaches for and grasps a cup. The sensorimotor mechanisms involved in experiencing this episode have been particularly closely studied, so it is possible to formulate a fairly detailed model of the processing that takes place when it is apprehended. On the syntactic side, the project focusses on the sentences which most directly report this example episode: transitive sentences such as *The man grabs the cup*, *He grabs it*, *L'homme prend la tasse*, *L'homme le prend* and so on.[1] My aim is to juxtapose a detailed model of the sensorimotor processes needed to apprehend the cup-grabbing episode against a model of the syntax of the example sentences, and look for formal similarities between the two models. If there are nontrivial similarities, this might provide some insight about how syntax reflects sensorimotor mechanisms, and ultimately, how syntax supervenes on neural mechanisms more generally.

The project is described at length in a recently published book (Knott, 2012). The book introduces both the sensorimotor and syntactic models 'from scratch', since there are few readers who are familiar with the details of both fields: as a result the key ideas are interspersed with a large amount of motivating background material. The purpose of the current paper is to present the key ideas in the book with a minimum of background material. I will begin in Section 3 by introducing the syntactic model, which is a version of Chomsky's Minimalism. In Section 4 I will introduce the overall form of the sensorimotor model I adopt, and in Section 5 I describe the form of the correspondence between syntax and sensorimotor processing which I envisage. In Sections 6 and 7 I introduce a detailed model of the sensorimotor processing through which a cup-grabbing episode is apprehended and stored. In Section 8 I make some proposals about how the syntax of the example sentences can be linked to the detailed sensorimotor model. In Section 9 I introduce a computational model of language processing and language acquisition which draws on these proposals. In Section 10 I consider how the proposals might extend beyond the example cup-grabbing sentences, and note some areas for further work.

3 A Syntactic Model of Transitive Cup-Grabbing Sentences

3.1 Choosing a Syntactic Formalism

To begin with, we need a syntactic formalism in which to describe the structure of the example cup-grabbing sentences. Of course, there are many formalisms to

[1] I will briefly discuss ergative languages and passive constructions in Section 10.

choose from. In the project I am describing, I use an early version of Chomsky's Minimalism (Chomsky, 1995b). In fact I will eventually adopt a heavily revised version of Minimalism which incorporates ideas from more empiricist syntactic frameworks, so I encourage non-Minimalist linguists not to stop reading here! However, I do want to adopt some of the key tenets of a standard Minimalist model. There are several reasons for this, which I will discuss in turn.

To begin with, Minimalism strongly emphasises simple general-purpose computations in the way advocated by Poeppel and Embick. The Minimalist programme aims to reduce to a minimum the amount of theoretical machinery required to generate the well-formed sentences in a language. Two simple operations, Merge and Move (or latterly Copy), do much of the work involved in generating a syntactic structure. I don't want to claim that Minimalism is the only formalism which adopts a small repertoire of basic computational operations. There are other formalisms which place an equal emphasis on computation, and posit an equally minimal repertoire of computations. For instance in categorial grammar, syntactic derivations are produced (largely) by two operations, 'functional composition' and 'type-raising'. And there is interesting research exploring the neural basis of these operations in non-linguistic domains; see for instance Steedman (2002). Nonetheless, Minimalism arguably meets Poeppel and Embick's criteria for linguistic formalisms: if we find neural mechanisms which plausibly implement Merge and Copy, we can make a substantive claim about how syntactic analyses refer to neural mechanisms.

A second reason for adopting Minimalism is that it allows strong claims about linguistic universals to be made. In Minimalism, there are two levels of syntactic analysis for a sentence: **phonetic form** (**PF**) represents its surface word order (among other things), and **logical form** (**LF**) represents the structure which the language processing mechanism delivers to a nonlinguistic semantic/conceptual system. The surface word order of sentences is obviously very different in different languages. One of the interesting claims in Minimalism is that there is a level of syntactic representation, namely LF, where many of these differences disappear, and where generalisations across languages are manifest. For instance, while sentences describing the cup-grabbing episode in different languages have very different PF structures, the claim in Minimalism is that they have the same structure, or at least very similar structures, at LF. Among current syntactic theories, this claim is unique to Minimalism. It is interesting, because it allows for particularly strong statements about the neural basis of syntactic structures: we can claim that an LF structure describes some neural process or representation which directly interfaces with language, and which is the same for all speakers. For instance, consider our project of studying the relation between syntax and sensorimotor processing in the cup-grabbing scenario. It is presumably uncontroversial that people the world over use the same sensorimotor mechanisms to apprehend a cup-grabbing episode regardless of the language they speak. Within Minimalism, we can frame a very strong hypothesis: that the LF of a cup-grabbing sentence (in any language) directly describes or recapitulates the sensorimotor processes involved in experiencing the episode it reports. It is only within a framework like Minimalism that we can express the idea that syntactic representations directly encode sensorimotor processes. Several theorists developing 'embodied' models of language have argued that entertaining the meaning of a concrete sentence involves a process of active sensorimotor simulation (see e.g. Gallese and Goldman, 1998; Jeannerod, 2001; Grèzes and Decety, 2001; Barsalou *et al.*, 2003; Feldman and Narayanan, 2004). In this paper I want to provide a syntactic framework for expressing this idea. But for the moment, my point is just that the Minimalist device of LF allows particularly strong, optimistic statements connecting syntactic structures and language-independent neural processes.

Some further reasons for adopting a Minimalist framework relate to the actual structure of LF as proposed within Minimalist models. I will briefly introduce this structure in the next section, and then describe why it has some appeal as the basis for an account of the neural mechanisms underlying syntax.

3.2 The LF Structure of a Transitive Cup-Grabbing Sentence

Minimalist LF representations have a strongly recursive structure. The primitive structural unit in LF representations is the **X-bar schema**, which is illustrated in Figure 1. X-bar theory (Chomsky, 1970; Jackendoff, 1977) long predates Minimalism,

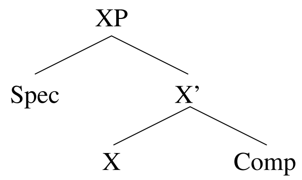


Figure 1: The X-bar schema

and forms an important part of many non-Chomskyan models of syntax, in particular HPSG (Pollard and Sag, 1994). The theory was originally introduced as a refinement of the general hypothesis that syntactic structure is lexical in origin, i.e. that the lexical items in a sentence contribute or 'project' their own elements of local syntactic structure. X-bar theory states this idea parsimoniously, positing that lexical items of all grammatical classes project a phrase with the same basic form, namely the XP structure shown in Figure 1. XP contains a position for the lexical item projecting the structure, called the head (X), and also two positions where semantic material required by the head can appear, the specifier (Spec) and the complement (Comp). These latter two positions can be recursively filled by other XPs.

In Chomskyan models, X-bar syntax has been extended in several directions. In current Minimalism, the X-bar schema is derived from applications of a still more basic structural operation, Merge, already mentioned in Section 3.1. For most of this paper I will retain the terminology of X-bar theory, though I will briefly discuss Merge in Section 10.1. The extension of X-bar theory I will focus on is the notion of **functional projections**, that features heavily in early versions of Minimalism, and is retained in modern Minimalism. Functional projections are XPs headed by non-lexical elements. In Minimalist analyses, the LF of a clause typically contains many functional projections as well as regular lexical projections, resulting in structures containing many more XPs than in other formalisms. These XPs attach to one another predominantly through adjunction to complement positions, creating what could be termed 'right-branching' LF structures, or more accurately, 'complement-branching' structures (since LF encodes hierarchical semantic structure rather than left-to-right linear order). These complement-branching structures, in conjunction with principles allowing movement of syntactic elements from one position to another, support a distinctively Chomskyan style of syntactic model.

The basic LF structure of the transitive sentence *The man grabs a cup* in the Minimalist model of Chomsky (1995b) is shown in Figure 2. As can be seen, the LF

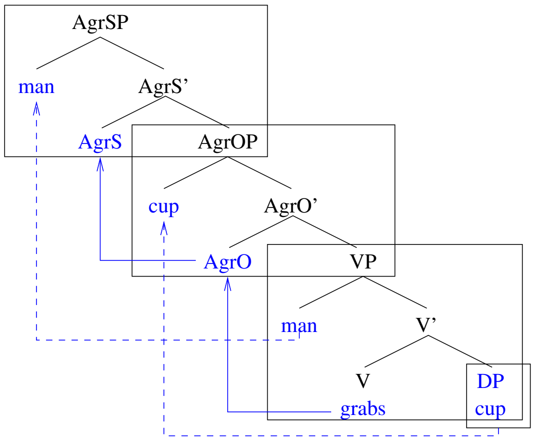


Figure 2: Basic LF structure of *The man grabs a cup*, in the Minimalist model of Chomsky (1995b). Each XP (including the complement of VP) is shown in a box. Head-raising operations are shown with solid lines; DP-raising operations are shown with dashed lines.

structure of a clause is strongly complement-branching. I have omitted the internal structure of noun phrases (which I will call 'determiner phrases' or 'DPs', following Abney, 1987), which introduce some measure of left-branching structure when they appear in specifier positions. But within a DP, Minimalist analyses again envisage a largely complement-branching structure of XPs (see e.g. Abney, 1987; Zamparelli, 2000; Alexiadou *et al.*, 2007). The idea that the LF structure of clauses and DPs is predominantly complement-branching is characteristic of Minimalism.

The most distinctive, and controversial, elements in the Minimalist model of a transitive clause are the functional projections, which are headed by non-lexical material. In Figure 2 there are two functional projections, AgrSP and AgrOP. Functional projections were actually introduced in the precursor to Minimalism, GB (Chomsky, 1981). A novel idea in GB was that XPs can be headed by morphological inflections as well as by whole words: for instance agreement inflections on verbs (and later, tense inflections) were assumed to introduce their own XPs, occupying specific positions in a clause. In Minimalism the idea of functional projections associated with inflections is retained, though these XPs are now headed by the semantic features signalled by inflections rather than by inflections themselves. AgrSP and AgrOP are headed by the main verb's 'subject agreement features' and 'object agreement features' respectively (in this case these are both third person singular). In fact in current Minimalism the LF of a clause standardly includes several further projections that do not feature in Figure 2, CP, TP, vP and several others: for most of this paper I will assume the simplified structure in Figure 2, but I will discuss some of these additional projections in Section 10.2.

A final adaptation of X-bar theory in Minimalism is that X-bar schemas impose a fixed temporal order on the surface form of sentences: the specifier of an XP appears to the left of its complement at PF (Kayne, 1994). For this reason I will continue to use the term 'right-branching' to describe the predominant complement-recursive structure in LF.

The LF structure shown in Figure 2 is generated through a process called **derivation**, which involves the step-by-step merging of XPs, interspersed with various movement operations, latterly reconstrued as Copy operations.[2] The key movement operations are shown in colour in the figure. During derivation, the inflected verb initially appears at the head of VP, but must raise to the heads of AgrOP and AgrSP in turn to 'check' its inflections. The verb's arguments (a subject and an object) initially appear at the specifier and complement of V, where they are assigned thematic roles (AGENT and PATIENT), but they must raise to the specifiers of AgrOP and AgrSP respectively, to be assigned 'Case' by the heads of these projections. At the end of the derivation, we have an LF structure holding multiple copies of the inflected verb, subject and object at different positions.

For Minimalists, the requirements that inflected verbs 'check' their features at higher heads, and that DPs are assigned 'Case' at higher specifier positions, are simply formal stipulations, which are mainly justified by the roles they play in the expression of a larger theory of syntax which economically accounts for a sizeable body of linguistic data. My aim in the current paper is to consider what these devices might correspond to in a model of neural mechanisms. But I will briefly outline the theoretical roles they play in the Minimalist model.

The LF structure produced by derivation illustrated in Figure 2 forms the basis for two accounts. One is an account of syntactic variations between languages—in particular of differences in the canonical ordering of constituents (e.g. subject, verb and object). The other is an account of nonlocal syntactic dependencies within a clause within any given language. The account of cross-linguistic differences turns on the idea that PF structures are formed by 'reading out' LF structures, with the constraint that only one copy of each moved constituent is pronounced. Differences in canonical constituent order between languages are modelled as differences in the conventions governing this read-out process: thus speakers of English (which has subject-verb-object or SVO order) learn to pronounce the subject 'high', and the verb and object 'low', while speakers of Māori (which has VSO order) learn to pronounce the verb 'high'. The Minimalist account of nonlocal dependencies explains agreement relationships between apparently distant syntactic items by appealing to unpronounced material at LF. For instance, in English, there is agreement between subjects and verbs, even though subjects appear outside the local domain of the verb at PF. In Minimalism agreement is explained by referring to the structure of LF, in which the verb is present at the head of AgrSP as well as at the head of VP.

---

[2] I will continue to refer to them as movement operations, though I will certainly adopt the copy theory of movement.

3.3 Prospects for a Sensorimotor Interpretation of Minimalist LF Structures

To summarise the previous section: Minimalist LF structures are formed from instances of the X-bar schema (including many functional projections), joined together into predominantly right-branching structures. They contain two kinds of nonlocality: one resulting from movement (copying) of DPs to Case-assigning positions, and one resulting from successive movement (copying) of Vs to feature-checking positions.

Both these properties of LF structures could potentially have interesting parts to play in an account of the relation between sensorimotor processing and syntactic structure. Firstly, LF structures are created from simple building blocks (XPs) combined in a simple general way. There are some sensorimotor models which argue that sensorimotor processing is likewise decomposed into simple building blocks, recursively combined to create right-branching structures. I will discuss one of these in Section 4. If any relationship can be found between the building blocks of LF and the building blocks of sensorimotor processing, this would be a very interesting discovery. Secondly, there is interesting evidence for multiple occurrences of certain representations during sensorimotor processing (see Section 6.3 for details). For instance, in the process of apprehending an episode in which a man grabs a cup, multiple distinct sensorimotor representations of both the agent (the man) and the patient (the cup) are activated. Perhaps these provide some basis for the Minimalist device of movements occurring during derivation of an LF structure. I will pursue this idea in Section 8.

4 A Proposal about the General Form of Sensorimotor Processing

An interesting proposal about the 'building blocks' of sensorimotor processing was made by Dana Ballard and colleagues (Ballard *et al.*, 1997), in the context of a general model of embodied cognition. In this section I will introduce Ballard *et al.*'s model.

4.1 Deictic Representations

Ballard *et al.* argue that at a certain timescale, which they term the 'embodiment level', cognitive processing is intimately connected to sensorimotor routines interfacing with the physical world. At this timescale, they suggest, cognitive processing engages in a special way with the moment-to-moment deployment of sensory and motor apparatus. The illustrations they provide mainly relate to the role of saccadic eye movements in cognition. Human agents make around 3 saccades per second throughout their waking lives. Each saccade an agent executes results in a very transitory fixation on some visual stimulus, and in a similarly transitory pattern of activity in the agent's visually-derived neural pathways. For instance, the activity in the 'ventral' visual pathway leading to inferotemporal cortex changes dramatically after every saccade (Freedman *et al.*, 2003); so does at least some of the activity in the 'dorsal' visual pathway through posterior parietal cortex (Colby and Goldberg, 1999). Ballard *et al.* call the cognitive representations which reflect the momentary deployment of an agent's sensorimotor apparatus **deictic representations**.

Ballard *et al.* argue that deictic representations play an important role in structuring cognitive processing at the embodiment level. For instance, deictic representations reflecting the current position of the eye can help plan the next saccade: when this is executed, it will generate a new set of deictic representations, so deictic representations and saccades are naturally organised into alternating sequences. To take another example, deictic representations play an important role in the organisation of motor actions. When we want to act on a target object, we typically make a saccade to the object first (Land and Furneaux, 1997; Johansson *et al.*, 2001). This creates deictic representations of the object in visual cortex and in parietal and premotor cortex (Geyer *et al.*, 2000, Murata *et al.*, 2000). Representations in the latter areas encode the object's motor affordances, and are used to select and eventually control a motor action (see e.g. Cisek and Kalaska, 2010). Note that motor actions are not only initiated by deictic representations—they also bring about new deictic representations: for instance, when an agent touches a target object, the tactile sensors of his hand are 'momentarily deployed' to the cup (see e.g. Goodwin and Wheat, 2004), in much the same way that saccades momentarily deploy the fovea to a particular point in the visual field.

4.2 *Deictic Operations*

Ballard *et al.*'s model is interesting because it identifies commonalities between attentional actions (e.g. saccades) and motor actions (e.g. reach movements). Actions of both types are modelled as **deictic operations**: cognitive operations which bring about updates in the agent's physical relationship with the environment, and also in his internal cognitive representations. When talking about deictic operations in general terms, Ballard *et al.* make use of a notion of **context**, which includes information about the agent's cognitive representations at any given time and also about the physical state of the agent and his environment at that time. To formalise their definition a little: a deictic operation is executed in an **initial context**, and results in the establishment of a **new context**; it also generates a **reafferent sensory signal** as a side-effect. The reafferent signal is a deictic representation which provides sensory feedback that the operation actually occurred. For instance, if the operation involves attending to a particular object in the environment, the natural reafferent signal confirming the operation has taken place is a sensory representation of the object. The general form of a deictic operation is shown in Figure 3.

Initial context	Deictic operation	Reafferent signal	New context
C_1	O_1	S_1	C_2

Figure 3: General form of a deictic operation

4.3 *Deictic Routines*

Ballard *et al.* propose that cognitive processing at the embodiment level is organised into sequences of deictic operations called **deictic routines**. A simple deictic routine associated with reaching-to-grasp is shown in Figure 4. This particular routine involves two **cycles**; each row of the table describes one of these. This example

Initial context	Deictic operation	Reafferent signal	New context
C_1	*attend_to_cup*	*cup*	C_2
C_2	*grab*		C_3

Figure 4: A simple deictic routine with two cycles

illustrates the recursive nature of deictic routines: the new context created by the first operation *attend_to_cup* (C_2) is also the initial context for the second operation *grab*.

Note that we need to talk about recursion rather than simple iteration. A point which Ballard *et al.* repeatedly stress is that deictic operations and representations typically cannot be interpreted by themselves: they tend to 'implicitly refer' to the operations and representations which preceded them. For instance, the deictic operation *grab* does not specify a target: the target is specified implicitly by the deictic representation in place at the time when *grab* is executed. According to this model, an agent selects the target of a motor action simply by attending to it, and then activating the general motor programme 'grab'. This motor programme has various free parameters: the location to reach for, the hand shape to form, and so on. The prior action of attention activates deictic representations in motor pathways which fix the values of these parameters. Thinking of the 'grab' programme as a deictic representation accurately reflects the neural mechanisms involved in actually controlling an action. To take another example, the deictic representation *cup* only provides partial information about a cup: it specifies its visual properties, but not its location in the world, since it is always centred on the retina, at the fovea. To establish where the fovea is currently directed, we must refer back to the deictic operation which positioned the fovea, namely *attend_to_cup*.

In the general case, to recover the meaning of a deictic representation, we must make reference not just to the immediately preceding deictic operation, but to some sequence of preceding deictic operations. For instance, in order to interpret the deictic representations at the end of a grab action, when the agent's hand is touching the cup (at context C_3 in Figure 4) we must refer back to the *grab* operation which directly led to these sensations, but interpreting the *grab* operation in turn requires reference back to the preceding *attend_to_cup* operation, as discussed above. The interpretation of deictic operations and representations is right-recursive: it requires recursive reference to a preceding deictic routine.

In Section 6 I will outline a detailed model of reaching-to-grasp which assumes the framework of deictic routines just sketched. But first I will introduce the basic proposal I want to make about the relation between syntax and sensorimotor processing, which is also expressed using the terminology of deictic routines.

5 A Sensorimotor Interpretation of LF Structures and of the X-Bar Schema

As outlined in Section 3, Chomsky's Minimalist model posits that sentences have an LF structure as well as a surface phonetic form: LF structure is relatively invariant over translations, and is composed of instances of the X-bar schema, connected in a largely right-branching way. As outlined in Section 4, Ballard *et al.* propose that sensorimotor processes are organised into deictic routines, which are composed of deictic operations in a similarly right-branching way. Having introduced a model of syntax and a model of sensorimotor processing, I will now formulate some proposals which link these two models. Recall that my proposals are only about the syntax of 'concrete sentences', which describe episodes that can be directly apprehended through sensorimotor processing.

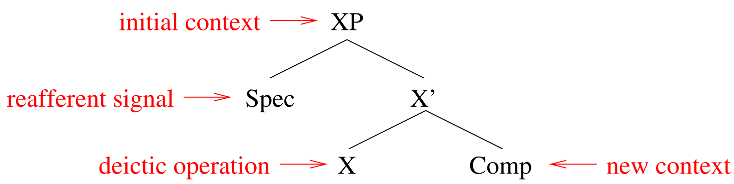
I will first make a general proposal linking syntactic structures to deictic routines.

> **Proposal 1** The LF of a concrete sentence reporting an episode E provides a description of the deictic routine through which E is experienced.

One way of thinking about Propoal 1 is as an expression of the kind of simulationist account of sentence meaning advocated by theorists like Gallese and Goldman (1998) and Feldman and Narayanan (2004). In Minimalism, the LF of a sentence represents its meaning—or at least, as much of its meaning as can be encoded syntactically. The novel thing about the proposal is that it expresses a simulationist account of meaning in a way which links recursively structured syntactic representations to recursively structured sensorimotor routines. This opens the way for a much stronger statement of the proposal, which links the basic *building blocks* of LF structures to the basic *building blocks* of sensorimotor routines. Ultimately, what I want to propose is a general sensorimotor characterisation of the X-bar schema from which LF structures are formed, as follows:

> **Proposal 2** Each X-bar schema in the LF of a concrete sentence describes a single cycle in the deictic routine described by the LF structure.

More specifically, I want to suggest that each constituent within the X-bar schema has a well-defined sensorimotor interpretation.

> **Proposal 3** The components of an X-bar schema describe a cycle of a deictic routine, as follows:
> - the maximal projection (XP) describes the initial context in which a deictic operation occurs;
> - the head (X) describes the operation itself;
> - the specifier describes the reafferent signal of the operation;
> - the complement describes the new context which the operation brings about. As shown in this figure:
>
> 
>

Proposal 2 and its extension Proposal 3 express very strong claims, as they link the basic structural element from which syntactic structures are formed to the basic element from which sensorimotor processes are formed. (Recall we are still restricting our attention to concrete sentences.) One of the attractive aspects of Minimalism is that it envisages simple building blocks for syntactic structures. One of the attractions of Ballard *et al.*'s model is that it envisages simple building blocks for sensorimotor process (at least at the embodiment timescale). I want to suggest that there is a relationship between the building blocks in each domain.

An important idea implicit in the above proposals is that the right-branching organisation of LF structures mirrors the right-branching structure of deictic routines. A corollary of Proposal 3 is that a right-branching structure of XPs describes a sequence of deictic operations, i.e. a deictic routine. This is illustrated in Figure 5, which shows a right-branching structure of two X-bar schemas, XP and YP. The

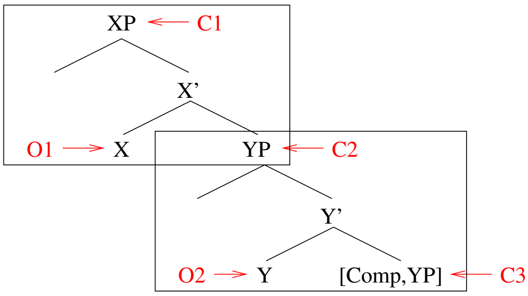


Figure 5: Sensorimotor interpretation of a right-branching structure of two X-bar schemas

red labels show some of the sensorimotor interpretations which are sanctioned by Proposal 3. XP describes the initial context ($C1$) for some deictic operation ($O1$). The complement of XP describes the new context ($C2$) brought about by this operation. Since YP appears at this position, it also describes $C2$, which is therefore also the initial context for a second deictic operation ($O2$), bringing about a third context ($C3$). Thus by Proposal 3, a right-branching strucure of X-bar schemas describes a deictic routine. (Note that if LF structures are in general right-branching, Proposal 1, that LF structures describe deictic routines, is also a corollary of Proposal 3.)

Because Proposal 3 is expressed in terms of the building blocks of LF structures, it makes very strong and specific predictions about the relationship between syntax and sensorimotor processing for *any* concrete sentence. The proposal predicts that we can take any sentence describing a concrete episode, and find a mapping of the right kind between the LF of this sentence and the structure of the sensorimotor processing through which the described episode is apprehended. Given that the accounts of LF structure and of sensorimotor processing are derived from completely separate data (well-formedness judgements vs data about neural processes), using completely different methodologies (syntactic argumentation vs experimental neuroscience), finding a mapping of the predicted kind would provide good empirical support for the generalisation expressed in Proposal 3.

Of course, there is more to LF than the right-branching structure illustrated in Figure 5. As already discussed in Section 3.2, the LF of a transitive sentence is regularly right-branching, but also contains various types of re-entrancy, reflecting movement operations: DP raising to Case-assigning positions and head-raising for semantic feature checking (see the illustration in Figure 2). If the LF structure of a concrete sentence really does describe sensorimotor processing, we should make several further predictions, namely that these movement operations can also be given sensible sensorimtotor interpretations. In the next three sections I will examine these predictions. In Sections 6 and 7 I will introduce a model of the sensorimotor processing required to apprehend a cup-grabbing episode and to store it in working memory. In Section 8 I will examine the relationship between this model and the LF model introduced in Section 3.2, considering both the overall right-branching structure of LF and the re-entrant structures associated with DP movement and head raising.

6 Sensorimotor Processing Involved in Apprehending a Cup-Grabbing Episode

In this section I will argue that the sensorimotor processing through which a cup-grabbing episode is apprehended takes the form of a deictic routine with three cycles, with the structure shown in Figure 6. The argument is given in much more detail in Knott (2012); I will just summarise it here.

Initial context	Deictic operation	Reafferent signal	New context
C_1	*attend_man*	*man*	C_2
C_2	*attend_cup*	*cup*	C_3
C_3	*grab*	*man*	C_4 / *cup*

Figure 6: Deictic routine involved in experiencing a cup-grabbing episode

6.1 Cycle 1: Attention to the Agent

The deictic routine shown in Figure 6 is somewhat more complex than the short routine used for illustrative purposes in Section 4.3. The main reason for this is that it provides an account of the apprehension of a whole *episode*, rather than just of the execution of a motor action. A cup-grabbing episode has an agent, as well as a motor action and a target. I will begin by arguing that whenever an observer apprehends an episode involving an agent through sensorimotor means, the agent must be attended to first.

As a starting point, note that the agent of the grab action—the man—could be someone external to the observer, but could also be the observer himself. In a sentence reporting the observed episode, this difference is reflected in the sentence's subject, and in the verb's subject agreement features (*I grab* vs *He grabs*). Interestingly, apart from these differences, the LF of the sentence is the same in these two scenarios. Proposals 1–3 make a strong prediction from this observation: that the sensorimotor processes through which an agent apprehends his own actions are largely the same as those through which he apprehends the actions of others.

There is certainly support for this general idea in neuroscience: it forms the basis for the well-known 'mirror system hypothesis' (see e.g. Gallese *et al.*, 1996). The hypothesis originated with the discovery of neurons in area F5 of the premotor cortex of macaque monkeys which respond to specific grasp movements whether they are initiated by the monkey itself or by an observed third party (di Pellegrino *et al.*, 1992). It has since been corroborated in many other studies finding common representations for observed and executed actions (see Fadiga *et al.*, 2005 for a summary). The hypothesis has been put to work in several ways in neural models of language, in particular in accounts of language evolution trading on the suggestion that area F5 is the macaque homologue of Broca's area (see e.g. Rizzolatti and Arbib, 1998; Arbib, 2005). But the main use I want to make of the hypothesis is in an account of the very first sensorimotor operation an observer must perform when apprehending an episode.

If we accept that agents use assemblies in premotor cortex to represent the actions of others as well as their own actions, this raises the question of how an agent is able to distinguish between his own actions and those of other people. There are several models of how this distinction is made. My model derives mainly from accounts by Haggard (2008) and Farrer and Frith (2002). The basic idea is that the neural circuitry activating action representations in premotor cortex must be *configured* for either action execution or action observation before representations in this area can be interpreted. There is good general evidence that the brain can switch between alternative general modes of connectivity, implemented by distinct large-scale neural networks (see e.g. Bressler and Menon, 2010). I argue that action execution and action observation require the establishment of different modes of connectivity within the mirror system. In 'execution mode', premotor representations are activated through a well-studied sensorimotor pathway which maps visual representations onto motor affordances (as discussed in Section 4), and in turn they generate overt motor actions. In 'perception mode', premotor representations are activated through a completely different neural pathway through the superior temporal sulcus, specialised for analysing biological movements (see e.g. Jellema *et al.*, 2000), and a mechanism is in place to prevent these premotor representations resulting in overt actions, so that the observer does not reflexively imitate the actions he sees.

The above account can be neatly expressed using the terminology of deictic routines. Action representations in premotor cortex are a variety of deictic representation: they cannot be interpreted in isolation, because in isolation we do not know whether to attribute them to ourselves or to some external agent. In order to interpret them, we must make reference to a prior mode-setting operation. The mode-setting operation can therefore be modelled as a deictic operation.

It is also useful to think of the deictic operation establishing action execution or action perception mode as an action of attention, to the *agent* of the action which will presently be represented in premotor cortex. The operation establishing execution mode is like an action of attention to onesself, since it is this operation which allows an agent to attribute the action subsequently evoked in premotor cortex to himself. There is good evidence that part of an agent's 'concept of self' is implemented in the neural mechanisms which initiate volitional actions; see e.g. Haggard, 2008; Chambon *et al.*, 2012. And there are good grounds for allowing for the possibility of actions of attention to onesself; see e.g. Damasio, 1999 and Critchley *et al.* (2003) for accounts of a 'pre-attentional' representation of the self from which higher-level representations of self can be selectively activated. The operation establishing perception mode is also plausibly attentional in origin. In the most obvious scenario it is triggered by attention to an external agent in the world, who presents a salient enough stimulus to the observing agent to cause him to engage his action perception circuitry. In this case, the deictic operation required to interpret activity in premotor cortex involves a regular action of visual attention *as well as* a mode-setting operation. The reafferent side-effect of this attentional action is a representation of the particular external agent who is attended to.

To summarise: when an observer apprehends an episode in which an agent executes an action, he must attend to the agent (generating a representation of the agent as a reafferent side-effect) before activating a representation of the action in premotor cortex. The agent attended to could be the observer himself or it could be an external agent, but in either case we can speak of an action of attention to this agent, resulting in activation of a representation of this agent.

6.2 Cycle 2: Attention to the Target

If the observed action involves reaching for a target object, the observer must also attend to this object before activating a representation of the action. As already noted in Section 4, an agent executing a reach action typically fixates the target object very early in the reach, probably before a detailed grasp motor programme has been activated (Land and Furneaux, 1997; Johansson *et al.*, 2001). So if the observer of the action is the agent, there is quite good evidence that the target must be attended to before a motor programme is activated. Interestingly if the observer is watching an external agent perform a grasp action, there is also evidence that attention to the target precedes activation of a motor representation. Observers watching an external agent reach for a target typically saccade to the target well before the agent's hand reaches it (Flanagan and Johansson, 2003; Webb *et al.*, 2010). They are able to infer the observed agent's intention, and then put themselves into the same attentional state as the agent while the action is still under way. This establishment of joint attention plays a crucial role in computational models of learning in the mirror system, because it allows agents to use visual representations of their own actions to learn how to recognise the actions of external agents (see e.g. Oztop and Arbib, 2002; Oztop *et al.*, 2005).

To summarise: in a deictic model of the perception of a reach-to-grasp action, two deictic operations must occur before the grasp action itself can be activated. First, the agent of the action must be attended to. This operation is either an action of attention to onesself, if one is executing the action, or an action of attention to an external agent. Second, the target of the action must be attended to. If one is executing the action, this operation is needed to fix the parameters of the motor action to be executed. If one is watching an external action, the operation is executed in order to establish the same attentional state as the observed agent. In either case, it is only once the agent and patient have both been attended to that a single motor programme can be selected.

6.3 Cycle 3: Action Monitoring and Completion

When the observer attends to the intended target object, a set of alternative motor programmes is activated in premotor cortex (the object's 'motor affordances') and these compete amongst one another until a winning programme is selected, either as the programme to be executed (see e.g. Fagg and Arbib, 1998) or as the winning hypothesis about the type of action being performed (Oztop and Arbib, 2002).

Once a motor programme has been selected in this fashion, the character of sensorimotor processing changes. In the case of action execution, a physical action is initiated, which is shaped in real time by the representations currently active in the agent's sensorimotor pathways (Cisek and Kalaska, 2010). The agent's neural pathways, motor effectors and physical environment together implement a dynamical system, which (if all goes well) moves towards an attractor state where the agent is holding the target. In the case of action perception, processing also has the character of a dynamical system, this time a simulated one (see again Oztop and Arbib, 2002; Oztop *et al.*, 2005).

Again it is useful to cast this processing within the framework of deictic routines. Activating the 'grab' representation in premotor cortex is an operation which initiates a dynamical system. This operation has its own reafferent sensory consequences, and also eventually results in a new context in which new deictic operations are possible. I will consider these two aspects of the operation in turn.

What is the reafferent sensory signal associated with the execution of or observation of an action? When we are in the process of executing an action, we are aware of our body moving: there is a characteristic 'match' betwen the pattern of outgoing motor signals and the pattern of incoming sensory signals, because the sensory signals are produced by our motor movements. This match signal appears to be involved in producing the sense of agency that we feel when we perform an action, through which we are able to attribute the action to ourselves (Farrer and Frith, 2002; Farrer *et al.*, 2008). Of course, our conception of self is highly multimodal. Different elements of this construct are activated at different stages of action execution. We have already seen that one component of an agent's conception of self is activated at an early stage during action preparation (Haggard, 2008, Chambon *et al.*, 2012). Another component is activated when an action as actually under way (Farrer *et al.*, 2008). (Chambon *et al.* refer to these two conceptions as 'propspective' and 'retrospective' concepts of self.) The circuits which control action execution must include some mechanism for binding together these different facets of our concept of self.

A very similar point can be made about action perception. As argued in Section 6.1, an observer initially attends to an external agent as a salient object in the world. But when the observer begins to monitor that agent's actions, the agent is represented differently, as a characteristic pattern of movement, rather than as a static object. (See Ramsey and Hamilton, 2010 for evidence that action observation involves dissociable representations of the observed agent as a token individual and as an animate entity.) Again, our conception of an observed agent is a multimodal construct, and different facets of the construct seem to be accessed at different points during action perception: the action perception mechanism must include a mechanism which integrates these different representations of observed agents.

In sum, I argue that an observer attends to the agent at two distinct points in the course of apprehending a reach-to-grasp action— once at the very start, when action perception or action execution mode is established, and once during the process of actively monitoring the action to completion—and that there must be a mechanism for binding together the different representations of the agent obtained through these two different actions of attention.

I now consider the new context brought about by the grasp motor action: namely the state in which the agent is holding the target object. As already noted in Section 4, this state is well modelled within the framework of deictic routines, because it has both physical and attentional components. Successfully executing a grasp action achieves a substantive physical change in the world: the agent's hand moves to achieve stable contact with the target. But it also deploys the agent's tactile senses to the target object, providing him with information about the target in a new modality. Just as there are two actions of attention to the *agent* in the course of apprehending a reach-to-grasp action, there are quite clearly two actions of attention to the *target*, again in different modalities, and again occurring at different times. The target is first attended to in the visual modality, as part of the process of preparing the grasp action, and it is later attended to in the haptic modality, when the action is complete. Again, the structure of the sensorimotor routine plays a critical in the formation of a cross-modal representations of the target object. At the point when a stable grasp is achieved on the object, the visual representation of the object axiomatically corresponds to one of its motor affordances (see e.g. Oztop *et al.*, 2004).

At this point, I have motivated the arguments for the deictic routine illustrated in Figure 6. In the initial context C_1, the observer attends to the agent (who is himself or an external agent), and as a reafferent consequence receives a representation of the agent (*man*) and establishes a new context C_2. In this context, objects in the agent's perispace compete for attention, and the observer attends to a cup, activating a representation of the cup as a reafferent consequence, and establishes a third context C_3. In this context, the agent activates the motor programme *grab*, and an action is dynamically monitored. While the motor programme is under way, a dynamic representation of the agent is activated as a reafferent consequence, and when the action is complete, we enter a final context C_4, in which a haptic representation of the cup is active.

A final thought: it might perhaps be thought that modelling the experience of a reach-to-grasp action as a deictic routine as illustrated in Figure 6 somehow understates the complexity of the action-monitoring process, which is construed as a single stage in the routine, with the same basic form as a simple attentional action. I certainly acknowledge that action monitoring, whether it involves execution or perception of an ongoing action, has greater complexity than an attentional action. Monitoring a reach-to-grasp action takes far longer than attending to an object, and it is well known that the process involves two largely separate neural pathways, one for reaching and one for grasping: see classically Jeannerod (1996) (and for a review see Knott, 2012). But my suggestion is that this type of compositional structure is not *visible to language*. My proposal is that the syntax of language engages with the discrete, *temporal* structure of a sensorimotor routine. And at this level, the process of monitoring an action may in fact be quite simple. A motor programme controlling the dynamics of the combined hand/arm system (whether actual or simulated) can bring about complex changes in this system without having complex dynamics itself; this is an important fact about dynamical systems. If language interfaces with representations that *control* the dynamics of movements, rather than with dynamical movements themselves, the additional complexity of action monitoring may not be visible in the sensorimotor signals that interface with language.

7 Working Memory Representations of Sensorimotor Routines

One outstanding question from Section 5 concerns what's meant by the assertion that an LF representation 'describes' a sensorimotor routine. Clearly I do not want to propose that sentences *directly* report sensorimotor processes, as these arise in real time. We can execute sensorimotor routines without engaging language, and we can produce concrete sentences which are unrelated to our current sensorimotor environment. The standard assumption in psycholinguistics is that sentences produced by a speaker reflect *representations* of events and states, held in that speaker's working memory (Levelt, 1989). I certainly want to adhere to that idea. I therefore need to provide an account of how the cup-grabbing episode is represented in working memory. Ideally, this should be framed within a more general account of episode representations in working memory. Ultimately, I need to re-express the proposals about LF made in Section 5 in a way which connects to an account of working memory episode representations.

In Sections 7.1 and 7.2 I make a suggestion about how concrete episodes are stored in working memory. In Sections 7.3 and 7.4 I discuss how this mechanism interfaces with language, and in the light of this, give a more precise interpretation of LF.

7.1 A Model of Episode Representations in Working Memory

My account of working memory is based on Alan Baddeley's recent model of working memory for episodes (Baddeley, 2000). Baddeley suggests that there is a working memory medium called the **episodic buffer**, which holds semantic representations of episodes, and which interfaces with the well-known **phonological buffer** (Baddeley and Hitch, 1974) in a way which supports language processing.

Baddeley's main argument for the episodic buffer hinges on the fact that experiencing a concrete episode often takes a significant amount of time. (Apprehending an episode involves monitoring it as it occurs; this may take several seconds, often longer.) In order to store the episode in long-term memory, it must be encoded in the hippocampus, as a preliminary to being consolidated in cortical long-term memory (McClelland *et al.*, 1995). But storing associations between representations in the hippocampus can only be achieved through the mechanism of long-term potentiation (LTP), which requires them to be active within about 100ms of each other (Abraham *et al.*, 2002). Baddeley concludes that experienced episodes must be buffered in a working memory medium, and then replayed to the hippocampus at a speed which allows them to be associated through LTP. Additional evidence for the existence of short-term memory representations of observed episodes is reviewed in Swallow *et al.* (2009).

Baddeley does not speculate much about the format in which episodes are encoded in the episodic buffer, beyond requiring that it supports them being transmitted to the hippocampus. There are many models of how episodes are stored in working memory (see e.g. Shastri, 2001, 2002; Chang *et al.*, 2002; Plate, 2003; van der Velde and de Kamps, 2006). I make a new suggestion, which is based on the assumption that concrete episodes are experienced as deictic routines. My suggestion is that a concrete episode like the cup-grabbing episode is stored in the episodic buffer as a *planned sequence of sensorimotor operations*, i.e. a planned deictic routine (see Takac and Knott, 2013 for an implemented model). This view of working memory episode representations is interesting for several reasons. Firstly, it offers a new solution to the question of how thematic roles are bound to participants in episode representations. Any neural model of episode representation must have a way of identifying the roles played by the different participants in an action (in our case, agent and patient). The deictic routine through which an episode is experienced distinguishes these roles clearly, because they are associated with different serial positions in the routine. Secondly, the neural mechanisms which support the preparation of sensorimotor sequences have been intensively studied, and we know quite a lot about them. If these prepared sequences are examples in a 'concrete' domain of working memory episode representations, then studying their properties may help us formulate a more general model of these representations which extends beyond concrete episodes. Thirdly, viewing working memory episode representations as prepared sequences fits well with their role in Baddeley's model of replay to the hippocampus. The hippocampus is often seen as specialised for storing sequentially structured information (Wallenstein, 1998; Eichenbaum *et al.*, 2004), and is known to support fast replay of sequences (see e.g. Foster and Wilson, 2006; Diba and Buzsáki, 2007). If working memory episode representations are prepared sensorimotor sequences, they can be communicated to the hippocampus by being replayed, in simulation, at high speed, in a mode where each of them activates an assembly in the hippocampus. Finally, thinking of working memory episode representations as supporting simulation accords well with the simulationist accounts of propositional meaning already mentioned (Gallese and Goldman, 1998; Feldman and Narayanan, 2004 and others).

In the remainder of this section I will refine my sensorimotor interpretation of LF in a way which makes reference to the model of working memory for episodes just outlined. The basic idea will be that an LF representation describes a deictic routine *as it is replayed from episodic working memory*, rather than as it occurs in real time. In Section 7.2 I discuss the mechanics of the working memory replay operation, and make a suggestion about the pattern of sensorimotor signals activated during a replayed sensorimotor routine. In Sections 7.3 and 7.4 I discuss the linguistic reflexes of these signals, and give a more precise sensorimotor definition of LF which makes reference to these.

7.2 Sensorimotor Signals Active During a Replayed Sensorimotor Routine

As just mentioned, we know a lot about the neural mechanisms which store prepared sequences of sensorimotor operations in working memory. These mechanisms are mainly in dorsolateral prefrontal cortex (see e.g. Barone and Joseph, 1989; Averbeck *et al.*, 2002; Averbeck and Lee, 2007) and the supplementary motor areas (Shima and Tanji, 2000).

One interesting property of prefrontal sequence representations is that while they support the execution of sensorimotor operations in sequence, they actually identify the different prepared operations individually, and *in parallel*. For instance, within the prefrontal representation encoding the cup-grabbing routine, it is possible to identify assemblies encoding each of the three prepared operations, *attend_man*, *attend_cup* and *grab* (see especially Averbeck *et al.*, 2002). I will make this explicit by designating the prefrontal sequence plan $plan_{attend\_man/attend\_cup/grab}$.

There is also evidence that planning representations remain tonically active while the planned sequence is being executed (see Averbeck and Lee, 2007; also the computational models of Rhodes *et al.*, 2004 and Takac and Knott, 2013). When a cup-grabbing episode is replayed from working memory, we therefore generate a mixture of sustained and transient signals to be activated, as shown in Figure 7. The sustained signals are all active in prefrontal areas; the transient ones occur in

Sustained signals	Transient signals			
	Initial context	Deictic operation	Reafferent signal	New context
$plan_{attend\_man/attend\_cup/grab}$	c_1	*attend_man*	*man*	c_2
$plan_{attend\_man/attend\_cup/grab}$	c_2	*attend_cup*	*cup*	c_3
$plan_{attend\_man/attend\_cup/grab}$	c_3	*grab*	*man*	c_4 / *cup*

Figure 7: The time course of signals occurring during the replay of the cup-grabbing episode in working memory

the sensorimotor areas which are active during actual sensorimotor experience, in accordance with simulationist models.

Another important property of the prefrontal sequence-planning mechanism is that it makes use of a representation of 'the current context' which is updated after each operation. Executing a planned sequence of operations relies on tonically active representations of the planned operations, but also on a dynamically changing representation of context. There are several different models of these context representations (see e.g. Dominey *et al.*, 1995; Houghton and Hartley, 1995; Beiser and Houk, 1998; Takac and Knott, 2013). I will remain neutral about their exact nature; in Figure 7 I just refer to them as c_1–c_4.[3]

7.3 The Interface Between Replayed Sensorimotor Sequences and Language

In Section 7.1 I proposed that an LF representation describes a sensorimotor routine as it is replayed from working memory. Before I consider how this proposal applies to the cup-grabbing sentence, I want to state it as concretely as possible, by providing an explicit proposal about how sensorimotor signals interface with linguistic representations at the level of neural circuits.

---

[3] Note that in earlier examples of deictic routines I used capitals to denote contexts ($C_1 \ldots C_n$). These were understood as denoting contexts as represented by a theorist watching an agent: they combine objective information about the agent's current physical state with information about his current sensorimotor representations. 'Subjective' contexts ($c_1 \ldots c_n$) are basically an agent's internal representations of objective contexts. In an account of how sensorimotor signals interface with language, we can only make reference to subjective contexts.

I first assume that there is an area of the human brain which encodes motor plans associated with linguistic actions (e.g. planned articulatory actions), which I will call the **premotor output area**. This roughly corresponds to what others have called the 'phonological output buffer', and have localised in parts of Broca's area and adjacent areas of prefrontal and articulatory premotor cortex (see e.g. Henson *et al.*, 2000; Sahin *et al.*, 2009).

I further assume that the human brain contains a number of neural circuits specialised for concrete language which I will call **interface circuits**, each of which links an area expressing sensorimotor signals to the premotor output area, so that activating sensorimotor signals in these areas can activate arbitrary motor plans (in particular articulatory plans). These circuits can hold learned associations between individual sensorimotor concepts and individual motor movements: they allow the production of atomic symbolic gestures, of the kind posited in Bickerton's (1995) account of 'protolanguage'. There may be several interface circuits, linking different classes of sensorimotor signal to the premotor output medium. I assume each of these circuits evolved separately, through an adaptive mutation allowing a particular class of sensorimotor concepts to be associated with overt expressive behaviours.

Now consider what happens when a sensorimotor routine stored in episodic working memory is replayed, and we generate a pattern of sensorimotor signals like the one shown in Figure 7. This will produce a pattern of signals in the premotor output area. Importantly, this pattern of output signals need not reflect the pattern of sensorimotor signals in every detail. There may not be an interface circuit for every sensorimotor area. And different circuits may express sensorimotor signals at different levels of detail. Thus when a prepared deictic routine is replayed, the pattern of signals evoked in the premotor output area will reflect the pattern evoked in sensorimotor areas, *as filtered* by the interface circuits. With this idea in mind, I will state more precisely how I suggest we can interpret the LF of a concrete sentence.

> **Proposal 4** The LF of a concrete sentence reporting an episode E provides a description of the pattern of signals activated when the deictic routine through which E is experienced is replayed from episodic working memory, including signals activated in the premotor output area.

7.4 The Replay Mechanism as the Basis for an Account of Head Raising

Proposal 4 suggests that LF reflects the structure of replayed sensorimotor routines, but also the properties of the interface circuits linking sensorimotor areas to the linguistic output medium. A key question now is: what are these interface circuits, and what are their properties?

As already mentioned, not all sensorimotor areas participating in a replayed deictic routine need have interfaces to linguistic outputs. Informationally speaking, there is a lot of redundancy in the pattern of signals active in a replayed deictic routine. During replay, each deictic operation is expressed in two different media, one encoding the operation as it is planned, the other encoding it as it is experienced (or rather, simulated). For instance, in the pattern shown in Figure 7 there is a representation of the motor operation *grab* in the planning medium (where it is tonically active), and another in the 'deictic operation' medium (where it is transiently active in Cycle 3). Likewise there are multiple representations of *man* and *cup*: these feature both as actions of attention and as transient reafferent object representations.

The presence of *sustained* representations of deictic operations in the pattern of activity created by a replayed deictic routine prompts an interesting suggestion about the neural mechanisms underlying the Minimalist account of head raising. Recall from Section 3.2 that an inflected verb generated at the head of VP must raise to the head of AgrOP and then the head of AgrSP, to check its 'agreement features'. AgrSP and AgrOP are headed by bundles of subject and object agreement features respectively, while VP is headed by the verb. But the verb is allowed to appear at higher head positions, and its agreement features are allowed to appear at lower head positions. The Minimalist account of head-raising allows, indeed requires, heads to appear 'out of position': it is through this device that a Minimalist analysis explains the extended syntactic domain of the verb and its agreement features.

If, as I am suggesting, a right-branching LF structure describes a temporal sequence of deictic operations, and if the head of each XP in an LF structure signals a deictic operation (see Proposal 3), then the principles which sanction head-raising must be interpreted as allowing heads to describe deictic operations *out of sequence*. They allow the head of an XP to report not just the deictic operation presently occurring, but *all* deictic operations in the deictic routine currently being rehearsed, including those which have already occurred and those which have yet to occur. A natural way of explaining this is to suggest that the linguistic signals of deictic operations come from the area where they are planned, and are therefore tonically active in parallel during replay, rather than from the areas where they are transiently expressed one by one. I will suggest a further elaboration of the 'sensorimotor' interpretation of an XP to this effect:

Proposal 5 The head (X) of an XP in the LF structure of a concrete sentence describes a deictic operation in a replayed deictic routine *as it is represented in the prefrontal area storing the deictic routine*.

Given that the prefrontal area holding prepared deictic routines represents their component deictic operations tonically, and in parallel, it follows from Proposal 5 that all the heads in a right-branching structure of XPs describe the same set of deictic operations.

To summarise: the suggestion made in Section 7.3, that an LF structure describes a *replayed* sequence of deictic operations (Proposal 4), recommends itself not only as a way of incorporating reference to working memory representations in a sensorimotor characterisation of LF. It also opens the way for a promising account of the neural basis of head-raising, which is grounded in known facts about the representation of prepared sequences in prefrontal cortex. If LF describes a temporal sequence of operations, then the Minimalist device of head-raising within LF can be understood as a way of encoding the presence of tonically active representations within the neural areas from which surface language forms can be produced. And we know that there are such representations when a prepared sensorimotor sequence is replayed.

Independently of these considerations, there is also good evidence that the processing of verbs and their inflections (the elements involved in head-raising) does indeed involve the dorsolateral prefrontal cortex. Left dorsolateral prefrontal cortex is an area associated with the production and interpretation of verbs (see Perani *et al.*, 1999; Tranel *et al.*, 2001; Pulvermüller *et al.*, 1999) and the processing of verb inflections (see Shapiro and Caramazza, 2003; Shapiro *et al.*, 2012).

8 A Sensorimotor Interpretation of the LF of the Cup-Grabbing Sentence

We can now consider how the general proposals just made about the LF structure of concrete sentences apply to our particular cup-grabbing scenario. We have a Minimalist model of the LF of a transitive sentence reporting a cup-grabbing episode (see Figure 2). And we have a model of the pattern of neural signals activated when the deictic routine through which a cup-grabbing episode is experienced is replayed from episodic working memory (see Figure 7). Proposal 4 asserts that the LF of the cup-grabbing sentence can be construed as a description of the pattern of signals activated during replay—i.e. that the LF structure in Figure 2 can be thought of as a description of the pattern shown in Figure 7. Proposal 2 asserts that in addition, each X-bar schema in the LF structure describes one cycle of the replayed deictic routine. And Proposal 3, extended by Proposal 5, suggests roles for each constituent in each X-bar schema. In this section I will examine these assertions, and in the process make some further general suggestions about how to interpret the Minimalist account of LF in sensorimotor terms.

8.1 General Alignment of the LF Structure and the Deictic Routine

At the level of X-bar schemas, the LF of a transitive cup-grabbing sentence aligns well with the deictic routine it is supposed to represent. The LF structure contains three XPs, and the deictic routine contains three cycles. We can therefore interpret each XP as describing a single cycle of the deictic routine. The predicted sensorimotor interpretations are shown in Figure 8. AgrSP describes Cycle 1 of the routine (an action of attention to the agent), AgrOP describes Cycle 2 (an action of attention to the patient), and VP describes Cycle 3 (execution/monitoring of the 'grab' motor programme). According to Proposal 3, each X-bar schema describes a context-update operation: XP describes the context in which a deictic operation is executed and its complement describes the new context which it brings about. This allows us to interpret the right-branching chain of XPs (AgrSP, AgrOP, VP, DP) as describing the four contexts c_1–c_4 in the deictic routine. The general form of the LF structure certainly corresponds well to the general form of the deictic routine.

In Sections 8.2 and 8.3 I will consider the interpretation of specific elements within the LF structure.

8.2 Sensorimotor Interpretations of Argument Positions

I will first consider the sensorimotor interpretations of the positions in the LF structure of the clause at which the verb's arguments appear. There are four of these.

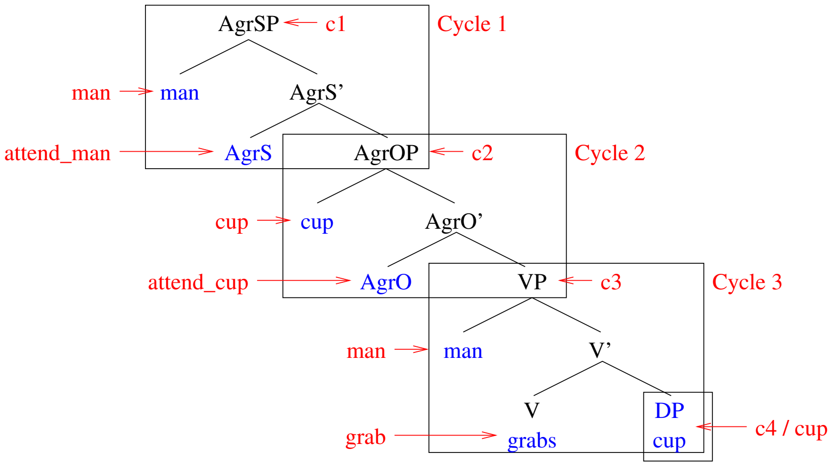


Figure 8: Sensorimotor interpretation of the LF structure of *The man grabs a cup*. Sensorimotor interpretations are shown in red.

The subject *man* appears at the specifier of AgrSP and the specifier of VP. The object *cup* appears at the specifier of AgrOP and the complement of VP.

If we assume the alignment of X-bar schemas and deictic routine cycles shown in Figure 8, and the sensorimotor interpretations of specifier and complement positions made in Proposal 3, these positions correspond exactly to the positions in the deictic routine where representations of *man* and *cup* are active. The specifier of AgrSP is predicted to describe the reafferent sensorimotor signal *man*—and the word *man* appears at this position. The specifier of AgrOP is predicted to describe the reafferent sensorimotor signal *cup*—and the word *cup* appears at this position. The specifier of VP is predicted to describe the second activation of the signal *man*, this time as a reafferent signal generated during monitoring of the motor programme—this is the other position where the word *man* appears. So far so good.

The remaining position, the complement of VP is predicted to describe the consequent state of the action, in which the agent is holding the cup. This is the other position where the word *cup* appears. This position is interesting syntactically, because it is the only place where an argument appears as a complement, rather than as a specifier. Within the sensorimotor model, we can make two predictions about it. On one hand, since the word *cup* occupies this position, we predict it will describe the reafferent signal *cup*. But equally, since it is a complement position, we predict it to describe the consequent state of the action taking place in the third cycle of the deictic routine, namely the *grab* action.

As discussed in Section 6.3, the consequent state of the *grab* action does indeed have a special dual status. This action is a substantive motor action, bringing about a change in the world. But it is also an attentional action, providing information about the cup in the modality of touch. The stable grasp state is *axiomatically* a state in which the cup representation is active, because it is at this point that the function mapping visual object representations onto goal motor states is trained. The dual status of the VP complement as an argument position *and* a complement position perfectly reflects this built-in identification of the target representation with the goal motor state.

In summary, the sensorimotor interpretations predicted for all four argument positions by Proposal 3 are quite plausible. In fact one might well go further, and suggest they may provide helpful insights into the Minimalist concepts of 'specifier' and 'complement'. Recall that the aim of the paper is to seek a linking theory connecting syntactic constructs to neural mechanisms. Without such a theory, syntactic constructs like 'specifier' and 'complement' are justified purely through the roles they play in a larger theory of syntax. Sensorimotor characterisations of these terms provide a way of thinking about them empirically which is entirely independent from the role they play in syntax. If the independent conceptions of 'specifier' and 'complement' line up with the conceptions motivated from syntactic theory, as they seem to in the cup-grabbing example, this has the character of an empirical result: it suggests that the empirical conceptions may be *the basis for* the theoretical conceptions, explaining why these conceptions play a useful role in a theory of syntax.

8.3 Sensorimotor Interpretations of Head Positions: V, AgrS and AgrO

I will next consider the sensorimotor interpretations of head positions which are predicted by Proposal 3. Proposal 3 asserts that the head of an XP denotes a deictic operation: thus the items appearing at V, AgrS and AgrO are predicted to describe the three deictic operations in the deictic routine through which the cup-grabbing episode is apprehended.

The case of V is quite straightforward: the word which appears at this position (*grabs*) is easily thought of as describing the third operation in the deictic routine, activation of the *grab* motor programme. The more interesting cases are AgrS and AgrO. In Minimalism, these heads hold the 'agreement features' of the subject and object respectively. Agreement features classify the subject and object into broad semantic categories along various dimensions, principally PERSON, NUMBER and GENDER, though exactly how categories are defined within these dimensions varies from language to language. The sensorimotor interpretation of LF sees these bundles of agreement features (e.g. 3RD PERSON SINGULAR, MASCULINE) as 'descriptions of attentional operations'. Does this proposal make sense?

To begin with, it should be noted that Agr projections as originally proposed in Minimalism were seen as carrying features with no real semantic content; Agr features were traditionally checked *and then erased*, so they were not visible in the LF structure presented to the semantic system. In my interpretation, Agr heads carry as much semantic conent as other heads; they describe sensorimotor operations, just like other heads: so they are certainly used somewhat nonstandardly in my interpretation. In any case, the question to consider is whether it is plausible to think of Agr features as describing attentional operations.

I will first consider the PERSON feature, and in particular the distinction between 1ST and 3RD person. DPs carrying the 1ST person feature (e.g. *I*) refer to the speaker; those carrying the 3RD person feature (e.g *he*, *she*, *the man*) refer to an external agent. At least in this case, I suggest that agreement features can be very well interpreted as descriptions of attentional operations. In the account of sensorimotor processing given in Section 6, the operation through which an observer attends to the agent of a cup-grabbing episode is also fundamental in implementing that observer's ability to distinguish between himself and external agents. Recall that action episodes are represented within the observer's mirror system. In my account, the circuitry in the mirror system has to be configured for execution of actions or for perception of external actions before signals in the system can be interpreted—and these mode-setting operations must also be construed as attentional operations, which direct attention to the agent of the observed episode. So at least with regard to the 1ST and 3RD PERSON features of AgrS, it makes perfect sense to think of agreement features as describing attentional operations.

Whether this idea extends to other agreement features, and to AgrO as well as AgrS, is a matter for further work. Certainly there is a plausible attentional basis for the distinction between 1ST and 3RD PERSON at AgrO. The attentional action through which an agent establishes himself as a reach target is very different from that which establishes an external target. There are specialised sensorimotor pathways controlling actions directed towards the self, and these actions have different dynamics from those directed externally (see e.g. Petreska and Billard, 2009; Ferri *et al.*, 2010). As regards the other PERSON feature, 2ND PERSON, attending to the addressee is substantially different from attending to a third party. When a speaker is producing an utterance, the addressee must already have been established as a focus of attention, so rehearsing an action of attention to the addressee involves a special kind of reattention which could plausibly be linguistically marked. There is also some reason to suppose that the distinction between SINGULAR and PLURAL is attentional in origin. For instance, the brain area which most plausibly encodes syntactic plurality, the left temporoparietal junction (Domahs *et al.*, 2012) is also activated by attentional operations parsing a visual stimulus as a group rather than as a single entity (Yamaguchi *et al.*, 2000). A computational model of the attentional origin of the singular-plural distinction is given in Walles (2010).

I will conclude with some comments about GENDER features. These features are much more open-ended semantically than PERSON and NUMBER, and much more language dependent (Corbett, 1991). Can these features be thought of as reflecting aspects of an attentional action? I think this is also plausible. The important thing to note is that attentional actions like *attend_man* or *attend_cup* involve top-down establishment of open-class object representations as well as saccades to external points in the world. When an observer executes *attend_cup*, he activates a representation of a cup as a search goal, which can be matched against object representations arriving bottom-up, so there is some way of knowing whether the action is successful (Tomita *et al.*, 1999; Hasegawa *et al.*, 2000; Hamker, 2004). I suggest that while most of the open-class properties of attentional operations are expressed through their reafferent sensory consequences (i.e. through nominal expressions), we can also read *some* of these properties from representations of the operations themselves, which are signalled by heads. Recall my assumption that there are 'interface circuits' linking areas evoking sensorimotor signals to a language-specific premotor output area (see Section 7.3). These circuits are allowed to have different capacities. I suggest that the circuit which generates linguistic reflexes of the open-class semantic properties of deictic operations has rather limited capacity, and that GENDER agreement features are generated through this channel.

In summary, there is some support for the sensorimotor interpretations of the head positions AgrS, AgrO and V. The interpretation of V is certainly plausible. The interpretation of the 1ST and 3RD PERSON features that appear at AgrS and AgrO is also very plausible, and again seems to provide some insight into the neural basis of these agreement features. But more work is needed to determine whether there are sensorimotor correlates of the other agreement features that can appear at AgrS and AgrO.

8.4 A Sensorimotor Interpretation of V-Raising and Agreement

In Section 7.4 I suggested that the heads of XPs should be understood as describing deictic operations 'as they are planned' rather than as they are evoked in sequence. This idea, taken together with known properties of the sequence-planning mechanism in prefrontal cortex, led to an interesting sensorimotor interpretation of the Minimalist device of head-raising. Now that we have sensorimotor interpretations for the AgrS, AgrO and V heads in the cup-grabbing sentence, it is useful to reconsider this sensorimotor account of head-raising, to see how it applies in this particular case.

Consider the English version of the sentence, *The man grabs the cup*. In the surface structure of this sentence, the subject *The man* appears outside the VP, and is therefore syntactically somewhat remote from the verb. Any account of syntax has to explain why the verb's inflection has to agree with the subject, even though the verb is not near the subject. The Minimalist account of head-raising explains this by positing that the inflection *-s* signals an agreement feature which actually 'belongs' at a position above VP, where it *is* near the subject. It is allowed to appear on the verb because at LF the verb 'covertly moves' up to the position where it really belongs.

In my proposed sensorimotor interpretation of head-raising, the subject agreement inflection on the verb is a signal of properties of the attentional action which established attention on the agent of the cup-grabbing action. (Specifically, it signals that this operation involved configuring the mirror system for action observation rather than action execution.) The reason why the inflection is allowed to appear on the verb is that linguistic reflexes of attentional actions are generated from the region where they are planned, and are therefore tonically active. According to this interpretation, the phenomenon of agreement is seen as reflecting the machinery through which episodes are stored in, and replayed from, working memory. The interpretation suggests a specific neural mechanism which accounts for the syntactic phenomenon of subject-verb agreement, as it is accounted for within the Minimalist model.

Most models of syntax include a device allowing agreement features to spread through a syntactic structure. For instance, this is achieved through unification operations in models like HPSG, Tree-Adjoining Grammar and Combinatory Categorial Grammar. Does my proposed sensorimotor interpretation of agreement apply equally well to the account of agreement features given in these other frameworks? I think there are two aspects of the Minimalist account which make it a particularly good vehicle for expressing this sensorimotor interpretation of agreement. Firstly, the Minimalist model envisages head movement taking place at a language-independent level of syntactic representation. The suggestion that deictic operations interface with language through a planning medium where they are tonically active makes no reference to particular languages; the Minimalist device of expressing movement at a language-independent level of syntactic representation is therefore particularly appropriate. Secondly, a Minimalist LF structure can be neatly interpreted as a description of a sequence of operations. It is therefore particularly suitable for expressing an account of agreement phenomena grounded in a neural model of prepared sequences.

8.5 A sensorimotor Interpretation of DP movement, Case and Thematic Roles

There are two kinds of movement at LF: movement of heads to higher head positions, and movement of argument DPs to Case-assigning positions. In this section I will consider what sensorimotor interpretation can be given to DP movement.

The Minimalist account of DP-raising supposes that the argument DPs of a verb initially appear at positions within the VP. In our example, the subject of *grabs* appears at the specifier of VP and its object appears at the complement. At these structural positions, the verb's arguments are assigned thematic roles: the specifier assigns AGENT role and the complement assigns PATIENT role. But Minimalism also requires arguments to be assigned 'Case'. Case can only be assigned by functional heads above VP: the heads AgrS and AgrO assign nominative and accusative Case to their specifiers respectively. So the subject and object DPs must raise to these specifier positions.

As with the other theoretical devices in Minimalism, the principle which requires arguments to raise to Case-assigning positions is justified purely through the formal role it plays in a complete model of syntax which neatly accounts for a large body of linguistic data. The idea that argument DPs must have Case is simply stipulated: there is no proposal about what Case 'is', in the same way as there is no proposal about what specifiers and complements 'are'. (Or rather, it is assumed that the principle 'DPs must raise to Case-assigning positions' corresponds to *some* neural mechanism, but there are no proposals about what this might be.) Does the sensorimotor interpretation of LF allow us to say anything about this principle?

Case Assignment I suggest that there is a very clear sensorimotor interpretation of the principle that DPs must raise to get Case. So far I have argued on several grounds that the functional projections AgrSP and AgrOP describe actions of attention to the agent and the target of the cup-grabbing episode, while the VP projection describes the monitoring of a motor programme. In sensorimotor terms, the requirement that the subject and object appear in AgrSP and AgrOP projections above VP translates as a requirement about the structure of sensorimotor routines—namely that an observer *must attentionally establish* the agent and target of a grab action before he can actively monitor a motor programme involving these individuals. Within a sensorimotor model, this requirement is completely justifiable in its own right. In fact this principle is at the heart of Ballard *et al.*'s conception of deictic routines. In order to monitor a motor programme involving multiple participants, in Ballard *et al.*'s model, an observer must first attend to these participants, one by one, to set up the deictic representations which instantiate the free parameters of the motor programme.

If we are thinking about LF in sensorimotor terms, we can now give a very clear account of the functional projections which assign Case to DPs. These XPs describe the attentional operations which establish the conditions under which the motor programme can be monitored. The general principle that Case is assigned 'by a functional head to its specifier' (which is an important part of the Minimalist account of Case) also has a very clear sensorimotor interpretation. A functional head describes an action of attention, and its specifier describes the deictic representation which this action results in. The deictic representation clearly *depends on* the action of attention. I suggest this dependence is the basis for the Minimalist idea that a specifier depends on, or is licensed by, its head.

Thematic Role Assignment Now that we have a sensorimotor interpretation of the 'higher' subject and object positions above VP, can we find an interpretation of the subject and object positions within VP, the specifier and complement of VP? In Minimalism, these are the positions where the verb's arguments receive 'thematic roles' (namely AGENT and PATIENT). What can we say in sensorimotor terms about these positions? Again, the sensorimotor interpretation of LF seems illuminating. The VP projection describes the cycle of the deictic routine in which the grab action is dynamically monitored. During this cycle, as discussed in Section 6.3, new representations of the agent and the target become active, in new modalities connected with the motor system. While action-monitoring is under way, a representation of the agent as an animate entity is activated as a reafferent side-effect. And when action-monitoring is complete, a representation of the target as a goal motor state becomes active. These points in the routine are described by the specifier and complement of the VP respectively. I suggest that the sensorimotor interpretations of these VP-internal positions explain why they assign the thematic roles they do—and also help us to understand the semantics of thematic roles by showing how they are grounded in sensorimotor representations (in concrete cases such as ours). The AGENT role is assigned by the specifier of VP because this position describes an animate representation of one of the action participants: and the meaning of 'AGENT' in this context comes largely from the nature of this animate representation. The PATIENT role is assigned by the complement of VP because this position describes an affordance-based representation of the other action participant: and the meaning of 'PATIENT' in this context likewise derives largely from the nature of this representation.

DP Movement Finally, we have to seek a sensorimotor interpretation of the Minimalist mechanism by which arguments 'move' from their VP-internal positions to their Case-receiving positions. In the Minimalist model this mechanism is entirely distinct from the mechanism by which verbs raise to higher head positions. We already have a sensorimotor account of verb raising. Can we give an account of DP raising?

Given the ideas suggested in this section, it is clear that the sensorimotor account of DP-raising should relate to the fact that both the agent and target are represented twice, in different modalities, in the course of the deictic routine involved in experiencing the cup-grabbing episode. None of these representations are tonically active during the routine, so their appearance at multiple positions in

LF is certainly due to a mechanism distinct from that underlying head-raising.

My suggestion is that the Minimalist device of DP-raising reflects the associative neural mechanisms through which representations of the agent and target in different modalities are tied together to form multimodal representations. There must be circuitry enforcing certain axiomatic correspondences between representations in different modalities. To create a multimodal agent representation, there must be circuitry linking the reafferent signal activated during action monitoring with the reafferent signal activated by the first action of attention. These two signals axiomatically represent the same object: the agent. To create a multimodal target representation, there must be circuitry linking the reafferent motor state active at the endpoint of the action with the reafferent signal activated by the second action of attention. Again these axiomatically represent the same object: the target. Note that these circuits must link representations purely in virtue of their structural positions in the deictic routine. I will not discuss how they might be implemented neurally, but I do suggest that there must be such circuits to explain how multimodal object representations are learned, and that these circuits provide a plausible neural basis for the Minimalist account of DP raising, which links particular structural positions within LF.

8.6 A Sensorimotor Interpretation of the Generative Mechanism

In Minimalism, LF structures are produced by a generative mechanism. The set of possible LF structures is infinite, so they cannot be enumerated: instead this set is defined indirectly, by characterising the mechanism which creates these structures. I have suggested a sensorimotor interpretation of complete LF structures. But it is also important to have some account of 'the mechanism which produces all possible LF structures'. What might this correspond to in the sensorimotor model?

It is hard to find a direct sensorimotor interpretation for the generative mechanism as it is proposed in Minimalism. The mechanism defined in the Minimalist model proceeds from the bottom up: the lowest projection at LF is created first, and higher XPs are successively adjoined to this. (In our example, the VP would be created first, and then merged successively with AgrOP and AgrSP.) If a right-branching LF structure describes the representation of a temporal sequence, as I suggest, then the Minimalist generative mechanism describes a process whereby this representation is assembled *in reverse*, beginning with the last elements. I cannot see anything in the sensorimotor model which corresponds to this process. The model includes an account of sensorimotor sequences being stored in working memory, but these sequences are stored, and replayed, in the order they are experienced.[4] While there is a good sensorimotor interpretation of the LF structure produced at the end of a derivation, I suggest there is no good interpretation of the generative process understood as a procedural mechanism.

To be clear: I do not want to say that the Minimalist generative mechanism 'does not describe neural processes'. I only want to say that it does not describe neural processes *when understood as a procedural mechanism*. The Mininalist gen-

[4] The only evidence I am aware of for *reversed* replay of experienced sequences is in studies of hippocampal representations of spatial location (Foster and Wilson, 2006; Diba and Buzsàki, 2007).

erative mechanism creates representations (LF structures) which I argue describe neural processes in considerable detail. And there are certainly components of the generative mechanism which pick out well-defined features of these processes. To take an example, consider the Minimalist idea that constituents 'move' from lower to higher positions within an LF structure while it is being derived. In one sense, I am saying that there is no such thing as 'movement' of constituents within LF. I do not think there is any neural mechanism corresponding to movement *as such*. But as already argued, I think there are good sensorimotor interpretations of the structures in LF which *result* from movement in the Minimalist account.

Of course, we still need to give a sensorimotor account of 'the mechanism which produces all possible LF structures'. This is a central component of the Minimalist model of grammar. But a sensorimotor account of this mechanism will be quite a radical departure from the Minimalist account. I will conclude this section by considering what this account will look like.

What are the constraints on possible LF structures, if these are thought of in sensorimotor terms? In the sensorimotor interpretation of LF, a right-branching LF structure describes a replayed sequence of sensorimotor operations. If we want to characterise the set of possible LF structures, we must specify *in general* what sequences of sensorimotor operations are possible. We have already seen that there are several general constraints on the sequences of sensorimotor operations an observer can execute. For instance, an observer cannot execute a motor routine without having attended to the participants involved (Section 8.5); an observer cannot attend to a target object before having attended to the agent (Section 6.1). These are the kinds of sequencing constraint which feature heavily in Ballard *et al.*'s model of deictic routines. So part of a sensorimotor account of the generative mechanism will probably involve enumerating constraints resulting from the embodied nature of cognitive processing, of the kind studied by Ballard *et al.* But there are also properties of LF structures which the sensorimotor model sees as reflecting properties of the working memory mechanism which allows an observer to store and replay the deictic routines he experiences (Section 7.4) and properties of the associative mechanisms which support the creation of multimodal object representations (Section 8.5). And finally, there are properties of LF structures which are suggested to reflect the nature of the interface circuits linking sensorimotor areas of the brain to a language output area (Section 7.3).

In summary, the 'sensorimotor' characterisation of the space of possible LF structures will be partly an account of the constraints on the sequential structure of deictic routines, partly an account of the neural mechanisms which store and replay these routines, and which exploit the structure of these routines to learn basic object representations, and partly an account of the neural interfaces between sensorimotor and language areas. Note that the first two parts of this account of LF structures are essentially accounts of embodied sensorimotor cognition: they do not make any reference to specifically linguistic representations or mechanisms at all. The only references to specifically linguistic representations are in the account of interface circuits.

9 A Minimalist-Inspired Model of Language Processing and Language Learning

9.1 The Minimalist Account of the LF-PF Interface

In Minimalism, the surface form of a sentence (its PF) is read from the terminal nodes of its LF structure during derivation, in a process called 'spellout'. The rules governing spellout are language-specific: an infant growing up in a particular language community has to learn a set of rules particular to this language. The rules to be learned relate to the positions at which constituents are pronounced. As discussed in Section 3.2, the LF of our example sentence contains two copies of the agent and patient and three copies of the inflected verb: at PF there is only one copy of each. The Minimalist proposal, in a nutshell, is that 'surface' syntactic differences between languages result from different conventions about which copy of these repeated elements is overtly pronounced.

The Minimalist account of the interface between LF and PF plays two related roles within the overall theory. Firstly, it contributes to a parsimonious model of the differences between languages. These differences are attributed to the mechanism which maps between LF and PF representations. Thus, for instance, we can give an account of languages with different constituent orderings (Subject-Verb-Object versus Verb-Subject-Object and so on) in a way which localises these differences to a single module of the grammar. Secondly, because differences between languages are localised to the LF-PF interface, we can tell a relatively simple story about the learning mechanisms which allow infants to acquire their native language. The mechanism responsible for creating LF structures is assumed to be largely innate. (We are allowed to assume this because LF structures are language-invariant.) In order to learn the syntax of their native language, infants only need to learn how to map LF structures to PF structures. The space of possible hypotheses to test is relatively small and well-structured: the infant just needs to learn the values of a small number of discrete parameters—for instance, whether to pronounce the subject 'high' or 'low'.

9.2 Problems with the Minimalist Account

While the Minimalist account of PF is neat in several respects, there are several well-known problems with it. I will mention three of these.

Firstly, it is hard to square the Minimalist account of PF with an account of sentence processing—that is, with an account of the psychological processes which take place in a speaker producing a sentence, or in a hearer interpreting a sentence. While research into sentence generation and interpretation is still at a fairly early stage, there is good reason to think that both processes are at some level 'incremental'—i.e. that syntactic representations are generated in roughly the order they are produced in (for generation) or heard in (for interpretation). For instance, there is evidence that speakers create representations of early constituents of a sentence first, so that they can begin talking while still in the process of planning later constituents (see e.g. Levelt *et al.*, 1999); likewise, hearers start to generate interpretations of a sentence as soon as the earliest constituents are heard (see e.g. Tanenhaus *et al.*, 1995). Minimalist derivations happen from the bottom up, as discussed in Section 8.6. Since the bottom of an LF structure corresponds to the end of a sentence at PF, and PF structures are read from LF structures, it is hard to see the derivational mechanism *understood as a procedure* as a representation of sentence processing mechanisms. Minimalist theorists frequently assert that Minimalist models describe neural processes (see e.g. Marantz, 2005; Hornstein, 2009), but these assertions tend to be about the general nature of structure-building computations in language rather than about the way these computations are ordered.

Secondly, Minimalism has no account of linguistic structures that are defined in the surface form of sentences. In several syntactic frameworks (see e.g. Fillmore *et al.*, 1988; Goldberg, 1995), descriptions of linguistic structures can include reference to particular surface words as well as to abstract syntactic categories. Some aspects of linguistic structure certainly seem well described in terms of abstract syntactic categories like VP and DP, which are defined recursively and hierarchically. But other aspects seem best modelled simply as patterns involving particular word forms. The clearest examples of such patterns in language are **idioms**: phrases like *how's it going* or *by and large*, which appear to deliver a meaning collectively, rather than individually. The main problem for Minimalism, as forcefully argued by Jackendoff (2002), is that idiomatic linguistic structures appear to interact with abstract grammatical structures in ways which are hard to model if all structure is assumed to be created at LF. For instance, it is sensible to analyse the verb phrase *take X to task* as a grammatical phrase featuring certain specific word forms, in virtue of which it receives a conventionalised meaning ('criticise X'). The phrase conforms to a regular syntactic pattern, and the position *X* can be occupied productively by any DP, but the word stem *take* and the words *to* and *task* cannot be varied productively: the phrase has the meaning it does because it features this pattern of specific words. Minimalist analyses have difficulty modelling partially idiomatic constructions of this kind, especially when the idiomatic elements are discontinuous, as in this example.

Thirdly, the Minimalist model of learning is unlike any other account of learning in current cognitive science. Theories of how humans learn in cognitive science are normally expressed as computational models: for instance neural networks or Bayesian reasoning systems. A central insight gained over the last twenty years or so is that these computational models are very powerful—certainly powerful enough to learn rich representations of surface structures in natural language with very little prior knowledge (see e.g. Pantel and Lin, 2002). A simple type of neural network called a **simple recurrent network** can learn rich representations of the sequential patterns in surface language (Elman, 1990, Christiansen and Chater, 1999). Experiments with computational learning systems lend support to 'empiricist' models of language acquisition, which posit that infants use general-purpose learning mechanisms to acquire syntax, rather than elaborate innate knowledge. The empiricist model of development is supported by evidence that infants' earliest syntactic constructions are defined around particular lexical items, and are therefore idiomatic in nature (see e.g. Lieven *et al.*, 1997; Tomasello, 2003). According to empiricist models, infants learn adult syntax by progressively abstracting away from concrete constructions featuring specific words (see e.g. Tomasello, 2003; Macwhinney, 2005).

To some extent, these difficulties facing the Minimalist model all stem from the way it construes the 'generative mechanism' defining the space of well-formed sentences. The fact that LF structures are generated from the bottom up makes the mechanism unsuitable as the basis for an account of sentence processing. The lack of an account of sentence processing makes it hard to express the Minimalist model of infant syntactic development as a computational mechanism. (Computational models of syntactic development are typically also processing models, which receive their training sentences incrementally, word by word.) The lack of a computational account of learning in Minimalism in turn limits what the theory can say about surface structures in language, since these are best analysed as the product of a learning mechansism.

I have argued that the Minimalist model of LF supports an interesting account of the neural basis of syntactic representations, grounded in an account of sensorimotor processing and working memory. But I have also argued (Section 8.6) that in order to formulate this account, we must abandon the Minimalist model of derivation, because it does not square with the sensorimotor interpretation of LF structures. This opens the way for an account of the relationship between LF and PF structures which is more compatible with models of sentence processing and syntactic development. In the next section, I will introduce a new computational model of sentence processing and syntactic development, whose form is inspired by empiricist models of language processing and language learning, but which also retains the Minimalist conception of LF—interpreted in sensorimotor terms—and the Minimalist idea of parameter-setting.

9.3 An Account of Language Processing and Language Learning

My sensorimotor interpretation of LF puts us in a position to address all three problems described in the previous section. Firstly, and most importantly, it provides an ideal basis for an account of sentence processing. Its central claim is that the LF of a sentence *describes a neural process*: namely the process of replaying an episode representation held in working memory. In this section I will propose that the neural mechanism which implements the described replay process is also the mechanism *through which the sentence is produced*—or at least, a central part of this mechanism. In the paper so far I have argued that thinking of LF as a describing a replayed sensorimotor sequence helps us express Minimalism's essentially declarative account of syntactic structure in a way which makes reference to neural mechanisms. I now want to suggest that a sensorimotor conception of LF has the *additional* advantage of supporting an interesting account of sentence processing.

Until now, the working memory mechanism which allows an experienced sensorimotor sequence to be stored and replayed has not been thought of in relation to language processing at all. The mechanism was introduced in Section 7.1 as part of a model of long-term memory for episodes: it provides a means for buffering experienced episodes so they can subsequently be stored more permanently in the hippocampus (and later still in cortex). It is possible to imagine this whole mechanism predating language altogether. I will begin by sketching an account of language evolution in which the replay mechanism did indeed predate language, and was co-opted by evolution for a new role in communication, supporting the production of word sequences and the learning of syntax. Then I will introduce a neural network model of the circuitry that evolved to co-opt the replay mechanism.

Background Assumption: Two Stages of Language Evolution In Section 7.3 I envisaged a point during human evolution when a collection of 'interface circuits' evolved, allowing agents to produce overt behavioural reflexes of their internal sensorimotor representations (see Section 7.3). These circuits allow agents to learn a vocabulary of atomic behavioural symbols. Several theories see the evolution of such circuits as the first major step in the evolution of human language (see e.g. Bickerton, 1995).[5] In many accounts, these circuits support the production of *sequences* of atomic behaviours, to enable an open-class vocabulary of behavioural symbols, and to permit the production of multi-symbol utterances (see in particular Jackendoff, 1999). I will assume interface circuits support sequential behaviours in this way. However, when interface circuits first evolved, I assume they were not used in any systematic way together with the working-memory episode replay mechanism. For instance, they may originally have permitted behavioural reflexes of an agent's *current* sensorimotor signals, rather than of sensorimotor signals retrieved from working memory. I now suggest that at some later evolutionary point, a second piece of language-related neural circuitry evolved, which allowed agents to produce sequentially structured behavioural signals conveying detailed information about *whole episode representations* rather than just about individual sensorimotor signals. I engisage that this circuitry co-opted the working memory replay mechanism. The replay operation generates a pattern of sensorimotor signals whose serial structure reflects the structure of the episode being replayed. Via the interface circuits, it also generates a sequential pattern of signals in the premotor output medium, as discussed in Section 7.3. These signals still need to be converted into overt motor movements. My suggestion is that the circuit which evolved to co-opt the replay mechanism for a communicative purpose transforms the sequence of signals evoked in the premotor medium during replay into an overt sequence of motor movements.

The Control Network The interface circuits which allow behavioural reflexes of sensorimotor signals generate *premotor* movement signals. These signals do not necessarily result in overt movements, but in general the most active premotor signal will be selected for overt execution (see e.g. Fagg and Arbib, 1998). However, even a strongly activated action signal in premotor cortex can be withheld, if the agent has learned a cognitive control strategy which demands this (see Cohen *et al.*, 2013 for a review). In the model I propose, the network which co-opts the replay mechanism for a communicative purpose learns a control strategy which selects just a *subset* of the premotor signals activated during replay for actual execution. I will call this network the **control network**.

The control network's purpose is to produce behavioural representations of replayed episodes which are short and efficient. Recall that there is considerable redundancy in the signals evoked in the premotor output medium when the cup-grabbing deictic routine is replayed from working memory. There are two activations of a signal reflecting the agent, two activations of a signal reflecting the

[5] There are also alternative theories; see e.g. Wray (1998).

patient and a constant activation of a signal reflecting the whole deictic routine as it is planned. Only one version of each of these signals needs to be expressed behaviourally, provided there is a convention about which versions are expressed. The role of the control network is to learn and implement these conventions.

In Minimalist terms, the control network can be thought of as the device which maps LF structures onto PF structures. The repeated activations of sensorimotor signals during replay of a working memory episode correspond to the multiple copies of subject, object and inflected verb at different positions in LF (see Section 8). In the Minimalist model, infants must learn which copy of each repeated element should be pronounced: this is exactly what is learned by the control network.

If we consider which brain region might plausibly implement the control network, an interesting candidate is Broca's area. There is good evidence that Broca's area and surrounding prefrontal regions play a role in syntactic processing, particularly during sentence generation (see e.g. Bookheimer, 2002). However, they also have a role in implementing cognitive control strategies. In fact, the clearest effect of damage to Broca's area itself is impaired cognitive control: 'an inability to override habitual or prepotent response behaviours' (Novick *et al.*, 2005). The role I envisage for the control network is precisely to override prepotent behavioural signals in premotor cortex, so it is natural to localise it in Broca's area.

A Neural Network Model of Sentence Generation I will now outline a model of the control network, and its interaction with the episode rehearsal mechanism and the interface circuits. The overall model is a neural network of sentence generation. A diagram showing its basic architecture is shown in Figure 9. The model is described in detail in Takac *et al.* (2012).

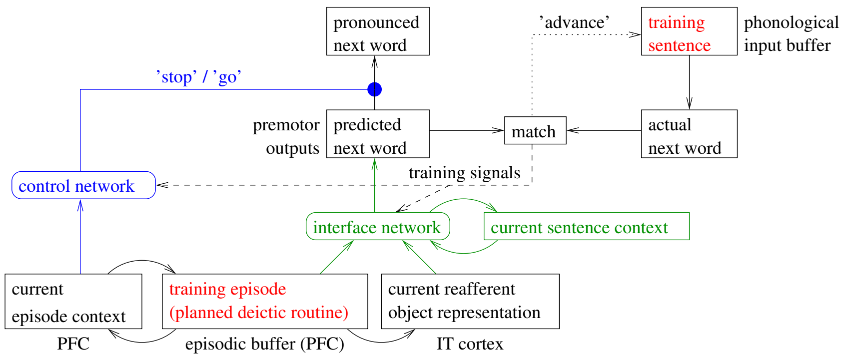


Figure 9: A neural network model of the episode rehearsal system, the interface circuits and the control network

The model is trained on pairs of episodes and sentences, which are shown in red in Figure 9. Each training episode is a representation of a transitive motor action, stored as a planned deictic routine. When replayed, this produces a sequence of episode context representations, and a sequence of reafferent representations of the agent and target of the action (see Section 7.2). Each training sentence is represented as a replayable sequence of words, stored in the phonological input buffer.

Thus episodes and sentences are stored in separate media in working memory, but these media interact, as envisaged by Baddeley (2000).

The interface circuits, here called the 'interface network', are shown in green. The interface network maps sensorimotor signals onto premotor behavioural signals. I will assume these premotor signals are articulatory plans—i.e. word forms, or 'words' for short. The interface network is trained by an error signal ('match' in the figure) which compares the words it predicts from its sensorimotor inputs to words in the training sentence replayed from the phonological input buffer. I assume words in the phonological input medium are also represented as articulatory plans (see e.g. Goldstein and Fowler, 2003) and can therefore be matched to articulatory plans in the output modality.

The interface network can learn to generate a single word from an input sensorimotor signal, but it can also learn to generate a sequence of words from a single sensorimotor input, as envisaged in the models of Bickerton and Jackendoff. This is achieved through the use of a recurrent context representation: the **current sentence context**. A recurrent context representation is used in some form in almost all neural network models of sentence processing. It holds a representation of the sequence of words processed so far, shaped by learning to support prediction of the next word. After each word is generated, the current sentence context is updated to reflect the generated word, and the phonological input buffer advances to the next word, giving the network an opportunity to learn the word which follows the generated word.

The control network, which learns to select a sequence of premotor outputs for overt execution, is shown in blue. This network learns to produce a binary control signal ('stop' or 'go') as a function of the current context representation in the episode rehearsal system—the **current episode context**. This is a representation of 'sensorimotor context' rather than sentence context: it is updated at each cycle of episode rehearsal as discussed in Section 7.2. The control network is also trained by the 'match' signal, but in a different way. If the predicted next word matches the actual next word in a given context, the control network learns to permit words to be overtly pronounced in this context, and causes the input buffer to advance to the next word. If the predicted word does not match, the network learns to refrain from pronouncing words in this context, and does not advance the input buffer. The control network learns language-specific conventions about when to 'read out' words. The conventions it learns depend on the constituent order (SVO, VSO, SOV etc) of the sentences it is trained on.

The full model alternates between two modes of iteration. In one mode, the control network iterates through episode contexts until it reaches a context where a word can be pronounced. The interface network then produces as many words as it can from the current sensorimotor signal, iterating to a new sentence context after each word produced. When it cannot confidently predict the next word, the control network takes over again. The effect of this interaction between the two networks is to combine a fairly traditional empiricist model of sentence processing and sequence learning with a recognisably Minimalist model of parameter setting. If we interpret LF structures as describing replayed deictic routines, then the control network can be interpreted very straightforwardly as a computational model of the mechanism through which infants learn to map LF onto PF representations.

(Note that the control network learns a function which makes no reference at all to the content of words—it only refers to context representations expressing the current state of an episode being replayed. The rules it learns are abstract structural rules, like those proposed in Minimalism.) At the same time, the recurrent component of the interface network allows it to learn rich representations of surface structure in the training data. The sentence context is updated after each word is produced. With training, the interface network can learn to map a single sensorimotor signal onto an idiomatic sequence of words. More interestingly, it also allows the learning of constructions involving a mixture of idioms and productive syntax, for instance 'discontinuous' idioms like *take X to task*. Another interesting point to note is that the model can be configured so that it learns item-specific idiomatic constructions first, and productive syntactic rules later, consistent with empiricist accounts of syntactic development (Lieven *et al.*, 1997; Tomasello, 2003; Macwhinney, 2005). The details of these results, as well as of the architecture and training of the network, can be found in Takac *et al.* (2012).

The model just described is very preliminary; it must of course be scaled up and extended in several directions. My main point in presenting it here is just to emphasise that thinking about LF as a description of a replayed deictic routine is not only helpful in suggesting a neural basis for the representations of syntactic structure proposed within Minimalism: it also allows the Minimalist model of syntactic structure to be stated in a way that is broadly compatible with empiricist accounts of sentence processing and language learning, and of the role of surface structures in language. The network presented here is an example of one such account.

10 Discussion

The aim of this paper is to express a 'linking hypothesis' connecting syntactic representations, as motivated within linguistic theory, to neural mechanisms, as motivated by experiments in psychology and neuroscience. The approach I have taken is to look in detail at a single example sentence, reporting a specific concrete episode, and at the sensorimotor process through which this specific episode is apprehended. My aim is to express a linking hypothesis which connects the *detail* of a sensorimotor model to the *detail* of a syntactic model. At the same time, I have expressed the hypothesis in very general terms, so that it makes predictions which extend beyond this particular example to other concrete sentences. As discussed in Section 5, in the domain of concrete sentences the linking hypothesis assumes that *any* right-branching LF structure describes a sequence of deictic operations. This hypothesis was extended in later sections. I proposed in Section 7.4 that any right-branching LF structure which is a domain for head-raising describes a sequence of deictic operations as it is replayed from working memory storage. And I proposed in Section 8 that any Case-assigning projection describes an attentional action establishing a participant in a sensorimotor routine, and that any instance of DP-raising reflects associative neural mechanisms involved in learning multimodal object representations. In fact, as discussed in Section 1, I am also committed to extending these general proposals beyond the domain of concrete sentences. The point of studying the sensorimotor domain first is simply to develop hypotheses about neural mechanisms in an area where these are relatively easy to study, so that we can later approach other domains with some idea of what we are looking for.

Pursuing these general proposals obviously forms the basis for a whole programme of research at the interface between theoretical syntax and neuroscience. There are many interesting directions to puruse, some of which are discussed in my book (Knott, 2012) and some of which form part of ongoing work. I will not discuss these here, but instead will conclude with some thoughts about the research programme as a whole: how the proposals square with recent developments in Minimalism, how they relate to existing cognitive interpretations of syntax, and how they can be extended beyond concrete sentences.

10.1 A Sensorimotor Interpretation of Merge?

As noted in Section 3.2, in modern Minimalism the X-bar schema is not the primitive recursive building block of LF structure; the structures formerly associated with X-bar schemas are now derived from applications of the more basic operation Merge (Chomsky, 1995a). Merge is an operation that combines two syntactic objects α and β into a single new object, and labels this object with the constituent α, as shown in Figure 10. In this operation, α plays the role of a head in X-bar

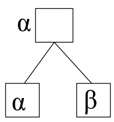


Figure 10: The Merge operation

theory. By applying two successive Merge operations, a structure akin to an XP schema can be created: the complement is joined to the head in the first Merge operation, and the specifier is joined to the result in a second Merge operation. A key difference is that in a Merge-based system, complements and specifiers are not primitives; rather they are defined as positions in structures created by particular combinations of Merge operations.

Is there a sensorimotor interpretation of Merge consistent with my proposed interpretation of the X-bar schema? My earlier suggestion was that an X-bar schema describes a context-updating deictic operation, activated as part of a deictic routine replayed from working memory: the operation is executed in an initial context, generates a reafferent sensory signal, and results in a new context. If there is a sensorimotor interpretation of Merge, it must reconstrue this replay process, identifying some of its more basic components. I should reiterate that since an LF structure describes a *process* in my interpretation, not a declarative mental representation, the structure formed by Merge will not be interpreted as describing a single static mental representation, constructed from two component mental representations. Rather it will be interpreted as describing a basic unit of spatiotemporal structure in a replayed deictic routine, and its constituents will be interpreted as describing specific sensorimotor signals activated in the course of such a routine. The key question, then is what these signals might be, and what relationships between them might be encoded by the minimal unit of structure created by Merge.

One possible way to interpret the structure formed by a Merge operation is with reference to the associative brain mechanisms that implement the storage of a sequence of deictic operations, and enable its replay. As discussed in Section 7.2, these mechanisms make use of a dynamically updating representation of context. One mechanism associates an initial context representation c with a deictic operation O, so that activating c triggers activation of the operation O. The other associates the operation O—in the current context c—with another sensorimotor signal S. These mechanisms interact: when c becomes active, this activates O; the combination of c and O in turn activates S. Now consider the structure in Figure 10, where a head constituent α is merged with another constituent β. One possible interpretation is that α describes a sensorimotor operation O, β describes some other sensorimotor signal S, and the constituent formed by merging α and β describes the context c, which triggers operation O, and then enables a subsequent association between O and the signal S. This constituent 'represents the combination of α and β' in that the context it describes enables an associative connection between the signals described by α and β. Its being 'labelled' with the head constituent α reflects the fact that the context it describes directly activates the signal described by α. This interpretation makes sense both in the case where the signal S is a reafferent consequence of the deictic operation O, and in the case where S is an updated context representation, which can in turn activate the next deictic operation.

It is beyond the scope of this paper to propose a detailed sensorimotor interpretation of Merge. My main suggestion is that thinking of the structure created by Merge as describing an element of structure in a cognitive *process*, rather than describing a static cognitive representation with component parts, may be helpful in characterising Merge in neural terms. The above interpretation of Merge is not completely satisfactory: it does not quite gel with the proposal that an X-bar schema is derived from two successive applications of Merge. (This would imply that an X-bar schema describes two successive deictic operations, while in the sensorimotor model I am assuming, the deictic operation that activates a reafferent signal is the *same* operation that triggers an updated context representation.) But for the moment, pursuing an improved interpretation will be left as a matter for further work.

10.2 Other Functional Projections in Minimalist Analyses

As also noted in Section 3.2, the LF structure of a clause contains several projections that do not appear in my simple model of LF structure: these include CP, which has a role in the syntax of questions, relative clauses and clausal complements, TP, whose head holds the semantic features signalled by the tense inflections of verbs or tense markers, and more recently vP, a projection headed by a light verb, that introduces the VP proper. If these projections do indeed feature in LF structure—and there is good evidence they do—my general sensorimotor interpretation of right-branching LF structures makes clear predictions that the process of apprehending a episode has additional stages to it, that appear at well-defined serial positions in relation to the stages I have described. I will briefly discuss some ongoing work exploring these predictions.

CP and TP are in the 'left periphery' of LF, above AgrSP; CP is higher than

TP. The existence of these projections predicts that experiencing an episode involves executing two sensorimotor operations in sequence *before attending to the agent*. My basic proposal is that there are several cognitive operations that must be executed in order to put the brain into a state where it is ready to evoke and rehearse a sensorimotor sequence. I will consider the operation corresponding to TP first. My suggestion is that the head of TP describes an operation that determines whether an episode representation is to be retrieved from memory or gathered directly from experience, through the sensorimotor system. As already noted in Section 6.1, there is good evidence that the brain can exist in several distinct modes of connectivity, implemented by large-scale brain networks (Bressler and Menon, 2010). There is a well-studied distinction between a mode associated with memory retrieval and a mode associated with attention to external stimuli (see e.g. Sestieri *et al.*, 2011). While the cognitive operations that establish these alternative modes are not yet well understood, they are good candidates for the operations described by TP, which encodes the distinction between present and past sentences. Turning to CP, I suggest that this projection may describe an even earlier mode-setting cognitive operation. My focus has been on CP as it appears in sentences with sentential complements, such as *X says [that] P*. My proposal is that the complementiser *that* describes a cognitive operation establishing a special 'verbal mode', in which concepts are linked to words rather than to the world. A model of this operation and how infants learn to engage it is given in Caza and Knott (2012); see also Knott (in press). In Knott (2012) I also briefly suggest an interpretation of CP as it appears in questions. In this context, I suggest CP describes an operation that engages the cognitive mode in which queries can be posed to episodic memory. (Again there is good evidence that such a mode exists.) Question formation is traditionally seen as involving the raising of an inflectional head to the head of CP; I interpret this type of head-raising as a reflection of the actual mechanisms through which a query is posed to episodic memory.

I conclude by considering vP. The proposal that VP is introduced by a vP projection headed by a light verb makes a prediction about the process of action monitoring that takes place at the end of a sensorimotor routine. It predicts that action monitoring is more complex than is posited in Section 6.3: rather than being a single continuous process, it should have two well-defined temporal stages that occur in succession. To explore this prediction I have focussed on causative light verbs, originally hypothesised as part of an account of the causative alternation. Verbs undergoing this alternation can appear as transitives but also as intransitives: an example is *John opened the door / The door opened*. A common account of this alternation posits that the LF of *John opened the door* involves two VPs: a higher vP headed by the verb *cause*, introducing a complement VP headed by *open* (thus *John caused [the door opened]*). There is very good evidence for causative actions in the motor system, i.e. for actions that are represented by the perceptual effects they bring about (see Hommel *et al.*, 2001 for a review). Lee-Hand and Knott (2013) present a neural network model of the learning and control of actions defined by their perceptual effects; in this model, executing such an action involves activation of a network that controls a causative action, *and then* monitoring of the perceptual consequences of this action. This sequence of processing corresponds very well to the sequence predicted by the dual-VPs analysis. What is more, the model also offers an interesting sensorimotor interpretation of the special type of head raising that allows the lower verb *open* to raise to adjoin to the light verb *cause*.

There are several other projections posited within LF that remain to be considered: VoiceP, AspP and many others. Testing predictions about these projections is a matter for further work.

I will conclude by revisiting the agreement projections AgrSP and AgrOP. Several theorists have suggested that these can be dispensed with given the presence of other functional projections in the clause: for instance, it has been proposed that the head of vP can check accusative Case and the verb's object agreement features, and that TP can check the verb's subject agreement features (see e.g. the discussion in Hornstein *et al.*, 2005:162–8). Given we do not yet have a well worked out sensorimotor interpretation of vP and TP it is premature to assess these proposals in any detail—but if the above proposals about vP and TP are on the right track, then I suggest it is unlikely the sensorimotor interpretations of these projections overlap with those of AgrSP and AgrOP. For instance, in Lee-Hand and Knott's model of causative actions, the sensorimotor routine involved in executing the action of breaking a cup involves an action of attention to the cup (corresponding to AgrOP) and *then* activation of the causative action network (corresponding to vP). Discrepancies of this kind can push in two directions. On one hand they can indicate problems for the proposed sensorimotor interpretation of LF. On the other hand they can prompt further efforts to motivate a separate AgrOP projection through syntactic argumentation. At present it is not clear which direction will predominate.

10.3 A Reductionist Model of Syntax

The linking hypothesis I propose has a reductionist flavour: I want to explain (some) syntactic structures in language as manifestations of nonlinguistic cognitive phenomena. This is a direction which many linguists have pursued, particularly within the field of cognitive linguistics (see e.g. Lakoff, 1987; Langacker, 1987, 2008). But there are also phenomena which appear to be irreducibly syntactic. A good discussion about the limits of reductionist accounts of syntax is given by Jackendoff (2002). To take a small example, verbs with apparently similar semantics can introduce different prepositions (e.g. *count* *on* vs *trust* *in*). It is hard to see such differences as reflecting semantic differences. To take a more substantial example, transitive verbs are able to express a wide range of semantic structures. My cup-grabbing sentence features a verb which takes an agent and a patient, but other transitive verbs take apparently different argument types: *like* takes an experiencer and a stimulus, *frighten* takes a stimulus and an experiencer, *own* takes an 'owner' and an 'ownee' and so on. These verbs all project the same syntactic structure as 'grab', featuring positions for a subject and an object. But do they share the same semantic structure?

To begin with I should note that my model of 'surface language' certainly allows for surface syntactic idiosyncracies which have no origin in semantics, and allows that these idiosyncracies play a large part in the grammar of a language. The sentence generation network described in Section 9 can certainly learn the kinds of arbitrary dependency which feature in phrasal verbs like *count on* and *trust in*.

The question about apparent semantic diversity in the structures projected by *like*, *frighten*, *own* etc is more telling. If these verbs are syntactically identical, and syntactic structures are understood to reflect semantic structures, what semantic characterisation of subject and object position can we give which is general enough to apply to all these semantically disparate verbs?

My main response here is that interpreting a syntactic structure as describing stages in a sensorimotor routine is not the same as seeing it as reflecting a specific semantic pattern. I would certainly agree that different transitive verbs describe eventualities of very different types, and even that these are apprehended through operations in different cognitive modalities. All I am proposing is that there are regularities in the temporal structure of these operations, and that these are reflected in the structure of working memory episode representations, and as a result, in syntax.

The idea that subjects and objects are best defined with reference to the perceptual processing of an episode rather than to its intrinsic semantic structure has been suggested several times before. A common idea is that subjects describe participants with higher 'attentional prominence'; see for instance Langacker (2008); Talmy (2000); also Dowty (1991). One difference in my proposal is that perceptual processes are seen as having strong sequential structure, and subject and object are defined in relation to this structure rather than simply in relation to prominence. (In my account, both agent and patient become prominent, but at different times.) Another difference in my proposal is that it characterises subject and object positions at a language-independent level of syntactic representation (LF) rather than in surface sentence structure. This is helpful in accounting for the argument patterns of stimulus-experiencer verbs like *frighten*, which pose problems for most attempts to characterise argument positions semantically. If our semantic characterisations of subject and object are about positions at LF rather than PF, then we can account for such cases by arguing that the surface object appears higher than the surface subject *at LF*—and there are some good arguments along these lines (see e.g. Pesetsky, 1995; Anagnostopoulou, 1999). Of course these arguments rescue a theory of argument linking at the expense of more complex hypotheses about LF structures—but at least in my approach these hypotheses make predictions about sensorimotor processing which can be independently tested.

10.4 The Idea of Language-Independent LF Structures

My sensorimotor interpretation of LF structures also appears to make a very strong claim about the language-independence of LF. Sensorimotor processes are uncontroversially language-independent, but no Minimalist would want to claim that LF structures are fully invariant over translation: to take a famous example, *John swam across the river* must translate in French to *John traversa le fleuve en nageant* ('John crossed the river swimming'), which has a clearly different LF structure. More relevant to my cup-grabbing example, there are languages where transitive motor actions are easily or even canonically expressed in passive constructions (languages with ergative characteristics like Māori sometimes have this character; see e.g. Harlow, 2007). Where does this leave my sensorimotor account of LF?

Of course even within a single language there are often several alternative ways of expressing an episode syntactically. Any model of language semantics must rely heavily on an inference mechanism, which identifies commonalities in the semantic contributions of sentences which paraphrase one another, such as active and passive versions of a sentence. In relation to the sensorimotor model, this inference mechanism can perhaps be identified with the mechanism which updates an agent's representation of the current context when a deictic routine is completed. I assume an agent's representation of context is a rich, high-dimensional structure, which cannot be directly expressed in language. (In my model, some of the signals which provide input to this update operation have direct linguistic reflexes, but the update operation itself is complex, and learned through long sensorimotor experience.) If we allow that several sensorimotor routines can bring about roughly the same context update operation, then perhaps we can account for cases where LF structures are not preserved across languages by positing that languages can encode conventions about the routines through which particular updates are communicated. It is likely that a sizeable portion of the grammar of any language would have to be made up of conventions of this kind—and this portion of the grammar will probably have an 'empirist', construction-based flavour. But note that the conventions encoded in any given language will not be entirely arbitrary; they will have their origins in the alternative sensorimotor routines through which a given episode can be experienced, and which result in a given update. So a study of the sensorimotor system is still of use in identifying the alternative constructions from which conventions can be formed.

10.5 Deictic Routines

Another foundational assumption of my proposal is that sensorimotor processes are all structured as deictic routines. Does this idea stand up to scrutiny? Are sensorimotor processes structured as sequences at a certain timescale?

Of course in many ways there is massive parallelism in sensorimotor mechanisms. The account of deictic routines which I propose is quite consistent with this. In my model, the deictic operations involved in experiencing an episode progressively extend a neural circuit in which there is continuous and parallel processing. For instance, in the cup-grabbing example, when the observer attends to the agent, he initiates processing in a neural circuit which tracks the agent; when he attends to the cup, he initiates processing in a second circuit which tracks the cup, which is active in parallel with the first circuit, and in which makes use of the representations it generates. When the observer monitors the grab action, this initiates processing in a third circuit, which uses the representations generated by the first two circuits and runs in parallel to them. I also assume that there is parallel activity in neural circuits before they are selected. For instance, when the observer is deciding whether to engage the action-perception circuit or the action-execution circuit, we expect there to be activity in both these circuits concurrently, representing their claims to be selected.

The model of deictic routines nonetheless proposes that there are discrete changes in the neural circuitry active during the apprehension of concrete events. In Ballard *et al.*'s original conception, this general idea was mainly supported by analyses of discrete elements of behaviour, in particular saccadic eye movements.

But it is also supported by the recent discovery of large-scale brain networks (Bressler and Menon, 2010), which are activated or deactivated as wholes by distinct neural operations (Sridharan *et al.*, 2008; Menon and Uddin, 2010), and which appear to have hierarchical structure (Doucet *et al.*, 2011).

10.6 Beyond Concrete Sentences

I have proposed a hypothesis linking the LF structure of concrete sentences to sensorimotor mechanisms. But there are abstract sentences with the very same LF as my example cup-grabbing sentence, for instance *The company acquired a subsidiary*: in this case, the LF structure cannot describe a sensorimotor process in any direct way. Clearly my hypothesis about concrete sentences commits me to some related claim about sentences like these.

A well-known approach taken by linguists interested in embodiment is to argue that abstract sentences aquire their meaning through metaphors grounded in concrete domains (see classically Lakoff and Johnson, 1980). Thus for instance if 'company' is metaphorically an agent and 'subsidiary' is metaphorically an object, then 'acquire' can have a meaning similar to the motor action 'take' or 'get'. If this idea is correct, then the sensorimotor interpretation of LF could possibly be extended to abstract sentences by proposing that the propositions they describe are apprehended and stored as deictic routines, operating over conceptualised or simulated objects rather than actual objects in the world. However, my feeling is that formulating a detailed hypothesis about how abstract propositions are grounded metaphorically in concrete domains first requires a thorough understanding of how concrete episodes are apprehended—which we are far from attaining. As a point of methodology, therefore, I think it may be premature to attempt a detailed metaphor-based account of abstract sentences.

Another way of looking beyond concrete sentences is to consider sentences that do more than report experienced episodes. Sentences can express desires, ask questions, report memories, and do many other things: these capabilities can be traced to particular XPs at LF, and ultimately my interpretation of LF should extend to these XPs too. In order to move in this direction, a natural strategy is to broaden the concept of a 'deictic operation', which currently applies only to sensory and motor actions, so that it includes other types of cognitive operation—for instance, operations which manipulate working-memory representations or which perform storage or retrieval of material in long-term memory. Our interpretation of the X-bar schema would then allow XPs to describe cognitive operations of this kind as well as sensorimotor operations. If there are XPs which can plausibly be interpreted as signalling purely cognitive operations, the LF structures in which they appear may provide interesting ways of thinking about the sequential organisation of these operations. The proposed interpretations of CP and TP discussed in Section 10.2 in fact move in this direction.

11 Summary

In this paper I have made a suggestion about how the rich and complex information received by the senses during the apprehension of a simple reach-to-grasp action is compressed into a linguistic representation. Obviously there is a huge amount of compression; my main suggestion is that the basis for this compression is the temporal structure of the sensorimotor processes—specifically, its structure as a deictic routine. Whether this idea can be successfully extended beyond the example cup-grabbing scenario is a matter for further work.

The proposal I have made about the interface between language and the sensorimotor is expressed in terms of Chomsky's Minimalist model. But the proposal requires some quite radical reinterpretations to Minimalism, particularly of its account of the derivation of LF structures. These revisions allow the Minimalist model to be supplemented with accounts of sentence processing, surface syntax and syntactic development derived from empiricist models of language which are normally regarded as Minimalism's competitors. I still maintain that the Mimimalist conception of LF provides a very helpful framework for a strong hypothesis about how syntactic structures relate to neural mechanisms. But at the same time, this hypothesis may help to restate some of the key insights of Minimalism in a way which is more compatible with alternative conceptions of syntax which are currently more widespread within cognitive science.

References

Abney, S. 1987. *The English Noun Phrase in its Sentential Aspect*. Cambridge, MA: MIT dissertation.

Abraham, W, B Logan, J Greenwood & M Dragunow. 2002. Induction and experience-dependent consolidation of stable long-term potentiation lasting months in the hippocampus. *Journal of Neuroscience* 22. 9626–9634.

Alexiadou, A, L Haegeman & M Stavrou. 2007. *Noun Phrase in the Generative Perspective*. Berlin: Mouton de Gruyter.

Anagnostopoulou, E. 1999. On experiencers. In A Alexiadou, G Horrocks & M Stavrou (eds.), *Studies in Greek Syntax*, 67–93. Dordrecht: Kluwer.

Arbib, M. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2). 105–167.

Averbeck, B, M Chafee, D Crowe & A Georgopoulos. 2002. Parallel processing of serial movements in prefrontal cortex. *PNAS* 99(20). 13172–13177.

Averbeck, B & D Lee. 2007. Prefrontal correlates of memory for sequences. *Journal of Neuroscience* 27(9). 2204–2211.

Baddeley, A. 2000. The episodic buffer: A new component of working memory? *TICS* 4(11). 417–423.

Baddeley, A & G Hitch. 1974. Working memory. In G Bower (ed.), *The psychology of learning and motivation*, 48–79. Academic Press.

Ballard, D, M Hayhoe, P Pook & R Rao. 1997. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20(4). 723–767.

Barone, P & J-P Joseph. 1989. Prefrontal cortex and spatial sequencing in macaque monkey. *Experimental Brain Research* 78. 447–464.

Barsalou, L. 2008. Grounded cognition. *Annual Review of Psychology* 59. 617–645.

Barsalou, L, W Simmons, A Barbey & C Wilson. 2003. Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences* 7(2). 84–91.

Beiser, D & J Houk. 1998. Model of cortical-basal ganglionic processing: Encoding the serial order of sensory events. *Journal of Neurophysiology* 79(6). 3168–3188.

Bickerton, D. 1995. *Language and Human Behavior*. London: University College London Press.

Bookheimer, S. 2002. Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25. 151–188.

Bressler, S & V Menon. 2010. Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences* 14. 277–290.

Caza, G & A Knott. 2012. Pragmatic bootstrapping: A neural network model of vocabulary acquisition. *Language Learning and Development* 8. 1–23.

Chambon, V, D Wenke, S Fleming, W Prinz & P Haggard. 2012. An online neural substrate for sense of agency. *Cerebral Cortex* 23. 1031–1037.

Chang, F. 2002. Symbolically speaking: A connectionist model of sentence production. *Cognitive Science* 26. 609–651.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. 1970. Remarks on nominalization. In R Jacobs & P Rosenbaum (eds.), *Reading in English Transformational Grammar*, 184–221. Waltham: Ginn.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1995a. Bare phrase structure. In H Campos & P Kempchinsky (eds.), *Evolution and Revolution in Linguistic Theory. Essays in honor of Carlos Otero*, Washington, DC: Georgetown University Press.

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

Christiansen, M & N Chater. 1999. Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science* 23. 157–205.

Cisek, P & J Kalaska. 2010. Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33. 269–298.

Cohen, J, E Berkman & M Lieberman. 2013. Intentional and incidental self-control in ventrolateral PFC. In D Stuss & R Knight (eds.), *Principles of Frontal Lobe Functions: Second Edition*, 417–440. Cambridge, MA: MIT Press.

Colby, C & M Goldberg. 1999. Space and attention in parietal cortex. *Annual Review of Neuroscience* 22. 319–349.

Corbett, G. 1991. *Gender*. Cambridge/New York: Cambridge University Press.

Critchley, H, C Mathias & R Dolan. 2003. Neuroanatomical basis for first- and second-order representations of bodily states. *Nature Neuroscience* 4(2). 207–212.

Damasio, A (ed.). 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.

di Pellegrino, G, L Fadiga, L Fogassi, V Gallese & G Rizzolatti. 1992. Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91. 176–180.

Diba, K & G Buzsàki. 2007. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience* 10(10). 1241–1242.

Domahs, F, A Nagels, U Domahs, C Whitney, R Wiese & T Kircher. 2012. Where the mass counts: Common cortical activation for different kinds of nonsingular-

ity. *Journal of Cognitive Neuroscience* 24(4). 915–932.
Dominey, P, M Arbib & J-P Joseph. 1995. A model of corticostriatal plasticity for learning associations and sequences. *Journal of cognitive neuroscience* 7(3). 311–336.
Doucet, G, M Naveau, L Petit, N Delcroix, L Zago, F Crivello, G Jobard, N Tzourio-Mazoyer, B Mazoyer, E Mellet & M Joliot. 2011. Brain activity at rest: a multiscale hierarchical functional organization. *Journal of Neurophysiology* 105. 2753–2763.
Dowty, D. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
Eichenbaum, H. 2004. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* 44. 109–120.
Elman, J. 1990. Finding structure in time. *Cognitive Science* 14. 179–211.
Fadiga, L, L Craighero & E Olivier. 2005. Human motor cortex excitability during the perception of others' action. *Current opinion in neurobiology* 15(2). 213–218.
Fagg, A & M Arbib. 1998. Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks* 11(7/8). 1277–1303.
Farrer, C, S Frey, J Van Horn, E Tunik, D Turk, S Inati & S Grafton. 2008. The angular gyrus computes action awareness representations. *Cerebral Cortex* 18. 254–261.
Farrer, C & C Frith. 2002. Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage* 15. 596–603.
Feldman, J & S Narayanan. 2004. Embodiment in a neural theory of language. *Brain and Language* 89(2). 385–392.
Ferri, F, G Campione, R Dalla Volta, C Gianelli & M Gentilucci. 2010. To me or to you? when the self is advantaged. *Experimental Brain Research* 203. 637–646.
Fillmore, C, P Kay & M O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64. 501–38.
Flanagan, J & R Johansson. 2003. Action plans used in action observation. *Nature* 424. 769–771.
Foster, D & M Wilson. 2006. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* 440. 680–683.
Freedman, D, M Riesenhuber, T Poggio & E Miller. 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorisation. *Journal of Neuroscience* 15. 5235–5246.
Gallese, V, L Fadiga, L Fogassi & G Rizzolatti. 1996. Action recognition in the premotor cortex. *Brain* 119. 593–609.
Gallese, V & A Goldman. 1998. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences* 2(12). 493–501.
Geyer, S, M Matelli, G Luppino & K Zilles. 2000. Functional neuroanatomy of the primate isocortical motor system. *Anatomy and Embryology* 202(6). 443–474.
Glenberg, A & P Kaschak. 2002. Grounding language in action. *Psychonomic Bulletin and Review* 9(3). 558–565.
Goldberg, A. 1995. *Constructions. A Construction Grammar approach to Argument Structure*. Chicago: University of Chicago Press.
Goldstein, L & C Fowler. 2003. Articulatory phonology: A phonology for pub-

lic language use. In A Meyer & N Schiller (eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, 159–207. New York: Mouton.

Goodwin, A & H Wheat. 2004. Sensory signals in neural populations underlying tactile perception and manipulation. *Annual Review of Neuroscience* 27. 53–77.

Grèzes, J & J Decety. 2001. Functional anatomy of execution, mental simulation, observation and verb generation of actions: A meta-analysis. *Human Brain Mapping* 12(1). 1–19.

Haggard, P. 2008. Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience* 9. 934–946.

Hamker, F. 2004. A dynamic model of how feature cues guide spatial attention. *Vision Research* 44. 501–521.

Harlow, R. 2007. *Māori: A Linguistic Introduction*. Cambridge, UK: Cambridge University Press.

Hasegawa, R, M Matsumoto & A Mikami. 2000. Search target selection in monkey prefrontal cortex. *Journal of Neurophysiology* 84. 1692–1696.

Henson, R, N Burgess & C Frith. 2000. Recoding, storage, rehearsal and grouping in verbal short-term memory: An fMRI study. *Neuropsychologia* 38(4). 426–440.

Hommel, B, J Müsseler, G Aschersleben & W Prinz. 2001. The theory of event coding (TEC): A framework for perception and action learning. *Behavioral and Brain Sciences* 24. 849–878.

Hornstein, N. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge, UK: Cambridge University Press.

Hornstein, N, J Nunes & K Grohmann. 2005. *Understanding Minimalism*. Cambridge, UK: Cambridge University Press.

Houghton, G & T Hartley. 1995. Parallel models of serial behaviour: Lashley revisited. *Psyche* 2(25).

Jackendoff, R. 1977. *X-Bar syntax: A study of phrase structure.* Linguistic Inquiry Monograph 2. Cambridge, MA: MIT Press.

Jackendoff, R. 1999. Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences* 3(7). 272–279.

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

Jeannerod, M. 1996. Reaching and grasping. Parallel specification of visuomotor channels. In W Prinz & B Bridgeman (eds.), *Handbook of Perception and Action Volume II: Motor Skills*, 405–460. Academic Press.

Jeannerod, M. 2001. Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage* 14. S103–S109.

Jellema, T, C Baker, B Wicker & D Perrett. 2000. Neural representation for the perception of the intentionality of actions. *Brain and Cognition* 44. 280–302.

Johansson, R, G Westling, A Backstrom & J Flanagan. 2001. Eye-hand coordination in object manipulation. *Journal of Neuroscience* 21(17). 6917–6932.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Knott, A. 2012. *Sensorimotor Cognition and Natural Language Syntax*. Cambridge, MA: MIT Press.

Knott, A. in press. How infants learn word meanings and propositional attitudes: a neural network model. In T-W Hung (ed.), *Language and Action*, Berlin/Heidelberg: Springer.

Lakoff, G. 1987. *Women, Fire and Dangerous Things*. Chicago and London: University of Chicago Press.

Lakoff, G & M Johnson. 1980. *Metaphors We Live By*. Chicago and London: University of Chicago Press.

Land, M & S Furneaux. 1997. The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London B* 352. 1231–1239.

Langacker, R. 1987. Nouns and Verbs. *Language* 63(1). 53–94.

Langacker, R. 2008. *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.

Lee-Hand, J & A Knott. 2013. A model of causative actions: from motor learning to syntactic structure. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, Berlin.

Levelt, W. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levelt, W, A Roelofs & A Meyer. 1999. A theory of lexical access in speech production. *Behavioural and Brain Sciences* 22(1). 1–75.

Lieven, E, J Pine & G Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24. 187–219.

MacWhinney, B. 2005. Item-based constructions and the logical problem. In *Proceedings of the Workshop on Psychocomputational Models of Human Language Acquisition*, 53–68.

Marantz, A. 2005. Generative linguistics within the cognitive neuroscience of language. *Linguistic Review* 22. 429–445.

McClelland, J, B McNaughton & R O'Reilly. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 102. 419–457.

Menon, V & L Uddin. 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function* 214. 655–667.

Meteyard, L, S Rodriguez Cuadrado, B Bahrami & G Vigliocco. 2012. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48. 788–804.

Murata, A, V Gallese, G Luppino, M Kaseda & H Sakata. 2000. Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology* 83(5). 2580–2601.

Novick, J, J Trueswell & S Thompson-Schill. 2005. Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective and Behavioural Neuroscience* 5(3). 263–281.

Oztop, E & M Arbib. 2002. Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics* 87. 116–140.

Oztop, E, N Bradley & M Arbib. 2004. Infant grasp learning: a computational model. *Experimental Brain Research* 158. 480–503.

Oztop, E, D Wolpert & M Kawato. 2005. Mental state inference using visual control parameters. *Cognitive Brain Research* 22. 129–151.

Pantel, P & D Lin. 2002. Discovering word senses from text. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 613–619.

Perani, D, S Cappa, T Schnur, M Tettamanti, S Collina, M Rosa & F Fazio. 1999.

The neural correlates of verb and noun processing—A PET study. *Brain* 122. 2337–2344.

Pesetsky, D. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge: MIT Press.

Petreska, B & A Billard. 2009. Movement curvature planning through force field internal models. *Biological Cybernetics* 100. 331–350.

Plate, T. 2003. *Holographic Reduced Representations* CSLI Lecture Notes Number 150. Stanford, CA: CSLI Publications.

Poeppel, D & D Embick. 2005. The relation between linguistics and neuroscience. In A Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, Hillsdale, NJ: Lawrence Erlbaum.

Pollard, C & I Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press.

Pulvermüller, F, W Lutzenberger & H Preissl. 1999. Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral Cortex* 9(5). 497–506.

Pulvermüller, F. 2010. Brain embodiment of syntax and grammar: Discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain and Language* 112. 167–179.

Ramsey, S & A Hamilton. 2010. Understanding actors and object-goals in the human brain. *NeuroImage* 50. 1142–1147.

Rhodes, B, D Bullock, W Verwey, B Averbeck & M Page. 2004. Learning and production of movement sequences: Behavioral, neurophysiological, and modeling perspectives. *Human Movement Science* 23. 699–746.

Rizzolatti, G & M Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21. 188–194.

Sahin, N, S Pinker, S Cash, D Schomer & E Halgren. 2009. Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326. 445–449.

Sestieri, C, M Corbetta, G Romani & G Shulman. 2011. Episodic memory retrieval, parietal cortex, and the default mode network: Functional and topographic analyses. *Journal of Neuroscience* 31(12). 4407–4420.

Shapiro, K & A Caramazza. 2003. The representation of grammatical categories in the brain. *Trends in Cognitive Sciences* 7(5). 201–206.

Shapiro, K, L Moo & A Caramazza. 2012. Neural specificity for grammatical operations is revealed by content-independent fMR adaptation. *Frontiers in Psychology* 3. Doi: 10.3389/fpsyg.2012.00026.

Shastri, L. 2001. Episodic memory trace formation in the hippocampal system: A model of cortico-hippocampal interaction. Tech. Rep. TR-01-004 International Computer Science Institute (ICSI), UC Berkeley Berkeley, CA.

Shastri, L. 2002. Episodic memory and cortico-hippocampal interactions. *Trends in Cognitive Sciences* 6(4). 162–168.

Shima, K & J Tanji. 2000. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *Journal of Neurophysiology* 84(4). 2148–2160.

Sridharan, D, D Levitin & V Menon. 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*

105(34). 12569–12574.
Steedman, M. 2002. Plans, affordances and categorial grammar. *Linguistics and Philosophy* 25. 723–753.
Swallow, K, J Zacks & R Abrams. 2009. Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General* 138(2). 236–257.
Takac, M, L Benuskova & A Knott. 2012. Mapping sensorimotor sequences to word sequences: A connectionist model of language acquisition and sentence generation. *Cognition* 125. 288–308.
Takac, M & Knott. 2013. A neural network model of working memory for episodes. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, Berlin.
Talmy, L. 2000. *Towards a Cognitive Semantics. Volume 1: Concept Structuring Systems*. MIT Press.
Tanenhaus, M, M Spivey-Knowlton, K Eberhard & J Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268. 1632–1634.
Tomasello, M. 2003. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
Tomita, H, M Ohbayashi, K Nakahara & Y Hasegawa, I Miyashita. 1999. Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401. 699–703.
Tranel, D, R Adolphs, H Damasio & A Damasio. 2001. A neural basis for the retrieval of words for actions. *Cognitive Neuropsychology* 18(7). 655–674.
van der Velde, F & M de Kamps. 2006. Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences* 29. 37–108.
Wallenstein, G, H Eichenbaum & M Hasselmo. 1998. The hippocampus as an associator of discontiguous events. *Trends in Neurosciences* 21. 317–323.
Walles, H. 2010. *A Perceptual Basis for Noun Phrase Syntax*: Department of Computer Science, University of Otago dissertation.
Webb, A, A Knott & M MacAskill. 2010. Eye movements during transitive action observation have sequential structure. *Acta Psychologica* 133. 51–56.
Wray, A. 1998. Protolanguage as a holistic system for social interaction. *Language and Communication* 18. 47–67.
Yamaguchi, S, S Yamagata & S Kobayashi. 2000. Cerebral asymmetry of the "top-down" allocation of attention to global and local features. *Journal of Neuroscience* 20(9). RC72.
Zamparelli, R. 2000. *Layers in the Determiner Phrase*. New York: Garland.
Zwaan, R & L Taylor. 2006. Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General* 135(1). 1–11.

Alistair Knott
Department of Computer Science
University of Otago
P.O. Box 56, Dunedin
New Zealand
alik@cs.otago.ac.nz

⋆ FORUM ⋆

From Comparative Languistics to Comparative (Bio)linguistics: Reflections on Variation

Evelina Leivada

1. Introduction

In *Me and Chomsky: Remarks from Someone Who Quit*, Sascha Felix writes about the nature and the orientation of current work in the field of (comparative) linguistics:

> In some sense I feel that much (but obviously not all) of current linguistic work displays a relapse to the spirit prevailing in pre-Chomskyan times. *Linguistics is about describing language data. Period. Beyond this there is no deeper epistemological goal. Of course, those who became linguists because they like to play around with language data could not care less, because they can pursue their interests under any development of the field, nowadays possibly with less pressure and stress.* Personally I felt that much of what I was offered to read in recent years was intolerably boring and that the field of linguistics was becoming increasingly uninteresting and trivialized". (Felix 2010: 71, emphasis added)

On the surface, many linguists claim their work to be driven by an interest to understand and describe the biological underpinnings of the faculty of language (henceforth, FL). However, a more careful look at the literature shows that this interest is not always reflected in the bulk of their work. Despite the fact that linguists are often quick to acknowledge an interest in core properties of FL, it seems that this interest fades away and the focus shifts from FL to particularities of grammar—described in highly technical detail—that would not mean much if the real focus was on FL, in the sense that the specific realizations of a grammatical phenomenon across languages might have a place in the grammar books dedicated to these languages, but not in a book about FL and human cognition. In other words, it seems that there is a divide between *linguistics* (or biolinguistics, with focus on FL) and *languistics* (with focus on detailed descriptions of grammars or what Felix calls 'language data')—a state of affairs reminiscent of the distinction between biolinguistics in the strong and biolinguistics in the weak sense (Boeckx & Grohmann 2007).

The biolinguistic enterprise, in its current state of development, aims to address five key issues, each of which can be formulated as a question. Boeckx & Grohmann (2007: 1), following Chomsky (1986) and in effect going back to Tinbergen (1963), reproduce the questions as follows:

I am grateful to Cedric Boeckx and Pedro Tiago Martins for comments and feedback. I acknowledge support through grant FFI2010-20634 from the Spanish Ministerio de Economia y Competitividad and a scholarship from the Cyprus State Scholarship Foundation.

ISSN 1450–3417

Biolinguistics 8: 053–066, 2014
http://www.biolinguistics.eu

(1) What is knowledge of language?
(2) How is that knowledge acquired?
(3) How is that knowledge put to use?
(4) How is that knowledge implemented in the brain?
(5) How did that knowledge emerge in the species?

Despite the frequently acknowledged interest in these questions, it seems to be the case that this conception of the discipline is not really depicted in discussions that deal exclusively with particular grammatical phenomena, and this comes at a cost for the discipline itself. The existence of this linguistics/ languistics divide is at times problematic when one seeks to establish truly interdisciplinary bridges between linguistics and neurobiology, due to a granularity mismatch between the primitives on which each discipline operates (Poeppel & Embick 2005). In Hornstein's (2013) words, it seems that "[t]here really is a linguistics/languistics divide that is quite deep, with a very large part of the field focused on the proper description of language data in all of its vast complexity as the central object of study. Though, there is no *a priori* reason why this endeavor should clash with the biolinguistic one, in practice it does".

The observed clash could be the result of linguists employing folk biology of language when discussing FL as a component of the human mind/brain. For example, linguists (at least those within the generative enterprise) have often followed Chomsky (2005) in assuming the three factors identified there as crucial components of language design. They also followed Chomsky in calling the first factor in language design 'Universal Grammar' (UG) and further describing it as the genetic endowment for language, again following Chomsky (2005).[1] It is highly likely that this narrow, genocentric vision of UG will prove problematic, particularly so when it comes to the integration and assimilation of results from linguistics into biology, which has progressively moved away from its genocentrism (Pigliucci & Müller 2010). Another reason for the clash Hornstein talks about could be the diversity of interdisciplinary insights that the two fields (comparative languistics and comparative biolinguistics) encompass: There are considerations about FL that are dealt with in a narrower way within the former than within the latter. For example, comparative languistics tends to favor a more narrow and restricted view of variation; what Benítez-Burraco & Boeckx (in press) refer to as "deal[ing] with variation 'at the surface' (dialects, languages, sociolects)". However, as they point out, a novel, comparative approach within the realm of (bio)linguistics should ask questions that aim to uncover the locus of variation (and its constraints) across genotypes, pathologies, or across species. This comparative biolinguistic approach entails the integration of various insights from the literature on evolutionary biology, genetics, paleoanthropology, clinical linguistics, and studies on externalization and variation across species.

More concretely, this novel approach seeks to bring the study of language within the fold of an Extended Synthesis in biology (Pigliucci & Müller 2010) and

---

[1] See Lorenzo & Longa (2009) for a list of studies that make reference to UG as 'blueprint', 'genetic endowment', or 'genetic equipment'.

to answer key questions about the nature of variation and its constraints across languages, pathologies, and species. This was the motivation behind organizing a two-day event (*Comparative Biolinguistics: An Exploratory Workshop*; henceforth, CBL) held in November 2013, at the University of Barcelona's Department of General Linguistics. The goal of CBL was to make progress with respect to establishing interdisciplinary linking hypotheses within a perspective of extended synthesis for FL. As the description of the talks below suggests, evolutionary biology, brain imaging, and clinical linguistics were the main points of departure for the presenters in this workshop.

2. Comparative Biolinguistics: An Exploratory Workshop

The goals of CBL were articulated in detail in the two talks that opened the event, delivered by *Antonio Benítez-Burraco* (Universidad de Huelva) and *Cedric Boeckx* (ICREA/Universitat de Barcelona).

The first talk was oriented towards providing the reasons for developing a new research program such as the one described above. The notion of variation was the main theme of Benítez-Burraco's presentation. He discussed the amount (and kind) of variation that we find in language and how we need to properly come to grips with it if we want to contribute to a real characterization of the biological foundations of language. In line with what is argued in Benítez-Burraco & Boeckx (in press), Benítez-Burraco suggested that a direct link between language features and the genome and a conflation of geneticism with nativism (i.e. first factor in language design = UG = linguistic genotype) are not likely to represent any progress in understanding the biological underpinnings of FL, because genes do not work this way. In his words, a direct link between the genotype and the phenotype is not only simplistic, but biologically untenable, given the way in which genes contribute to developmental processes and how development actually takes place. Genes are not blueprints and developmental processes also depend on non-genetic factors.

Under these assumptions, variation was argued to be constrained, with only some of the possible pathological phenotypes being actually realized. In other words, some aspects of language processing seem to be vulnerable in all related pathological conditions, while others seem to be preserved across pathological conditions. For instance, inflectional morphology is problematic across different pathologies, whereas operations of narrow syntax are never shown to be problematic (see Benítez-Burraco & Boeckx, in press, for discussion and references). The take-home message was that it is evident that those interested in discussing the notion of variation from a comparative biolinguistics perspective should rely on certain key Evo-Devo concepts (e.g., canalization, development plasticity, robustness, evolvability, adaptive landscapes, etc.).

In the second talk, Boeckx discussed possible tools for comparative biolinguistics. He argued that in order to advance the new research program, attention has to be paid to the tools one uses when establishing the relevant comparisons and pinpointing the limits of variation. Focusing on the Chomsky hierarchy, the three main conclusions drawn in his talk were: (i) choosing to use this tool entails ignoring the fact that the Chomsky hierarchy does uniquely characterize FL since no region of the hierarchy can be identified as unique to FL; (ii) most comparative biolinguistic experiments using the Chomsky hierarchy are artificial language experiments that target the learning of purely syntactic patterns, stripped off semantic consequences. However, natural languages do not work this way, since there are no syntactic patterns without semantic patterns, (iii) there are inherent limitations of the Chomsky hierarchy in capturing the constraints of variation. Boeckx's point here is that even if we discovered that other species were capable of mastering a mildly context-sensitive language, by running an artificial language experiment, this would not be very informative, because it would not provide any information about the algorithm that was used when learning the patterns in question. All in all, this talk highlighted the fact that linguists have some translation work to do before they are able to use the right linguistic tools into a broader comparative framework.

The next set of talks, by *Aritz Irurtzun* (IKER/CNRS) and *Maia Duguine* (University of the Basque Country), combined different linguistic perspectives in an effort to understand the locus and limits of variation in language. Irurtzun focused on prosodic constraints in linguistic theory. Having reviewed a great number of linguistic representations from a variety of languages, he argued that cross-linguistic prosodic variability escapes any surface-based generalization and that this amounts to a constraint on the (bio)linguistic theories we construe. He also stressed the fact that a potential problem of going abstract (i.e. substance-free)[2] is that we may lose the external plausibility for the restrictions we may pose; a fact that should not be taken as a problem if the goal is to describe FL. Duguine talked about the nature of parametric variation, and her discussion targeted a specific parameter: *pro*-drop. In the field of theoretical linguistics, the locus of variation has been frequently described in the form of syntactic or lexical variants that are encoded in the initial state of FL (i.e. a parametric UG, following Chomsky 1965 *et seq*.). Since Rizzi's (1986) work on null subjects, languages are usually classified either as allowing null subjects (*pro*-drop languages such as Italian) or not (non-*pro*-drop languages such as English). Duguine, however, showed how the *pro*-drop phenomenon cannot be reconstructed as a lexical parameter. Going through data from different languages, she suggested that there is no set of formal features that forms a class which would effectively identify and group the relevant items together. In her words, lexical parameters are defined by the formal features of functional categories, subject to cross-linguistic variation, but when one asks what is the property (i.e. formal feature) that sets apart Italian, Catalan, Japanese, etc., from French, English, German, etc., the answer that surfaces is that there is no such property. The reasonable conclusion to draw would be that likewise there is no such parameter as *pro*-drop. Based on Duguine's arguments about *pro*-drop, one might wonder whether the point she makes can be valid for other parameters as well; and it probably is. Put differently, *pro*-drop being one of the few standard textbook examples of a (lexical) parameter, one wonders whether there really exists in the literature a single

[2] Substance-free approaches (e.g., Hale & Reiss 2000) in phonology argue for a light, simple phonological component of UG, plausibly deprived of phonetic biases, consisting of a core computational system that is ready to naively receive input and manipulate it.

example of a lexical parameter that can be accurately classified as such. The answer seems to be negative. This paves the way for a unified approach to the *pro*-drop phenomenon in phonological terms based on a PF-deletion analysis of dropping (Duguine 2013), and then for extending an analysis along these lines to other 'lexical' or 'syntactic' parameters identified as such in the literature.

The next four talks of the first day of CBL aimed to contribute different insights to the research agenda of comparative biolinguistics by sharing different, yet related, points of departure. *Lluís Barceló-Coblijn* (Universidad de Murcia) talked about hominins through communication and language. This was related to how language emerged in the species (question (5) above), more than to what language consists of in linguistic terms (as in the two previous talks), what are the sources of variation within FL and how the issue of variation should be dealt with in the new research agenda. Barceló-Coblijn brought up an important aspect of CBL (i.e. the difference between us and extinct hominins) and discussed the possibility of figuring out if extinct hominins had aspects of our language faculty.

In the evolution literature, admixture between other hominins and anatomically modern humans has recently been confirmed (e.g., modern humans and late Neanderthals, as in the case of a child from Lagar Velho; Duarte *et al*. 1999). However, there is no direct link between genotype and phenotype; development is influenced by factors other than the genes (Oyama 2000), and the observed developmental trajectories (and plausibly the cognitive abilities they finally support) are modeled by other factors (as mentioned above, canalization, development plasticity, robustness, evolvability, adaptive landscapes, etc.). Eventually, it can be claimed that "other hominins could have had a 'linguistic system' […], [h]owever, the available data suggests that the 'languages' they plausibly spoke would have lacked some defining properties of human languages, particularly, complex syntax" (Benítez-Burraco & Barceló-Coblijn 2013: 241).

Tobias Scheer (Université de Nice), in perhaps the most interactive talk of the event, engaged in an extensive commentary on the outline of the comparative biolinguistics program as sketched in Benítez-Burraco & Boeckx (in press) and in the first two talks of CBL. He dealt with a variety of topics such as the mind–brain relation, the variability of FL, modularity, language universals, language pathologies, and different arguments for or against scenarios on the adaptive emergence of language. Lastly, he addressed one of the most important topics on the current (bio)linguistic agenda: third factor patterns. Given the exploratory character of CBL and also the at times different perspectives held by the participants, his direct engagement with the material presented by Antonio Benítez-Burraco and Cedric Boeckx provided the means for sharpening the understanding of the linking hypotheses that are to be established between the different disciplines that are to be integrated into the new research program. His basic departure from Benítez-Burraco & Boeckx (in press) was his dualist position; his emphasis was on us being able to focus on the mind, without however having a clear definition on what the mind is.

Rie Asano (University of Cologne) and *Uwe Seifert* (University of Cologne) brought into the discussion yet another perspective in addressing the relation between biolinguistics and (bio)musicology. Asano focused on theoretical and neuroscientific considerations that are relevant when comparing syntax in music and language as well as the implications that this comparative approach carries for the evolution of language. Seifert's talk was oriented towards offering an evolutionary framework for comparative research on the functional architecture of the musical mind. Both talks can be related first to biolinguistic concerns about the uniqueness of language or, better, about identifying the uniquely human core mechanisms (say, syntactic mechanisms) and second to understanding whether these mechanisms are indeed unique to FL (see Hauser *et al*. 2002 and Fitch *et al*. 2005 on the potential candidates for this uniqueness) or whether they can exhibit parallels to music or other domains of human cognition. This issue is a frequently addressed topic in linguistics, particularly so ever since Hauser *et al*. (2002) introduced the distinction of FL in the narrow and in the broad sense, and argued that syntactic recursion (i.e. Merge) is a potential candidate for FLN(arrow). It is important to highlight here the points of commonality between language and music. Both are innate and universal, they constitute part of the great leap forward, they seem to have parallels in other species, there seems to be structure and hierarchy in both, etc.

On the second day, CBL had a focus on pathologies. Most of the talks presented experimental results from studies on typical or atypical language acquisition/performance in child and adult populations. Aiming to provide robust comparisons across a variety of pathological conditions, the idea was to solicit contributions that discuss impairments in speakers of the same language, so as to keep the reported predictions and/or results as comparable as possible. Coming from the Cyprus Acquisition Team, *Kleanthes K. Grohmann* (University of Cyprus), *Maria Kambanaros* (Cyprus University of Technology), *Eleni Theodorou* (University of Cyprus), and *Elena Papadopoulou* (University of Cyprus) started the second day of CBL with specifying aspects of Specific Language Impairment (SLI) in Cypriot Greek. In their talk, they highlighted four areas in the domain of syntax: clitic placement, comprehension and production of relative clauses, sentence repetition performance, and *wh*-questions. The main focus of their investigations were atypically developing children, in particular the identification of SLI; however, the discussion was not restricted to that. Instead, their findings were put in perspective by comparing SLI with other syndromes and by administering similar experimental tools to patients with Broca's aphasia, traumatic brain injury, and dementia of the Alzheimer's type.

In the next three talks of that day, different studies were presented, each of them focusing on different atypical populations and/or experimental tasks. First, *Christiana Christodoulou* (University of Cyprus & MIT) and *Kleanthes K. Grohmann* discussed the grammar of Down Syndrome. In line with Christodoulou (2011), the aim was to see whether the differences in Down Syndrome grammars and typical grammars are (i) syntactically, (ii) morphologically, or (iii) phonologically and phonetically conditioned. The reported results are based on a broad variety of tasks (i.e. combinations of visual and audio stimuli with guided production, elicited imitation, and storytelling) and touch upon all aspects of the grammar under investigation. Findings point out phonetically conditioned differences between the groups, with a small residue of morphologically and phonologically conditioned differences. These findings seem to grant experimental support to the view of variation that is entertained in Benítez-Burraco & Boeckx (in press): It seems to be the case that variation is constrained, with only some of the possible pathological phenotypes actually realized. In fact, the conclusion to be drawn based on the experimental work presented at CBL, but also from a broader literature review across pathologies, is that morphophonology might be impaired, but syntax is preserved. Put differently, we do not know of an atypical population that is unable to syntactically combine different elements or that manipulates syntactic objects in an atypical way that is not a licit option in unimpaired syntax (e.g., to add negation by moving the third element of the clause to the sentence-initial position), whereas we repeatedly find atypical patterns of various aspects of morphophonology across disorders.

Maria Kambanaros and *Kleanthes K. Grohmann* talked next about verb–noun dissociations across language-impaired populations. They presented the results obtained from administering the Greek Object and Action Naming task (Kambanaros 2003) to three different populations: (i) adults with aphasia, (ii) adults with schizophrenia, and (iii) children with SLI. This task assesses lexical retrieval of object and action names, and the results showed a verb–noun dissociation across the populations under study (see Kambanaros *et al*. 2010, 2014 for detailed presentations of the results). It is interesting to note here that the reported errors are similar: Different populations produced similar results in the sense that they employed circumlocutions (light verbs, e.g., 'make a house' instead of the target 'build') or superordinate terms ('tool' for 'hammer').

Crucially, the types of substitutions that were generated by this research clearly indicate a by now recognizable pattern: preserved syntactic abilities (as suggested by the overt realization of the phase head when a light verb construction is produced instead of a single verb) and preserved semantic abilities (as suggested by the ability to activate semantically relevant lemmas that might not, however, correspond to the target word). Kambanaros and Grohmann interpreted their findings by proposing difficulties in accessing the target phonological representations. In their words, children with SLI do not have a strong enough phonological representation or strong enough links between the semantic and phonological representations in the output lexicon to support correct retrieval, whereas aphasic patients were argued to have greater difficulty accessing the morphophonological representation or lexemes of verbs.

In this context, the next talk by *Maria Kambanaros* on instrumentality and the neurological underpinnings of verb processing fit nicely into the already reported pattern. The research objective was to see whether semantically complex verbs are easier or more difficult to retrieve across language-impaired populations. The naming task was administered to four populations: (i) adults with aphasia, (ii) adults with schizophrenia, (iii) adults with multiple sclerosis, and (iv) children with SLI. Results in adult populations suggest a negative effect of instrumentality on verb retrieval in Broca's aphasics and a positive effect in bilingual anomic aphasics (in both languages) as well as in patients with multiple sclerosis, and no effect of instrumentality on verb retrieval in schizophrenics. With respect to child populations, non-instrumental verbs were significantly better retrieved than instrumental verbs for all children (SLI and controls with typical language development). Children with SLI were found to perform better in instrumental verbs with a name relation compared to instrumental verbs without a name relation. These findings are compatible with a theory that would want different populations to be unable to resolve the competition at the phonological level.

The next talk by *Valantis Fyndanis* (University of Athens, University of Potsdam & Technological Educational Institute of Patras) dealt with evidence from Standard Modern Greek on the morphosyntactic production in agrammatic aphasia and probable Alzheimer's disease. This talk added a flavor of variation in the landscape that emerges from the literature on pathologies, as the different status of the functional heads Asp(ect), T(ense), and Agr(eement) in terms of impairment/preservation (e.g., in Fyndanis *et al.* 2012, Asp was found significantly more impaired than T, and T significantly more impaired than Agr in Greek-speaking agrammatic patients) was explained in the talk as the result of differential demands: Agr is easier than T/Asp, because Agr requires implementation of only grammatical knowledge, whereas T/Asp require integration of extralinguistic/conceptual information as well. T further requires reference to temporal entities (e.g., event time, speech time; see Wenzlaff & Clahsen 2005) which makes them more costly in pragmatic terms. With respect to the group with probable Alzheimer's disease, the Asp < T < Agr pattern of impairment was retained and Fyndanis argued that the impairment in T is due to participants' difficulty in retrieving the verbal morphology matching the sentence-initial T adverbials. Overall, in this case too, it seems that variation is either extralinguistic (i.e. integration of pragmatic/contextual resources) or, if within grammar, it amounts to morphophonology, but never to syntax.

This was the idea pursued in the last talk of CBL (*Evelina Leivada*, Universitat de Barcelona). I sketched out a state of affairs that deals with the notion of variation across languages and pathologies in a way that brings forward a strong parallel between what is reported in the literature that comes from the study of specific languages and the literature that deals with clinical findings from work on various pathologies. More specifically, I suggested that (i) the same loci of variation can be identified across the two research programs (comparative languistics/variation across languages and comparative biolinguistics/variation across pathologies) and (ii) certain pathologies can inform our understanding of variation in the comparative biolinguistics domain through administering a specific linguistic task on the semantics–pragmatics of quantification to specific populations, such as schizophrenics. With respect to the first point, all syntactic or lexical ('parametric') variation can be reconstructed in non-syntactic terms by viewing parameters as externalization-related epiphenomena that amount to morphophonological variants rather than as UG primitives. In other words, variation is constrained to one domain of grammar, because syntax is invariant (cf. Boeckx's 2011 Strong Uniformity Thesis according to which principles of narrow syntax are not parametrizable) and lexical semantics of course varies—but following Ramchand & Svenonius (2008), these differences were taken to be arbitrary and not reducing to any obvious discrete parametric system.

The same picture emerges from the literature on pathologies, as already suggested above for a variety of atypical populations that were discussed at the workshop. The only difference is that in this literature it is unclear what happens with respect to the semantics–pragmatics interface in some cases, which is where the need of administering specific linguistic tasks that disentangle particular aspects of impaired patterns arises. For example, literature on language comprehension and/or production in schizophrenics makes reference to a variety of language-related impairments: (i) failure to segregate phonological engrams due to the lack of the typical left hemisphere advantage for language (Angrilli *et al*. 2009), (ii) failure to use words in a semantically acceptable way (Oh *et al*. 2002), (iii) impaired referential and lexical cohesion (Ragin & Oltmanns 1986, McKenna & Oh 2005), and (iv) impaired pragmatics and inability to interpret things by making use of linguistic context (Chapman & Chapman 1973, Kuperberg *et al*. 1998). With the exception of (i), all the other deficits listed above seem to pertain to the semantics-pragmatics interface. At the same time, though, it is not clear whether some of these impairments really boil down to semantics *per se*: For example, the inability to use words in a semantically acceptable way might not be a semantic deficit that reflects a problem in conceptual-semantic stores, but instead a matter of impaired word retrieval abilities. In the talk by Kambanaros and Grohmann on verb–noun dissociations across impaired populations, the results of Kambanaros *et al*. (2010) on object and action naming in schizophrenia were presented. These authors suggested a retrieval issue when discussing their findings. More specifically, the absence of dissociation in comprehension of action and object names coupled with semantic errors in naming for both these two classes is taken by them to suggest intact conceptual-semantic stores, but difficulties with mapping semantics onto the lexicon, that is, access/retrieval problems.

Fine (1999: 85) describes this state of affairs by arguing that "to assess language use and its relationship to a psychiatric entity such as schizophrenia requires that the context be carefully taken into account and that the semantic resources related to contexts be considered". This view that makes reference to both semantic resources and contextual variables (i.e. pragmatics) suggests a need to disentangle the semantics-pragmatics interface in schizophrenic language in order to understand which aspects of language are impaired in these patients. Doing so requires a testing tool that involves both semantically felicitous/infelicitous and pragmatically felicitous/infelicitous test items, such as the 'Cave-girl Task' used to test the semantics and pragmatics of quantifiers in child and adult populations across 25 languages (COST Action A33, 2006–10; Katsos *et al*. 2012). As I suggested in my talk, if schizophrenics have intact semantic stores but show impaired use of pragmatic/contextual variables, the prediction is that they would perform more accurately in identifying the semantically infelicitous utterances than the pragmatically infelicitous ones. Testing this prediction will shed further light first on the nature of variation across pathologies and second on the grammar of schizophrenia (Leivada, in progress).

All the cases of pathological conditions discussed in CBL are particularly telling when the aim is to delimit variation and to determine which aspects of grammar show up as impaired more often than not.[3] The case of schizophrenia is

---

[3] Of course pinpointing these aspects is also subject to the theory of FL one endorses, but in line with mainstream generativist assumptions that accept Merge as the operation that lies

of particular interest to such inquires: Although it has been argued that "at the level of syntactic processing, schizophrenic patients' speech is usually normal" (Marini *et al*. 2008: 145, referring to Andreasen 1979 and Covington *et al*. 2005), a case for reduced syntactic complexity can also be made (Morice & McNicol 1986, Fraser *et al*. 1986, Thomas *et al*. 1990, discussed in Marini *et al*. 2008). One possible way to go about describing the relevant findings is to suggest "a disruption of executive function and pragmatics, perhaps with impairment of the syntax-semantics interface" (Covington *et al*. 2005: 85). However, if we are able to observe reduced syntactic complexity, we are also able to observe some syntactic complexity. When this is the case, it entails that the (syntactic) operation responsible for recursion is preserved. Once more, we do not find an inability to syntactically combine different elements or a manipulation of syntactic objects in a way that is not licit in unimpaired syntax (e.g., to add negation by moving the third element of the clause to the sentence-initial position). Therefore, another way to describe reduced syntactic complexity has to be found. For example, it is possible that grammatically unacceptable utterances and reduced syntactic complexity are the cumulative result of the existence of a number of features typical of schizophrenic production (see Andreasen 1979 and more recently McKenna & Oh 2002 for a list of such features) such as clanging, derailment, and semantic paraphasias which arise due to the nature of the schizophrenic semantic network where "loose associations are caused by an unrestrained associations-chain in semantic memory" (Lerner *et al*. 2012: 5).

The roundtable discussion that closed the event revolved around concerns that relate to the tools to be used in the new research program of comparative biolinguistics, to testing concerns when it comes to informing our theories of FL on the basis of atypical populations, to the steps that need to be taken next in order to provide a more solid agenda for comparative biolinguistics and, above all, to the issue of variation, which was arguably the notion that figured more prominently throughout most of the discussion periods that followed the talks. Moreover, some problems regarding the feasibility of biolinguistics were brought up, and some of the more foundational issues of the divide between linguistics and other sciences were discussed, such as whether the mind can be studied separately from the brain, and whether it is acceptable for linguistics to evolve independently from allied disciplines and to not be solidly grounded in biology.

3. Outlook: The Beginnings of a New Research Agenda

All in all, CBL managed to bring together different aspects of the five key questions in biolinguistics that were given in (1)–(5). The fact that all the talks were followed by lively discussion periods, despite the diversity of perspectives (and perhaps even theoretical persuasions) of the participants, can only suggest that dialogue is possible and that building the right type of bridges between linguistics and interfacing disciplines is doable.

In Paris of 1866, all discussion on the origins of human language (an issue

---

in the core of syntax, it is a rather straightforward claim that observing non-typical use of lexical items (i.e. impaired semantics), retrieval issues, or missing morphological markers cannot qualify as impaired syntax.

related to question (5) above) was famously banned by the Linguistic Society of Paris. Almost one and a half centuries later, it seems to be the case that adequate progress has been made and that we have accumulated enough knowledge from the various disciplines mentioned above to make an attempt to provide linking hypotheses across these different disciplines through a novel, comparative biolinguistics perspective.

Crucially, this novel perspective does not intend to dismiss or neglect the progress made over the last decades within the comparative languistics approach (i.e. variation across languages); instead it can benefit from this progress and make use of the relevant findings. However, it does require that the findings, tools, and primitives that survive the passage from one discipline to the other are able to inform on the somewhat larger frame of the new enterprise. I believe that it also requires that they are linked more robustly with the five key questions of the biolinguistic agenda. Put differently, linguistic representations, when used, have to go hand in hand with interdisciplinary linking hypotheses that say something novel about FL, rather than being followed by highly technical discussions that provide no explanatory adequacy at all because they exhaust themselves on describing how construction A is realized in language B. In Felix's (2010: 68) words, once more, "[i]f you, like Chomsky, are primarily interested in cognitive psychology, your specific perspective on the entire generative enterprise might be somewhat different from the one of someone who is just interested in language and language data".

References

Andreasen, Nancy C. 1979. Thought, language, and communication disorders: II clinical assessment, definition of terms, and assessment of their reliability. *Archives of General Psychiatry* 36, 1315–1321.

Angrilli, Alessandro, Chiara Spironelli, Thomas Elbert, Timothy J. Crow, Gianfranco Marano & Luciano Stegagno. 2009. Schizophrenia as failure of left hemispheric dominance for the phonological component of language. *PLoS ONE* 4(2), e4507.

Benítez-Burraco, Antonio & Lluís Barceló-Coblijn. 2013. Paleogenomics, hominin interbreeding and language evolution. *Journal of Anthropological Sciences* 91, 239–244.

Benítez-Burraco, Antonio & Cedric Boeckx. In press. Universal Grammar and biological variation: A program for Comparative Biolinguistics. *Biological Theory*.

Boeckx, Cedric. 2011. Approaching parameters from below. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 205–221. Oxford: Oxford University Press.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The *Biolinguistics* manifesto. *Biolinguistics* 1, 1–8.

Chapman, Loren J. & Jean P. Chapman. 1973. *Disordered Thought in Schizophrenia*. New York: Apple-Century-Crofts.

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

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

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

Christodoulou, Christiana. 2011. Cypriot Greek Down Syndrome. Their grammar and their interfaces. Doctoral dissertation, The University of British Columbia.

COST Action A33. 2006–10. *Crosslinguistically Robust Stages of Children's Linguistic Performance*. European research network. Brussels: COST Office.

Covington, Michael A., Congzhou He, Cati Brown, Lorina Naçi, Jonathan T. McClain, Bess Sirmon Fjordbak, James Semple & John Brown. 2005. Schizophrenia and the structure of language: The linguist view. *Schizophrenia Research* 77, 85–98.

Duarte, Cidália, João Maurício, Paul B. Pettitt, Pedro Souto, Erik Trinkaus, Hans van der Plicht & João Zilhão. 1999. The early Upper Paleolithic human skeleton from the Abrigo do Lagar Velho (Portugal) and modern human emergence in Iberia. *PNAS* 96, 7604–7609.

Duguine, Maia. 2013. Null arguments and linguistic variation: A minimalist analysis of *pro*-drop. Doctoral dissertation, The University of the Basque Country, Vitoria-Gasteiz.

Felix, Sascha W. 2010. Me and Chomsky: Remarks from someone who quit. In Thomas Hanneforth & Gisbert Fanselow (eds.), *Language and Logos: Studies in Theoretical and Computational Linguistics, Festschrift for Peter Staudacher on His 70th Birthday*, 64–71. Berlin: Akademie Verlag.

Fine, Jonathan. 1999. On the puzzle of language, pragmatics, and schizophrenia. *Psychological Reports* 84, 84–86.

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179–210.

Fraser, W.I., K.M. King, P. Thomas & R.E. Kendell. 1986. The diagnosis of schizophrenia by language analysis. *The British Journal of Psychiatry* 148, 275–278.

Fyndanis, Valantis, Spyridoula Varlokosta & Kyrana Tsapkini. 2012. Agrammatic production: Interpretable features and selective impairment in verb inflection. *Lingua* 122, 1134–1147.

Hale, Mark & Charles Reiss. 2000. Phonology as cognition. In Noel Burton Roberts, Philip Carr & Gerald Docherty (eds.), *Phonological Knowledge: Conceptual and Empirical Issues*, 161–184. Oxford: Oxford University Press.

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

Hornstein, Norbert. 2013. LSA Summer Camp. Retrieved from http://facultyoflanguage.blogspot.com/2013/07/lsa-summer-camp.html.

Kambanaros, Maria. 2003. Verb and noun processing in late bilingual individuals with anomic aphasia. Doctoral dissertation, Flinders University, Adelaide.

Kambanaros, Maria, Kleanthes K. Grohmann, Michalis Michaelides & Eleni Theodorou. 2014. On the nature of verb–noun dissociations in bilectal SLI:

A psycholinguistic perspective from Greek. *Bilingualism: Language and Cognition* 17: 169–188.

Kambanaros, Maria, Lambros Messinis, Vassilis Georgiou & Panagiotis Papathanassopoulos. 2010. Action and object naming in schizophrenia. *Journal of Clinical and Experimental Neuropsychology* 32, 1083–1094.

Katsos, Napoleon, Maria-José Ezeizabarrena, Anna Gavarró, Jelena Kuvač Kraljević, Gordana Hrzica, Kleanthes Grohmann, Athina Skordi, Kristine Jensen de López, Lone Sundahl, Angeliek van Hout, Bart Hollebrandse, Jessica Overweg, Myrthe Faber, Margreet van Koert, Chris Cummins, Nafsika Smith, Maigi Vija, Sirli Parm, Sari Kunnari, Tiffany Morisseau, Manana Rusieshvili, Kazuko Yatsushiro, Anja Hubert, Spyridoula Varlokosta, Katerina Konstantzou, Shira Farby, Maria Teresa Guasti, Mirta Vernice, Ingrida Balčiūnienė, Jūratė Ruzaitė,Helen Grech, Daniela Gatt, Arve Asbjørnsen, Janne von Koss Torkildsen, Ewa Haman, Aneta Miękisz, Natalia Gagarina, Julia Puzanova, Darinka Andjelković, Maja Savić, Smiljana Jošić, Daniela Slančová, Svetlana Kapalková, Tania Barberán Recalde, Duygu Özge, Saima Hassan, Heather van der Lely, Uli Sauerland & Ira Noveck. 2012. The acquisition of quantification across languages: Some Predictions. In Alia K. Biller, Esther Y. Chung & Amelia E. Kimball (eds.), *Proceedings of the 36th Annual Boston University Conference on Language Development*, 258–268. Somerville, MA: Cascadilla Press.

Kuperberg, Gina R., Philip K. McGuire & Anthony S. David. 1998. Reduced sensitivity to linguistic context in schizophrenic thought disorder: Evidence from on-line monitoring for words in linguistically anomalous sentences. *Journal of Abnormal Psychology* 107, 423–434.

Leivada, Evelina. In progress. The limits of variation across languages and pathologies: From linguistics to biolinguistics. Doctoral dissertation, Universitat de Barcelona.

Lerner, Itamar, Shlomo Bentin & Oren Shriki. 2012. Excessive attractor instability accounts for semantic priming in schizophrenia. *PLoS ONE* 7(7): e40663.

Lorenzo, Guillermo & Víctor M. Longa. 2009. Beyond generative geneticism: Rethinking language acquisition from a developmentalist point of view. *Lingua* 119, 1300–1315.

Marini, Andrea, Ilaria Spoletini,, Ivo Alex Rubino, Manuela Ciuffa, Pietro Bria, Giovanni Martinotti, Giulia Banfi, Rocco Boccascino, Perla Strom, Alberto Siracusano, Carlo Caltagirone & Gianfranco Spalletta. 2008. The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research* 105, 144–155.

McKenna Peter J. & Tomasina M. Oh. 2005. *Schizophrenic Speech*. Cambridge: Cambridge University Press.

Morice, Rodney & Don McNicol. 1986. Language changes in schizophrenia: A limited replication. *Schizophrenia Bulletin* 12, 239–251.

Oh, Tomasina M., Rosaleen A. McCarthy & Peter J. McKenna. 2002. Is there a schizophasia? A study applying the single case study approach to formal thought disorder in schizophrenia. *Neurocase* 8, 233–244.

Oyama, Susan. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. Durham, NC: Duke University Press.

Pigliucci, Massimo & Gerd B. Müller (eds.). 2010. *Evolution — The Extended Synthesis*. Cambridge, MA: MIT Press.

Poeppel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Ragin, Ann Barnett & Thomas F. Oltmanns. 1986. Lexical cohesion and formal thought disorder during and after psychotic episodes. *Journal of Abnormal Psychology* 95, 181–183.

Ramchand, Gillian & Peter Svenonius. 2008. Mapping a parochial lexicon onto a Universal Semantics. In Theresa Biberauer (ed.), *The Limits of Syntactic Variation*, 219–245. Amsterdam: John Benjamins.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17, 501–557.

Thomas, P., K. King, W.I. Fraser & R.E. Kendell. 1990. Linguistic performance in schizophrenia: A comparison of acute and chronic patients. *The British Journal of Psychiatry* 156, 204–210.

Tinbergen, Nikolaas. 1963. On the aims and methods of ethology. *Zeitschrift für Tierpsychologie* 20(4), 410–433.

Wenzlaff, Michaela & Harald Clahsen. 2005. Finiteness and verb-second in German agrammatism. *Brain and Language* 92, 33–44.

Evelina Leivada
Universitat de Barcelona
Department of Linguistics
Gran Via de les Corts Catalanes, 585
08007 Barcelona
Spain
evelina@biolinguistics.eu

Signs and Offline Brain Systems in Language Evolution

Bouchard, Denis. 2013. *The Nature and Origin of Language*. Oxford: Oxford University Press.

by Gonzalo Castillo

Denis Bouchard's book is a refreshingly new take on the old problem of detailing the processes by which humans became linguistic creatures, a puzzle that researchers of varying disciplines have been attempting to solve since long before the inception of modern biolinguistics. It is perhaps not surprising that the book, divided into four parts, starts with several chapters that call to our attention the apparent failure to provide a definite answer to this question. Bouchard argues that, for a start, language cannot be explained scientifically if linguistics receives a treatment or a status that is different from the other sciences, a mistake that he finds evidenced by the scarcity of principled explanations in the literature.

A principled explanation is one that considers the object of study as dependent on logically prior elements from which it arises. Since language can be considered as a system that links percepts and concepts, or representations of sound and meaning, the principled elements of language should be those studied by the sciences of meaning and perception. Explaining the evolution of language, in sum, is determining how the systems that produced concepts and percepts changed in the brains of our ancestors so that their products could be linked and become *signs*. The existence of signs is, therefore, "the only special property of language" (p. 97).

Parts II and III of the book, introducing Bouchard's own Sign Theory of Language (STL), invite us to consider the evolutionary implications of assuming that language is just a system of signs. But first, what is a sign? According to Saussure (1916), a sign can be defined as a relation between a representation of a sound pattern (a *signifier*, e.g. /dɔg/), and a representation of a chunk of cognition (a *signified*, e.g. the concept of dog). Two special properties of signs are crucial to understanding their nature: abstraction and arbitrariness. Signs are abstract because they are detached from any brain-external stimuli or immediacy. Signs are arbitrary relations because there is nothing in any of the properties of their parts that justifies their linking.

Perhaps controversially, Bouchard argues that abstraction was the change in the conceptual system of our ancestors which ended up granting humans their unique cognitive suite. Not only language, but also complex imitative abilities, theory of mind, episodic memory, and object permanence could be traced back to

biolinguistics

ISSN 1450–3417
Biolinguistics 8: 067–072, 2014
http://www.biolinguistics.eu

the emergence of *Offline Brain Systems* (OBS) in the brain of our ancestors. Since these systems did not appear from nothing, but were a new functionality that occurred in preexisting mirror systems, perhaps the term is a bit misleading. It should not be taken, however, as anything else than the claim that some parts of the brain started to represent things in the absence of external stimuli, while inhibiting any motor actions that would follow from actual perception. This change is linked by the author to an increase in the number of neurons in the human cortex, which, added to a tendency towards a more globular shape of the cranium, led to an increase in connectivity and internal activity (pp. 115–119).

Of course, signs would not be possible without OBS. Abstraction is the key difference that separates human concepts from animal categories or percepts. By taking a percept (for instance, a sound pattern like /dɔg/) and abstracting it from the immediate environment, we are able to entertain a signifier. Since signifiers have the same psychological nature as signifieds (both are conceptual), linking them to create a sign becomes an easy step. Bouchard thinks that the relation that exists between signifier and signified is one of reciprocal predication. It should be noted that the author follows Hurford (2007) in considering predication as a trait with evolutionary ancient roots, present in the perceptual systems of complex animals. This pre-existing property is then employed to fulfill the role of sign formation in a similar way to how a specific color pattern is attributed to a leopard.

The second property of signs, arbitrariness, arises as a side effect of the completely different perceptual origins of signifiers and signifieds. Arbitrariness has important structural implications for language. Since the link between signifier and signified is generally unmotivated, the range of possible signs, in theory, is unbounded, and the links between their parts can change rapidly both across time and individuals. This would lead to a chaotic, random system, if it were not for the fact that language is a biological function that arises under the physical and cognitive constraints of the human mind. In chapter 6, Bouchard introduces the notion of *epigenetic self-organizing constraints*, borrowed from biology (Jacob 1982, Erwin 2003), to account for the language-external properties of both parts of the sign.

Regarding external constraints on signifiers, we find that vocalizations and gestures are naturally perceived as discrete segments, and that the phonemic repertoire is organized towards a balance between ease of articulation and ease of perception. As a consequence of this balance, the phonemic repertoire tends to cluster into a small set of percepts that maximizes contrast between its elements. Given that the set of phonemes ends up being very small, and that the set of possible meanings is much larger, phonemes are not enough to match all of the latter, so they start to combine with each other (together with other biologically common vocal resources like stress, length, and intonation), building words. This process is made possible by OBS, which can reinterpret a chain of distinctive phonemes as a discrete unit, granting it the status of a new sign. Similar to phonemes, words are limited in complexity by their frequency of use (cf. Zipf 1949/1965) and, of course, human memory. This proposal that I just summarized is also a justification for the existence of signifiers. We could perfectly well imagine an internal language that lacks the phonological component but is still able to discretely combine signifieds. However, Bouchard argues that this hypothetical language would "lack the triggering elements for combinatorial processes to emerge" (p. 160).

Signifieds are also affected by language-external constraints. The author, again following Hurford (2007), points at perception being organized around objects and properties, a distinction manifested in language by the subject-predicate and head-dependent relations. Similar to what happened in the case of signifiers, although OBS can take any perceptual episode and encode it as a signified, the meanings that usually cluster into signs are those that are activated more frequently and within a broader array of contexts. Since our perception is based on categories, it is expected that the words of a language can be employed to refer to sortals, abstract qualities, etc., and not exclusively to specific perceptual events.

At the end of Part III (pp. 169–179), the author introduces an original view of syntax that is also based on the sign. More specifically, syntax is defined as a small set of signs (*C-signs*) that are in charge of relating words to produce complex meanings. The signifier of a C-sign can be either part of a paradigm that is stored in memory (such as in case marking), or a syntagmatic relation of elements (such as the subject–verb–object relation). Its signified, on the other hand, is based on the most distinctive feature of perception: property attribution (predication). The relation between signifier and signified, being arbitrary, produces cross-linguistic variation depending on the way languages randomly match specific predicative relations with oral or gestural resources such as juxtaposition of elements, intonation, stress, length, or morphological markers.

Since both C-signs and unit signs are based on the same pre-existing properties of the sensorimotor and conceptual interfaces, Bouchard argues against the existence of a long proto-language stage, defending that all that is needed to have syntax is a brain equipped with OBS and subject to external self-organizing constraints (ch. 7). Thus, the origin of C-signs is the result of predication being a ubiquitous aspect of perception, which ends up creating a pressure to match it with the gestural or oral percepts available to a language. Since storing words is costly in terms of memory, and perception is organized around an object–property dichotomy (Hurford 2007), it is not plausible that a language will develop a lexicon that treats each attribution of properties with a different root, so compositional processes come into play. Syntax, therefore, does not emerge *for* communication or the organization of thought (although it might as well produce such benefits), but as a regularizing side effect of the chaotic system that is triggered by OBS.

The final part of the book (Part IV) can be described as an open letter specifically addressed to linguists. These chapters analyze some formal, UG-based models of linguistic phenomena under the cognitive umbrella of Bouchard's STL, and so they are more technical and less interesting from a multidisciplinary perspective. I will only highlight that the author, faithful to the idea of looking for principled explanations, rejects many milestones of generative theory (Principles and Parameters, *wh*-movement, c-command), extending his notion of C-signs to specific cases that would require a very detailed elaboration, far beyond the scope of this review.

So far, I have provided a summary for what I consider the main ideas discussed in the book. Although the picture of language that they present is coherent and promising, they will only triumph after some questions are some day answered, so I would like to finish this review by introducing a few of them. The first thing I would like to talk about is abstraction being the clear-cut line that separates humans from other animals. There is no doubt that humans excel at offline thought processes, but it is, in my view, too early to claim that these are a unique result of the evolutionary path followed by hominins. In fact, it is currently possible to find evidence suggesting that many species have to a certain extent some of the functions attributed to OBS. For instance, Osvath & Osvath (2008) show that great apes can select tools that will only be useful in the future in a not currently perceived location, indicating that immediate needs do not play a role in their behavior. Other experiments on animal episodic memory such as Clayton & Dickinson (1998) and Hoffman *et al*. (2009) show a capacity to integrate and remember information on *what*, *where*, and *when* in scrub jays and rhesus monkeys respectively.

This indicates that offline thought processes may not be human-specific, something that the STL would have to take into account to explain why language is. If the solution is that abstraction turns out to be language-enabling when it becomes frequent enough, then we would still need to consider why less-refined abstract thought processes cannot confer a less-refined linguistic stage, provided that abstraction was indeed all that is needed to acquire signs. In other words, if abstraction is gradual, we should ask ourselves either why within this gradation a critical point that facilitated the sudden acquisition of signs could have appeared, or if the acquisition of signs could have also been a gradual process. If the latter was the case, sign systems would not only have evolved by means of self-organizing constraints, but also in parallel with the evolution of OBS.

Whatever the answer turns out to be, my view is that the emergence of abstraction is simply not enough to explain the human-specific usage of signs, and that a means to reliably control abstractive processes, such as an improvement in executive control (Baddeley 1996, 2002), is also an important requirement, and perhaps the key to those who would like to claim that language is a relatively recent and sudden cognitive revolution. The investigation on OBS could benefit from the copious amount of research that has been conducted so far on Working Memory, since both are reaching similar conclusions about the nature of humans' cognitive uniqueness (e.g., Coolidge & Wynn 2005).

Additionally, a distinction should be made between the cognitive properties that are brought about by OBS, and those that are just an indirect result of them, being caused by the existence of signs. The most relevant example for the latter is cross-modularity, where representations produced by different modules can be joined together under the same word or sentence, generating what appear to be new conceptual capacities. In a series of experiments by Hermer & Spelke (1994, 1996) and Hermer-Vasquez *et al*. (1999), it was shown that masking the linguistic abilities of adult humans by making them repeat what they are listening to can impair their orientation skills to the level of rats and prelinguistic infants, since they are unable to use expressions such as *to the right of X* in their thoughts. Similarly, numerical words seem to be a combination of a system for subitizing individuals and a system for representing large, inexact numerosities (Dehaene 1997).

Since these capacities do not seem to be part of our prelinguistic stock, it is not completely accurate to claim that language is a mere reflection of a shared conceptual interface plus abstraction, and that integration only pertains to the origin of OBS. On the contrary, it seems that the sign awakens new integrative processes, affecting thought in a unique way that still needs to be investigated as an enterprise on its own. To do so, we do not need to abandon the search for principled explanations. On the contrary, we should follow Bouchard's suggestion by looking for a possible relation between the sensorimotor aspects of signs and the enhancement of cognition. As a suggestion, we can claim that the arbitrary relation that exists between the signifier and signified allows that multiple signifieds can be associated with the same signifier, thus strengthening the connections between different, unrelated systems.

All in all, Bouchard's Sign Theory of Language has the advantage of building a bridge from other disciplines to the usually isolationist realm of linguistics. Unlike other multidisciplinary attempts, this one makes a conscious effort to prevent as much as possible an almost inevitable simplification or vagueness, providing testable hypotheses and original perspectives for the solution of old problems. A quote on the back cover of the book by Christopher Petkov says: "Denis Bouchard's theory may be exactly what is needed to take linguistics and neuroscience in exciting new directions." Whether one of these directions will effectively take us to a full explanation of language evolution, we cannot yet foresee, but one thing we can depend on is that the advancement of science rarely sticks to the same road for long.

References

Baddeley, Alan D. 1996. Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A* 49, 5–28.

Baddeley, Alan D. 2002. Fractionating the central executive. In Donald T. Stuss & Robert T. Knight (eds.), *Principles of Frontal Lobe Function*. New York: Oxford University Press, 246–260.

Clayton, Nicola S. & Anthony Dickinson. 1998. Episodic-like memory during cache recovery by scrub jays. *Nature 395*, 272–274.

Coolidge, Frederick L. & Thomas Wynn. 2005. Working memory, its executive functions, and the emergence of modern thinking. *Cambridge Archaeological Journal* 15, 5–26.

Dehaene, Stanislas. 1997. *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

Erwin, Douglas. 2003. The Goldilocks hypothesis. *Science* 302, 1682–1683.

Hermer, Linda & Elizabeth S. Spelke. 1994. A geometric process for spatial representation in young children. *Nature* 370, 57–59.

Hermer, Linda & Elizabeth S. Spelke. 1996. Modularity and development: The case of spatial reorientation. *Cognition* 61, 195–232.

Hermer-Vasquez, Linda, Elizabeth S. Spelke & Alla S. Katsnelson. 1999. Sources of exibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology* 39, 3–36.

Hoffman, Megan L., Michael J. Beran & David A. Washburn. 2009. Memory for "what", "where", and "when" information in rhesus monkeys (Macaca mulatta). *Journal of Experimental Psychology: Animal Behavior Processes* 35, 143–152.

Hurford, James R. 2007. *The Origins of Meaning: Language in the Light of Evolution*. Oxford: Oxford University Press.

Jacob, François. 1982. *The Possible and the Actual*. New York: Pantheon.

Osvath, Mathias & Helena Osvath. 2008. Chimpanzee (Pan troglodytes) and orangutan (Pongo abelii) forethought: Self-control and pre-experience in the face of future tool use. *Animal Cognition* 11, 661–674.

Saussure, Ferdinand de. 1916. *Cours de Linguistique Générale*. Critical edition prepared by Tullio de Mauro. Paris: Payot.

Zipf, George K. 1949/1965. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Hafner.

Gonzalo Castillo
Universitat de Barcelona
Department of Linguistics
Gran Via de les Corts Catalanes, 585
08007 Barcelona
Spain
g.castillo@gmx.com

Some Problems for Biolinguistics

Derek Bickerton

Biolinguistics will have to face and resolve several problems before it can achieve a pivotal position in the human sciences. Its relationship to the Minimalist Program is ambiguous, creating doubts as to whether it is a genuine subdiscipline or merely another name for a particular linguistic theory. Equally ambiguous is the relationship it assumes between 'knowledge of language' and the neural mechanisms that actually construct sentences. The latter issue raises serious questions about the validity of covert syntactic operations. Further problems arise from the attitudes of many biolinguists towards natural selection and evo-devo: The first they misunderstand, the second they both misunderstand and overestimate. One consequence is a one-sided approach to language evolution crucially involving linguistic 'precursors' and the protolanguage hypothesis. Most of these problems arise through the identification of biolinguistics with internalist and essentialist approaches to language, thereby simultaneously narrowing its scope and hindering its acceptance by biologists.

Keywords: *biolinguistics; covert movement; evolution; internalism; Minimalist Program*

1. Introduction

In this article I try to deal with some issues that seem to me to be crucial if biolinguistics is to achieve the centrality in the human sciences to which its subject-matter surely entitles it. One or two of these may be issues that involve image and perception, but most are much more substantive, involving brain–grammar relations, understanding of old and new aspects of evolutionary biology, the process of language evolution, and fundamental issues in the philosophy of biology. Some issues result from still unclarified aspects of the relationship between biolinguistics and generative grammar, but all of them, to a greater or lesser extent, prejudice the unification of biolinguistics with other biological fields.

Fears widespread among both linguists and non-linguists that 'biolinguistics' may turn out to be merely a more scientific-sounding term for generative minimalism are reinforced by the way the distinction is made between 'strong' and 'weak' senses of biolinguistics by Boeckx & Grohmann (2007: 2). They define "the strong sense" of the term as "provid[ing] explicit answers to questions that necessarily require the combination of linguistic insights and insights from

ISSN 1450–3417
Biolinguistics 8: 073–096, 2014
http://www.biolinguistics.eu

related disciplines (evolutionary biology, genetics, neurology, psychology". They define the "weak sense" as "refer[ing] to 'business as usual' for linguists, so to speak, to the extent they are seriously engaged in discovering the properties of grammar, in effect carrying out the research program Chomsky initiated in *Syntactic Structures*. Emphasizing that by their use of the words 'strong' and 'weak' they are not proposing some two-tier system of "superior" and 'inferior' biolinguists, they point out that work "focusing narrowly on properties of the grammar… has very often proven to be the basis for more interdisciplinary studies" (loc. cit.).

It is difficult to avoid the conclusion that adhering to the latest version of generative grammar is indeed a prerequisite, not perhaps for simply attempting to engage in biolinguistics, but certainly for being taken seriously by serious biolinguists. Granted, the authors try to forestall this conclusion by claiming that "minimalism is an approach to language that is largely independent of theoretic-cal persuasion" (Boeckx & Grohmann 2007: 3). But I suspect that even a long-time Yakuza member could number the minimalist works by non-generativists on the fingers of one hand. In fact, too many biolinguists have taken over without questioning a number of assumptions made within generative grammar at one time or another, many of which pre-date the Minimalist Program (MP), none of which have any necessary connection with it, and some of which are orthogonal, even prejudicial, to the achievement of MP goals.

2. The Problem of Knowledge versus Neural Mechanisms

2.1. *'Knowledge of Language'*

Let's begin by examining some versions of the famous 'Five Chomsky Questions' (Chomsky 1986, 1988, Jenkins 2000, Boeckx & Grohmann 2007, Di Sciullo *et al.* 2010). As indicated by the dates of these citations, the questions precede the efflorescence of biolinguistics but are now routinely repeated in one form or another by authors of programmatic statements about the field. It is interesting (and very relevant) to compare the wording of Question 4 in three versions of the questions. That of Boeckx & Grohmann (2007) adheres most closely to Chomsky's 'knowledge of language' formula:

(1) How is that knowledge [of language—DB] implemented in the brain?

'Knowledge' in this context has long provoked the ire of empiricist philosophers, but my objection is quite different; use of the term gives a highly misleading picture of the nature of syntax. Although syntax is often regarded as part of cognition, its operations are automatic and out of reach of conscious awareness. We are no more aware of how our brains construct sentences than we are of how our stomachs digest food or our hearts circulate blood. No-one who proposed to study our 'knowledge of digestion' or 'knowledge of circulation' could hope to be taken seriously. Granted, one says informally things like "Does he know Russian?", whereas nobody ever said "Does he know digestion?"—but there are many languages, and only one digestion. The problem here arises, I think, simply from the ambiguity of the term 'language', as opposed to the French distinction between *langage* (the faculty) and *langue* (an individual language). I know English and I have (hopefully reliable) intuitions about what are, or are not, grammatical sentences in (some variety of) English, as do all other speakers of that language. But if I did not have years of professional training and experience I would be as unable to explain the basis for those intuitions as is any naïve speaker, and I have no intuitions whatsoever about what is grammatical in Russian. It is surely significant that the anonymous reviewer who queried my treatment of 'knowledge of language' admitted that "speakers' internalized linguistic capabilities are 'about' *one or another particular* grammar" (my emphasis) and not about *langage* at all. But it is surely *langage* and not *langue* that we must be talking about if we are asking the "five questions".

Whether guided by some awareness of this or for other reasons, three years later Question 4, like the other four questions, was rephrased to excise 'knowledge' (Di Sciullo *et al*. 2010):

(2) How is language implemented in the brain?

But it is perhaps even more revealing to see how Jenkins (2000) produces yet a third variant of the question:

(3) What are the relevant brain mechanisms?

All of these formulations, fortuitously or otherwise, avoid one of the most crucial issues that biolinguistics should be resolving—the relationship between grammars and how the brain actually produces sentences.[1] Consider Chomsky's (1988) version of Question 4.

(4) What are the physical mechanisms that serve as the material basis for this system of knowledge and for the use of this knowledge?

Chomsky's formulation presupposes two distinct and separate mental objects: a system of knowledge and a system for executing that knowledge. It is an astonishingly dualist claim from someone who has consistently adhered to monism, but let that pass. Taking (4) at its face value, there could obviously be two ways of describing syntax. One would provide maximal coverage of the empirical data while simultaneously achieving maximal levels of elegance,

[1] I am well aware of work by Embick & Poeppel (2005a, b) and associates on the relationship between neurobiology and linguistics. Though the issues may seem the same, I approach them from a different direction with different assumptions and different goals. Consequently we see different problems and different solutions. To discuss these differences would take us too far from present topics, but some flavor of them maybe found in this quotation from Poeppel *et al*. (2012: 14130): "By connecting the brain science of language to formal models of linguistic representation, the work decomposes the various computations that underlie the brain's multifaceted combinatory capacity." Poeppel and his associates seem to believe that, given the right granularity level, current analyses in linguistics and neurobiology can be matched without substantive change to either. I do not.

simplicity, and explanatory power. The other would adhere, as far as possible, to a literal description of what the brain actually does in order to produce sentences. Would those two descriptions be isomorphic? Not necessarily. The first, constrained solely by the linguistic data, could legitimately use whatever devices might help it achieve its goals of simplicity, elegance, and comprehensiveness, regardless of how its solutions related to what brains actually do. *Should* those two descriptions be isomorphic? Obviously yes. To the extent that they differed, one would simply be wrong, and if they prove instead to be isomorphic, one is redundant. But which is redundant, the knowledge model or the mechanistic model? There can be no question that the former is redundant, since without the latter, there would be nothing to describe.

The standard objection to this sort of argument is to say, "We don't yet know enough about the brain to let it influence the construction of grammatical theories". If that is still true, something much less obvious than it was a decade or so ago, we are not yet ready for biolinguistics, and 'the weak sense' is the only one that might be applicable. However, we can't afford to sit on our hands and wait for neurobiologists to do our work for us. We might find ourselves waiting a long time. The only course is to kick-start the procedure by beginning to think about and discuss what, given all we already know or can reasonably surmise, the brain might be expected to do. What the brain seemed likeliest to do in order to meet its own goals of economy would then become a default hypothesis for the grammar, to be maintained unless or until valid reasons (linguistic or neurological) for abandoning it became manifest.

2.2. *Covert Movement*

There are, of course, serious obstacles to any rapprochement between linguistics and neurobiology, of which the 'granularity mismatch' discussed by Embick & Poeppel (2005a, b) is perhaps the best known. Here I will suggest that differences in the levels at which the analytic units of linguistics and neurobiology respectively apply may be far from all that is involved here. It may be that there are also serious mismatches between the *types* of process envisaged by linguistic analysis and the processes the brain actually uses when it forms sentences. This is not a pressing problem yet, since few if any proposals specific enough to evoke it have so far appeared, but it will surely become one, and very soon, if biolinguistics is to go on developing.

As noted above, most biolinguists subscribe to the MP, which increases the likelihood that at least the basic assumptions of the MP, and likely also the kinds of syntactic process that it employs, will serve as a basis for any attempts to achieve the desired rapprochement. One of the most ubiquitous features of MP analyses is covert movement.

Covert movement differs from overt movement in the following respect. In overt movement, the same syntactic unit is associated with (at least) two positions in the same sentence, but is pronounced in only one of them. The reality of the unpronounced unit can be linked to empirical findings such as the blocking of *want-to* contraction where the unit is subject of the embedded clause. Given the copy theory of movement, overt movement is unproblematic for the brain; two instantiations of the same item are present during the brain's assembly of sentence materials, but only one is actually uttered. Covert movement is another matter altogether.

Covert movement has been invoked to explain a variety of phenomena, from differences in the positions of verbs and adverbs in French and English to variability in quantifier scope. In contrast to overt movement, covert movement does not link with any empirical finding; its motivation is theory-internal. Processes that involve covert movement include verb raising (Emonds 1985, Pollock 1989), quantifier raising (May 1977), VP shells (Larson 1988), subject-raising (Koopman & Sportiche (1985), and more. The theory-internal nature of these can be readily demonstrated. For instance, VP shells, which involve initially merging direct and indirect objects into positions where the former will c-command the latter and then re-merging them to yield the English surface order, were motivated by a desire to preserve the relation of c-command and thus avoid the apparent violations of Principle A of the Binding Theory first pointed out by Barss & Lasnik (1986). In the case of verb raising, subjects are supposed to end up in SpecIP, but the latter being a functional projection, they cannot be theta-marked there, and must consequently be assumed to have acquired their theta-marking within the maximal projection of V (Burton & Grimshaw 1992) before being raised.

Readers will have noticed that all the citations for covert movement given above come from pre-MP versions of the grammar, and that the analyses provided therein presuppose versions of X-bar theory that according to Chomsky (1995) should not be included in the strong version of the MP. Yet covert movement lives on and is if anything more frequently invoked than ever, as constituents are required to move covertly for purposes of feature-checking, and, as in the case that follows, also to satisfy the requirements of the Linear Correspondence Axiom (Kayne 1994), among which is that preceding constituents asymmetrically c-command following ones. Take the following derivation of a simple sentence from Hornstein (2009: 31, ex. 22).

(5)
a. Merge *her* with *likes*: [her likes]
b. Merge *v* with [her likes]: [v [her likes]]
c. Copy *likes* and merge with [v [her likes]]: [likes + v [her likes]]
d. Merge *John* with [likes + v [her likes]]: [John [likes + v [her likes]]]
e. Merge T^o with [John [likes + v [her likes]]]: [T^o [John [likes + v [her likes]]]}
f. Copy *John* and merge with TP: [John [John [likes + v [her likes]]]]][2]

[2] I have not copied Hornstein's derivation letter for letter, because it has undergone inadequate editing. The first bracketed segment in (b) is given as "likes her", although "her likes" is given as the consequence of the merge—an obvious error, though natural enough in view of (6). The use of italicization in the original is also inconsistent; I have repaired this by using italics for unmerged items and normal lower-case for merged items throughout. On a different level entirely, one might question the status of T^0 and TP, given Chomsky's proposal that "any structure formed by the computation" should be "constituted of elements already present in the lexical items selected for N; *no new objects are added in the course of computation* apart from rearrangements of lexical properties (in particular, no indices, bar levels in the sense of X-bar theory, etc." (Chomsky 1995: 228, emphasis added).

Copies are retained and originals deleted throughout. Thus a three-word sentence requires six operations, whereas on a naïve view of Merge, two would suffice:

(6) a. Merge *her* and *likes*: [likes her]
b. Merge *John* and [likes her]: [John [likes her]]

Why should not (6), rather than (5), be the way the brain does things? How likely is it that in the course of constructing sentences, the brain should have to move constituents repeatedly into new configurations? (6) is simpler, shorter, and requires less energy, and one of the things we do know about the brain is that it consumes an enormous amount of energy, maybe as much as a quarter of human energy, despite the fact that it forms only a small fraction of body mass. Indeed, the take-home message for biolinguists from the work of Cherniak (1994, 2005, Cherniak *et al*. 2002) and his associates should be not so much "non-genomic nativism" (interesting and reassuring though that may sound) as the fact that the brain's optimization of wiring patterns is driven precisely and exclusively by its own need for energetic economy.

At this stage, biolinguists are likely to respond, "But it's ridiculous to change tried and trusted analyses just because of vague intuitions about what the brain can and can't do." I agree, it would be, but that's not what I'm saying. All I'm saying is that if we are serious about biolinguistics we should start asking ourselves (and one another) whether it's okay to unquestioningly accept analyses whose motivation is mainly if not wholly theory-internal and which in many cases originated before anyone had started thinking about evolution or brain mechanisms and before there was even a hint of the MP. While some recent works such as Balari & Lorenzo (2013) show a commendable effort to explore physiological and computational foundations for the language faculty, such work is still at a fairly abstract level as compared with the kind of nuts-and-bolts, neurology-friendly description of what core grammatical computations of the specificity of (5) and (6) above really look like that, sooner or later (preferably sooner) must be a task for any adequate biolinguistic theory. A good place to start might be to determine which of the formal proposals of the MP would best fit such a theory; covert movement does not look like a promising candidate.

3. Problems with Biology (Old and New)

3.1. Natural Selection

A more immediately pressing issue concerns ways in which biolinguists understand (or misunderstand) biology. In the first place, they have problems with the notion of natural selection, up to and including a total failure to comprehend what it is and how it works. Typical is the following statement from the abstract for Longa (2001): "Natural selection is claimed to be the only way to explain complex design. The same assumption has also been held for language. However, sciences of complexity have shown, from a wide range of domains, the existence of a clear alternative: self-organisation, spontaneous patterns of order arising from chaos."

Natural selection could not 'explain' complex design, even if Pinker & Bloom (1990), Dennett (1995), and others who are not biologists think it does. In fact, natural selection does not provide a single one of the factors that go into creating design. As its name suggests, it selects, and that's all it does. The only sense in which it contributes to complex design is by (a) selecting certain alleles to fix in a population, (b) narrowing the search space by its successive choices, and (c) not undoing its own work, so that a ratchet effect preserves each step towards a better adaptation, forming a secure base for subsequent steps. What natural selection selects from—that is, where design and everything else come from—is variation, and many different factors generate that variation: the consequences of assortative mating, genetic mutations of several kinds, variation in gene expression, interactions between genes and genes and between genes and environment, and more. Self-organization is simply one of those factors, albeit a very potent one where the brain is concerned (Bickerton 2014).

Thus Longa is attacking a straw man, and his claim that any process is an 'alternative' to natural selection is simply a category mistake. All the processes that he and others treat as alternatives to natural selection are in fact suppliers of the materials without which natural selection could not even exist. Natural selection simply preserves whichever of these materials works best for a particular species in a particular situation. Whether the result of such preservation increases or reduces complexity depends entirely on the species and the situation concerned: the same force that resulted in eyes for formerly eyeless lineages may lead to blindness in others that formerly had eyes. Natural selection may best be conceptualized not as a designer (of complexity or anything else) but simply as a test that every biological development (not excluding the 'Promethean' mutation of Chomsky 2010: 59) has to pass. How else does Chomsky suppose that his mutation was "transmitted to offspring, coming to predominate"?

Even while they reject the 'creative' role so often attributed to natural selection by non-biologists such as Pinker and Dennett, evo-devo specialists, unlike their linguist aficionados, continue to recognize the centrality and ubiquity of natural selection. In an article specifically claiming that evo-devo represents not a mere addition but an alternative paradigm to neo-Darwinism, Laubichler (2010, 207) asserts that "[t]he developmental system determines whether or not a new phenotype is produced in the first place. Natural selection, *of course*, then decides its future fate" (emphasis added). The filtering (but exceptionless) role of natural selection is clearly expressed by de Robertis (2008: 194): "In sum, several types of mutations, some acting on the function of conserved developmental gene networks, provide the variation on which natural selection acts." The overall position taken by most, if not all, specialists in evo-devo is well expressed by Arthur (2011: x): "However, [the process of development] is seen as being important as well as, not instead of, changes in gene frequency caused by Darwinian natural selection. This is a crucial point, because some previous approaches to evolution advocated a dismissal of population genetics and a denial that micro-evolutionary changes within species form the basis of most long-term evolution; this denial is now seen to be mistaken."

Before we leave natural selection we should note a striking irony. Longa's 'alternative' to natural selection—"spontaneous patterns of order arising from chaos"—is virtually identical with the claims of computational linguists who oppose the whole idea of an innate universal grammar and promote iterated learning models in its place (Batali 1998, Kirby 2001, Brighton 2002, Christiansen & Ellefson 2002, etc.). If self-organization can single-handedly produce from chaos a brain capable of constructing language, why couldn't it take a short cut and directly produce language itself?

3.2. *Evo-Devo*

Biolinguistic problems extend to more recent developments in biology. In biolinguistics generally, evo-devo (the union of evolutionary and developmental biology) is routinely invoked in any discussion of evolutionary issues by biolinguists (Chomsky 2005, 2007, 2010, Berwick & Chomsky 2011, Boeckx 2006, Balari & Lorenzo 2009, Uriagereka 2011, etc.). Most of these are long on programmatic statements and short on detailed proposals with empirical support. For instance, Berwick & Chomsky (2011: 27) claim that in development "very slight changes can yield great differences in observed outcomes", but the sole example they offer involves pelvic spines in sticklebacks—hardly on a par with the emergence of what Maynard Smith & Szathmary (1995) classified as one of only eight major transitions in the whole of evolution.

Berwick & Chomsky (2011: 26) seem to suppose that such developmental changes can occur in an ecological vacuum, without any prompting from environment factors, hence they see language as resulting from some purely organism-internal factor, perhaps "absolute brain size" or "some minor chance mutation". But this runs counter to a large consensus among evo-devo specialists who repeatedly indicate that developmental changes, even if not directly provoked by external factors, can only take place if there is intensive interaction between genetic or epigenetic events and the environment and ecology of the organisms concerned. Nowhere is this better understood than in the field of ecological and evolutionary developmental biology ('eco-evo-devo'). For example, Ledón-Rettig & Pfennig (2011: 391) recommend taking the spadefoot toad, a species whose tadpoles show extensive phenotypic variation in response to "diverse environmental stimuli", as "a model system for addressing fundamental questions in ecological and evolutionary developmental biology (eco-evo-devo)." The authors go on to declare that "By characterizing and understanding the interconnectedness between *an organism's environment, its development responses, and its ecological interactions* in natural populations, such research promises to clarify further the role of the environment in not only selecting among diverse phenotypes, but also *creating such phenotypes in the first place*" (emphasis added; see also Blute 2008, Gilbert & Epel 2008, etc.).

But what is perhaps the most authoritative statement on the true relationship between internal and external forces is made in one of the most influential and most frequently cited treatises in the evo-devo paradigm (West-Eberhard 2003: 20), which deserves citation at some length. "First, environmental induction is a major initiator of adaptive evolutionary change. The origin and evolution of adaptive novelty do not await mutation; on the contrary, genes are followers, not leaders, in evolution. Second, evolutionary novelties result from the reorganization of existing phenotypes and the incorporation of environmental elements. Novel traits are not de novo constructions that depend on a series of genetic mutations."

Where does all this leave stickleback pelvic spines? According to Berwick & Chomsky (2011: 27), "[t]here are two kinds of stickleback fish, with or without spiky spines on the pelvis. About 10,000 years ago, a mutation in a genetic 'switch' near a gene involved in spine production differentiated the two varieties, one adapted to oceans and one adapted to lakes." Not only does this claim (like the associated suggestion that language evolution could have been triggered by a 'minor chance mutation') run directly counter to West-Eberhard's formulation, it is based on a serious distortion of the very papers that the authors cite as primary sources.

The primary sources the authors cite (Colossimo *et al*. 2004, 2005) have nothing at all to say about the presence or absence of spines in sticklebacks; both papers concern differing quantities of armored plates on oceanic and lacustrine varieties of the species in question (known as the "three-*spined* stickleback"), and in Colosimo *et al*. (2005) there are 126 references to these plates as against one mention of spines. The authors can only have derived the notion that the varietal differences involve spines rather than armor from a popular account of evo-devo in the *New Yorker* (Orr 2005) that they also cite. Furthermore, the notion that the change was due to a single mutation, unrelated to environment or ecology, is not supported by either of the primary sources. The very first sentence of one of these reads: "Particular phenotypic traits often evolve repeatedly when independent populations are *exposed to similar ecological conditions*" (Colosimo *et al*. 2005: 1928, emphasis added). Indeed, while mutation could have contributed to the physiological changes, Colossimo et al. note the occurrence of "repeated selection on the standing genetic variation already present in marine ancestors" and conclude that "the presence of a shared haplotype in most low-plated populations suggests that *selection on standing variation is the predominant mechanism* underlying the recent rapid evolution of changes in lateral plate patterns in wild sticklebacks" (Colosimo *et al*. 2005: 1932, emphasis added).

In other words, evo-devo factors are constrained by the resources of pre-existing phenotypes and internal developments are typically not stochastic processes but responses triggered by external (ecological, environmental) events. This is a hard pill for dedicated internalists to swallow (and a devotion to exclusively internal processes is key to most of biolinguists' problems with biology, as we will see) but swallow it they must if they want to engage in substantive dialog with biologists. The pill would be swallowed more easily if biolinguists were as cognizant of the other radical innovation in twenty-first century biology—niche construction theory (Odling-Smee *et al*. 2003, Laland & Sterelny 2006)—as they are of evo-devo. Indeed, the two areas complement one another (Laland *et al*. 2008) by showing precisely how developmental factors interact with environmental ones to bring about evolutionary innovations. The central thesis of niche construction theory is that animals whose livelihood is threatened by some environmental change may respond by trying to carve out a new ecological niche for which they are not genetically pre-adapted, whereupon the target of selection shifts to any traits that support exploitation of that niche, and both genetic and epigenetic factors combine to produce phenotypes that are progressively better adapted to the new niche. But niche-construction theory is mentioned once in the biolingusitic literature for every ten or even hundred times that evo-devo is mentioned.

Why this difference in the treatment of two equally radical and equally influential revisions of the 'Modern Synthesis' (MS) that is so often a target for biolinguistic disapproval? The answer perhaps lies in the dreaded word 'environment'. Generative grammar has been virtually from its beginning an internalist theory, allowing only endogenous factors to play a role in the development of the language faculty. More will be said on this score when we come to deal with biolinguistic assumptions about language evolution. Here, I would merely note the disproportion in the amount of attention given to evo-devo and niche construction as illustrating a tendency among biolinguists to cherry-pick biology for researchers whose work supports, or may be presented as supporting, traditional generative positions. This leads them to exaggerate both the extent and the significance of the changes biology is currently undergoing (e.g., "a multiplicity of stunning advances in biology and in evolutionary theory in the last several years have… completely reshaped the standard neo-Darwinian picture"; Piattelli-Palmarini 2008: 185). There is little doubt that within the next decade or two the MS of neo-Darwinism will undergo a substantial revision; the first shots have already been fired (Pigliucci & Müller 2010). There is equally little doubt that this revision will not amount to the kind of gross paradigm shift that many biolinguists hope for and expect--one that would sideline and demote, if it did not banish entirely, the specter of natural selection. For generativists, the title of the Pigliucci & Müller volume ("*Evolution: The Extended Synthesis*") will recall the Extended Standard Theory (EST) of Chomsky (1973). More than a mere similarity of names is involved. They should find it helpful to note that the relationship between the Extended Synthesis and the MS is very similar to that between the EST and the Standard Theory, in that in both cases the former is an extension rather than a replacement of the latter.

While appeals to evo-devo are ubiquitous in biolinguistic work on language evolution, I know of only two works by evo-devo biologists that directly and substantively address this topic. One is Scharff & Petri (2011), but this paper offers cold comfort for biolinguists. In the first place, it makes no reference to anything in the biolinguistic literature except for the Hauser *et al*. (2002) program of seeking precursors of language components in other species (see below). In the second place, its focus is on "discussing the evolution of language *in the context of animal vocalizations*" (emphasis added), thereby ruling out any consideration of the syntactic (recursion, etc.) or the semantic (mind-dependent concepts that Merge computes over) aspects of language, as well as invoking a notion of communicative continuity that is anathema to most biolinguists. The main body of the paper devotes itself to summarizing the present situation with regard to comparative animal studies and discussing the possible functions of FoxP2. In the third place, it culminates with the depressing finding that while FoxP2 obviously has *some* connection with language, it is still far from clear what that connection is. One of the few things the authors are sure of is that only two amino acids distinguish the human version of FoxP2 from the chimpanzee version and that these acids were very likely not the target of the selective sweep that affected human FoxP2 in the last few hundred thousand years. This means, of course, that the (so far) most plausible candidate for a recent recursion-enabling mutation looks likely to turn out a non-starter.

The second paper, Dor & Jablonka (2010), offers even colder comfort. This paper presents "a social-developmental, innovation-based theory of the evolution of language", at the core of which lies "the understanding that language itself, the socially constructed tool of communication, culturally evolved before its speakers were specifically prepared for it on the genetic level" (Dor & Jablonka 2010: 136). Jablonka is, of course, also co-author of one of the major treatises of the evo-devo paradigm (Jablonka & Lamb 2005), and in this context it is revealing to consider the reaction of a review of this book in the journal *Biolinguistics* (Piattelli-Palmarini 2008). Piattelli-Palmarini highly praises the overall evo-devo approach of the volume, but is deeply shocked when its authors seemingly abandon this approach in the case of language evolution, substituting a gradualist, culturally-driven account. He does not consider an alternative explanation: that biolinguists in general may have misunderstood evo-devo, distorting and exaggerating its emphasis on organismal-internal development, and that in consequence, when it comes to language evolution, evo-devo is no more friendly to orthodox biolinguistic accounts than it is to gradualist-externalist ones.

4. The Problems with Language Evolution

4.1. 'Design Features' and 'Precursors'

One might have hoped that when real biologists came on board, so to speak, biolinguists might have acquired a better understanding of modern biology. Unfortunately, this was not to be the case. Chomsky's collaboration with two biologists, Marc Hauser and Tecumseh Fitch, gave rise to a paper (Hauser *et al.* 2002) that most biolinguists treat with reverence as a classic example of "Science's Compass" (the section of *Science* in which the article originally appeared), pointing the way to all subsequent investigators of language evolution. Unfortunately, discussion of this paper has focused almost exclusively on quibbles about what is, and what is not, to be included in FLN (the faculty of language, narrowly conceived) as opposed to FLB (the totality of mechanisms involved in language—see Pinker & Jackendoff 2005, Fitch *et al*. 2005 etc.). Commentators failed to notice much more important and troubling aspects of the paper that related to biology rather than to linguistics.

The evolution of language must have taken place during the evolution of humans, as a part of that evolution, and indeed, given its importance in their subsequent development, as arguably the most important part of that evolution. In fact, surprisingly little of the literature, biolinguistic or other, makes any serious attempt to place language evolution in the context of human evolution. But even in that company, Hauser *et al*. (2002) stands out as being perhaps the only work on the evolution of language that includes not a single word about how humans evolved. (Imagine a paper about the evolution of dam-building without a word about how beavers evolved.)

The resultant space is filled with abuses of the comparative method. These involve decomposing language into component features or functions and then seeking other species where these components can allegedly be found. Ironically, this approach was pioneered by Hockett (1960) and Hockett & Altman (1968), while Hockett was developing what his *Wikipedia* entry describes as his "stinging criticisms of Chomskyan linguistics". Hockett's work was praised by Hauser (1996) and gave rise to the methods pursued by Hauser *et al*. (2002), which differed from Hockett's only in that "design features" such as 'semanticity' and 'duality of patterning' were replaced by more functional-sounding components such as "vocal imitation and invention", "capacity to acquire non-linguistic conceptual representation" and "imitation as a rational, intentional system". Such components were to be sought among species as diverse as whales, macaques, and starlings. It is assumed without argument throughout the paper that once these 'precursors of language' have been found and analyzed, language evolution has been definitively explained (except perhaps for recursion, unless this too can be found somewhere else in the animal kingdom). Subsequently, biolinguists have accepted, still without argument, that "building blocks of language" (Lorenzo 2012: 289) lie scattered across a wide range of species, just waiting to be assembled in the human brain.

I know of no species other than humans for which such a procedure has even been suggested, let alone put into practice. Standard texts in comparative evolutionary biology such as Harvey & Pagel (1991: 1) give as examples of the kinds of question comparative biologists might try to answer as "How much molecular evolution is neutral? Do large genomes slow down development? Is sperm competition important in the evolution of animal mating systems? What lifestyles select for large brains? Are extinction rates related to body size?" Nowhere is it suggested that any complex trait in a given species can be explained by breaking it into components and studying those components regardless of phylogenetic distance or ecological context.[3]

For example, studies of the evolution of echolocation in bats (Zentali 2003, Neuweiler 2003, Jones & Holderied 2007, Li *et al*. 2007) never look outside bats for explanations, even though a number of other species—whales, dolphins, oilbirds,

[3] It should be noted, however, that any set of components must be arbitrary and subjective, since the Hockett and Hauser *et al*. lists differ in every particular. I do not doubt that a third and equally disjoint set could be easily assembled. An anonymous reviewer sees the study of FLB as licensing the kind of comparative studies that I criticize "if one assumes the dichotomy" of FLN/FLB. But regardless of whether one assumes it or not, this isn't part of the solution—it's part of the problem! Such comparative studies are legitimate if the dichotomy is legitimate and the dichotomy is legitimate if the comparative studies are legitimate, but both are assumptions, and assumptions, moreover, that entail one another—if you think language divides in this way you must think that most language components are spread across other species, and conversely. This, though a blatant circularity, might be excusable if there weren't any other possible assumptions. But in fact at least one assumption is more plausible: niche construction theory strongly suggests that a novel trait with all its essential components (as distinct from mere pre-requisites) evolves in place, as a structured whole rather than a collection of mostly pre-existing attributes.

swiftlets, shrews, tenrecs—have developed different types of echolocation. Yet one of these sources, Li *et al*. (2007), suggests bat echolocation as a precursor of human language! This approach has been sharply criticized by some comparative psychologists. In reviewing studies that claim similarities between non-human traits and components of human language, Rendell *et al*. (2009: 238) state that "the loosely defined linguistic and informational constructs […] are problematic when elevated beyond metaphor and pressed into service as substantive explanation for the broad sweep of animal-signaling phenomena". According to Owren *et al*. (2010: 762) the procedure becomes abusive when "characteristics of signaling in an array of species are routinely tested for possible language-like properties, thereby turning the normal evolutionary approach on its head", and incidentally taking an approach to the comparative method that is not only "more a distraction than a boon to serious scientific inquiry" (Owren *et al*. 2010: 763) but also "both teleological and circular" (Rendell *et al*. 2009, loc. cit). Such an approach presupposes that humans are somehow special and should therefore be treated differently from other species. It is almost as if human language constituted the goal towards which animals were constantly striving but were as constantly falling short.

However, treating humans as special is far from the only failing of Hauser *et al*. (2002). Let us give the article the benefit of the doubt and assume that a novel and highly complex trait *could* have emerged in a single species through the accumulation of component parts from a large number of different species.[4] Would this explain how and why language evolved in humans and only in humans? Not really—in fact, not at all. Even if we make an additional leap of faith, assuming that all the components of the language faculty stem from "deep homologies" (Shubin *et al*. 2009), so that the same genetic and developmental mechanisms underlie vocal imitation in human and whales, constraints on rule learning in humans and macaques, and discrimination of sound patterns in humans and starlings, the real problems remain. How did all these components come together in a single species? Why did this happen in humans but not in any other species, some of which must have shared many, if not all, of the same components? When they came together, how and why did they form a single module

[4] Two anonymous reviewers found fault with my claim that a componential approach to the evolution of language is illegitimate and that biolinguists who adopt it are thereby misusing the comparative method. One provided me with a list of biology textbooks showing that "reusing and recombining pre-existing resources" is the default explanation for evolutionary novelties. This, as I was well aware, is indeed the case—where physiological form is concerned! But form is not behavior, and the texts I was referred to deal exclusively with form; not a single behavior is analyzed in this way. If and when biologists successfully decompose orb-web spinning, echolocation, bowerbird nest construction, or—to bring things closer to home—hymenopteran communication systems, I will be happy to reconsider my position. As things stand, to extrapolate from form to behavior is simply wishful thinking. The same reviewer also cited the work of Lynn Margulis as an indication that complex biological traits could derive from separate components, but this is again comparing apples with oranges. Margulis's work is concerned exclusively with whole prokaryotes that absorbed one another to form eukaryotes. Nobody is (I hope) claiming that humans emerged when a whale swallowed a macaque and a starling (three species Hauser *et al*. 2002 mention as possessing language precursors), but that is the only kind of process that might be analogous in the present context.

devoted to language? Why didn't they simply go on doing what they had done in other species, which by definition (since language is unique to humans) must have been things that had nothing to do with language?

Even stating all these questions does not exhaust the problems. In biology generally, it is assumed that novel evolutionary developments can be driven only by a particular set of circumstances that changes the selective pressures operating on the species in question. Hauser himself, when not associating with Chomsky, fully recognizes this, indeed takes it for granted. For instance, he asks: "What *special problems* do bats confront in their environment that might have selected for echolocation?" (Hauser 1996: 154, emphasis added). Similarly, he points out that "[t]he goal [in dealing with possible analogies—DB] isn't to mindlessly test every species under the sun, but rather, to think about the ways in which even distantly related species *might share common ecological or social problems*, thereby generating *common selective pressures* and ultimately, solutions given a set of constraints" (Hauser *et al*. 2007: 108; emphasis added). Most biologists would unquestioningly agree with this, but Hauser *et al*. (2002), like most of the biolinguistic literature, simply ignore any connection between novel traits and special external problems.

4.2. *Protolanguage*

If language didn't evolve to solve any special problem but emerged as a result of organism-internal developments, there need not be anything you could call proto-language. I can think of nothing more likely to create a barrier between bio-linguists and a majority of biologists than the former's insistence that language emerged ready-made, "pretty much as we know it today" (Boeckx 2012: 495). For most biologists it is axiomatic that any complex evolutionary trait has *real* pre-cursors, that is to say not separate alleged components in other species but im-mature versions of the complete trait, in the species concerned or its immediate ancestors, that would have similar functions but lack some of the mature trait's features, or have them only in some partially developed form, or both. Among biolinguists, however, protolanguage denial is the norm,[5] and possible real pre-cursors, as distinct from the illegitimate ones described in previous paragraphs, are often explicitly dismissed.

[5] An anonymous reviewer complained that s/he had found no protolanguage deniers among biolinguists apart from Berwick and Chomsky. I find this remark extraordinary in light of the fact that two more are cited in this section of this paper: Boeckx (see his remark that language emerged ready-made, "pretty much as we know it today" cited earlier in this pa-ragraph) and Piattelli-Palmarini (see below), who in the article there cited, without explicitly denying the possibility of a protolanguage, renders one effectively impossible by denying the possibility of a medium with words but without syntax and rejecting the belief that lang-uage could have evolved in a series of steps. The assumption of a sudden and rapid evo-lution of language some 50,000 to 100,000 years ago, shared by Hornstein (2009) and numer-ous other biolinguists, also entails that there cannot have been a protolanguage, regardless of whether this is explicitly claimed or not. Note also the absence of any discussion of proto-language (how it might have been constituted, or how it might relate to language) from virtually all biolinguistic accounts of language evolution apart from Fitch (2010). Bio-linguists for the most part do not even go to the trouble of denying protolanguage. Despite the number of authors that have discussed it, they simply assume it doesn't exist and is therefore not even worth talking about.

Consider the following, from Berwick & Chomsky (2011: 31): "Notice that there is no room in this picture for any precursors to language—say a language-like system with only short sentences. There is no rationale for postulation of such a system: to go from seven-word sentences to the discrete infinity of human language requires emergence of the same recursive procedure as to go from zero to infinity, and there is of course no evidence for such protolanguages." This echoes in slightly different words Chomsky's (2010: 53) claim that "There are many proposals involving precursors with a stipulated bound on Merge: for example, an operation to form two-word expressions from single words, perhaps to reduce memory load for the lexicon; then another operation to form three-word expressions, etc. Clearly there is no evidence from the historical or archaeological record for such stipulations…"

It is surely significant that though there have been many coherent arguments for the necessary existence of a protolanguage (Bickerton 1990, Jackendoff 1999, Fitch 2010, among others), none of them are answered or even mentioned here. In place of rational answers we find straw men or even outright falsehoods. Chomsky cites not a single example of the "many proposals" for protolanguages with stipulated sentence lengths, for the simple reason that there are none. No-one has suggested even a language with "short sentences", because utterances in protolanguage have never been claimed to be sentences. Sentences of natural language (and I know of no other kind) are propositions with syntax; protolinguistic utterances are propositions without syntax. As for absence of "evidence from the historical or archaeological record", protolanguage had disappeared tens if not hundreds of thousands of years before there was any 'historical record', while the 'archeological record', throughout history and prehistory alike, remains stubbornly silent on length and complexity of any utterance, whether in protolanguage or language.

Part of the problem is that Chomsky does not accept the existence of any way to put words together except through Merge, which is nothing if not a full-fledged syntactic process. His position here seems to me entirely irrational. Its full flavor cannot be grasped without quoting from a correspondence we had on this precise issue. When I wrote, "[p]rotolanguage consists of A + B + C…, i.e. there is no Merge," Chomsky replied, "[t]hat's commonly believed, but it's an error. A sequence a, b, c… that goes on indefinitely is formed by Merge: a, {a, b}. {{a, b}, c}, etc. (or some other notation, it doesn't matter). If we complicate the operation Merge by adding the principle of associativity, then we suppress {, } and look at it as a, b, c…. So a sequence is a special case of Merge, with added complications" (Noam Chomsky, p.c. , 16 March 2006).

The principle (more frequently described as 'property') of associativity is what makes processes like addition and subtraction apply to sequences like 1 + 3 + 6 regardless of the order in which the operations are carried out (in other words [1 + 3] + 6 yields an identical result to 1 + [3 + 6]). It follows that the order in which integers are arranged—1, 3, 6: 6, 3, 1; 3, 1, 6…—is equally immaterial. This is precisely true of the examples of types of protolanguage for which we do have historical records. For example, we have Nim Chimpsky's utterance (Terrace 1979):

(7) Give orange me give eat orange me eat orange give me eat orange give me you.

The propositional meaning of (7) is clear—Nim wants someone to give him an orange to eat—but the same meaning is conveyed by any arrangement of the constituents. This contrasts sharply with the English equivalent, where of the sentences in (8), only (8a) arranges its constituents in an order acceptable to English speakers:

(8) a. Give me [an] orange [to] eat.
b. *Me give [to] eat [an] orange
c. *Me [to] eat [an] orange give
d. * [To] eat [an] orange give me, etc.

Similarly we have utterances of pidgin speakers, many of which are semantically more opaque than Nim's utterances (note that for the sake of comprehensibility I have adjusted the phonology, interesting but irrelevant here, to fit English spelling conventions):

(9) a. And then, white meat tuna, three hundred seventy-five dollar, one ton—that's why, white meat kind, us go get 'em, no? (Japanese pidgin speaker, Hawaii)
b. And too much children, small children, house money pay, very hard time, no more money—poor. School children, my children go school, take house money pay, everything poor, too hard, that's why Korea Kim name one more time me marry. (Korean pidgin speaker, Hawaii)
c. Inside *lepo* (dirt) and *hanapa* (to cover) and blanket. (Filipino pidgin speaker, Hawaii)

Clearly, norms of constituent structure found in any natural language do not hold in (early-stage) pidgins such as that used in Hawaii from 1788 to the emergence of creole around 1900 (Roberts 1998), and subsequently by any adult immigrants who had arrived before the first creole speakers reached adulthood and began to influence the rest of the population (Bickerton & Odo 1976). If a medium lacks any consistent constituent structure, the most likely (perhaps the only possible) reason is because that medium lacks syntax—no principle or rule-governed process, certainly not Merge, determines the order in which words are strung together. Yet if Chomsky is correct, both pidgin speakers and Nim the chimp must first be applying Merge, then the 'principle of associativity' to undo any combinatorial properties peculiar to Merge, and then presumably some 'principle of distributivity' to arrive at the variable orderings shown in (9). Why anyone would have to go to such lengths when they have the obvious alternative of just stringing the words together anyhow is something only Chomsky, if anyone, can explain.

The impossibility of protolanguage is supported, from a different albeit complementary position, by Piattelli-Palmarini, who claims that there cannot be any form of language that has words but no syntax: "Words are fully syntactic entities and it's illusory to pretend that we can strip them of all syntactic valence to reconstruct an aboriginal non-compositional protolanguage made of words only, without syntax" (Piattelli-Palmarini 2010: 160). But stripping words of some or all of their syntactic valence is exactly what both Nim and the pidgin speakers do, in their rather different ways, in (7) and (9) respectively. Words to them, and presumably to the original pre-human protolanguage speakers, simply were not the same as words today. But if one is committed to essentialism, that assumption is impermissible.

The first words can't have had syntactic valences because there was no syntax to provide those valences. They were mere lexical shells, vocal or gestural forms that could carry a meaning of some sort (perhaps vaguer and more general than that carried by natural-language words) but little else. I call such things 'words' because what else could you call them—proto-words? They are not 'calls' or 'signals' in the animal-communication sense of those terms. They are symbolic, but they are more than mere symbols; a cross on a map may be a symbol for 'church', but you can't insert that into a conversation, however crude and simplified. All one has to do is accept that words, like language itself, evolved over time. To claim otherwise commits one to essentialism. And it is as a result of the intersection of essentialism and internalism that the most serious problems for biolinguistics arise.

5. Essentialism + Internalism = Anti-Biologism

Essentialism is anathema to most biologists for a variety of reasons, the "population thinking" of Mayr (1963) being the most frequently mentioned and the issue of speciation being perhaps the most relevant here. Regardless of whether changes occur in a species over long periods or in a rapid cascade (the likelier procedure under niche construction theory) there comes a time when some individuals descended from species X can no longer be regarded as members of species X but must be assigned to a new species, species Y. If it were possible to draw a hard and fast line anywhere in the process—if for instance each differentiating feature, from a new means of exploiting food sources to sterility and ultimately impossibility of hybridization, occurred simultaneously and instantaneously on the flipping of a set of developmental switches—essentialism might make some sense. But things don't happen that way. What is misleadingly characterized as a 'speciation *event*' may take hundreds of thousands if not millions of years to complete (Foley & Lahr 2005). In light of these facts, a sudden birth of words with all their current properties seems far less likely than a developmental process, giving time for a variety of influences, both external and internal, that would have progressively added to and refined properties of the original lexical shells.

Internalism runs equally counter to most biological thinking. Even those biologists who join with biolinguists in rejecting the MS (see, e.g., Dor & Jablonka 2010, Laubichler 2010 as cited above) concur with supporters of the MS in conceding that external forces and events are almost always instrumental in triggering evolutionary developments, as shown be the numerous citations of evo-devo authors in preceding sections. The consensus is most forcibly stated by

Müller (2010: 314) who gives a fully explicit statement of the sequence of events as perceived by evo-devo: "…the 'behavioral change comes first' position also gained new support from developmental psychology. Behavioral flexibility based on developmental plasticity is argued to result in behavioral neophenotypes, which in turn cause morphological innovation followed by genetic integration."

But internalism seems to be entailed by essentialism. Essentialism paints biolinguists into a corner by imposing a strict time limit: if language is deprived of true precursors, forms intermediate between animal communication and full language that arose in the two-million year history of the genus Homo, it cannot be older than the species that possesses it (~200 kya, at the most) and any universal grammar cannot be younger than the start of the human diaspora (90-60 kya). This gives insufficient time for any prolonged interaction with the environment or for any complex new traits to develop. In the words of Boeckx (2012: 495), "[t]he recent emergence of the language faculty is most compatible with the idea that at most one or two evolutionary innovations, combined with the cognitive resources available before the emergence of language, delivers our linguistic capacity pretty much as we know it today." Logically, only internal developments could bring this about in the narrow time-window available. Logically, but not biologically—the notion that a single mutation, or even a rapid cascade of mutations, could precipitate one of the eight major transitions in evolution is something that the geneticist Rebecca Cann has dismissed as "magical thinking" (Diller & Cann 2010).

Painting oneself into a corner always has negative consequences, and the essentialist-internalist corner is no exception. Chomsky (2010: 57) was perhaps the first to clearly spell out one of the most crucial differences between language and animal communication. The latter refers directly to what Chomsky called "mind-independent entities"—things out there in the world—whereas language does so only indirectly, having as its primary reference the "mind dependent entities" of categorical concepts. It does this by a process of lexicalization: by providing each of these concepts with an associated word. Words form a common currency that "mixes conceptual apples and oranges in virtue of them all being word-like things", as Boeckx (2012: 498) insightfully observes.

But where do words come from? The inability of biolinguistics (so far) to deal with this question is clearly shown by the fact that one leading biolinguist has published in the same year two contradictory explanations. Berwick, as second author in Miyagawa *et al*. (2013), commits himself to the opinion that the alarm-calls of vervet monkeys constitute "the simplest lexically based system" suggesting that "non-human primate calls may be construed as lexical" and thus formed precursors of "lexical structure" that only required to be joined with a computational component for full human language to emerge. But as lead author in Berwick *et al*. (2013) he takes a much more pessimistic (and realistic) view, noting that primate calls "lack key properties of human words" and that consequently "there is scant evidence on which to ground an evolutionary account of words" (p. 93).

Other biolinguists are equally baffled. While emphasizing that only words, not animal signals, are accessible to the operation Merge, Chomsky (2010) has nothing to say on their origin. Boeckx (2012: 499) does try to grapple with the issue, but he can only offer three possibilities: "random mutation", "an inevitable spandrel", or "we will never know". Another serendipitous mutation on top of the one for recursion is too much to swallow. A spandrel immediately prompts the question "spandrel of what?" for which no immediate answer is forthcoming. "We will never know" is a counsel of despair that has been uttered countless times in human history and almost as frequently refuted by the advances of the natural sciences.

What is, from a scientific perspective, totally unacceptable about this treatment of word origins is that there already exists, and has been in print for the last four years, a fully-developed explanation of how words could have originated (Bickerton 2009; see now the much fuller exposition in Bickerton 2014) that is nowhere discussed or even mentioned in the sources cited above. This is, moreover, an explanation explicitly licensed by Hauser *et al.*'s (2002: 1572) proposal of "the extension of the comparative method to all vertebrates (*and perhaps beyond*)" (emphasis added) as well as by the already-cited adjuration of Hauser *et al*. (2007: 108) to "think about the ways in which even distantly related species might share common ecological or social problems, thereby generating common selective pressures…" This explanation may be, as a reviewer remarked, "plain radical externalism", but so what? For any unbiased inquirer, this would no more exclude it from consideration than its being 'plain radical internalism', especially if internalist accounts had failed to supply any explanation at all. It is precisely this tendency among biolinguists to prejudge issues along ideological lines that I am objecting to.

More than a decade ago, Lyle Jenkins (2000) stated that the major goal of biolinguistics was to become integrated into the natural sciences. Alas, practices like those described here are taking it not nearer but further from that goal. Some biolinguists may react defensively to what I have written here. I think that would be a mistake, because this paper is not an attack and was never intended as an attack. I have merely tried to take an objective view of beliefs and practices that may have been held and carried out without full realization of their consequences. Unless biolinguists really wish to become isolated from other biological sciences, they should as a minimum think much more carefully than they have done to date about the issues raised here.

References

Arthur, Wallace. 2011. *Evolution: A Developmental Approach.* Oxford: Wiley Blackwell.

Balari, Sergio & Guillermo Lorenzo. 2009. Computational phenotypes: Where the theory of computation meets evo-devo. *Biolinguistics* 3, 2–60.

Balari, Sergio & Guillermo Lorenzo. 2013. *Computational Phenotypes: Towards an Evolutionary Developmental Biolinguistics.* Oxford: Oxford University Press.

Barss, Andrew & Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17, 347–354.

Batali, J. 1998. Computational simulations of the emergence of grammar. In J. R. Hurford, M. Studdert-Kennedy & C. Knight (eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases*, 405–426. Cambridge: Cambridge University Press.

Berwick, Robert. 2011. *Syntax facit saltum redux*: Biolinguistics and the leap to syntax. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The Biolinguistic Enterprise*, 65–99. Oxford: Oxford University Press.

Berwick, Robert & Noam Chomsky. 2011. The biolinguistic program: The current stage of its development. In Anna Maria di Sciullo & Cedric Boeckx (eds.), *The Biolinguistic Enterprise*, 19–41. Oxford: Oxford University Press.

Berwick, Robert, Angela D. Friederici, Noam Chomsky & Johan J. Bolhuis. 2013. Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, 17(2), 89–98.

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

Bickerton, Derek. 2009. *Adam's Tongue*. New York: Hill & Wang.

Bickerton, Derek. 2014. *More than Nature Needs: Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.

Bickerton, Derek & Carol Odo. 1976. *General Phonology and Pidgin Syntax*. University of Hawaii (mimeo).

Boeckx, Cedric. 2006. *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

Boeckx, Cedric. 2012. The emergence of language from a biolinguistic point of view. In Maggie Tallerman & Kathleen R. Gibson (eds.), *The Oxford Handbook of Language Evolution*, 492–501. Oxford: Oxford University Press.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The biolinguistics manifesto. *Biolinguistics* 1, 1–8.

Blute, Marion. 2008. Is it time for an updated 'Eco-Evo-Devo' definition of evolution by natural selection? *Spontaneous Generations* 2, 1–5.

Brighton, H. 2002. Compositional syntax from cultural transmission. *Artificial Life* 8, 25–54.

Burton, Strang & Jane Grimshaw. 1992. Co-ordination and VP-internal subjects. *Linguistic Inquiry* 21, 305–315.

Cherniak, Christopher. 1994. Philosophy and computational neuroanatomy. *Philosophical Studies* 73, 89–107.

Cherniak, Christopher. 2005. Innateness and brain-wiring optimization: Non genomic nativism. In Antonio Zilhao (ed.), *Evolution, Rationality and Cognition*, 103–112. New York: Routledge.

Cherniak, Christopher, Z. Mokhtarzada & Uri Nodelman, 2002. Optimal wiring models of neuroanatomy. In Giorgio A. Ascoli (ed.), *Computational Neuroanatomy: Principles and Methods*, Vol. 1, 71–82. New York: Humana Press.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory*(2), 113–124.

Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2(2), 137–167.

Chomsky, Noam. 1972. *Language and Mind*, enlarged edn. New York: Harcourt, Brace, Jovanovich.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart and Winston.

Chomsky, Noam. 1982. *Noam Chomsky on the Generative Enterprise: A Discussion with Riny Huybregts and Henk van Riemsdijk*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Chomsky, Noam. 1988. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.

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

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

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

Chomsky, Noam. 2010. Some simple evo-devo theses: How true might they be or languge? In Richard K. Larson, Viviane Deprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 45–62. Cambridge: Cambridge University Press.

Christiansen, M. H. & M. R. Ellefson. 2002. Linguistic adaptation without linguistic constraints: The role of sequential learning in language evolution. In Alison Wray (ed.), *Transitions to Language*, 335–358. Oxford: Oxford University Press.

Colosimo, Pamela F., Catherine L. Peichel, Kirsten Nereng, Benjamin K. Blackman, Michael D. Shapiro, Dolph Schluter & David M. Kingsley. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol* 2(5): e109.

Colosimo, Pamela F., Kim E. Hosemann, Sarita Balabhadra, Guadalupe Villarreal Jr., Mark Dickson, Jane Grimwood, Jeremy Schmutz, Richard M. Myers, Dolph Schluter & David M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307, 928–933.

Deacon, Terrence. 2010. A role for relaxed selection in the evolution of the language capacity. *Proceedings of the National Academy of Science*, Supplement 2: 9000–9006.

De Robertis, E. M. 2008. Evo-devo: Variations on ancestral themes. *Cell* 132, 185 195.

Diller, Karl C. & Rebecca L. Cann. 2010. The innateness of language: A view from genetics. In Andrew D. M. Smith, Marieke Schouwstra, Bart de Boer & Kenny Smith (eds.), *Proceedings of the 8th International Conference on the Evolution of Language*, 107–115. Singapore: World Scientific.

Di Sciullo, Anna Maria, Massimo Piattelli-Palmarini, Robert C. Berwick, Cedric Boeckx, Lyle Jenkins, Karin Stromswold, Lisa L. Cheng, Andrew Wedel, James McGilvray, Elly V. Gelderen & Thomas G. Bever. 2010. The biological nature of human language. *Biolinguistics* 4(1), 4–34.

Dor, Daniel & Eva Jablonka. 2010. Plasticity and canalization in the evolution of language communication: An evolutionary developmental approach. In

Richard K. Larson, Viviane Deprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 135–147. Cambridge: Cambridge University Press.

Embick, David & David Poeppel. 2005a. Mapping syntax using imaging: Prospects and problems for the study of neurolinguistic computation. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn. Oxford: Elsevier.

Embick, David & David Poeppel. 2005b. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-First Century Psycholinguistics: Four Cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Foley, Robert & Marta M. Lahr. 2005. The origins of modern humans: Insights into speciation. Paper presented at the conference on "Rethinking the Human Revolution", Cambridge, UK, September 2005.

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

Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179–210.

Gilbert, Scott F. & David Epel. 2008. *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution.* Sunderland, MA: Sinauer Associates.

Hall, Brian K. 1999. *Evolutionary Developmental Biology*. Dordrecht: Kluwer.

Harvey, Paul H. & Mark D. Pagel. 1991. *The Comparative Method in Evolutionary Biology*. Oxford: Oxford University Press.

Hauser, Marc D. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.

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

Hauser, Marc D., D. Barner & T. O'Donnell. 2007. Evolutionary linguistics: A new look at an old landscape. *Language Learning, and Development* 3, 101–132.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203, 89–96.

Hockett, Charles F. & Stuart Altmann. 1968. A note on design features. In Thomas A. Sebeok (ed.), *Animal Communication: Techniques of Study and Results of Research*, 61–72. Bloomington: Indiana University Press.

Hornstein, Norbert. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.

Jablonka, Eva & Marion J. Lamb. 2005. *Evolution in Four Dimensions.* Cambridge, MA: MIT Press.

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

Jenkins, Lyle. 2000. *Biolinguistics: Exploring the Biology of Language*. Cambridge: Cambridge University Press.

Jones, Gareth & Marc W. Holderied. 2007. Bat echolocation calls: adaptation and convergent evolution. *Proceedings of the Royal Society*, B 7, 274, 905–912.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax.* Cambridge, MA: MIT Press.

Kirby, S. 2001. Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation* 5(2), 102–110.

Koopman, Hilda & Dominique Sportiche. 1985. The position of subjects. *Lingua* 85, 211–258.

Laland, Kevin & Kim Sterelny. 2006. Seven reasons (not) to neglect niche construction. *Evolution* 60, 1751–1762.

Laland, Kevin N., F. John Odling-Smee & Scott F. Gilbert. 2008. Evo-devo and niche construction: Building bridges. *Journal of Experimental Zoology* 310B, 549–566.

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

Laubichler, Manfred D. 2010. Evo devo bio offers a significant challenge to the neo-Darwinian paradigm. In Francisco J. Ayala & Robert Arp (eds.), *Contemporary Debates in Philosophy of Biology,* 199–212. Malden, MA: Wiley Blackwell.

Ledón-Rettig, Cris C. & David W. Pfennig. 2011. Emerging model systems in eco-evo-devo: The environmentally responsive spadefoot toad. *Evolution and Development* 13, 391–400.

Li, Gang, Jinhong Wang, Stephen J. Rossiter, Gareth Jones & Shuyi Zhang. 2007. Accelerated FoxP2 evolution in echolocating bats. *PLoS ONE* 2(9): e900.

Longa, Víctor M, 2001. Sciences of complexity and language origins: An alternative to natural selection. *Journal of Literary Semantics* 30, 1–17.

Lorenzo, Guillermo. 2012. The evolution of the faculty of language. In Cedric Boeckx, Maria del Carmen Horno-Cheliz & Jose-Luis Mendivil-Giro (eds.), *Language from a Biological Point of View: Current Issues in Biolinguistics*, 263–289. Newcastle upon Tyne: Cambridge Scholars Publishing.

May, Robert. 1977. The grammar of quantification. Cambridge, MA: MIT dissertation.

Maynard Smith, John & Eors Szathmary. 1995. *The Major Transitions in Evolution*. Oxford: Freeman.

Mayr, Ernst. 1963. *Animal Species and Evolution*. Cambridge, MA: Belknap Press of Harvard University Press.

Miyagawa, Shigeru, Robert C. Berwick & Kazuo Okanoya. 2013. The emergence of hierarchical structure in human language. *Frontiers in Psychology* 4, 71.

Müller, Gerd B. 2010. Epigenetic innovation. In Massimo Pigliucci & Gerd B. Müller (eds.), *Evolution: The Extended Synthesis*. Cambridge, MA: MIT Press.

Neuweiler, Gerhard. 2003. Evolutionary aspects of bat echolocation. *Journal of Comparative Physiology A*, 189, 245–256.

Odling-Smee, F. John, Kevin N. Laland & Marcus. W. Feldman. 2003. *Niche Construction: The Neglected Process in Evolution.* Princeton, NJ: Princeton University Press.

Orr, H. Allen. 2005. Turned on: A revolution in the field of evolution? *New Yorker*, October 24.

Owren, Michael J., Drew Rendall & Michael J. Ryan. 2010. Redefining animal symbolism: Influence versus information in communication. *Biology and Philosophy* 25, 755–780.

Piattelli-Palmarini, M. 2008. Novel tools at the service of old ideas. *Biolinguistics* 2, 237–246.

Piattelli-Palmarini, M. 2010. What is language, that it may have evolved, and what is evolution, that it may apply to language. In Richard K. Larson, Viviane Deprez & Hiroko Yamakido (eds.), *The Evolution of Human Language: Biolinguistic Perspectives*, 148–162. Cambridge: Cambridge University Press.

Pigliucci, Massimo & Gerd B. Müller. 2010. *Evolution: The Extended Synthesis*. Cambridge, MA: MIT Press.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.

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

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20, 365–424.

Rendell, Drew, Michael J. Owren & Michael J. Ryan. 2009. What do animal signals mean? *Animal Behaviour* 78, 233–240.

Roberts, S. J. 1998. The role of diffusion in the genesis of Hawaiian Creole. *Language* 74, 1–39.

Scharff, Constance & Jana Petri. 2011. Evo-devo, deep homology and FoxP2: Implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society B 27* 366(1574), 2124–2132.

Shubin, Neil, Cliff Tabin & Sean Carroll. 2009. Deep homology and the origins of evolutionary novelty. *Nature* 457, 818–823.

Terrace, Herbert S. 1979. *Nim*. New York: Knopf.

Uriagereka, Juan. 2011. *Syntactic Anchors: On Syntactic Structuring*. Cambridge: Cambridge University Press.

West-Eberhard, Jane. 2003. *Developmental Plasticity and Evolution.* Oxford: Oxford University Press.

Zentali, Thomas R., Edward A. Wasserman, Olga F. Lazareva, Roger K. R. Thompson & Mary Jo Rattermann. 2008. Concept learning in animals. *Comparative Cognition and Behavior Reviews* 3, 13–45.

Derek Bickerton
University of Hawai'i at Mānoa
Department of Linguistics
569 Moore Hall
1890 East-West Road
Honolulu, HI 96822
USA
derbick@hawaii.rr.com

The Bird Is the Word

Bolhuis, Johan J. & Martin Everaert, eds. 2013. *Birdsong, Speech, and Language: Exploring the Evolution of Mind and Brain*. Cambridge, MA: MIT Press.

by Pedro Tiago Martins

The idea that birds might have something related to language that humans also seem to have has gone full circle: After the developments of linguistics and psychology during the 20th century put the 'uniquely human' in the center stage, with the help of failed or misled language experiments with animals, it now seems that perhaps birds have something to tell us after all. Even though the study of our closest cousins still very much dominates the understanding of our own biological and behavioral traits and tendencies, current, cutting-edge theories of language evolution now give a great deal of importance to the study of birds and their vocal abilities. It is not the case of course that scientists nowadays think that birds have 'human language' (they don't, as the reader will also have concluded, if he has ever been around birds and tried to have a conversation). Instead, what has happened is that recent developments in various fields have made the study of birds a perfectly fine component of any serious approach to the unveiling of the nature of language.

Indeed, the study of birdsong is now an emerging trend in the biolinguistic sciences. In recent years, many papers, talks, and some books have been devoted to the subject. Not surprisingly, most of the work on birdsong in the context of language studies has come from non-linguists, who are more in touch with the methodology and literature on animal studies and biology in general. The degree to which birdsong has at least fascinated linguists, however, is arguable at best. Phonology is the obvious core area of language study that should pay attention to it, but the subject is a rare sight in the phonology literature (with some exceptions by, for example, some of the contributors to the volume under review). Moreover, I suspect that the idea of even approaching it will seem ludicrous to most working phonologists today. Morris Halle's endorsement on the back cover of the book under review is somewhat revealing regarding this point, as he says that "[b]oth humans and birds produce and react to acoustic signals, but they do so in ways that have some similarities and many obvious differences" (emphasis mine). It is true that there are many obvious differences, but it is also true that for the most part we haven't been able to uncover and appreciate the similarities, partially because of a lack of interest. Of course, it is not the case that phonologists should personally be interested in a subject that apparently does not have

Part of this work has been supported by research funds from the Fundació Bosch i Gimpera.

ISSN 1450–3417

Biolinguistics 8: 097–107, 2014
http://www.biolinguistics.eu

much to do with what they were trained to analyze and explain, that is, I do not claim that the lack of discussion on birdsong represents a blatant omission in the phonological literature.[1]

Instead, I claim that this would be a good time and opportunity for phonologists and other linguists to reassess their claims about innateness and biology (UG) in light of a broader, evolutionary picture, of which birds and birdsong are also (a very important) part, on the one hand, and apply their knowledge of structural analysis to this domain, on the other. Perhaps phonologists even have it better than other linguists regarding the latter point; after all, other core areas of the study of language cannot rely on cues as concrete as speech or sign. As put by Philip J. Monahan and colleagues:

> [That less attention has been paid to the biolinguistics foundations of phonological systems than to those of syntactic ones] is surprising because we believe that there are a number of reasons that biolinguistic inquiry into this domain [phonology] should be more tractable. First fewer levels of abstraction separate the fundamental representations of phonology from the basic sensory input representations. This means that knowledge about how basic auditory information is represented and processed in both humans and animals is more likely to provide important insights into how phonological information could be represented and processed. […] Second there already exists an extensive literature to build on from cognitive psychology that has investigation the extent to which "speech is special" (Liberman 1996). […] Third, on most linguistic theories, phonological representations are the basic unit that connects sensory input and motor output. Therefore, by investigating the biological basis of phonological knowledge, we can benefit from existing evidence from other cognitive domains on the biological basis for sensory-motor translation, such as is needed for visually guided reaching.
>
> (Monahan *et al*. 2013: 233–234)

One would expect that the tools and methods developed in phonology could be used to describe and analyze the vocalizations of animals (namely, birds), but that has only very rarely happened. While the basic assumptions of phonology,

---

1 The lack of discussion on human biology tout court, which—one expects—is more fundamental than discussions on birds, is what we should see as a blatant omission. Surely, many linguists claim that generative grammar studies the biology of language, but in practice the field stays far apart from it. Witness, for example, the following quote by Martin Everaert and Riny Huybregts, in the first chapter of the volume under review: "Generative linguistics is biolinguistics and deals, for example, with properties of the genetic endowment of a human biological system for language (UG)" (p. 13). Quotes of this sort are a staple in the generative tradition; after all, the intention of Noam Chomsky in 1950s and 1960s was precisely to study the biological foundations of language, and be done with the mechanistic description of particular languages. However, the results of the generative program obviously didn't do justice to that intention. It is customary in the literature to repeatedly state the biological aspirations of generativism, and surely some major, paradigm-defining conceptual arguments have been put forward by generativists, but in reality the fingers of one hand might be enough to count the works in the generative tradition that actually and objectively have dealt with "properties of the genetic endowment of a human biological system for language." Thus, the general claim that "generative linguistics is biolinguistics" cannot be supported. Perhaps some of the work in generative linguistics more accurately falls within what Boeckx & Grohmann (2007) call "biolinguistics in the weak sense", that is, a concern for biolinguistic issues, but no real commitment to them, in the sense that real, biological explanations for biolinguistic phenomena are not sought.

as in linguistics, pertain to humans (usually, the rules, constraints and units of language are ascribed to human biology), there is nothing about the tools themselves that makes them only applicable to human phonology. Sure, some tweaking is in order, but as soon as intuition tells us that birdsong has structure, regardless of what accounts for it, we should be tempted to use the tools we have amassed and developed over the course of a century and apply them to this new world of sound. One could even hope to refine the theories and tools used in human phonology by applying them elsewhere, as it would be a good test for what is intrinsically human in phonology (accidentally or not) and what pertains to more general constraints of the animal or, even more interestingly, the physical world.

Choosing this path would most likely blur the line between phonology and phonetics, but this should not deter anyone. In fact, in recent years that line has been moved quite a distance. For example, Blevins (2004) has very convincingly shown that many phenomena usually considered phonological could in fact have a phonetic explanation, relying on natural rules of sound change and the way the production and perceptual systems work. In the same vein, although from a different angle, phonologists such as Samuels (2011) have put forward accounts of what remains phonological after those now external factors and whittled away.

More generally, part of the mission of biolinguistics should be the blurring of conceptual, epistemological, methodological and classificatory lines. Let me explain: Any science or program should abide by very rigorous definitions of the objects of its study and the entities it cares about, but I have the impression that very often, by focusing on the division of specific factors, and the ascription of any one phenomenon to each of them individually—say, as in the case of the three factors in language design (Chomsky 2005), or the FLN/FLB distinction (Hauser *et al.* 2002), often appealed to as if they were not closely related and intertwined in many ways—, scientists ignore the important interactions from which phenomena and ultimately explanations might arise. The same could perhaps be said about biological dichotomies such as continuity/discontinuity, adaptation/exaptation, or nature/ nurture, which so often take charge of discussions of evolution. This is not to say that deriving dynamic evolutionary explanations is easy, specially if the "trait" in question is language. Martin Everaert and Riny Huybregts appeal to Chomsky's pessimistic stance on language evolution in the first chapter:

> Chomsky addresses the question of why one would want to work on language evolution, and comes to a negative conclusion on the basis of considerations like the following. There are many simpler questions that are scarcely investigated, such as the evolution of communication in the hundreds of species of bees, because they are regarded as much too hard. (p. 19)

It might be true that studying language evolution is an extremely challenging task, but I think that this quote can be countered by using one of Chomsky's own mantras: that we should allow ourselves to be puzzled by the world. Chomsky has lived up to it: He as co-authored papers on language evolution (Hauser *et al.* 2002, Berwick *et al.* 2013), and birdsong has not been left out. With this state of mind in place, I will move on to the contents of the book itself.

This volume is divided in six parts: 'Introduction', 'Acquisition of Birdsong and Speech', 'Phonology and Syntax', 'Neurobiology of Song and Speech', 'Genes, Song, Speech, and Language', and 'Evolution of Song, Speech, and Language'. While these titles seem well delineated, the structure of the books feels somewhat looser: One must bear in mind that this is a collection of contributions to a meeting ("Birdsong, Speech and Language: Converging Mechanisms" in Utrecht, 2007), and some decisions as to where each belongs had to be made. For this reason, different foundational and basic notions of both the study of language and birdsong come up more than once in different parts of the book. Since there are so many contributors to this volume, I could not possibly analyze each of them in detail, so I will instead go over each chapter briefly.

Part I is devoted to the introduction of some of the pillars of the current biolinguistic study of language. Everaert and Huybregts start off the main matter of the book by offering an outline of the generative enterprise, along with a discussion on some of the ways that classical linguistic notions have been used in the study of birdsong. Apart from some brief remarks—such as the ones I already called attention to above—, not much is objectable, and the text will seem fairly straightforward, the goal clearly being the familiarization of the non-linguist reader with the kind of things that the linguist usually does and cares about.

Conversely, to familiarize the non-biologist, Tecumseh Fitch and Daniel Mietchen treats us to a subject very much at the heart of comparative biology: homology. As the authors define it:

> Unadorned, the term homology today denotes a character shared by two taxa by virtue of inheritance from a common ancestor, regardless of current form or function. Homologies are typically used by systematicists to construct taxonomies, and in phylogenetic analysis to reconstruct ancestral traits. (p. 45)

This definition is sufficient for capturing the gist of what it means to say structures X and Y are homologues, but as the authors point out, there is more to it than that. Homology may refer to different things, both historically and in current use. For example, Richard Owen, who coined the term, saw homologues as structures within and across species that were similar in form, while for Darwin they were structures that descended from a common ancestor (something Owen would call special homology). The authors discuss various types of homology, with a focus on the very interesting notion of deep homology: "[T]raits in two widely separated species [...] generated by one or more genes or genetic networks that are homologous." (p. 48). Fitch and Mietchen put forward that the famous FoxP2 gene might be a case of deep homology in the behaviour of birds and humans: This gene plays a very important role in vocal abilities in both groups.

Gary F. Marcus' contribution on the nature of trees as a way of mentally representing structural relations is no doubt an interesting read, but its eminent relevance for the volume is not obvious, and actually no references to birds are made throughout the text. It surely is interesting to question long-held assumptions about the way humans organize information in general and sentences in particular, but indeed the engagement with the main topic leaves something to be desired.

Neurobiologist Erich Jarvis lays out a very good summary of the topic he has become known for: brain pathways for vocal learning. He goes through the brain pathways that seem to be involved in vocal learning in both birds and mammals (although with restrictions on the mammalian side, due to the ethical concerns that prevent large-brained mammals to be the subjects of certain experimental procedures), and puts forward a motor theory of vocal learning, according to which there might be deeper constants that shape the development of vocal learning systems in distant species, which goes back to the topic of Fitch and Mietchen's chapter, deep homology.

Part II deals with the acquisition models of birdsong and language. Sanne Moorman and Johan J. Bolhuis lay out a very brief and to the point chapter on the similarities and differences between birdsong and human speech, by outlining some important characteristics of birdsong and later comparing them to the analogous human behavior. We learn that songbirds, like humans, imitate and learn their songs from their parents, and that some more nuanced behavior also occurs in birds, such as a tendency to imitate their conspecific song when more inputs are available. There also seems to be a 'sensitive period' for song acquisition in birds, which gives support to a more general notion of the critical period for language acquisition in linguistics and may help explain it. Other similarities are existence of different learning phases and the importance of auditory feedback, both humans and songbirds start by taking in the characteristics of their parents' song/speech very early on, and only later start imitating them, eventually perfecting their production also with the help of their own audition. Another similarity birdsong and human speech, comes from the structure behind them: Like human language, birdsong reveals syntactic structure, even though, as the authors conclude, the connections are not so clear in this case, as birds lack a lexicon and presumably also semantics.

Neil Smith and Ann Law choose to look at parametric variation as applied to birds. This choice strikes me as odd, since parametric variation as a biologically plausible or useful notion has been convincingly disputed (reference). The authors go on to briefly summarize Principles & Parameters theory (Chomsky 1981 *et seq*.) and to identify a number of criteria for determining whether parametric variation is true for a given system, namely birdsong. The conclusion is not entirely clear, to me and the relevance for the study of the relation between birdsong and human language does not seem clear, either. This contribution is closer to a formal exercise than an investigation into the nature of variation.

Olga Fehér and Ofer Tchernichovski try to answer the following question once put by Partha P. Mitra:

> Is it experimentally feasible to have a songbird colony established by an isolate founder ad then test if and how the improvised song produced by such isolate (ISO) birds would evolve toward wild-type (WT) song over generations without any external influence? (p. 144)

They set up an experiment by establishing "an 'island' bird colony with an isolate founder, and, in addition, performed a series of experiments where exposure to songs was controlled across 'generations' of song tutoring: ISO songs were imitated by unrelated juvenile birds who, when adults, trained another generation of birds." The authors concluded that juveniles are born with biases toward WT song, yet they must be exposed to songs to imitate before these biases take them there.

Frank Wijnen suggests in his chapter that there might be a general learning mechanism for the acquisition of linguistic categories. By pointing to some of the brain structures that underlie this mechanisms, his work opens way for cross-species experiments, although no considerations regarding this point are made.

Part III is devoted to the comparison between the phonological and the syntactic components of human language and birdsong. Moira Yip's contribution is a clear use of phonological notions developed in linguistics within the domain of birdsong. The author tries to find in birdsong parallels for all of the major sound units linguists are used to (from the syllable to the intonational phrase). Upon close inspection, one finds structure in birdsong that could reach up to six hierarchical levels, but it seems much more constrained in what can happen within that template than what it is the case in human language. Still, as the author points out, there is no doubt that the application of phonological tools to the domain of birdsong can only help understand both domains.

Eric Reuland offers a chapter on recursivity in language, a topic much discussed ever since Hauser *et al*. (2002). After a fairly straightforward overview of the subject, in which Reuland discusses the different types of reactions that the Hauser *et al*. (2002) have triggered—"(i) there is much more in language that is unique; (ii) recursivity is not just the basis of syntax, but recursivity is also—or even primarily—a property of the other components of the language system, notably the conceptual system; and (iii) manifestations of recursivity are also found in other species (p. 219)—, he sets out to assess the validity of (iii), which is clearly the most interesting in the context of this book. On the basis of work carried out by Gentner *et al.* (2006) on starlings, Reuland concludes that birds probably can't tell us much about recursivity, since it appears that they can only differentiate between patterns by making use of a good memory system, but no not internally represent those patterns in a recursive fashion.

Kazuo Okanoya looks at the syntax of birdsong. The term 'syntax' is used here to mean structure and hierarchy of sound: "Each birdsong note has specific acoustical properties; these song notes are ordered according to rules that are typically referred to as 'syntax'." Okanoya concludes that despite the lack of meaning and compositionally in birdsong, its syntax is a perfectly fine model of human language, since they have so much else in common (such as different stages of acquisition or similar brain mechanisms at play).

Carel ten Cate, Robert Lachlan, and Willem Zuidema present in their chapter the perfect follow up to previous two, by going over the phonological and syntactic structure of human speech and birdsong and coming to conclusions similar to the previous authors.

Irene M. Pepperberg, goes over some data from her experiments with Alex, a grey parrot, and showed that birds might have a sense oh phonology, or at least of some rudimentary phonotactics. Recent developments in the study of mechanisms underlying phonological awareness now give Pepperberg new grounds for supporting some of the results she has obtained over decades of experimentation. These results, even though highly publicized, may not have received the attention they deserve, always at odds with the idea that language and its most important parts are uniquely human.

Part IV is the devoted to the neurobiology of song and speech. Sophie K. Scott, Carolyn McGettigan, and Frank Eisner show speech perception and production, while very obviously related, also show a fair degree of dissociation, namely at the brain level, with the dorsolateral temporal lobes being associated with perception and the bilateral motor and premotor cortex, the left anterior insula and the left posterior-medial auditory cortex with production. These areas might be activated in various patterns by different behaviours, and not always the most intuitively obvious (for example, the movement of the articulatory apparatus as if to produce a word, even if this word is not vocalized, can activate the motor cortex). Once again, no connections are made with birdsong, but understanding the brain mechanisms behind simple characteristics of speech may open the way to cross-species investigations.

A good companion to the previous chapter, Sharon M. H. Gobes, Jonathan B. Fritz, and Johan J. Bolhuis' contribution, and review the literature on the neural mechanisms underlying vocal learning in songbirds in mammals, and find actually the neurological models based on birdsong are the ones that most help us understand the human case, since both at the genetic and neurological level birds come closer to humans than non-human primates when vocal learning abilities are considered.

Christopher Pallier offers a review of neurological data in support of the critical period hypothesis, ultimately concluding that this is till a very prolific area of research. Once again, basing the investigation not on linguistic data but on neurological discoveries allows for investigation in other domains and species, such as birds and their song.

Hermann Ackerman and Wolfram Ziegler discuss the components of human language that most consensually can be looked at (and for) in birds: the phonological components, leaving aside other, so-called syntactic components. The authors note that, despite the apparent lack of meaning (in the human, semantic sense) of bird vocalization, its development bears some similarities with human speech, an idea already discussed in the volume by, for example, Kazuo Okanoya. That is, while in the end there might be some human component(s) that, with speech, constitute human language, speech alone and birdsong are strikingly similar at various levels. The authors go on to discuss various cerebral structures, mechanisms and pathways which show that, despite some obvious differences, humans and birds (and other animals) have much in common.

In one of the chapters that I found the most interesting, Michale S. Fee and Michael A. Long give us a well-crafted summary of experimental work that, according to the authors, goes on to show that different time-scales of birdsong—notes, syllables, phrases, song, etc.—are not due to different time-scales of different brain mechanisms and circuits ('oscillations'), but rather from the execution of different 'behaviour modules' in succession. According to the model the authors propose:

> Each syllable is generated by different synaptically connected chains of neurons in HVC [an essential brain area in songbirds for song production]. Each

> chain forms a 100ms behavioral module that can be activated by the thalamic nucleus Uva. (p. 369)

This is extremely interesting in that it puts into perspective a great deal of literature on the importance of brain oscillations for the determination of sound units. Moreover, it gives the thalamus a very important role, and this applies not only to speech and language, but to the generation of complex hierarchical behaviors, as the authors explain.

Jonathan F. Prather and Richard Mooney discuss mirror-neurons, a very hot topic, specially since the collaboration between Giacommo Rizzolatti and Michael Arbib that started in the 1990s. The authors show—based on experiments in their research group—that some specific neurons in the swamp sparrow brain exhibit auditory-vocal correspondence, that is, they fire both when the swamp sparrow produces or listens to the same vocal gesture, as well as to similar gestures of other species. They go on to defend that the activity of these cells while singing is important for vocal learning itself, which renders these neurons an important component of vocal learning in birds.

Gabriël J. L. Beckers provides in his chapter a comparison between peripheral mechanisms of vocalization in birds and humans. As the author explains, interest in the mechanisms of bird vocalization have long been studied, but comparisons with humans have only more recently been studied, after the realization that, even though their vocal apparatuses are different, there is ample room for comparative approaches between the two. Beckers goes through years of research in the physical principles of vocal production (common to all tetra-pods), the role of respiration (both birds and humans mostly vocalize during expiration), the voice organ (different in humans and birds, since the latter make use of their specialized syrinx, instead of the larynx, for vocalizing), the mechanisms of voice production (which tends to periodic, rich sound waves in humans, and sinusoidal, pure tone sound waves in birds), or vocal tract filtering (much more dynamic in humans than songbirds, although also important in the latter). The general conclusion is that there is much still much work to be done in bird vocalization, since there is no way of generalizing the known mechanisms to all 9,000 species, by the author's count. One can hope that renewed interest in bird-song and its relation to language will inspire researchers to carry out more work of this sort in the upcoming years. Of course, all of the aspects mentioned also display important differences in birds and humans, which Beckers is right to point out.

Part V is devoted to genetics. Simon E. Fisher offers a general overview of what is known about the FOXP2 gene. The main lesson to take from this chapter is that, as the author rightfully acknowledges, FOXP2 is not a 'language gene':

> The investigations of FOXP2's potential role in human evolution have led to something of a revival of the 'speech gene'/'language gene' tag, particularly in the media. Is it worth reiterating here that it is unlikely that any single gene is responsible for the emergence of the complex suite of skills that allows members of our species to acquire spoken language. (p. 447)

Fisher goes on to explain how FOXP2 should be interpreted in the context of language evolution:

> FOXP2 is an ancient gene and is found in similar form in nonspeaking vertebrates, where we suspect it affects plasticity of circuits involved in sensorimotor integration and motor-skill learning. Perhaps the alterations of FOXP2 in the human lineage were important in enhancing these processes, at time points when spoken language was emerging an evolving (driven in part by other genetic and nongenetic factors). Such modifications may have had wider ramifications, beyond facilitating sequence of articulatory movements, if FOXP2 also plays roles in neural plasticity during procedural learning, for example. This fits in with the idea that our speech and language skills did not appear fully formed and out of the blue, instead involving recruitment and refinement of existing anatomical, physiological, and neurological systems (Fisher & Marcus, 2006). (p. 447)

Is it definitely worthwhile to go through Fisher's chapter and understand what FOXP2 can do, what other genes it's closely related with (e.g. CNTNAP2, and what it cannot do.

A great follow-up, Constance Scharff and Christopher K. Thompson's chapter do for birdsong what Fisher does for human speech and language. After covering some of same ground, the authors highlight the development and expression of FOXP2 in birds and the effects it has both for their vocal abilities and for the study of vocal learning in general.

Franck Ramus devotes his chapter to the way language disorders can inform our understanding of the genetic basis of language, and thus our (at least partial) understanding of its evolution (for a very recent take on this topic, see Benítez-Burraco & Boeckx 2013). More specifically Ramus focuses on developmental dyslexia, a disorder of reading acquisition. He describes the disorder's cognitive and neurological phenotypes, and reviews the genetic findings related to it, establishing some links with it and SLI and Speech Sound Disorder. The rigorous study of the genes that enter into this and other language disorders might prove to be essential in the unveiling of the genetic basis of different components of the language faculty: a language genetics.

Part VI is the devoted to the evolutionary models. Tecumseh Fitch presents a modern version of a Darwinian model for language evolution: musical protolanguage. This model gives pride of place to vocal control, and as such it is well place in the comparative, cross-species landscape. Under this model, music and speech, which require similar (or the same) brain mechanisms and genetic basis become, become two evolved versions of a more general phenomenon in the animal kingdom.

Kazuo Okanoya reviews several experimental results that seem to indicate that the evolutionary path of birdsong goes to various stages of complexity. Okanoya focuses on the Bengalese finch, and offers a scenario according to which complex song-note transitions became the object of sexual selection, later toned down by environmental needs and costs of several kinds, and need for a certain degree of 'simplicity' in order for a species member to be recognized. Later, domestication eliminated most selective and environmental pressures, allowing for the genetic basis of song complexity to materialize.

The final chapter, by Irene M. Pepperberg, offers and avian model for the evolution of vocal communication. Pepperberg uses the bellbird mirror neuron system, a species which appears to be at an intermediate stage of vocal learning, as a model for what could innate and learned, and more specifically for a model of what needed to have happened non-linguistic primates to homo sapiens.

This volume is perhaps the only one available that offers a state-of-the art perspective on birdsong and its relation to language. All chapters are written by acclaimed figures of their respective fields, offering mostly what is the result of their own work or their associates, which is a guarantee for an authoritative view on the subject. For the linguist, perhaps some of the chapters will seem too technical for a book of this sort, with extensive use of abbreviations of names of brain areas, making argumentation at times hard to follow. This obstacle notwithstanding, taking the extra step to understand and learn what each author—and field—is trying to tell us is ultimately rewarding. However, as I have pointed out above, some chapters, namely some the more linguistically oriented ones, seem odd in the context of birdsong and language, offering no angle that makes them a better fit for this volume than for any standard textbook in linguistics. Still, this is a very minor fault, and surely some interesting insights can also be derived from those chapters if the reader is willing to do the work.

But perhaps even more importantly than providing a state-of-the-art, this book gives the reader information about whose and which lines of work to pay attention to: Remember that these contributions come from a conference that took place almost seven years ago, and even though some important work published after that is often cited throughout the book, each of the topics covered and most of the authors have produced fresh literature in the meantime. Also, some chapters are reworked versions of work published after the meeting took place (this is either inferable from the text or explicitly acknowledged in most cases). This being said, perhaps this volume would have made a bigger impact two or three years ago, had it taken less time to put together.

In my opinion, *Birdsong, Speech, and Language* is a clear example of biolinguistics in the strong sense (cf. Boeckx & Grohmann 2007), with real biological explanations of biolinguistic phenomena. Even though we are in our comparisons obviously dealing with human language as one side of the equation, I think it's important to forget about its specificity, which many times results in very strict methodological limitations. After all, biology does not 'know' what language is, and the biological processes that lead to language do so because it so happens; the idea that everything or most things about human language are special and unique is no longer valid, and it might even turn out that nothing about it is unique, except perhaps for the fact that all of its components are in place in humans. Some of these components are present in other species, and some of these species are not closely related to humans. Birds are a very important and the most widely studied example of this subset of species, and the convergence of neurological, genetic and behavioral discoveries of recent years that are brought together in this book will only help solidify our understanding of human biology as one possible result of the biological processes that guide all animal life.

As a whole, the different contributions are a lesson for those who think that Chomskyan linguistics and biolinguistics are one and the same thing. This reaction is common from both Chomskyans and non-Chomskyans, and it has different but related consequences: The Chomskyan linguist will not try to go beyond what has become the canonical, descriptivist modus operandi of the field, while the non-Chomskyan will outright reject anything with 'Biolinguistics' in its name. Upon noticing that Noam Chomsky and Robert Berwick wrote the foreword for this book, I urge Chomskyans and anti-Chomskyans alike not to take their respective positions for granted, but rather to open the book, read the chapters and realize that perhaps things are not so black and white.

Birdsong, Speech, and Language is recommended not only to anyone who is interested in the foundations of birdsong and its relation to human language and speech, but also to anyone who wants to take a look at where biolinguistics is hopefully heading.

References

Benítez-Burraco, Antonio & Cedric Boeckx. 2013. Language disorders and language evolution: Constraints on hypotheses. *Biological Theory*, 1–6.

Berwick, Robert C., Angela D. Friederici, Noam Chomsky & Johan J. Bolhuis. 2013. Evolution, brain and the nature of language. *Trends in Cognitive Sciences* 17, 96–98.

Blevins, Juliette. 2004. *Evolutionary Phonology.* Cambridge: Cambridge University Press.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The biolinguistics manifesto. *Biolinguistics* 1, 1–8.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris.

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

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

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

Monahan, Philip J., Ellen F. Lau & William J. Idsardi. 2013. Computational primitives in phonology and their neural correlates. In Cedric Boeckx & Kleanthes K. Grohmann (eds.), *The Cambridge Handbook of Biolinguistics*, 233–256. Cambridge: Cambridge University Press

Samuels, Bridget. 2011. *Phonological Architecture: A Biolinguistic Perspective*. Oxford: Oxford University Press.

Pedro Tiago Martins
Universitat de Barcelona
Department of Linguistics
Gran Via de les Corts Catalanes, 585
08007 Barcelona
Spain

Centro de Linguística da
Universidade do Porto
Via Panorâmica s/n
4150-564 Porto
Portugal
martins@letras.up.pt

★ REVIEWS ★

From Wallace's Problem to Owen's Solution

Sergio Balari & Guillermo Lorenzo

Derek Bickerton. 2014. *More than Nature Needs. Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.

Out of the ground
Into the sky
Out of the sky
Into the ground
'Very Ape', Nirvana (1993)

1. Introduction

Nineteenth century biological treatises are almost unexceptionally beautifully crafted works of art, whose closing *perorations*—respecting the essentials of the classical rhetorical cannon—are a never-ending source of literary joy. Darwin's (1859) words at the end of *On the Origin of Species* are well remembered for this reason, but nowadays readers would also certainly be greatly surprised by other contemporary masters of the genre, were their works as accessible as Darwin's still are. We are particularly thinking of Richard Owen: For example, the final pages of the third and last volume of *The Anatomy of Vertebrates* (1868), where he confesses himself a tough-minded materialist concerning such delicate matters as the essence of life or the nature of mind; and, above all, the final section of *On the Nature of Limbs* (1849), a beautiful literary exercise that includes something alike to a twin-earth thought experiment. Owen speculates there that given the astronomers' certainty that basic conditions of other planets and their associated satellites might be as beneficial to the proliferation of life forms as those actually benefitting it in the Earth, such forms should be not very different from the ones we presently know here, as these are constrained by laws of variation affecting a restricted array of basic forms that define the realm of the organic. Let us quote Owen's beautiful prose at large:

> The naturalist and anatomist, in digesting the knowledge which the astronomer has been able to furnish regarding the planets and the mechanisms of the satellites for illuminating the night-season of the distant orbs that revolve round our common sun, can hardly avoid speculating on the organic mechanism that may exist to profit by such sources of light, and which

This work has been funded by the Ministerio de Ciencia e Innovación of the Spanish Government (FFI2010-14955—SB & GL), and the Generalitat de Catalunya (grant 2009SGR1079 to the Centre de Lingüística Teórica, Universitat Autònoma de Barcelona—SB).

ISSN 1450–3417

Biolinguistics 8: 108–119, 2014
http://www.biolinguistics.eu

> must exist, if the only conceivable purpose of those beneficent arrangements is to be fulfilled. But the laws of light, as of gravitation, being the same in Jupiter as here, the eyes of such creatures as may disport in the soft reflected beams of its moons will probably be organized on the same dioptric principles as those of the animals of a like grade of organization on this earth. And the inference as to the possibility of the vertebrate type being the basis of the organization of some of the inhabitants of other planets will not appear so hazardous, when it is remembered that the orbits or protective cavities of the eyes of the Vertebrata of this planet are constructed of modified vertebræ. Our thoughts are free to soar as far as any legitimate analogy may seen to guide them rightly in the boundless ocean of unknown truth. And if censure be merited for here indulging, even for a moment, in pure speculation, it may, perhaps, be disarmed by the reflection that the discovery of the vertebrate archetype could not fail to suggest to the Anatomist many possible modifications of it beyond those that we know to have been realized in this little orb of ours.
>
> The inspired Writer, the Poet and the Artist alone have been privileged to depict such. (Owen 1849 [2007]: 83–84)

It is a happy coincidence—as Richard Owen has never been a point of reference of Derek Bickerton's lifelong project of figuring out the origins and evolution of language—that Bickerton's own *peroration* in his last book also contains a reflection along the same speculative, almost dreamlike mood. Let us also quote (from here on, we will refer to Bickerton's book as *MTNN*, with page or chapter numbers added when appropriate):

> [T]here is still a strong possibility that, on any planet that hosts life forms, some species that has reached the chimpanzee-dolphin-crow level of cognitive capacity will eventually adopt a niche similar to that occupied by ants, bees, and human ancestors on this planet. […] [T]hen other planets with "intelligent life" become perhaps unavoidable.
>
> In the week these lines were written, Harvard astronomers, analyzing new data from the Kepler telescope, estimated that there might be as many as 17 billion Earth-size planets in the Milky Way alone, a sizable percentage of which would have orbits within a zone congenial to life […]. Earth, far from being the galactic anomaly many previously believed, is as ordinary a planet as the last common ancestor of apes and humans was an ordinary primate. To speculate further is premature, but these findings strongly suggest that the array of life forms on these planets may differ little in their cognitive spread from those found here, and that consequently "intelligent life," far from being a rare or even unique aberration, may have multiple loci throughout the universe. (*MTNN*: 273)

This review of *MTNN* is mostly devoted to show that what strikes us as a happy coincidence is more than a simple matter of literary style.

2. The return of another hopeful monster (or the raise of neo-Owenian biolinguistics)

Surely enough, readers may have guessed that Owen and Bickerton are defending exactly the same stance in their respective texts above: Namely, that given patterns of organization unavoidably linked to particular environmental conditions, the array of conceivable paths of diversification from such patterns must be constrained (and radically so) in a law-like manner. Granted that Owen is referring to the constrained plasticity of vertebrae to compound structures like cavities or limbs, while Bickerton is writing about brains and their associated cognitive capacities. But it is an additional happy coincidence that Owen also devoted similar reflections to the case of brains. It can be witnessed in passages like the following one, which in the original is preceded by an exposition of the major morphological distinguishing features of the varyingly complex versions of the mammalian brain:

> In Man the brain presents as ascensive step in development, higher and more strongly marked than that by which the preceding subclass was distinguished from the one below it. Nor only do the cerebral hemispheres overlap the olfactory lobes and cerebellum, but they extend in advanced of the one and further back to the other. Their posterior development is so marked that anthropotomists have assigned to that part the character and name of a 'third lobe:' it is peculiar and common to the genus *Homo*: equally peculiar is the 'posterior horn of the lateral ventricle' and the 'hippocampus minor,' which characterize the hind lobe of each hemisphere. The superficial grey matter of the cerebrum, through the number and depth of the convolutions, attains its maximum of extent in Man.
>
> Peculiar mental powers are associated with this highest form of brain, and their consequences wonderfully illustrate the value of the cerebral character; according to my estimate of which, I am led to regard the genus *Homo* as not merely a representation of a distinctive order, but of a distinct subclass, of the Mammalia, for which I propose the name *ARCHENCEPHALA*.
>
> (Owen 1859: 25–26)

We have only to lament that Owen's ideas in passages like this one have been historically distorted to the point of making them a gross caricature of their real import. For Owen's theses were very clear: (1) that the human brain is but a particular variant of the same organ in different species (or "homologues"), the diversity of which may be described by means of a few distinctive criteria; (2) that, as a matter of observation, this particular variant diverges from its closest homologues more than the latter diverge from their corresponding ones down in a scale of morphological complexity; and (3) that such a morphological contrast correlates with the gulf between the "mental powers" of humans relatively to that of the chimpanzee and other quadrumana, as "intelligence" (using now Bickerton's word) is nothing but the organic activity of brains (Balari & Lorenzo 2013a). More than a century and a half ago, Owen was as convinced as Bickerton is today that the "specialness" of humans ("peculiar mental powers are associated with this highest form of brain," Owen 1859: 26; "humans have large brains with unusual computational capacities," *MTNN*: 45) was a matter of fact, not of faith. No wonder their solutions to such a defying biological question are constructed along very similar lines. As defenders of a self-assumedly Owenian project aimed at disentangling some of the most recalcitrant difficulties in order to frame the evolutionary understanding of cognition and language within normal explanatory parameters (Balari & Lorenzo 2013b), we cannot but welcome Bickerton's new and highly promising attempt in a congenial direction.

We want to stress from the start that *MTNN* is a quantum leap relatively to Bickerton's previous *Adam's Tongue*—a failed effort to explain "how humans made language, and language made humans," as the book promised in its subtitle (Bickerton 2009). As Bickerton self-acknowledged after a review that we targeted to the book (Balari & Lorenzo 2010a), *Adam's Tongue* was "about the transition from the alingual state that characterizes all other species to something that might qualify as a genuine precursor of language" (Bickerton 2010: 128), but it had almost nothing to offer beyond that, in clear contradiction with its declared purpose. In *MTNN* Bickerton offers an honest diagnosis of why *Adam's Tongue* was a flawed project and also a very interesting plot to overcome its many shortcomings.

As for the diagnosis, Bickerton coins in the book the suggestive name of "Wallace's Problem" as a shortcut for referring to the problem that threatened to leave his original project into a dead end. This apt name reminds us Alfred Russel Wallace's conclusion that such typically human skills as language or music, but also some physical characters like the loss of hair or the shape of the hand, were far beyond the explanatory scope of Natural Selection (see, for example, Wallace 1870, and the valuable comments in Shermer 2002: Ch. 8), the idea that Wallace conceived independently of Darwin (Darwin & Wallace 1858). Some decades after originally formulating his point, Wallace expounded it with the following words:

> In ["The limits of natural selection as applied to man"] I apply Darwin's principle of natural selection, acting solely by means of "utilities," to show that certain physical modifications and mental faculties of man could not have been acquired through the preservation of useful variations, because there is some direct evidence to show that they *were not* and *are not* useful in the ordinary sense, or, as Professor Lloyd Morgan well puts it, not of "life-preserving value," while there is absolutely no evidence to show that they were so. In reply, Darwin appealed to the effects of female choice in developing these characteristics, of which, however, not a particle of evidence is to be found among existing savages races. (Wallace 1908: 212–213)

"Wallace's problem" was certainly Bickerton's problem in *Adam's Tongue*, for he was incapable to invent a just-so story there able to make sense of the releasing of early humans from the state of proto-linguistic communication that, according to Bickerton's tale, living in a confrontational scavenging niche had left them into. We will not assess here the plausibility of this latter thesis. It has already been the target of enough criticism (Balari & Lorenzo 2010a, Arbib 2011, Clark 2011), which Bickerton has decided to plainly ignore in the new book.

As for the plot, the main novelty of *MTNN* is that Bickerton strongly adheres to the idea now that no adaptive narrative is capable of dealing with the complexities and subtleties not just of full-fledged language, but even of interlanguage phenomena like early child language or pidgins in the way of becoming creoles. Obviously enough, this does not entail (contra Wallace) that for Bickerton, no biological explanation can be offered to such natural developments. As a further happy coincidence, Bickerton's line of argumentation again converges with that of Owen here, who after wielding some very thoughtful criticisms against the power and the scope of Natural Selection (Owen 1860, 1866:

Preface, 1868: Ch. XL), largely ignored by today's historians and philosophers of science, articulated an alternative based on the inherently self-organizing capacity of organic matter (Balari & Lorenzo, in progress).

Bickerton's recipe for solving Wallace's problem involves disassembling particular languages in order to take apart, on the one hand, a universal component made of very basic computational operations and, on the other hand, grammar particular rules that complete and repair the former in areas of under-specification when applied in the communicative uses of language. According to Bickerton, the fact that evolutionary linguistics had hitherto been incapable of understanding the difference between the former (biological) stratum of languages and the latter (cultural) one is the main responsible of the state of stagnation of the field, mostly devoted to Byzantine debates between nativist and empiricist oriented opinions. In his own words:

> In retrospect it seems bizarre that nobody, throughout this debate, proposed a principled and systematic distinction between those parts of syntax that were biologically given and those that had to be acquired through acculturation into one of the many thousands of speech communities.
>
> But what you have just read includes, and to a considerable extent depends on, the first coherent theory of syntax that makes such a distinction. (*MTNN*: 274)

Fair enough; but we cannot escape pointing out that the statement in the first paragraph is, to say the least, inaccurate: Prior to Bickerton's book, the distinction was at least very clearly established, and along very similar lines, in Balari & Lorenzo (2013b), also as a necessary point of departure for an evolutionary understanding of language. Let us quote:

> This problem, which we will refer to as the "dual nature of language," concerns a pervading confusion, both in linguistics in general and in evolutionary linguistics in particular, between the biological/computational system with which all members of the human species are endowed and the psychological/cultural systems every member of the species is capable of developing—namely, for the lack of a better term, "grammars." […] The issue can be summarized very briefly by stating that it is one thing to investigate the origin of FL (with the specific technical meaning we reserve for this term here) and another, different thing to speculate on the process, contexts, and contingencies that favored the emergence of grammatical systems. We regard the former as a strictly organic question, whereas the latter concerns the interaction between biology and culture. Our contention here is that the former, the organic aspect of language, is basic—actually, prior—if we want any just-so story about the invention of grammars to make any sense at all, but also if we want to achieve a better understanding of language as a whole. (Balari & Lorenzo 2013b: 7–8)

Let us also clear up that we are not raising this point here as a question of property rights—incidentally, a very common issue in Victorian biology (Richards 1987). We simply want to underlie it as a further reason to align Bickerton's new ideas with the neo-Owenian current of thought defended in our own book. Obviously enough, differences also exist between the respective approaches: For instance, Bickerton's biological level is very close to the barebones of current minimalist analyses (Chomsky 1995, and subsequent works), while our main source of inspiration are the barest essentials of the computational/representational theory of mind of approaches like that of Pylyshyn (1980). As for the cultural level, the two approaches are congenial in underscoring its value both in individual acquisition and use, but while Bickerton seems to emphasize the latter, in Balari & Lorenzo (2013b), we adhere to a primacy thesis theoretically favoring the former (Minelli 2003, Balari & Lorenzo, submitted). In this sense, Bickerton's ideas run parallel to other minimalist framed approaches to language evolution, like Longa *et al.* (2011), where it is also stressed the pressure of vocabulary explosion in the emergence of a first stratum of computational complexity (Longa *et al.* 2011: 601–604), as well as the needs of exteriorization as the main driving force leading to a further complexity stratum underlying phenomena like long-distance dependences, crossed paths, parallelisms, and so on (Longa *et al.* 2011: 610–615). But leaving aside details like these, the two approaches are coherent enough as to deem them both as neo-Owenian, for they equally give support to the idea that brain evolution as to be the site of "unusual computational capacities" (Bickerton) could not possibly be driven by Natural Selection, the alternative being certain spontaneous capacities for reorganization. In Bickerton's words:

> Indeed whether the brain shows diversity within uniformity or uniformity within diversity seems to depend not on any kind of external pressure but solely on the brain's ability to optimize its own resources. (*MTNN*: 119)

In one of the passages where Owen confronted the means by which the complementary mechanisms of Natural Selection and of Derivation (Owen's name for his suggested principle) worked, he wrote the following:

> 'Derivation' holds that every species changes, in time, by virtue of inherent tendencies thereto. 'Natural Selection' holds that no such change takes place without the influence of altered external circumstances educing or selecting such change. (Owen 1868: 808)

Clearly enough, Bickerton's is a derivative rather than a selective theory—also in concurrence with Balari & Lorenzo (2013b) here. Again, there exist some differences between one approach and the other as in, for example, the role the latter concede to standard evo-devo mechanisms in driving the process, where the former shows much more confidence in 'third factor' (Chomsky 2005) principles of sorts alone to the same effect. But, as an aside, let us say that we understand (even if we do not share) Bickerton's (*MTNN*: 51–53) distrust in the application of Evo-Devo to the case of language, for it is actually the case that it has hitherto been vacuously appealed to more than truly applied in the evolutionary explanation of the language faculty (as an example see Chomsky 2010, and for a critical appraisal Benítez-Burraco & Longa 2010). We however think that Balari & Lorenzo (2013b: Ch. 6) proves that this is not necessarily so.

3. A brief note on the primacy of the developmental role

In the preceding paragraphs, we have highlighted a number of points of contact between Bickerton's proposals and the ones set forth by ourselves in Balari &

Lorenzo (2013b). Our main motivation, so far, was to emphasize that the solution to Wallace's problem is more Owenian than Darwinian, both in form and in spirit and, hence, closer to Evo-Devo contentions than Bickerton actually suspects. In doing so, we have mostly enumerated those elements where Bickerton's work and ours either agree or diverge, without paying much attention to the details. In the following, however, we would like to concentrate on one particular aspect of *MTNN* which, in our opinion, is perhaps the most interesting and welcome of the whole book: Bickerton's approach to variation, acquisition, and creolization.

In the first chapter of our *Computational Phenotypes*, right after expressing the need to recognize the dual nature of language in the terms illustrated by the quotation above, we delineated a very sketchy proposal as to how grammatical systems might have emerged, with some hints also as to how the problem of variation could be dealt with. Our proposal boiled down to two main assumptions: (1) that no selective theory based on the idea of the optimization of communication could ever be able to explain the presence of such grammatical features as case o agreement markers in some languages but not in others; and (2) that such features could nevertheless be understood as accidental products of the process of acquiring language, for which they acted as scaffolds and which were later preserved in the adult system for no specific purpose apart perhaps from this developmental role (Balari & Lorenzo 2013b: 15–20).

To be sure, our proposal did not go beyond this promissory note and the (implicit) indication that this was a topic for future research. *Computational Phenotypes* was, after all, a book intended to provide a detailed account of the origins of language as a computational system. In *MTNN*, however, Bickerton, in his clear bet for breadth rather than depth, devotes the second part of his book to outline a much more articulated proposal to explain the emergence of grammars and variation (*MTNN*: Chs. 6–8). Now, if we took the two or three pieces of the puzzle that we set out on the table and then added those contributed by Bickerton, the whole would conform a rather coherent and consistent image, with just a little distortion perhaps at the edges.

To summarize Bickerton's position, he doesn't believe either in the adaptive value of a single grammatical feature, although he nonetheless does believe that a collection of such features inserted in an otherwise greatly underspecified grammar would make of it a better tool for communication that would, in turn, "confer enhanced fitness" (*MTNN*: 153) to those possessing it. Thus, grammatical features would have emerged in order to repair the radical underspecification left by the biological component of language through what Bickerton describes as a speaker–hearer arms-race of sorts, where speakers struggle to minimize costs during the production of utterances and hearers wanting a maximal precision in order to prevent misunderstanding.

At first glance, this looks like a model where grammatical features are an *adult* invention introduced by horizontal diffusion (Labov 2010: Ch. 15) through several rounds of more or less fruitful acts of communication. But this cannot be so if we take into account the chapter Bickerton devotes to creoles. In chapter 8 of *MTNN*, Bickerton presents a revival of his Language Bioprogram Hypothesis (Bickerton 1981), now refurbished to fit into the more streamlined conception of UG he presents in the first part of the book. Refurbishments are minimal and

Bickerton makes a strong case for his original idea that creoles are the product of children acquiring a pidgin as their first language in a process that is completed in a single generation. We are no creolists and we can therefore not fully evaluate the details, but it is certainly the case that the picture Bickerton portrays in *MTNN* fits much better than the alternative offered by his adversaries into the well-grounded assumption that variation and change are to a large extent the product of vertical transmission from parents to children (Ringe *et al*. 2002, Labov 2010). Note that this position is accepted even by those, like Peter Trudgill, for example, who see in culture a strong factor influencing grammatical structure (Trudgill 2004, 2011).

Accordingly, Bickerton's story would run like this: Grammatical systems would have emerged through transmission of a pidgin-like protolanguage from parents to children as repair strategies to fill in the gaps left by a highly underspecified UG component. It remains an open question whether, as in the case of creoles, this would have occurred in a single generation or in subsequent transmission rounds as suggested by Kirby's Iterated Learning models (e.g., Kirby 2013); but this is just a secondary question that would certainly not undermine Bickerton's model as it is quite likely that the sociolinguistic and biological context of our ancestors was not comparable to the one giving rise to creoles.

We would like however to point out what we believe to be a flaw in this scenario. Bickerton throughout the second part of *MTNN* seems to hesitate between settling on a "facilitation-of-processing" role versus a "facilitation-of-efficient-communication" role for grammatical features. Granted, these are perhaps two sides of the same coin, but we suspect that both introduce an ecological factor in the explanation that runs the risk of seriously undermine Bickerton's explanation of the nature of linguistic variation. The point is simple and can be summarized with the following question: If what motivated grammatical features was just the need for much more efficient communication and understanding, why don't we observe what in general is observed when similar ecological conditions occur, namely convergence; or, in other words, why don't all languages have case, agreement, etc.? Indeed, why is there variation at all?

Obviously enough, the question would deserve a full-length monograph, but we would like to suggest here what we see as the key to solve it one day: Grammatical features emerged (where they emerged) accidentally in the course of language development during vertical transmission. Where they were present, they acted as "ontogenetic adaptations" (Oppenheim 1981, 1984) capable of "scaffolding" (Caporael *et al*. 2014) later stages of the process to the point that some became "generatively entrenched" and to the extent that some downstream features depended upon them (Wimsatt 1986). For example, an interpretation along these lines may be appropriate to explain some observations concerning children learning German, who do not master the intricacies of V2 phenomena until full completion of the agreement paradigm (Clahsen 1986). Moreover, well-known developmental delays selectively affecting the agreement system seem to have a similar cascading effect in children at older ages, with signs of recovering showing up after intensive therapy exclusively focused on agreement (Clahsen & Hansen 1997; see Balari & Lorenzo, submitted, for an interpretation). This "primacy of developmental role" view does not suffer from the same shortcomings as the "ecological" approach of Bickerton. Well-understood development does not entail "development toward," but "development from" certain previous conditions; its conceptualization must emphasize "processes" instead of "outcomes" (Moore 2003). So it is not aimed at certain pre-established goals (Thelen & Smith 1994), but it is constructed upon mere stages contingently paving the way for further stages (Oyama *et al*. 2001). Within a viewpoint like this, the expectation is not that of convergence toward predictably optimal solutions, but rather patterns of ramifications of unexpected scaffolding effects.

It goes without saying that this is an extremely sketchy idea, but we think that it ought to serve to open a salutary debate within a framework in which the basics have already more or less been agreed upon. When properly worked out, it hopefully could serve to rescue Bickerton from his present state of distrust and dissatisfaction with the idea of disentangling how language and languages could possibly have emerged through the intertwined action of biological and cultural development.

4. Conclusion

Bickerton's book describes a really fascinating journey out of the ground of proto-linguistic communication into the sky of mental computations, and out of the sky of mental computations into the ground of the vagaries of linguistic communication as dealt with by grammars. While we strongly disagree with the starting point (i.e. a scavenging related protolanguage) and only partially agree with Bickerton's favored explanations for the underlying motivation of the endpoint (i.e. full-fledged grammatical systems), we sincerely believe that the journey is worth pursuing and, in any event, discussing. We like *MTNN*. This may come as a surprise to those who followed our debate around *Adam's Tongue* (Balari 2010a, 2010b, Bickerton 2010); maybe to Bickerton himself, who confesses in the book to be ready for strong criticisms (*MTNN*: 271). But we actually like *MTNN* for the same reasons that we disliked *Adam's Tongue*: i.e. for scientific, and not personal or clannish reasons. With its blind spots, *MTNN* is a well-argued and well-written book, in which Bickerton displays great doses of expertise in the fields of grammatical analysis, language acquisition theorizing and creole studies. So we feel proud of underscoring the continuity between our own efforts in Balari & Lorenzo (2013b) and related works and Bickerton's new project. From a practical perspective, Bickerton's is certainly a good and accessible introduction to this way of looking at the origin and the evolution of language—and of languages—and a most welcome contribution to what above we termed the neo-Owenian current in biological thought.

We hope that Richard Owen would also be proud of all us.

References

Arbib, Michael. 2011. Niche construction and the evolution of language. *Interaction Studies* 12, 162–193.

Balari, Sergio & Guillermo Lorenzo. 2010a. Specters of Marx: A review of *Adam's Tongue* by Derek Bickerton. *Biolinguistics* 4, 116–126.

Balari, Sergio & Guillermo Lorenzo. 2010b. Incidental biology: A reply to Derek Bickerton's response. *Biolinguistics* 4, 132–136.

Balari, Sergio & Guillermo Lorenzo. 2013a. Richard Owen on the mind/body problem. *Theoretical Biology Forum* 106, 131–146.

Balari, Sergio & Guillermo Lorenzo. 2013b. *Computational Phenotypes. Towards an Evolutionary Developmental Biolinguistics*. Oxford: Oxford University Press.

Balari, Sergio & Guillermo Lorenzo. In progress. The Derivative Hypothesis and the roots of Evo-Devo. Ms., Universitat Autònoma de Barcelona & Universidad de Oviedo.

Balari, Sergio & Guillermo Lorenzo. Submitted. The end of development. Ms., Universitat Autònoma de Barcelona & Universidad de Oviedo.

Benítez-Burraco, Antonio & Víctor M. Longa. 2010. Evo-Devo—Of course, but which one? Some comments on Chomsky's analogies between the biolinguistic approach and Evo-Devo. *Biolinguistics* 2, 225–232.

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

Bickerton, Derek. 2009. *Adam's Tongue: How Humans Made Language, and Language Made Humans*. New York: Hill and Wang.

Bickerton, Derek. 2010. Response to Balari & Lorenzo. *Biolinguistics* 4, 128–132.

Bickerton, Derek. 2014. *More than Nature Needs. Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.

Caporael, Linda R., James R. Griesmer & William C. (eds.). 2014. *Developing Scaffolds in Evolution, Culture, and Cognition*. Cambridge, MA: MIT Press.

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

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

Chomsky, Noam. 2010. Some simple evo devo theses: How true might they be for language. In Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds.), *The Evolution of Language. Biolinguistic Perspectives,* 45–62. Cambridge: Cambridge University Press.

Clark, Brady. 2011. Scavenging, the stag hunt, and the evolution of language. *Journal of Linguistics* 47, 447–480.

Clahsen, Harald. 1986. Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics* 24, 79–121.

Clahsen Harald & Detlef Hansen. 1997. The grammatical agreement deficit in specific language impairment: Evidence from therapy experiments. In Myrna Gopnik (ed.), *The Inheritance and Innateness of Grammars,* 141–160. Oxford: Oxford University Press.

Darwin, Charles & Alfred Wallace. 1858. On the tendency of species to form varieties; and on the perpetuation of varieties by natural means of selection. *Journal of the Proceedings of the Linnean Society of London. Zoology* 3, 45–50.

Kirby, Simon. 2013. Language, culture, and computation: An adaptive systems approach to biolinguistics. In Cedric Boeckx & Kleanthes K. Grohmann (eds.), *The Cambridge Handbook of Biolinguistics,* 460–477. Cambridge: Cambridge University Press.

Labov, William. 2010. *Principles of Linguistic Change*, vol. 3: *Cognitive and Cultural*

Factors. Malden, MA: Wiley-Blackwell.

Longa, Víctor M., Guillermo Lorenzo & Juan Uriagereka. 2011. Minimizing language evolution: The Minimalist Program and the evolutionary shaping of language. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism,* 596–616. Oxford: Oxford University Press.

Minelli, Alessandro. 2003. *The Development of Animal Form*. Cambridge: Cambridge University Press.

Moore, Celia L. 2003. Evolution, development, and the individual acquisition of traits: What we've learned since Baldwin. In Bruce H. Weber & David J. Depew (eds.), *Evolution and Learning: The Baldwin Effect Reconsidered*, 115–139. Cambridge, MA: MIT Press.

Oppenheim, Ronald W. 1981. Ontogenetic adaptations and retrogressive processes in the development of the nervous system and behaviour: A neuroembriological perspective. In Kevin J. Connolly & Heinz F.R. Prechtl (eds.), *Maturation and Development: Biological and Psychological Perspectives*, 73–109. Philadelphia, PA: Lippincott.

Oppenheim, Ronald W. 1984. Ontogenetic adaptations in neural and behavioural development: Toward a more 'ecological' developmental psychobiology. In Heinz F.R. Prechtl (ed.). *Continuity of Neural Functions from Prenatal to Postnatal Life*, 16–30. Philadelphia, PA: Lippincott.

Owen, Richard. 1849. *On the Nature of Limbs. A Discourse* [ed. by Ron Amundson]. Chicago, IL: The University of Chicago Press, 2007.

Owen, Richard. 1859. *On the Classification and Geographical Distribution of the Mammalia*. Cambridge: Cambridge University Press, 2009.

Owen, Richard. 1860. Darwin on the origin of species. *Edinburgh Review* 3, 487–532.

Owen, Richard. 1866. *On the Anatomy of Vertebrates. I. Fishes and Reptiles*. London: Longmans, Green, and Co.

Owen, Richard. 1868. *On the Anatomy of Vertebrates. III. Mammals*. London: Longmans, Green, and Co.

Oyama, Susan, Paul E. Griffiths & Russell D. Gray (eds.). 2001. *Cycles of Contingency: Developmental Systems and Evolution*. Cambridge, MA: The MIT Press.

Pylyshyn, Zenon W. 1980. Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences* 3, 111–132.

Richards, Evelleen. 1987. A question of property rights: Richard Owen's evolutionism reassessed. *British Journal for the History of Science* 20, 129–171.

Ringe, Donald, Tandy Warnow & Anne Taylor. 2002. Indo-European and computational cladistics. *Transactions of the Philological Society* 100, 59–129.

Shermer, Michael. 2002. *In Darwin's Shadow: The Life and Science of Alfred Russel Wallace — A Biographical Study on the Psychology of History*. Oxford: Oxford University Press.

Thelen Esther & Linda B. Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: The MIT Press.

Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitableness of Colonial Englishes*. Oxford: Oxford University Press.

Trudgill, Peter. 2011. *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.

Wallace, Alfred Russel. 1870. The limits of natural selection as applied to man. *Contributions to the Theory of Natural Selection* (April), 332–371.

Wallace, Alfred Russel. 1908. *My Life: A Record of Events and Opinions*, new edn. (condensed and revised). London: Chapman & Hall.

Wimsatt, William C. 1986. Developmental constraints, generative entrenchment, and the innate-acquired distinction. In William Bechtel (ed.), *Integrating Scientific Disciplines*, 185–208. Dordrecht: Martinus Nijhoff.

Sergio Balari
Universitat Autònoma de Barcelona
Departament de Filologia Catalana
Edifici B, Campus UAB
08193 Bellaterra (Barcelona)
Spain
Sergi.Balari@uab.cat

Guillermo Lorenzo
Universidad de Oviedo
Departmento de Filología Española
Campus El Milán
33011 Oviedo
Spain
glorenzo@uniovi.es

Conceptual and Methodological Problems with Comparative Work on Artificial Language Learning

Jeffrey Watumull, Marc D. Hauser & Robert C. Berwick

Several theoretical proposals for the evolution of language have sparked a renewed search for comparative data on human and non-human animal computational capacities. However, conceptual confusions still hinder the field, leading to experimental evidence that fails to test for comparable human competences. Here we focus on two conceptual and methodological challenges that affect the field generally: 1) properly characterizing the computational features of the faculty of language in the narrow sense; 2) defining and probing for human language-like computations via artificial language learning experiments in non-human animals. Our intent is to be critical in the service of clarity, in what we agree is an important approach to understanding how language evolved.

Keywords: artificial language learning; faculty of language in the narrow sense; recursion

1. Introduction

Within the past several decades, starting with the synthetic reviews of Lieberman (1984), Bickerton (1990), and Pinker & Bloom (1990), there has been increasing interest and empirical study of the evolution of language (e.g., Fitch 2012, Tallerman & Gibson 2012). Nevertheless, considerable confusion remains regarding the central theoretical issues and core concepts to be engaged, leading to empirical studies that are sometimes far off the mark.

Perhaps nowhere has this confusion been greater than in reaction to the issues raised by Hauser *et al*. (2002), and this is especially the case with respect to comparative studies of artificial language learning in animals (Fitch & Hauser 2004, Gentner *et al*. 2006, Murphy *et al.* 2008, Abe & Watanabe 2011, Rey *et al*. 2012). Here we focus on two problems that have hindered work in this area, especially its potential contribution to linguistics, cognitive science, neuroscience, and evolutionary biology.

First, despite broad interest in the mechanisms underlying the capacity for language, and especially what is unique to humans and to language, studies with non-human animals are often not appropriately designed to answer questions

ISSN 1450–3417

Biolinguistics 8: 120–129, 2014
http://www.biolinguistics.eu

about these mechanisms; running artificial language learning experiments is non-trivial (Reber 1967). In particular, several studies focus too narrowly on the problem of syntactic-like embedding as *the* defining feature of our uniquely human capacity. But this approach is flawed: Embedding is neither necessary nor sufficient for a full description of human language. Furthermore, but far more peripherally, many have incorrectly suggested that Hauser *et al*.'s (2002) thesis about the evolution of language places center-embedding as a core process in human linguistic competence. Since several comparative studies of animal computation focus on this work, it is important to get it right: Hauser and colleagues specifically suggested that what is unique to humans and unique to language (the Faculty of Language in the Narrow Sense, FLN) is *recursion and its mappings to the sensory-motor and conceptual-intentional systems*.

Second, standard methodology in this research area—massive training of captive animals with reward for 'correct' behavior—bears little resemblance to experimental child language research, or to child language acquisition (Wanner & Gleitman 1982, Ambridge & Lieven 2011); studies of children explore acquisition by means of spontaneous methods, using passive exposure or habituation–discrimination. Consequently, animal researchers cannot so easily draw conclusions about either the trajectory of human language development or its computational-representational properties.

Our central aim, therefore, is to clarify these conceptual and methodological issues, and then end with a few suggestions on how empirical work in this important area might progress.

2. Testing for Uniquely Human Mechanisms of the Language Faculty

Given the broad set of factors that enter into language, empirical research is only tractable by first defining a narrow subset of core linguistic properties. This was one motivation for Hauser *et al*. (2002) to define the language faculty in the narrow sense (FLN) as "the abstract linguistic computation system alone, independent of the other systems with which it interacts and interfaces" (p. 1571) in the language faculty defined broadly (FLB). FLN comprises "the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces with [conceptual-intentional and sensory-motor systems]" (p. 1573). The FLN/FLB distinction leaves open, of course, what factors are part of FLN and FLB, while establishing methodological criteria for future investigation.

The FLN/FLB distinction was also developed as a conceptual guide. FLN characterizes linguistic *competence* in the form of recursive (computable) functions that generate a discrete infinity of structured expressions, formally analogous to the procedure for the inductive generation of sets and so the natural numbers. The set of linguistic expressions and the set of natural numbers are thus *effectively computable* in that, though infinite, they are "calculable by finite means" (Turing 1936: 230). For example, a function—a finite representation—can be specified to generate the infinite, nonrandom decimal expansion of π. Because this expansion is infinite, it cannot be physically represented as such. However this is an independent—and arbitrary—fact of the *performance* mechanisms that implement the finite function; π does not cease to be a computable number if the physical resources required to calculate it are exhausted, or even nonexistent. It is in this same sense that FLN is a competence system—a system of recursively generated discrete infinity—logically isolable from the performance systems with which it interfaces to form FLB.

FLN qua recursive function is thus typified by three essential properties (see Watumull *et al*. 2014 for further discussion): (i) *computability*, (ii) *definition by induction*, and (iii) *mathematical induction*. Computability is reflected in a procedure—equivalent to a type of *Turing machine*, discussed below—that generates new and complex representations by combining and manipulating discrete symbols. The computable function must be defined by a sophisticated form of induction: Outputs must be carried forward and returned as inputs to generate a hierarchical structure over which can be defined complex relations (e.g., syntactic, semantic, phonological, etc.). In technical terms, the function *strongly generates* structures corresponding to *weakly generated* strings (e.g., the weakly generated string *the boy saw the man with binoculars* is one string with (at least) two syntactic structures, {{the, boy}, {saw, {the, man}}, {with, binoculars}} and {the, {boy, {saw, {the, {man, {with, binoculars}}}}}}, corresponding to (at least) two different semantic interpretations). Finally, mathematical induction is realized in the jump from finite to infinite, as in the projection from a finite set of words to an infinite set of sentences.

Given this specification of FLN, it is false to conflate recursive *functions* with center-embedded (CE) *patterns* of the form $a^n b^n$ (e.g., *the antelope* [a1] *the lion* [a2] *ate* [b2] *ran like a snail* [b1]). The most recent example of this error is by Rey *et al*. (2012) in experiments with baboons: "[T]he central claim [of Hauser *et al*. 2002 is] that the ability to process CE structures is a critical cognitive feature distinguishing human from nonhuman communication" (p. 180). Following this line of argument, 'success' by non-human animals in processing CE structures leads to the overly strong conclusion that, "[c]ontrary to the commonly accepted claim that recursion is human specific[,] CE structures produced by humans could have their origins in associative and working memory processes already present in animals" (p. 182–183).

As noted, this conclusion is problematic because it falsely equates center-embedding with recursion, and more narrowly, attributes to Hauser *et al*. (2002) the incorrect thesis that the ability to process CE patterns is what defines FLN. The correct thesis is that FLN characterizes the uniquely human character of language. To repeat, Hauser et al. proposed that FLN comprises "the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces with [conceptual-intentional and sensory-motor systems]" (p. 1573). Recursion, as noted in this hypothesis, was understood in the standard mathematical sense given above. Expressions generated by this system may be (center-)embedded or not; whether a function is recursive or not is independent of the form—or even existence—of its output. Theorems from the formal sciences (e.g., the work by Rice, Ackermann, and others) demonstrate that in general it is exceedingly difficult to infer anything truly germane as to the nature of a computational mechanism from patterns in its outputs. Consequently, testing for the ability to process (center-)embedding does *not* constitute a test of the FLN claim, contrary to what is claimed by Rey *et al*. (2012) and studies on which it builds. Here we work through the Rey *et al*. study as an illustration of these problems, but note that they arise in other work as well (e.g., Gentner *et al*. 2006, Murphy *et al*. 2008, Abe & Watanabe 2011).

In the Rey *et al*. (2012) experiments, captive baboons were conditioned to associate pairs of visual shapes a_ib_i to test whether they would order selection of those shapes in a 'center-embedded' $a_ia_jb_ja_i$ pattern. Rey *et al*. summarize their results: "[B]aboons spontaneously ordered their responses in keeping with a recursive, centre-embedding structure" (p. 180). They then conclude that "the production of CE structures in baboons and humans could be the by-product of associative mechanisms and working memory constraints" (p. 183). In other words, *neither baboons nor humans are endowed with FLN*—a surprising and unevidenced result in the case of humans. This *non sequitur* derives from the failure to distinguish associative processes from recursive computations.

Association is indeed the most parsimonious explanation of the baboon results: intensive, repetitive, conditioned associative learning that is ubiquitous in the animal kingdom, from invertebrates to vertebrates (Gallistel 1990). As Rey *et al*. observe, "the [baboon's] preference for producing CE structures requires (1) the capacity to form associations between pairs of elements (e.g., *a*1*b*1 or *a*2*b*2) and (2) the ability to segment these associations and maintain in working memory the first element of a pair (*a*1) in order to produce later its second associated element (*b*1). [T]hese two requirements are satisfied in baboons and are sufficient for producing CE structures having one-level-of-embedding" (p. 182). Two implications follow.

For Rey *et al*., the 'language' to be recognized is strictly finite, in the form $a_ia_jb_ja_i$ for $i, j = 1, \ldots, 6$ (with i, j distinct). As such, it is unnecessary to posit any embedded structure—let alone any underlying grammar—to correctly recognize this language. Furthermore, such a result runs precisely counter to the original aim of the study: Instead of showing that baboons are endowed with a capacity that parallels the characteristic unboundedness of human language, it shows that baboons display a finite, bounded processing ability. Second, if association suffices for 'one-level-of-embedding', this in turn implies that the extension of such an ability to process two levels of embedding would demand extensive additional training (i.e. listing additional associations), a result that has been amply demonstrated as fatal in connectionist networks (Berwick 1982, Elman 1991), and is fundamentally different from human language acquisition.

Another way in which Rey *et al*. err can be seen in the fact that the linguistic patterns that are 'easy' and 'difficult' for people to process do not align well with center-embedded word sequences and their possible foils—such patterns are both too strong and too weak. As noted by Rogers & Hauser (2010), while people find language patterns in the form a^nb^n difficult to process (e.g., *people*^{n} *left*^{n} (e.g., *people people people left left left*)), their corresponding paraphrased forms (*people who were left* (*by people who were left*)^n *left* (e.g., *people who were left by people who were left left*)) seem easier for people to analyze; several authors, including Rey *et al*., assume that these latter patterns are within the reach of non-human animal abilities. Notably, the processing of center-embedded structures in humans is known to be limited by working memory, a point acknowledged by Rey *et al*., as well as from the classic studies by Miller & Isard (1964). But memory by itself is not an ability or competence. As Rey *et al*. acknowledge, it is simply the work-space within which particular procedures are executed. Human performance in such cases can be extended indefinitely without any change to the internal 'program' (competence) if time and access to external memory are increased (Miller & Chomsky 1963); and far from being unfalsifiable (see Gentner *et al*. 2006 for such a claim), the independent existence of a particular linguistic competence can be demonstrated by varying performance as a function of computational complexity. In contrast, this effect has not been demonstrated in baboons, nor is it obvious how one would run the relevant tests.

Rey *et al*. conclude that "increasing the levels-of-embedding could be too demanding for baboons" (p. 182), and then speculate that "[a]lthough the present results indicate that baboons are not qualitatively limited in processing CE structures, their performance could be limited quantitatively to the processing of one or two embeddings" (p. 182). But this is misleading. Rey *et al*. provide no evidence to indicate that the qualitative limits do not simply reduce to quantitative limits, that is, that an unlimited competence underlies the baboons' limited performance. Finally, as Rogers & Hauser (2010) observe, center-embedded a^nb^n patterns correspond to the 'simplest' possible kind of embedding structure. For example, they allow for Sentences embedded within other Sentences (e.g., *John knows that the baboon learned language)*, but not Sentences embedded within Noun Phrases, as in relative clauses (e.g., *the baboon who learned language*), let alone many other constructions in human language. In short, a^nb^n patterns—the proxy for center-embedded structure—are simply not what is essential to FLN; they are not good 'human language detectors', being both too simple and too complex. This critique holds independently of the method used to demonstrate how individuals acquire such patterns, a point we explore below.

To think that human linguistic competence can be reduced to association and working memory reveals a misunderstanding of the critical difference between a look-up table—a finite list of associations—and a Turing machine—a mathematical model of computation represented by a control unit of stored instructions and an unbounded read/write memory tape enabling unbounded computation. If one takes the computational theory of the mind/brain seriously, it is the Turing machine (or one of its formal equivalents) that serves as the natural model for human cognition, including language; the look-up table is a nonstarter (see Gallistel & King 2009).

The distinction between finite and infinite memory, more specifically the independence of assumptions about working memory from those about syntactic competence, has proved fruitful for the bulk of research in human syntax during the past sixty or so years. While it is true that the human brain is finite, and so *could* be represented as a (large) finite-state machine or look-up table, this isn't relevant. The set of outputs a human can generate is in principle unlimited and, importantly, *non-arbitrary*, (i.e. the set of outputs is nonrandom, inclusion in the set being determined by the generative function). It is infinite models of these finite systems that yield scientific insight (see Turing 1954 on the generate/look-up distinction).

Consider human arithmetical competence. Here, the finite/infinite distinction seems so clearly necessary that the cognitive science literature assumes without question that this competence is somehow internalized (perhaps not transparently) in the form of some finite set of rules; it further assumes that these rules, unmodified for any particular arithmetic task, determine an infinite—and non-arbitrary—range of outputs. Here, performance may be 'truncated' by working memory, among many other factors, in recognizable ways (e.g., Hitch 1978, Dehaene 1999, Trotzke *et al*. 2013). Indeed, multiplication cannot even be carried out by a finite-state machine. What is required for multiplication is something similar to a Turing machine with a potentially unbounded input/output tape, so that intermediate results can be written to an external tape and carried forward ('recursed') to later stages of the computation. Any purely association-based method must fail at some point. Yet no one doubts that people have internalized the rules for multiplication (operating on an internal tape). Nor is there any confusion that the same holds for any physically realizable computer, like a laptop. Unsurprisingly, in all cases, the infinite model yields the proper theory for the physically realized device.

Arithmetical competence corresponds in many important respects with linguistic competence. As observed above, both arithmetic and language are systems of digital infinity, each enumerating inductively a potentially infinite and non-arbitrary set of discretely structured objects via computable functions. As Chomsky (1959) noted, the grammar for generating a set of linguistic expressions can be characterized as a function mapping the integers onto this set. As hypothesized for FLN (Watumull 2012), the discrete elements of a syntactic expression (e.g., words) are read as input and, as instructed by internalized linguistic rules (principles and parameters, etc.), combined into sets (e.g., phrases) and written onto the memory 'tape' to be carried forward as 'intermediate results', serving as inputs to subsequent computations. This enables the unbounded combination of words into phrases, and phrases into sentences, and sentences into discourses.

The generative process just described is essentially the "iterative conception of a set", with sets of discrete objects, linguistic or arithmetic, "recursively generated at each stage" such that "the way sets are inductively generated" is formally equivalent to "the way the natural numbers […] are inductively generated" (Boolos 1971: 223). Thus both language and arithmetic draw on similar generative procedures, a point reiterated in Hauser *et al*. (2002). Though non-human animals appear to be able to carry out some arithmetical operations using analog quantity representations, or perhaps subitizing for small integers, there seems to be no evidence for anything resembling the computable rule systems sketched above or the inductive generalization to an unbounded domain of structured arithmetical expressions. Even when animals are taught the Arabic integers through reinforcement, they never acquire anything remotely like the successor function, generalizing beyond the trained input (Kawai & Matsuzawa 2000). Moreover, and of direct relevance to the methodology of most animal studies in this area including the artificial language studies discussed here, the research on animal integer processing also demonstrates that this capacity is entirely different from children's development of arithmetical competence: Animals never exhibit the kind of inductive leap (the best evidence for discrete infinity) that all children take once they have acquired knowledge about the first few integers (Leslie *et al*. 2008, Carey 2011). What is required is some way to carry forward arbitrarily large, inductively generated intermediate results, say by means of an arbitrarily long input/output tape (mentally represented), as in the multiplication and syntactic examples described earlier.

3. Methodology for Experiments with Non-Human Animals

Understanding how behavior is acquired is essential to comparative inquiry. It is particularly important in work on artificial language learning because children do not acquire language by means of massive, long-term training. Further, a hallmark of virtually all aspects of language acquisition is the inductive aspect of recursion: Once a particular component of linguistic competence develops, it rapidly generalizes to a virtually limitless range of possible expressions. In the case of most work on artificial language learning, whether on birds, rodents, or primates, the method entails massive training with little evidence of anything remotely resembling unbounded generalizations. The animals seem merely to be compiling a list—a look-up table—rather than internalizing rules. Thus, even if one were to grant that animals exhibit certain linguistic-like behaviors, their mode of acquisition is nothing like that evidenced by human children, and whatever has been acquired appears extremely bounded in its expressive power.

A counter-example to this approach is the original study of finite-state and phrase-structure grammars by Fitch & Hauser (2004) with cotton-top tamarins, and pursued in a slightly different way by Abe & Watanabe (2011) in Bengalese finches. Here, the method paralleled those used by researchers working on artificial language learning in human infants, and in particular, a familiarization-discrimination technique. In brief, this technique exposes subjects in a passive listening context to the relevant input, and then follows with presentations of exemplars that match the input as well as exemplars that are different in some fundamental way. If subjects have picked up on the pattern inherent in the familiarization phase, they should respond more intensely to the exemplars in the discrimination phase that are different than to those that are the same.

Though this technique captures the spontaneity of processing that is characteristic of language processing, it suffers from at least two problems. First, unlike the training techniques that involve highly objective and robust behavioral measures (e.g., touching a button), the familiarization-discrimination techniques involve a more subjective and ambiguous response: looking time or looking orientation. Despite methods designed to provide relatively high inter-observer reliabilities, these remain relatively fragile techniques, due in part to the often small differences in response measures across conditions (often a matter of a couple of seconds). Second, and more importantly, in studies of non-human animals, where the test population is extremely limited and small, it is necessary to run different conditions with the same population. This is not the case in studies of human infants where different conditions are tested on different populations. Given the limited test population, animals often habituate to the general test environment, and further, are exposed to many different conditions, thereby changing their experience over multiple conditions.

We are thus left with a spontaneous method that cannot deliver the requisite information about processing capacities that are like child language acquisition, or a training method that can potentially identify an ability, but one that may well be fundamentally different from what is in play for human children during acquisition. In other words, even if a training study shows that an animal can 'compute' center-embedded patterns, the underlying representations are likely to be entirely different because of the procedures used to demonstrate this capacity. In any event, such methods have, thus far, failed to demonstrate the unboundedness that is required of human linguistic computation.

4. Conclusion

What results with non-human animals might challenge the claim that the language faculty is uniquely human? And more narrowly, what evidence might refute the hypothesis proposed by Hauser *et al*. (2002) regarding the composition of FLN? With respect to the generative component of their thesis, and in particular, its focus on recursive mechanisms, it would be necessary to show that animals *spontaneously* respond to stimuli that are characterized by (i) *computability*, (ii) *definition by induction*, and (iii) *mathematical induction*—the three properties typical of linguistic recursion that we briefly noted above. Computability requires proof of a procedure that generates new and complex representations by combining and manipulating symbols, as in human language; this *productive* process is to be contrasted with the *retrieval* of representations from a look-up table (finite and innately specified or memorized), as in non-human primate calls. The computable function must be defined by a sophisticated form of induction: Outputs must be carried forward and returned as inputs to *strongly generate* a hierarchical structure over which can be defined complex relations (e.g., syntactic, semantic, phonological, etc.); this also implies the discreteness of representations. Lastly, mathematical induction is seen in the jump from finite to infinite. This can be demonstrated by significant generalization beyond the exposure material (e.g., counting indefinitely beyond the training set) and by revealing an unbounded competence underlying bounded performance.

In conclusion, to advance this important field, greater conceptual and methodological clarity is necessary (for recent discussions, see Fitch & Friederici 2013, Zuidema 2013). Conceptually, it is necessary to understand the formal aspects of recursive functions in order to capture the fundamental generative and unbounded properties of all natural languages (where embedding is an interesting but incidental phenomenon). Experiments should focus on all aspects of the Turing-like architecture of the faculty of language in its narrow sense: aspects of the enumeration by finitary procedures and read/write memory of a non-arbitrary digital infinity of hierarchically structured expressions and relations. Devising such tests may prove difficult, but this is the critical challenge for a theoretically rigorous and empirically grounded approach to the evolution of language.

References

Abe, Kentaro & Dai Watanabe. 2011. Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience* 14, 1067–1074.

Ambridge, Ben & Elena V. M. Lieven. 2011. *Child Language Acquisition: Contrasting Theoretical Approaches*. Cambridge: Cambridge University Press.

Berwick, Robert C. 1982. Locality principles and the acquisition of syntactic knowledge. Cambridge, MA: MIT dissertation.

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

Boolos, George. 1971. The iterative conception of set. *The Journal of Philosophy* 68, 215–231.

Carey, Susan. 2011. *The Origin of Concepts*. Oxford: Oxford University Press.

Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and Control* 2, 137–167.

Dehaene, Stanislas. 1999. *The Number Sense: How the Mind Creates Mathematics*. Oxford: Oxford University Press.

Elman, Jeffrey L. 1991. Distributed representation, simple recurrent networks, and grammatical structure. *Machine Learning* 7, 195–225.

Fitch, W. Tecumseh & Angela D. Friederici. 2012. Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B*, 1933–1955.

Fitch, W. Tecumseh & Marc D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303, 377–380.

Gallistel, C. R. 1990. *The Organization of Learning*. Cambridge, MA: MIT Press.

Gallistel, C. R. & Adam Philip King. 2009. *Memory and the Computational Brain: Why Cognitive Science Will Revolutionize Neuroscience*. New York: Wiley-Blackwell.

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

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

Hitch, Graham J. 1978. The role of short-term working memory in mental arithmetic. *Cognitive Psychology* 10, 302–323.

Kawai, Nobuyuki & Tetsuro Matsuzawa. 2000. Numerical memory span in a chimpanzee. *Nature* 403, 39–40.

Leslie, Alan M., Rochel Gelman & C. R. Gallistel. 2008. The generative basis of natural number concepts. *Trends in Cognitive Science* 12, 213–218.

Miller, George A. & Noam Chomsky. 1963. Finitary models of language users. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology II*, 419–491. New York: Wiley and Sons.

Miller, George A. & Stephen Isard. 1964. Free recall of self-embedded English sentences. *Information and Control* 7, 292–303.

Murphy, Robin A., Esther Mondragón & Victoria A. Murphy. Rule learning by rats. *Science* 319, 1849–1851.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13, 707–784.

Reber, Arthur S. 1967. Implicit learning of artificial grammars. *Verbal Learning and Verbal Behavior* 5, 855–863.

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

Rogers, James & Marc D. Hauser. 2010. The use of formal language theory in studies of artificial language learning: A proposal for distinguishing the differences between human and nonhuman animal learners. In Harry van der Hulst (ed.), *Recursion and Human Language*, 213–232. Berlin: De Gruyter Mouton.

Tallerman, Maggie & Kathleen R. K. Gibson (eds.). 2012. *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press.

Trotzke, Andreas, Markus Bader & Lyn Frazier. 2013. Third factors and the performance interface in language design. *Biolinguistics* 7, 1–34.

Turing, Alan M. 1936. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42, 230–265.

Turing, Alan M. 1954. Solvable and unsolvable problems. *Science News* 31, 7–23.

Wanner, Eric & Lila R. Gleitman (eds.). 1982. *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press.

Watumull, Jeffrey. 2012. A Turing program for linguistic theory. *Biolinguistics* 6, 222–245.

Watumull, Jeffrey, Marc D. Hauser, Ian G. Roberts & Norbert Hornstein. 2014. On recursion. *Frontiers in Psychology* 4, 1–7.

Zuidema, Willem. 2013. Context-freeness revisited. *Proceedings of CogSci 2013*, http://mindmodeling.org/cogsci2013/papers/0309/paper0309.pdf.

Jeffrey Watumull
University of Cambridge
Theoretical & Applied Linguistics
Cambridge CB3 9DA
UK
jw647@cam.ac.uk

Marc D. Hauser
Risk-Eraser
PO Box 376
West Falmouth, MA 02574
USA
mdh102559@gmail.com

Robert C. Berwick
Massachusetts Institute of Technology
Electrical Engineering & Computer Science
Cambridge, MA 02139
USA
berwick@csail.mit.edu

★ FORUM ★

Self-Organization and Natural Selection: The Intelligent Auntie's Vade-Mecum

Víctor M. Longa & Guillermo Lorenzo

1. Introduction

This paper is aimed at clarifying one particular aspect of Derek Bickerton's recent contribution to *Biolinguistics* (Bickerton 2014a), where he contends that biolinguists tend to emphasize the specifics of certain non-standard evolutionary models in order to prejudicially avoid the theory of natural selection. According to Bickerton (2014a: 78), "they [biolinguists] have problems with the notion of natural selection, up to and including a total failure to comprehend what is and how it works". This is the most understandable, also according to Bickerton, because even evolutionary psychologists and philosophers like Pinker and Dennett, who have devoted well-known papers and books to explaining and applying natural selection to the case of cognition and language, have failed to understand the real import of Darwin's idea: "Natural selection could not 'explain' complex design", claims Bickerton (2014a: 79), "even if Pinker & Bloom (1990), Dennett (1995), and others who are not biologists think it does. In fact, natural selection does not provide a single one of the factors that go into creating design".

Bickerton's comments in the *Biolinguistics* piece are specifically targeted at the model of 'self-organization' associated to complexity sciences, which is introduced in Longa (2001) as potentially capable of dealing with some recalcitrant problems of the evolution of language. Bickerton (2014a: 79) writes that Longa's attacks point to "a straw man", and that his claim that self-organization is an alternative to natural selection is "a category mistake", for self-organization is simply one of the factors that generates the variation that natural selection selects from. So, according to Bickerton, natural selection and self-organization must be conceptualized as two complementary mechanisms that operate in a coordinated manner to bring about complex biological designs.

In this response we want to explain that this is a wrong conclusion supported on wrong premises. For that purpose, we first document that biologists generally agree on the idea that natural selection creates design; second,

This work has been funded by the Ministerio de Ciencia e Innovación of the Spanish Government (Biolingüística: evolución, desarrollo y fósiles del lenguaje; FFI2010-14955). We wish to thank Sergio Balari and an anonymous reviewer for their comments to early drafts of this note.

ISSN 1450–3417
Biolinguistics 8: 130–140, 2014
http://www.biolinguistics.eu

we explain that self-organization is primarily concerned in the stability and robustness of form rather than in introducing variation; and, finally, we systematize the differences between self-organization and selection as explanatory paradigms. Considering all these pieces of evidence altogether, we conclude that they cannot be conceptualized as coordinated evolutionary strategies, save at the price of making one or another devoid of its original meaning. As a matter of fact, we think that this is exactly the position of Bickerton, to whom 'natural selection' boils down to the idea of the 'survival of the fittest'. But if so, we agree with Fodor & Piattelli-Palmarini (2010: 139ff.) that natural selection becomes a platitude: "There survive those that survive".[1] At the end of this response we offer a good illustration of the possibility of respecting the difference between natural selection and self-organization for explanatory purposes, curiously enough taken from Bickerton (2014b).

2. Natural Selection: The Biologists' View

According to Bickerton's presentation, natural selection does not create design: Natural selection simply selects among designs independently created by other means—self-organization being just one (Bickerton 2014a: 79). Consequently, we ought not to present natural selection and, for example, self-organization as rival evolutionary mechanisms, because they are complementary pieces or a single multifaceted process. This is not what the relevant specialized literature shows (see Table 1), for there one can easily find what is characterized as the 'creative view' of natural selection (Razeto-Barry & Frick 2011, Razeto-Barry 2013). According to Razeto-Barry and Frick's presentation, natural selection "is a creative force because it can generate new traits by the cumulative selection that makes probable a combination of mutations which are necessary for trait development and that would not probably be combined together without natural selection" (Razeto-Barry & Frick 2011: 344). As a matter of fact, such a characterization is an unavoidable one if we take the neo-Darwinist dissection of the evolutionary process at face value. As Gould (2002: 141-146) explains, the variation on which natural selection acts is small, copious and isotropic (i.e. insensitive to direction). Consequently, "variation only serves as a prerequisite, a source of raw material incapable of imparting direction or generating evolutionary change by itself" (Gould 2002: 155). In other words, such raw material is only creatively cooked by selection.[2]

---

1 According to an anonymous reviewer, this phrase is an epitome of the creationists' creed, a fact we were not aware of when we originally wrote it. Be as it may, we consider this observation irrelevant to the point. Some biologists have previously defended the thesis that natural selection entails a tautology (for example, Vallejo 1998), without aiming it at supporting creationism. For that matter, one might also correctly say that Darwin's concept of adaptation was continuous with that of theologians. Obviously enough, from this fact one cannot derive an argument supporting Darwin's intimate beliefs. The same reviewer notes that Bickerton nowhere mentions 'survival' or 'fitness' in his paper. This is correct, but if it means that Bickerton sees these concepts alien to the theory of natural selection, readers may wonder what natural selection actually boils down to for him.

2 A reviewer suggests that supporters of the creative view clearly reify and anthropomorphize evolution, envisioning it as an agent with abilities proper of intentional minds. This

According to Gould, it was one of Darwin's key postulates "the claim that natural selection acts as the creative force of evolutionary change" (Gould 2002: 583). We think that Gould (and Darwin, for that matter) is not suspect of being one of those non-biologists that failed to comprehend what natural selection is and how it works.

"Selection molds the separate units of heredity into a coordinated whole, a process as truly creative (although of course not planned or directed) as the combination of separate bricks into a building."	(Simpson *et al*. 1957: 413)
"All evolution is due to the accumulation of small genetic changes, guided by natural selection."	(Mayr 1963: 586)
"Natural selection is at one and the same time a blind and a creative process."	(Dobzhansky 1973: 126)
"Darwin's theory of evolution by natural selection is satisfying because it shows us a way in which simplicity could change into complexity, how unordered atoms could group themselves into ever more complex patterns."	(Dawkins 2006 [1976]: 12)
"We start from the presumption that natural selection is the only plausible explanation for adaptive design."	(Maynard-Smith & Szathmáry 1995: 290)
"Selection thus acts as a creative force that has made possible biological organizations that would otherwise have been highly improbable."	(Strickberger 2000: 136)
"It is the cumulative selection ('adding up,' in Darwin's terms) of variation that forges complexity and diversity."	(Carroll 2006: 32)
"As a consequence of natural selection, organisms exhibit design, that is, exhibit adaptive organs and functions."	(Ayala 2007: 8570)
"Complexity cannot evolve except by natural selection."	(Futuyma 2009: 296)

Table 1. Natural selection: The creative view

3. Self-Organization: The Complexity Scientists' View

Complexity Sciences aim at discovering laws of form capable of offering models for patterns of order and regularities found in nature. The laws of concern are alien to external pressures (as for example, to adaptive pressures), but obey intrinsic generative principles that induce organization on matter in a self-sufficient way. They are typical of dynamic complex systems, composed of an intricate net of interacting elements, capable of abruptly and spontaneously reaching ordered patterns of organization. From a logical point of view, such systems could attain many different positions within a space of possibilities, but they place themselves in a well-defined area ('at the edge of chaos'), where self-organization arises.

argument has been recently elaborated and directed against contemporary Darwinists in Fodor & Piattelli-Palmarini (2010) and Richards (2012) explains that it is correct at least of Darwin's original formulations of the idea of natural selection. From the Darwinist side, philosophers however argue that the very properties emphasized by Gould guarantee that it is a 'stupid' process (Dennett 1995), incapable of planning, looking ahead, and other intelligent qualities, notwithstanding being creative.

This short presentation may be enough to appreciate that self-organization is not really the best of the allies of variation, for one of its main properties is anisotropy (i.e. directionally biased). Besides, self-organization produces steady and robust patterns of combined elements, not in competition with slightly different concurrent patterns. Obviously enough, forms thus generated may not be particularly fitted to survive in a given environment, and so they may be rejected and disappear from it. But this is not natural selection positively acting on raw materials, but a negative filter disposing of independently cooked ones.

Readers can contrast the accuracy of this short presentation, enough for our purpose, in the following sources: Lewin (1992), Waldrop (1992), Goodwin (1994), Kauffman (1995), Stewart (1999), Solé & Goodwin (2000), Camazine et al. (2001), Johnson (2001), Longa (2001), Gribbin (2004), Edelmann & Denton (2007), and Heylighen (2008).

4. Self-Organization and Natural Selection: A Short Summary of Differences

The following is a list of a total of eight differences between natural selection and self-organization that reflects that they cannot be conflated into a unique mechanism: They are complete explanatory frameworks on their own, each incompatible with the other in particular applications (see Table 2). Of course, they may be thought of as particularly fitted to different aspects of organic designs (see section 5), but successfully applying one of them in a particular occasion automatically renders the other inadequate for the same goal.

Natural selection (NS)	**Self-organization (SO)**
NS is gradual.	SO is abrupt.
NS is positively creative.	Selection after SO is negatively rejective.
Order is accidentally induced by tinkering.	Order is induced by intrinsic inertias.
NS is an externally guided process.	SO is an internally guided process.
NS acts on passive matter.	SO happens in active matter.
NS's outcomes are open.	SO's outcomes are fixed.
NS is historically contingent.	SO is generatively necessary.
NS is gene centered.	SO is epigenetic.

Table 2. Natural selection and self-organization: A case of incompatibility of character.

The table is eloquent enough and justifies Edelmann & Denton's conclusion:

> Self-organization is […] totally different in essence from cumulative selection as a causal agent of bio-complexity. If self-organization is in fact widely exploited by organisms to generate adaptive complexity […] then this does indeed provide a serious challenge to the Darwinian claim that cumulative selection is the major creative agency in evolution.
>
> (Edelmann & Denton 2007: 598)

Let us to dwell a little more on the list.

4.1. *Gradual vs. Abrupt Character*

According to Gould, gradualism "may represent the most central conviction residing both within and behind all Darwin's thought" (Gould 2002: 148), for otherwise variation itself, and not selection, should be deemed the true agent of evolutionary change. From the point of view of Darwinism, "creativity must reside in the summation of [the tiny increment of each step]," with natural selection acting "as the agent of accumulation" (Gould 2002: 150). Dawkins also points out that denying gradualism entails "to deny the very heart of [Darwinian] evolution theory" (Dawkins 1986: 318).

Self-organization operates on a radically different basis, which Edelmann & Denton explain in the following way:

> The realm of self-organized complexity is an unpredictable realm of sudden spontaneous emergent complexity that is generated by non-linear interactions via something like a phase transition. This is a realm where saltation, emergence, spontaneous sudden change and bifurcations rule; a realm in which the concepts of intermediacy, gradualism and continuity, so central to the Darwinian, no longer apply.
> (Edelmann & Denton 2007: 585)

4.2. *Positive vs. Negative Selection*

Natural selection acts positively inducing order and consistency upon a material that would otherwise diversify to the point of making populations amorphous collections of mutually unrecognizable individuals. On the contrary, self-organization guarantees similar outcomes without the need of selecting among competing designs. This does not entail that a parallel guarantee exists that self-organized designs are automatically sanctioned to overcome the perils of every imaginable environment. However, self-organized structures are simply 'selected', not 'naturally selected', for they are subject to a negative or filtering process of rejection, different from the source that independently creates them. The idea can be traced back, for example, to the works of Richard Owen, who favored an idea of 'natural rejection' along these lines as an alternative to Darwin's natural selection (Owen 1860). The following fragments offer more recent formulations of the idea:

> [Selection] does not have a lot to do except act as a coarse filter that rejects the utter failures. (Goodwin 1994: 157)

> Self-organized material patterns may be selected by, but not created by natural selection. (Edelmann & Denton 2007: 598)

4.3. *Tinkering vs. Generative Inertias*

The expression 'tinkering' is customarily used to express the opportunistic character of natural selection, which manages to take advantage of any haphazardly occurring variant within the range of a population. Tinkering is thus the resource that natural selection has at hand to accidentally impose order where otherwise "there would be nothing but incoherent disorder"

(Kauffman 1995: 8). Instead, self-organization derives order from inner generative laws, internal to the organism, so from this perspective, "vast veins of spontaneous [not accidental] order lie at hand" (Kauffman 1995: 8; the insert is also from Kauffman, same page).

4.4. *External vs. Internal Guidance*

Within the framework of natural selection, the isotropic character of variation determines that the directionality of change is a function of the external pressures acting on individuals. In other words, the evolution of populations follows the path of the proliferative superiority of their fittest representatives (Strickberger 2000, Gould 2002, Ayala 2007, Futuyma 2009). According to Futuyma, natural selection boils down to "any consistent difference in fitness among phenotypically different classes of biological entities" (Futuyma 2009: 283). Contrary to this, self-organization is an internalist framework, where complexity comes for free and it is attained by means of internal dynamics alone. This contrast is well captured in the following quote from Gould:

> "Under internalist theories of evolution, environment, at most, holds power to derail the process not behaving properly […]. Under Darwinian functionalism, however, environment becomes an active partner in both the modes and directions of evolutionary change." (Gould 2002: 161)

Or in the words of Edelmann & Denton (2007: 588): "Self-organized order is spontaneous from within; the order of selection is additive from without".

4.5. *Passive vs. Active Character of Matter*

This difference follows from previous ones:

> From the externalist viewpoint, living matter is a passive and a non-intrinsically ordered entity that needs an external factor (natural selection) to acquire from. From the internalist perspective, living matter is an active entity capable of exhibiting order spontaneously.
> (Linde Medina 2010: 25)

4.6. *Open vs. Fixed Outcomes*

The isotropic character of variation and the instability of environmental conditions determine that natural selection can lead to any result within a given space of design (Dennet 1995). As illustrated by Goodwin (1994: 87): "Small variations are such that almost anything can happen—organisms can take any form, have any color, and eat any food, subject only to very broad constraints". In the case of self-organization, systems robustly point to a specific point within a space of logical possibilities, one in which the attractor captures and stabilizes it.

4.7. *Historical Contingency vs. Generative Necessity*

Natural selection connects biology with history, while self-organization connects it with physics or chemistry. The point of view of the former is that of "historical narratives: Which species come from which ancestors under which circumstances" (Goodwin 1994: 88). The latter aims at explaining biological phenomena like other sciences "in which principles of organization allows one to understand the […] world in terms of regularities and general principles" (*ibid.*).

4.8. Genocentrism vs. Epigeneticism

Natural selection is commonly associated to the idea that evolutionary change is ultimately and chiefly anchored on genes (but see Okasha 2006). Thus, Dobzhansky's (1937: 11) classic definition of evolution as "a change in the genetic composition of populations" still pervades Darwinian thought (to wit, see Futuyma 2009). Self-organization limits the centrality of genes, for self-generated patterns of organization cannot be said to be a matter of genetic pre-specification. Accordingly, "self-organized order is indeed genuinely epigenetic and not necessarily in the genes at all" (Edelmann & Denton 2007: 587).

5. From Adam to Wallace: An Illustration of the Difference

We want to conclude this clarification note with a particularly nice example taken from the field of evolutionary linguistics, where the suggestion is made that the evolution of language was originally bootstrapped thanks to a process of a selective character, but lastly accomplished through self-organization at the brain level. The case has been raised in two successive books by Bickerton (2009, 2014b), which offer a perfect illustration of what we have been trying to explain and document above.

According to Bickerton, "the transition from the alingual state that characterizes all other species to something that might qualify as a genuine precursor of language" (Bickerton 2010: 128) could only have happened as an adaptive response to some particular need of some hominid species. Consequently, Bickerton elaborates an historical narrative that reads approximately like this: There was a time when some human ancestor entered a confrontational scavenging niche, where announcing one's sightings and asking for help were imperative. Then some individuals accidentally developed the capacity of producing some noises while the image of their sightings still reverberated in the head. The capacity was inherited and accidentally more and more elaborated and successively inherited by the progeny of those individuals, until it became species typical. In due time, human ancestors were endowed with an inborn full-fledged capacity for displaced communication by means of a protolanguage (Bickerton 2009, 2014b: Chap. 4).

Bickerton's proposal has been subject to strong criticism for different reasons (Balari & Lorenzo 2010a, Balari & Lorenzo 2010b, Arbib 2011, Clark 2011), but this is not what is at issue here. What we want to emphasize is the value of Bickerton's idea as a 'textbook case' of the application of natural selection to a particular aspect of the evolution of human mind: It presents the earliest stages of language evolution as due to a process of "long, slow gestation" (Bickerton 2009: 212), that succeeded because it worked as an "evolutionary adaptation, just as much as walking upright, shedding body hair, or getting and opposable thumb" (p. 103)—in this particular case for "recruitment, that turns out to be the key word in the birth of language" (p. 132). Even if they do not occur in a vacuum, but in a niche, adaptations are "genetic changes [that] can improve the ability of organisms to survive, reproduce, and, in animals, raise offspring" (p. 110). In dealing with processes like this one, "there's no recourse […] but to tell just-so stories" (p. 218).

Bickerton explains the evolution of language proper from protolinguistic communication very differently, for the former shows properties that defy any clear adaptive motivation—Bickerton basically assumes Chomsky's thesis of the underspecification of language for communicative uses and of the never ending array of communicative and non-communicative uses of language. He also explains that the most defying features of language regarding its concatenative properties do not require a long story of evolutionary development by small increments. So he opts in this case for a solution inspired by the alternative model of self-organization: "The tasks that were required [for protolanguage to become true language] lie well within the brain's powers of self-organization" (Bickerton 2014b: 117; insert also from Bickerton, same page), and they were executed without "any kind of external pressure" (p. 119).

Bickerton's complete account of the process is this:

> Protolanguage emerged because of triggering external events: confrontational scavenging led to the need for recruitment, which in turn necessitated displaced communication, which eventually sufficed, in social animals with large brains, to create a crude and structureless protolanguage—all that nature needed. However, these processes necessarily caused symbolic items to be stored in the brain, and […] brain-internal processes […] were directly initiated by the brain's need to deal with such items. (Bickerton 2014b: 115)

It is Bickerton himself who emphasizes that "syntactic infrastructure resulted from self-organizing activity within the brain itself" (Bickerton 2014b: 106) and that "such changes do not need to be triggered by natural selection" (p. 107).

Bickerton's goal in his last book is to explain how it is possible that languages seem universally to be so far away, from a formal point of view, of any imaginable human particular need, an aspect of what he refers to as 'Wallace's Problem'. His suggested solution is a multi-staged model of language evolution: One of these stages resulted from "particular selective pressures operating specifically on human ancestors," which were capable of releasing these people from the strictures of animal communication; another stage "consisted of purely brain-internal operations responding to unusual phenomena" that previous evolution had originated, but now with "no relation to the ecological needs of humans" (Bickerton 2014b: 262) and opened to them the never ending possibilities of language recursion. Let us conclude this note by simply noting how scrupulously Bickerton respects in this project the distinction that he simultaneously questions in the *Biolinguistics* piece (Bickerton 2014a).

6. Conclusion

The aim of this paper has not been to defend any personal stance regarding the advantages of auto-organization over selection in explaining the evolution of the language faculty, but to correct Bickerton's (2014a) misconception of the former, as if it were an evolutionary mechanism at the service of the latter. According to Bickerton, auto-organization auto-organizes variation, that selection further selects—our phrasing, of course. Here we have tried to show that fortunately enough, this has little to do with the status that current biological theories attribute to the said mechanisms. They rather conceptualize them as alternative mechanisms, a consequence of which is that the door is open to apply them separately to different aspects of a particular organism or organic system, as Bickerton actually does in his latest book, *More than Nature Needs*.

References

Arbib, Michael. 2011. Niche construction and the evolution of language. *Interaction Studies* 12, 162–193.

Ayala, Francisco. 2007. Darwin's greatest discovery: Design without designer. *Proceedings of the National Academy of Sciences USA* 104, 8567–8573.

Balari, Sergio & Guillermo Lorenzo. 2010a. Specters of Marx: A review of *Adam's Tongue* by Derek Bickerton. *Biolinguistics* 4, 116–126.

Balari, Sergio & Guillermo Lorenzo. 2010b. Incidental biology: A reply to Derek Bickerton's response. *Biolinguistics* 4, 132–136.

Bickerton, Derek. 2009. *Adam's Tongue: How Humans Made Language, How Language Made Humans*. New York: Hill & Wand.

Bickerton, Derek. 2010. Response to Balari & Lorenzo. *Biolinguistics* 4, 128–132.

Bickerton, Derek. 2014a. Some problems for biolinguistics. *Biolinguistics* 8, 73–96.

Bickerton, Derek. 2014b. *More than Nature Needs: Language, Mind, and Evolution*. Cambridge, MA: Harvard University Press.

Camazine, Scott, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz & Eric Bonabeau. 2001. *Self-Organization in Biological Systems*. Princeton, NJ: Princeton University Press.

Carroll, Sean B. 2006. *The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution*. New York: W.W. Norton & Co.

Clark, Brady. 2011. Scavenging, the stag hunt, and the evolution of language. *Journal of Linguistics* 47, 447–480.

Dawkins, Richard. 2006 [1976]. *The Selfish Gene*, 30th anniversary edn. Oxford: Oxford University Press.

Dawkins, Richard. 1986. *The Blind Watchmaker*. London: Longman.

Dennett, Daniel C. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.

Dobzhansky, Theodosius. 1937. *Genetics and the Origins of Species*. New York: Columbia University Press.

Dobzhansky, Theodosius. 1973. Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher* 35, 125–129.

Edelmann, Jonathan B. & Mike J. Denton. 2007. The uniqueness of biological self-organization: Challenging the Darwinian paradigm. *Biology & Philosophy* 22, 579–601.

Fodor, Jerry A. & Massimo Piattelli-Palmarini. 2010. *What Darwin Got Wrong*. New York: Farrar, Strauss and Giroux.

Futuyma, Douglas. 2009. *Evolution*, 2nd edn. Sunderland, MA: Sinauer Associates.

Goodwin, Brian. 1994. *How the Leopard Change Its Spots. The Evolution of Complexity*. New York: Charles Scribner's Sons.

Gould, Stephen J. 2002. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press.

Gribbin, John. 2004. *Deep Simplicity: Chaos, Complexity and the Emergence of Life*. London: Allen Lane.

Heylighen, Francis. 2008. Complexity and self-organization. In Marcia J. Bates & M.M. Maack (eds.), *Encyclopedia of Library and Information Sciences.* vol. 2, 1215–1224. Oxford: Taylor & Francis.

Johnson, Steve B. 2001. *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. New York: Scribner.

Kauffman, Stuart A. 1995. *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. New York: Oxford University Press.

Lewin, Roger. 1992. *Complexity: Life at the Edge of Chaos*. New York: MacMillan.

Linde Medina, Marta. 2010. Natural selection and self-organization: A deep dichotomy in the study of organic form. *Ludus Vitalis* XVIII/34, 25–56.

Longa, Víctor M. 2001. Sciences of complexity and language origins: An alternative to natural selection. *Journal of Literary Semantics* 30, 1–17.

Maynard-Smith, John & Eörs Szathmáry. 1995. *The Major Transitions in Evolution*. Oxford: W.H. Freeman.

Mayr, Ernst. 1963. *Animal Species and Evolution*. Cambridge, MA: Harvard University Press.

Okasha, Samir. 2006. *Evolution and the Levels of Selection*. Cambridge, MA: Harvard University Press.

Owen, Richard. 1860. Darwin on the origin of species. *Edinburgh Review* 3, 487–532.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences*, 707–727.

Razeto-Barry, Pablo. 2013. Complexity, adaptive complexity and the Creative View of natural selection. *Studies in History and Philosophy of Biological and Biomedical Science* 44, 312–315.

Razeto-Barry, Pablo & Ramiro Frick. 2011. Probabilistic causation and the explanatory role of natural selection. *Studies in History and Philosophy of Biological and Biomedical Science* 42, 344–355.

Richards, Robert J. 2012. Darwin's principle of divergence and natural selection: Why Fodor was almost right. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43, 256–268.

Simpson, George G., Colin S. Pittendrigh & Lewis H. Tiffany. 1957. *Life: An Introduction to Biology*. London: Routledge & Kegan Paul.

Solé, Ricard V. & Brian Goodwin. 2000. *The Sign of Life: How Complexity Per-*

vades Biology. New York: Basic Books.

Stewart, Ian. 1999. *Life's Other Secret: The New Mathematics of the Living Word*. New York: John Wiley & Sons.

Strickberger, Monroe W. 2000. *Evolution*, 3rd edn. London: Jones and Bartlett.

Vallejo, Fernando. 1998. *La tautología darwinista y otros ensayos de biología*. México: Universidad Nacional Autónoma de México.

Waldrop, Mitchell. 1992. *Complexity: The Emerging Science of the Living World*. New York: John Wiley & Sons.

Víctor M. Longa
Universidade de Santiago de Compostela
Departamento de Literatura Española,
Teoría da Literatura e Lingüística Xeral
Praza Isabel a Católica 2, 2ºE
36204 Vigo (Pontevedra)
Spain
victormanuel.longa@usc.es

Guillermo Lorenzo
Universidad de Oviedo
Departmento de Filología Española
Área de Lingüística General
Campus El Milán s/n
33011 Oviedo
Spain
glorenzo@uniovi.es

★ REVIEWS ★

At the Interface of (Bio)linguistics, Language Processing, and Neuropsychology

Sanz, Montserrat, Itziar Laka & Michael K. Tanenhaus (eds.). 2013. *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures*. Oxford: Oxford University Press.

by Diego Gabriel Krivochen

This book, part of the Oxford Studies in Biolinguistic series, presents a state-of-the-art overview of the field, more specifically, on psycho- and neurolinguistics and their relation to models of syntax, semantics, and morpho-phonology, while advancing its limits with cutting-edge research. A distinctive feature of the piece is the strong presence of interdisciplinary work and the internal coherence of the volume, integrating computational science, cognitive science, neurology and psycholinguistics, as well as syntax, semantics, and morpho-phonology; an integration that is most welcomed as it triggers debate and productive revisiting of the machinery assumed within all aforementioned sub-disciplines of linguistics. The volume is organized around the notion of garden path sentences, relative clauses, and their relations at the processing level; this includes major problems of natural language processing and the relations between syntax, semantics, and morpho-phonology from a more general point of view as well.

The editors have chosen to open the book with a reprinted article by Thomas Bever, from 1970 (which becomes a recurrent *motif* to which the contributors refer once and again as a departing point, thus giving structural and thematic unity and coherence to the book as a whole), a *locus classicus* for the psycholinguistic and neurocognitive approaches to ambiguity resolution, parsing (sentence perception, at the moment) strategies, and so-called 'garden path sentences' (GPS), the best known example being *The horse raced past the barn fell*, even if, as Tanenhaus claims in the Afterword, none of those is the prime theme of the work (but it is mostly about the relation between language and general cognitive strategies, an early plea for holism). The opening seems appropriate, since it provides the reader with an overall perspective on the studies of *language* as a concept analogous to those of "species or organ, as they are used in biological science" (p. 2). The article makes a case of distinguishing *language* as a mental/biological entity from *language* as a behavior; but, crucially, language structure and development are not to be isolated from the development of other cognitive capacities. Choosing this particular article is a statement in itself: Perceptual mechanisms, cognitive structures (including counting and number approximation, visual patterns and 2-D/3-D illusions), and linguistic structures (grammatical role assignment, abstraction of a structural pattern like 'active' or

ISSN 1450–3417

Biolinguistics 8: 142–162, 2014
http://www.biolinguistics.eu

morphological/syntactic 'passive'), trying to abstract common cognitive routines (taking the term from computational science) and statistically valid parsing strategies (where one of the most important features of the article rely), are analyzed in their interactions and complexity, without limiting the scope to narrow linguistic mechanisms (cf. Hauser, Chomsky & Fitch 2002 and their narrow concept of 'syntax'), but adopting a holistic approach to cognition. The properties of perceptual systems affect language acquisition and, therefore, adult grammar. The other way around, once the neurophysiological substratum of perceptual/behavioral systems is found, the question to be asked is how that substratum organizes and computes information provided by perceptual systems, via different kind of 'strategies' involved in acquisition and maturation of mental organs. Such a perspective, I think, should be taken into account more often in current biolinguistic studies, particularly given the very active role the relations between language, cognition, and brain have in Bever's piece (and throughout the volume).

The perspective put forth by Bever is reinforced and actualized in Chapter 1, by Montserrat Sanz, Itziar Laka & Michael Tanenhaus. Of particular interest is the claim that, if some structures do not appear, it is due to the fact that they might not be learnable (p. 81), which sets a strong empirical challenge to be addressed in upcoming years (apart from the attention it has received since Bever's foundational piece). The historical perspective adopted in this chapter (relating Bever's research with previous experiments by Piaget on development and learning) is essential not only for non-linguists who might be venturing into the field from a Biolinguistic stance, but also for scholars working within the field, as the chapter helps situating historically, justifying methodologically, and demystifying some pervasive claims in the field. Developmental psychology, as well as cognitive science, is revealed as a foundational stone for linguistic theories of acquisition, and more recently language processing research, essentially focused on the computational and neurophysiological nature of parsing. Some of Bever's strategies are summarized and discussed, and a partial classificatory typology is established. Within the limits of a book article, the piece provides a well-informed and wide historical scenario, including the aspects of past research that have had major impact on current research (including, but not limited to—and here lies one of the major contributions of the book in terms of wide potential readership—the Minimalist Program advanced over the past 20 years by Noam Chomsky and related scholars).

The book often looks back at itself and provides the reader with means to contextualize some specific papers (as is the case of chapters 8 and 12 with respect to the syntax–semantics interface), and the inclusion of opposing views is more than welcomed: For example, chapters 18 and 19 offer different interpretations of neurocognitive evidence regarding the existence of a set of uniquely linguistic elements (a 'Faculty of Language in the Narrow Sense' or FLN, in terms of Hauser, Chomsky & Fitch 2002), ranging from a defense of FLN to the claim that FL does not contain unique elements that cannot be found in other cognitive domains. This self-referential nature, and the pervasive interconnectivity between chapters is best explicated by a useful (although a bit confusing at first sight) diagram, which makes connections between chapters explicit, in terms of themes and methodologies (p. 109). To the best of my knowledge, this is a rare feature in this kind of contributions to linguistic investigations, and I think the focus on such features should be encouraged regardless the personal opinion the reader might have with respect to the theory or theories entertained in each contribution.

Gerry Altmann builds on Bever's contributions in his 1970 paper in Chapter 2, while doing a review of the development of psycholinguistics from 1980 to this day. Empirically, his focus is set on the interpretation of GPS like (1), mentioned above:

(1) The horse raced past the barn fell.

The parsing that takes [the horse raced past the barn] as the matrix clause is misleading, as the introduction of a finite V [fell] at the end of the sentence requires a complete change of perspective when assigning a structural representation to (1). The chapter describes with clarity the evolution of thinking about garden path and structural ambiguity resolution (of the 'flying planes…' kind), from focus on the process itself to research on the cognitive procedures *before* parsing. Bever's contribution is taken as a departing point to discuss a number of approaches that emerged during the late '70s, and during the '80s and '90s, including the author's own research vinculating syntactic and semantic processing (as opposed to Bever's dissociative view, p. 115) via linking rules. Interesting perspectives blurring the distinction between syntactic and semantic processing, as well as competence and performance, are introduced; although references might seem a bit outdated if the reader wants to follow up to the chapter. The discussion about connectionist networks and their mainly statistical approach to meaning (based on Elman's work) in section 2.3 is clear and concise, but only one recent reference addressing the issue (Altmann & Mirkovic 2009) is mentioned, which I think is somehow anticlimactic. (The same actually occurs in subsequent sections: Relatively recent references are almost always limited to the author's own works, the only exception being the reference to a special issue of *Trends in Cognitive Science* from 2010.) This chapter is eminently descriptive/ explicative (briefly introducing ideas and authors, and summarizing effectively three decades of psycholinguistic research while acknowledging the impact Bever's work had on computational linguistics and neural network research), and does not engage on independent argumentation or raise new questions: Its place in the book seems to me to be well chosen (as an introductory chapter), but it might disappoint the reader looking for original research.

Chapter 3, by Maryellen McDonald, also stems from Bever's considerations about garden path sentences, but confronting them with sentences like (2):

(2) The boy that the girl likes hit the man.

Both GPS and sentences with relative clauses containing an overt C [that], like (2), have been addressed from psycholinguistic points of view. However, and this is one of McDonald's points, seldom have they been discussed in a single piece, contrastively and comparatively. Like Lin in chapter 4, McDonald rejects the proposal that GPS and complex, but unambiguous, sentences are processed under different assumptions: The goal is to bridge the division of the field of sentence comprehension by unifying those two apparently different kinds of processes. McDonald sheds light over Bever's initial assumptions regarding a constraint-based system of ambiguity resolution, and relativizes the equation noun + verb = subject + verb by including not only syntactic patterns into account, but also lexical information and extra-linguistic resources (e.g., speaker identity, visual environment, etc.; p. 132). It is relevant to point out that the constraint-based approach has not been applied to relative clause interpretation (at least not widely), since they are traditionally believed to be unambiguous, and constraint-based systems have usually been associated to ambiguous sentences. Work in relative clause processing, summarized in chapter 4, has mostly appealed to a two-stage system, in which syntactic processing precedes semantic effects. Chapter 3 has as its goal to apply Bever's general ideas, expressed through a constraint-based model, to object relative clauses (i.e., relative clauses in which the *wh*-operator is the object of the embedded V). The exposition takes into account computational limitations in human minds, as well as factors such as memory limitations and interference, when rejecting purely structural accounts (i.e., accounts based primarily on independent syntactic processing) in favor of a constraint-based system as a model for general comprehension, which implies a simpler and more coherent theory. The argument requires development of research of ambiguity within relative clauses, which is provided in a complete and clear subsection (pp. 135ff.), providing recent and relevant references to the interested reader. The discussion of the author's own work towards integration of both phenomena (GPS and relative clause processing), as announced, takes into account the animacy feature of the relative clause's antecedent as an important cue for interpretation, thus resorting to lexical/semantic factors as well as structural information (e.g., the antecedent is coindexed with the object of the relative clause). To complete the chapter, a discussion of *production* models is provided, with which the offered perspective is even wider, even if production is addressed almost exclusively from a statistical perspective which takes into account tendencies regarding animate and inanimate antecedents. The references in this section are mostly the authors', and the paper concludes almost surprisingly, with section 3.5. A conclusion section, summarizing the highlights of the piece, would have been welcomed. However, it does advance some lines of current and future research.

Chapter 4, by Chien-Jer Charles Lin, is intimately related to chapter 3, as it deals with relative clause processing. However, unlike chapter 3, it draws heavily on Chomskyan generative grammar, which leads to claims of the kind "they [relative clauses] demonstrate three critical formal properties of human language: recursivity, the existence of empty categories (e.g., traces), and constraints on dependencies related to those categories" (p. 142). These are problematic claims, insofar as no independent evidence is provided, nor are alternative accounts discussed. For example, there is a debate about the role of recursive procedures in natural language (which are obviously not the same as recursive functions as defined by Gödel 1931 [1986]) which the author overlooks; the same happens with the existence of empty categories in the sense of Chomsky (1981), namely *wh*-trace, NP-trace, *pro*/PRO, which require a strong burden of independent proof (as it is possible to conceive internally coherent frameworks without the need to resort to traces/copies) and has been challenged from more than one front (e.g., Culicover & Jackendoff 2005). This is not to say that the author's arguments are to be rejected; it just means that the argumentation does not take into account alternative frameworks and is in this way limited. The discussion on relative clause processing relies crucially on the interpretation of *gaps*, and the establishment of *filler–gaps* dependencies. Notice that the notion of *gap* as presented by Mainstream Generative Grammar requires, as Culicover & Jackendoff (2005: 16) put it, 'hidden levels of syntax', related by means of structure mapping operations. In my opinion, no sufficient *independent* evidence is provided in this piece to accept that conclusion. The author proposes parallel syntactic and thematic (semantic) processing strategies (the latter comprising 'templates' which are activated automatically, in relation to Bever's N + V = Subject + V procedure), but does not specify whether thematic information can be used by the syntactic parser and vice versa (cf. chapter 3), which is, I think, a gap in the argument. Moreover, both representations must match (p. 144), but no details are provided with respect to how exactly this process takes place. Multiple recent references are provided, and, should the reader accept the initial assumptions, the discussion is internally coherent and consistent. My objection to the structure and content of this chapter stems precisely from the lack of justification for those initial assumptions, too strong to be taken for granted. Empirically, this chapter provides comparative evidence from English and Chinese, a most welcomed strategy, and processing asymmetries in production and comprehension of relative clauses are justified with this comparative evidence. Unfortunately, no future prospects are provided, and there is no independent conclusion section.

The next chapter focuses on English data, and relates to chapters 3 and 4 insofar as its object of inquiry is the processing asymmetries of subject and object relative clauses, and complexity issues related to this processing. While not committing themselves to any particular theory of syntax (unlike the previous chapter), a substantial amount of literature is provided by the authors Edward Gibson, Harry Tily & Evelina Fedorenko in each point of the discussion, and cross-linguistic studies are also mentioned (although, as clearly stated in the chapter title, concrete cross-linguistic data is not discussed). The chapter is organized in three main parts, corresponding to three main theories about complexity issues arising in the processing of extraction effects: Section 5.2 is devoted to *reanalysis theories*, according to which an incorrect parsing (e.g., interpreting [raced] as the main V in (1), with [the horse] as its subject, following the N+V = Subject +V procedure) is to be somehow repaired (although no details is provided about how this is performed, even within the discussion section 5.5); section 5.2 discusses *experience-/surprisal-based theories*, according to which interpretation is a statistical function of previous experience with similar input (e.g., more frequent words are easier to process), both at the word-level and the phrasal level. This section presents more discussion, and brief reports of different tendencies within the general approach, something that would have been desirable for reanalysis theories as well. Finally, *memory-based theories* take into account *working memory* capacity, and predict more complexity in ORC processing because they imply both more *storage cost* (maintaining a dependency active, as would be a relative operator and its associate base position) and more *retrieval cost* (retrieving word-meaning from LTM, for instance). Once again, a concise discussion about alternatives is provided, and in the last two sections the reader can find recent relevant references. Section 5.5 summarizes the predictions each group of theories would make, and provides evidence for each. The discussion is extremely neat and very well organized, the arguments can be followed with no difficulty, and for each empirical problem a substantial load of references is available. Section 5.6 problematizes the predictions made by each theory, with the interesting and thought provoking claim that no theory by itself can explain all considered phenomena; followed by section 5.7 in which two experiments including the relevant data, considering as a variable the possibility that a single NP can fulfill or not several grammatical functions with respect to the verbs, are carefully reported (including participants, methods, and results) and discussed from each viewpoint, spelling-out the predictions each theory makes applied to the experiment in question. The permanent discussion between theories and the specific predictions they make in different cases (with particular focus on variants of memory-based and experience-based theories) is one of the highlights of the chapter, and is very welcomed as a methodology for the presentation and contextualization of both the frameworks and the data. In the conclusion, prospects for the application of retrieval-based theories (a subtype of memory-based theories) to languages with different basic word order than English (SVO), like Japanese, Korean, or Chinese are presented, and constitute a further challenge for the retrieval-based framework the authors mostly support, while always arguing convincingly and with various sources in favor of mixed approaches and not relying on a single mechanism to explain such a complex phenomenon as RC interpretation.

Chapter 6 is built upon the concept of *psycholinguistic chain* (P-chain), which relates production, prediction (error), and (consequent) acquisition in a novel form, with modifications presented in the chapter. The first approximation to the P-chain would be as follows (p. 175):

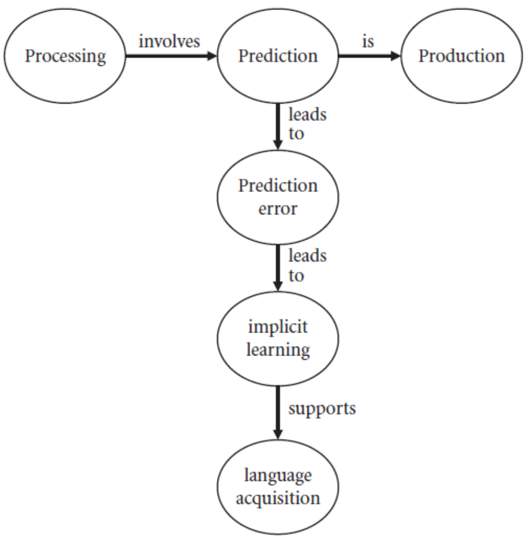


Figure 1: Psycholinguistic chain

The authors (Gary Dell & Audrey Kittredge) argue their way through the P-chain in a clear and concise manner, providing theoretical and experimental support for their claims, and relating language with other cognitive capacities (e.g., eye-tracking), as well as stressing the mutual relation between processing and production, as well as implicit learning and prediction (as loops in the chain), making the model more dynamic and powerful. Section 6.2 is devoted to full explication and argumentation of the claims involved in the P-chain framework, dedicating particular subsections to each of the 'links' so that the reader can have a general idea of what the framework is about, and how it impacts on empirical research. The framework is contextualized historically (in relation to previous theories) as well as epistemologically (in relation to contemporary alternatives), which is a great help for readers with a limited background on psycholinguistics and its development. The chapter presents a most interesting view of acquisition, in which partial errors in structural/lexical prediction lead to changes in the linguistic system (a perspective that is perfectly compatible with recent advances in complexity theory and language as a complex system), resulting in the readjustment of the system (acquisition). Overtly independent of syntactic frameworks like Generative Grammar, the article assumes the existence of "an 'innate' architecture that has the ability to learn sequences and to represent meaning" (p. 179), which, exposed to input (sentences, their meaning being inferred from the context), corrects activation weights in neural networks and the system thus acquires a structure *and* its meaning, based on a 'trial and error' basis. While objectionable (particularly from the viewpoints of formal syntax and semantics, as the computational/neurocognitive nature of syntactic structure is not made explicit), the framework is a dynamic attempt to coherently relate three essential processes in language use, external or not. It is to be noticed that the case study presented by the authors involves phonotactic learning, which does not have the articulated structure that is currently theorized for syntactic representations or semantic structures. It should be a challenge (and a desirable and exciting development) for the P-chain theory presented here to try to accommodate the RRCC data presented in previous chapters and the acquisition of discontinuous Operator-Variable syntactic dependencies, to give some examples.

Chapter 7 gets back to Bever's article as a foundational stone for many issues of present relevance for psycholinguistics. David Townsend focuses on four claims deeply related to Bever's paper (in fact, reformulations of Bever's strategies) and provides theoretical and empirical support for them (p. 184):

- Comprehenders form linguistic structures.
- Linguistic elements project structure.
- Common representations interact.
- Grammar checks the adequacy of projected structures.

Given the strength of some of those claims and/or the theoretical and empirical consequences they have, it would have been nice to define some key concepts, like 'structure formation' or 'projection', of crucial relevance in current theoretical linguistics, particularly within generative grammar (from which the author takes concepts, like 'traces' in p. 191).

The author compares and confronts the program advanced by Bever with previous theories (e.g., Derivational Theory of Complexity), providing evidence in favor of the former. Each of the claims above is expanded on in its own section, which also features Bever's original interpretative strategies for reference and clarification. Garden path effects are compared with other kinds of parsing ambiguities (e.g., homonymy), and argument alternations (e.g., The dancer tripped John/The dancer tripped), which also affect the assignment of a structural description to a certain string. The author introduces concepts about verb semantics (like 'bounded verbs') and the requirements they establish for co-occurring arguments which are not formalized or made totally explicit, however, the argumentation can be followed without problems. Section 7.4 clarifies the somehow obscure *sentential* 'common representations interact', by making a case for the interaction of syntax and semantics, and how comprehension makes use of structural and semantic information simultaneously, while comparing and contrasting competing theories with experimental basis (although little independent evidence is mentioned at this point). Section 7.5 explores the means by which grammar and meaning interact via patterns that provide provisory semantic representations to be refined in real time, although no clear definition of what 'grammar' comprises is given (sometimes it seems to be used as a synonym of 'syntax', but that is not clarified). The section is carefully argued, and extends on the mechanisms via which comprehenders anticipate meaning and structure in terms of conceptual and/or linguistic representations.

Chapter 8, by Robert Berwick, is more narrowly linguistically oriented. Taking as a 'cornerstone' of Bever's seminal article to highlight Chomsky's (1986) distinction between *knowledge of language* and *use of that knowledge*, the author attempts to provide a synthesis between internalist and externalist models of language. External modeling is identified with statistical methods in corpus linguistics (and part of computational linguistics as well, including insights from information theory), having as its aim to be able to predict the next element to appear in a string, in turn assuming a certain model of comprehension based on memory-retrieval, as we saw above. Internal modeling (which the author identifies with Chomskyan generative linguistics, which is a perspective I find quite limited, given the amount and quality of alternative formal internalist approaches), on the contrary, focuses on simplicity at the time of formulating generalizations about the mentally represented knowledge a speaker has in its mind-brain. The author proceeds to discuss formal grammars in section 8.4, assuming that the subset relations expressed in the so-called Chomsky hierarchy hold (but see Krivochen 2014 for a critique of such a claim both theoretically and empirically grounded). The argument expands on that made in Chomsky (1957) about the inadequacy of Markov models to account for linguistic structure, as they are based on linear relations (a claim which has to be, at best, relativized to portions of natural language grammars, as I show in Krivochen 2014), which is useful for the reader more familiar with neurolinguistic literature than with the foundational texts of generative transformational grammar, but adds little if anything to the discussion about the adequacy of certain formal grammars to generate structural descriptions for natural languages or particular segments of them. A main concern of the author seems to be to establish a comparison procedure for grammars, understood as sets of (generative + transformational) rules, which favors phrase structure grammars (a claim traceable back to Chomsky's early writings). The author points to an apparent tension between statistical methods and linguistic description, given by the fact that constituency tests do not always coincide with statistical preferences (p. 201), such that, for instance, the syntactic constituents [VP walk [PP on ice]] are differently chunked when it comes to statistical prediction, and the P [on] is more likely to appear with [walk] that [walk on] is to appear with [ice]. The proposed solution is highly theory-dependent, and consists on substituting a standard phrase structure grammar with Bare Phrase Structure (BPS), in which there are no labels (for details, see Chomsky 1995). However conceptually appealing the proposal might seem, there are to the best of my knowledge no neurocognitive accounts that support such theory, what is more, performance compatible models (e.g., Sag & Wasow 2011) provide more explicit accounts of the apparent tension between knowledge and use than the alternative proposed here, strictly tied to the Minimalist Program. Weakening the conditions for predictive models is not, in my opinion, an advisable methodological step. The announced synthesis consists on taking BPS as higher-level instructions that generate particular instructions (knowledge put to use), appealing to the (digital) computer analogy, and acknowledging the combination of different information sources (a recurrent motif throughout the volume). In spite of the multiple theoretical biases we find here when it comes to internalist linguistics, the argumentation is clean and neat, and the reader who is not familiar with the field of formal grammars will find a nice introduction to some old but still relevant arguments.

Connected with the linguistic concerns of chapter 8, Chapter 9 also introduces the discussion of Center-Embedded Clauses (CEC) within the framework of Chomsky's version of the Minimalist Program. Janet Fodor links Bever's strategies (which are not language-specific) to the (methodological) desire to minimize the specificity of the Language Faculty and allocate as many properties of language as possible in other cognitive systems, so-called 'external systems'. CEC present discontinuous dependency patterns like (3):

(3) The dog the cat the fox was chasing was scratching was yelping.
(Bever 1970: 334)

Fodor presents recursion and movement (transformational rules) as two *facts* about human language, for the second time in the volume (cf. chapter 4), thus restricting her theoretical framework (and the syntactic representations she uses) to Minimalism (without acknowledging much discussion about the nature and properties of recursion, including problems concerning its very definition). She describes CECs as particularly difficult to parse while not presenting differences with other kinds of clauses in terms of multiple embedding or movement rules. In opposition to previous accounts based on structural subjacency, Fodor reformulates Bever's *Strategy J*, regarding the relative roles of NPs in NP, NP, …, V configurations and the assignment of grammatical functions within their respective clauses. She holds the threshold of two levels of embedding as providing particular processing problems, while deriving it from the syntax–prosody interface. Her account relates the assignment of structural descriptions to units based on the prosodic contour of local units, in turn relying on the idea of local phonological cycles and the necessity to wipe the working memory clean of structure as soon as possible, several times during a derivation. Fodor provides cross-linguistic variation patterns of RC attachment preferences, based on derivational chunks ('packages', similar in spirit to Chomskyan *phases* but psycholinguistically supported) and the difficulty of the parser 'looking into' an already finished package, which is subjected to interpretation as a unit. The packaging mechanism would not be directly derived from memory issues, but from prosodic patterns (which are present even in silent read, the so-called 'Implicit Prosody Hypothesis'), thus cross-linguistic variation can be accounted for without the need to suggest different speakers of different languages have different working memory capacities. The article has important consequences for the theory of phases and syntactic locality in general, although a discussion of the implications this theory has for semantic cycles would be necessary in order to implement the model within a wider program. The author integrates phonological and lexical information, but it is not clear whether the packaging occurs on the meaning side as well as on the sound side, an interesting challenge for the theory presented here to address. The prosodic interface and inner structure of intonational phrases are however described with great detail, and even if a one-sided (i.e., taking into account only the sound interface) explanation of the phenomena involved does not seem plausible to me, the evidence is carefully presented and the arguments follow from the initial claims with no gaps, should one accept the path taken by Fodor. The author herself provides a discussion of non-prosodic explanations in section 9.5, focused on syntactic accounts mainly worried about structural distance between dependent constituents for memory reasons (which do not coincide with the "distances that matter for prosodic parsing", p. 228), and giving arguments in favor of the superiority of the prosodic account in terms of predictions. In relation to the previous chapter, a mention of Markov models for phonological structure would have been a nice link (as there is a mention of 'flattening' structure at the syntax-phonology interface, p. 217), but it is a task left for the reader to undertake.

Chapter 10 takes the reader back to neurocognitive issues, drifting away from generative linguistics. Brian McElree & Lisbeth Dyer focus on the role of working memory in deriving linguistic expressions in real time (a topic explicitly left aside in generative grammar), and how linguistic processing is limited by non-language-specific constrains on the amount of structure that can be processed at any given time. The authors, advancing Bever's (1970) inquiries on the role of memory representations during comprehension, its nature, and the factors determining the success or failure of memory-involving processes; provide a much needed gap-filling, since there has been surprisingly little research on working memory and its relation to real-time language processing, particularly when facing long-distance dependencies between constituents. The authors review previous theories, problematizing tricky notions like 'processing complexity' (which are taken for granted in many narrow-syntactic works) and critically evaluating their impact on different accounts of memory-limitations approaches to comprehension, impairments (e.g., as result of brain injuries), and reduced processing capacities, providing the reader with a fair amount of relevant bibliography on each section. The very notion of working memory as opposed to a retrieval-only Long Term Memory (LTM) is challenged, insofar as empirical evidence has not been conclusive enough to postulate two separate systems, and a more dynamic system is argued for, in which WM does not have a fixed capacity (cf. Miller's 'plus-minus seven units'), and information retrieval does not seem to be privileged or faster for elements predicted to be in the WM with respect to elements predicted to be in LTM. Therefore, a fixed approach to WM is inadequate, the authors claim, and a dynamic real-time approach is necessary to account for comprehension phenomena. Section 10.3 is devoted to information retrieval in language comprehension, comparing predictions about the respective roles of WM and LTM with exemplified experiments (only essential details and general discussion). Retrieval models are relevant insofar as cue-driven retrieval can account for both rapid access to information as well as failure to pick the relevant piece of data out (importantly, this is not limited to language, but applies to "any complex cognitive skill", p. 238), if the cue does not point towards the required information with enough specificity, what is called 'retrieval interference', there being the possibility of overlapping between cues. This model can provide the flexibility that fixed WM accounts lack, and the authors carefully argue their point. The claims about memory and retrieval possibilities are adequately exemplified, with clear cases and a concise account of each, without adhering to any particular grammatical formalism (which is a positive note, insofar as the reader can translate the results to the framework of his preference). The reader is lead through the discussion gently, with numerous and recent bibliographical sources.

The next chapter, by Ina Bornkessel-Schlesewsky & Mathias Schlesewsky, touches on a crucial point for both grammatical theories (including syntax, semantics, and morpho-phonology) and neurocognitive approaches to language: the role and identity of universals. The authors begin with an overview of the concept, and the (relative) dissociation between linguistic universals and cognitive universals (already drawn by Bever), as well as the present difficulty of finding real universals, particularly after the tremendous growth of typological studies, which often force theoreticians to relativize universals into tendencies. The authors' goal is to combine neurocognitive research with linguistic typology, in so-called 'neurotypology', a most interesting aim and certainly welcomed gap filling in (non-UG-driven) research on universals. The enterprise is based on a dynamic approach to the relations between language, brain organization, and (general principles of) cognition. Specific linguistic characteristics would be given by a direct relation between properties of the brain and properties of language, without mediation by cognition. The authors discuss a number of related proposals which address topics underlying the aforementioned tripartite relation in section 11.2; and address the issue of inter-linguistic variation in section 11.3, providing evidence of different neurological responses to form-meaning conflicts in different languages, supported with a good deal of references and brief, but effective, experiment reporting. This section makes a point of qualitative inter-linguistic variation from a neurological point of view, which the authors attempt to derive within a framework based on Bever's strategies, particularly the NP–V–

NP pattern as *actor–process–undergoer*, and the properties of a given language in terms of cognitive categorization and decision-making. The notion of 'cue', which has already appeared in the volume, is of key relevance here, as languages seem to rely of different cues when assigning semantic roles to arguments in a syntactic construal (p. 247). In spite of differences, it seems that the cognitive system prefers prototypical actors and actor-initial argument orders, in order to identify roles as quickly and effectively as possible, a conclusion supported by inter-linguistic electrophysiological studies involving languages belonging to different families (Germanic, Altaic, Romance) and typologies (accusative vs. ergative languages). Asymmetries between actors and undergoers (patients–themes in theta-theoretical terms) are also attested, always providing neurophysiological evidence (which considerably strengthens the argument), which the authors interpret as a 'competence' for the actor role (p. 249), depending on the prototypicality (animate, human, definite, nominative/absolutive) of the competing arguments, a competence which does not arise for other argument roles. This competition, the authors claim, is a plausible universal of linguistic processing (which is not to be confused with a proper linguistic universal of the kind advocated for in UG-based proposals or even Greenberg-type universals). The authors introduce the category of 'neural attractor' for the actor role, insofar as it is that role which triggers the competence between arguments in processing. Moreover, this competence could be modeled by means of attractor networks, which is in itself an exciting empirical challenge for the neurotypological enterprise (and its collaboration with related disciplines, particularly mathematical modeling of complex systems) in future years.

Chapter 12 focuses on the syntax-semantics interface (arguably, a topic also present in chapter 11, taking into account its concern for role assignment in processing), and the respective takes of formal syntax (assuming the Minimalist Program, in detriment of alternatives which are not even mentioned) and psycholinguistics. Montserrat Sanz provides a useful *racconto* of the takes on thematic structure from GB to Minimalism, and problematizes the mapping between semantic construal (in which notions like 'event' are core) and syntactic construal (which works with formal, semantic, and phonological features, in the theory assumed by the author). Sanz claims that linguistics deals with competence, whereas psycholinguistics deals with performance, a claim that leads to justifying Chomsky's seminal distinction. However, the distinction has blurred in several occasions, and unifying theories have been proposed (some in this very same volume, but see also Sag & Wasow 2011 for an alternative outside transformational generative grammar), whose discussion would have been welcomed. The author's take on the syntax-semantics interface is heavily influenced by the strong role features play in the Minimalist program, and parsing is also tackled from this stance. This perspective, while not extent of problems (particularly given the difficulty of assigning neurocognitive reality to formal features, which are at the very core of Minimalism), is novel and therefore welcomed; and whereas the concept the author has about what constitutes the syntax-semantics interface can be discussed, it is a pushing-forward development of the initial Minimalist desire to explain properties of language in terms of output conditions established by the C-I and S-M interface systems. Moreover, it critically discusses both exclusively syntactic and exclusively semantic attempts at explaining the parsing difficulties of GPS, which strengthens the interface approach to linguistic phenomena. Assuming the Minimalist Program (Chomsky 1995 *et seq.*), then, the article is not only reader-friendly and well-grounded, but also provocative (at least for the most orthodox takes on the role of the interfaces in parsing). The author is very careful in including multiple factors into account (e.g., Aktionsart, lexical syntax/semantics, verb typology, etc.) when arguing in favor of a particular take on GPS processing, and provides relevant references for each factor. Considering these variables, Sanz argues that there is a gradient of difficulty for GPS processing, which is more easily accounted for assuming aspects of lexical semantics and compositional properties of the Verb Phrase and the event it denotes (including Aktionsart) than ignoring those factors. The discussion turns highly technical when the author considers the possibility of including an Event Phrase as a functional projection in the syntactic construal, whose interpretable features are read off at the semantic component. The author claims that thematic roles are *parsing necessities*, not grammatical necessities (*contra* those approaches within Minimalism, like Hornstein 2003, that consider theta roles as features to be checked *before* the derivation reaches the interfaces): This claim has potentially interesting consequences not only at the empirical level, but also when considering the 'design' problem for the Faculty of Language (for those approaches that assume such a notion). Within the theoretical limits imposed by the Minimalist Program, Sanz makes a valuable contribution, and advances the ground in relevant and little explored aspects of the (lexicon-)syntax-semantics interface(s).

Massimo Piattelli-Palmarini tackles the issue of Platonistic vs. Cartesian approaches to language, which has been the topic of recent debate (see, e.g., Watumull 2013 for an attempt of unification, and Behme 2014 or Krivochen 2013 for a critical view and extended discussion). The main issue is whether linguistic objects can be abstract and biological or not, and the ontology of derivations and the generative system in each case. Piattelli-Palmarini makes an assessment of the role of abstraction in linguistic theory, clearly aligned with the Chomskyan view that there is no 'knowledge of language', as 'language' itself is the knowledge a speaker has in its mind-brain, UG and the grammar of a particular language (p. 264). The article is clear and well-organized, although key notions are left undefined (as in most of the papers constituting the realist-conceptualist debate), 'abstract object' being perhaps the most important. The author argues his point with empirical evidence regarding so-called 'conservativity', a set-theoretical property according to which $A \cap B = (A \cap B) \cap A$. This is particularly revealing of the framework he assumes, since the Minimalist generative operation Merge forms *sets* (Shieber's 1986 Unification also works with sets, but the resulting object is not characterized as identical to either of the terms involved in the operation, but as the union of the feature matrices involved). This property is said to hold for determiners (a linguistic label) *in all natural languages* (references are provided for this claim, but only few examples outside English are analyzed), however, the examples actually involve existential and universal quantifiers and their logical properties, p. 266 (regardless their materialization, namely morpho-phonological form, this is an important distinction). Piattelli-Palmarini compares English determiner 'the' and its properties with imaginary determiners for which the property of 'conservativity' does not hold, and claims that non-conservative determiners would invert the logical properties of the word class as a two-place predicate. In my opinion, there is a mix between linguistic and logical properties of determiners (a trend traceable back to Russell 1905), which makes the discussion a bit hard to follow if one distinguishes logics from (formal) grammars of natural languages. It is also not clear whether the property arises at the semantic interface or is relevant in the so-called 'narrow syntax', although the second option is hinted at—insofar as feature checking considerations are mentioned, but the exact *locus* of conservativity relevance is not formulated explicitly: The mention of 'syntactico-semantic structures' (p. 268) does not clarify the matter, as 'semantic structures' is left undefined. After the consideration of the examples, and discussion from a Minimalist stance, section 13.3 dwells with the status of this 'universal', and goes back to the initially mentioned debate between realism and conceptualism: The author's hypothesis is that the universal follows from a property of the Language Faculty, even if the motivation for this property is "outside the domain of linguistics" (p. 270), which is at least an anticlimactic claim: Why should it be outside the boundaries of linguistic inquiry, particularly considering the 'biolinguistic enterprise'? Sections 13.4 and 13.5 (the latter, a conclusion) focus on the properties of 'abstraction' as a cognitive operation, but the exact nature of 'abstract objects', central to the realist-conceptualist debate, is never clarified. Nor is the highly problematic notion of 'virtual conceptual necessity', introduced at the very end (p. 271). The article is reader-friendly, particularly for those who are familiar with the assumptions and axioms of the Minimalist program, but is very likely to leave other readers asking questions about the overt and covert assumptions that guide and underlie the argumentation.

Chapter 14 follows on the topic of determiners and their role in the core syntax, but from an *empirical* stance, which is most welcomed. Methodologically, Virginia Valian chooses to begin by justifying her selection of determiners as the object of study, which is a rare and welcomed feature in innateness-related studies. Moreover, the author (briefly) addresses the issue of what is innate and what is learned/acquired, again a point in favor of the methodology followed in this article. The author's concern to make her assumptions fully explicit before entering data discussion (limited by space reasons) is ostensive and clarifying. The study includes careful analysis of determiner acquisition timing: Experiments, rather than being fully explained including methods, participants, and results followed by discussion, are directly discussed, even though relevant results are incorporated in tables. Empirical predictions stemming from the claim that schematic representations of determiners are innate are spelled out, and they involve continuity on the developmental trajectory (p. 276). Evidence in favor of a continuity approach is provided, including the crucial notion of *underspecification* and its role in acquiring the relevant elements; as well as *equivalence classes* (how the class of 'determiners' is abstracted from the data, and, conversely, how elements in speech stream are assigned a class). After discussing what the author hypothesizes to be innate, and providing experimental evidence in favor of her hypothesis; she proceeds to discuss what is left to be learnt in section 14.5. Three factors are identified here: prosodic templates, knowledge of specific lexical items, and controlled processing including several sources of information.

These factors are succinctly explained, and it would have helped the reader to have a separate conclusion summarizing the main points of the piece, as well as including future prospects and empirical challenges. The article is nevertheless a valuable piece, as it relies on experimental data to make assertions, and predictions are clearly spelled out, which makes them falsifiable and thus scientifically interesting.

In Chapter 15, Simona Mancini, Nicola Molinaro & Manuel Carreiras analyze the concept of morphological *agreement*, from the perspective of the Minimalist program, in which feature agreement triggers operations like Move and even, in some versions of the theory, Merge (e.g., Wurmbrand 2014). Despite considering features (of the kind [*value*, Dimension], as in [*Past* Tense]) "the basic building blocks of a derivation" (p. 282) as in orthodox versions of Minimalism (a claim shared with models like HPSG and LFG) but an assumption that has not remained unchallenged (e.g., by Cognitive Grammar, Construction Grammar, and the like), the authors provide a novel perspective over so-called φ-features, a bundle including *person, number,* and *gender*, from a psycholinguistic stance. The thesis is that, since there are differences in the processing of each kind of information, those features should not be treated as a single unit. The authors distinguish between morphosyntactic information conveyed by a feature (φ-value) and the semantic-pragmatic information concerning the *denotatum* (σ-value). That is an interesting distinction insofar as φ-features are traditionally thought to be uninterpretable by LF, and the notion of σ-value makes a point of the semantic relevance of those features, at the cost of introducing yet another distinction in a theory that is already quite far from 'minimalist'. The discussion is clear, and examples are clarifying, particularly given the fact that new notions are introduced, like σ-value or 'interpretative anchor' of a feature (its σ-value). Theoretically, there is a further complication represented by matching operations between both sets of values, but if the enterprise pays off empirically, the complication will have been justified. The article does not present so detailed an analysis that we can be certain of this, but it is a challenge to be addressed in future research. However, section 15.3 tests the validity of the approach against psycholinguistic evidence, including ERP patterns for some of the Spanish examples cited, which reinforces the point, as well as analyses of N–A and V–N agreement patterns, including acceptable morphological mismatches and their subsequent explanation in terms of the model presented here, both theoretically and via psycholinguistic evidence in processing. It is not clear that the data cannot be accounted for via different agreement patterns, as the alternative is not considered (which opens the door for future simplifications of the theory), but the level of descriptive adequacy is reached, and, should one adhere to feature-driven operations in the syntax, so is explanatory adequacy. A point in favor of the article is that there are comparisons drawn between what would be expected in mainstream models of comprehension (involving φ-features as a single bundle) and the 'anchor' model presented in the article, which helps situating the proposal within the field, in relation to alternatives. The dissociation between φ- and σ- values allow the authors to explain qualitatively different patterns in neurocognitive studies between person mismatches and number mismatches, which advances the ground with respect to orthodox agreement research. The notion of σ-values and anchors could have been developed further, and future prospects for the theory could as well have been given, but the perspective of the piece is overall interesting and novel within Minimalist assumptions.

Colin Phillips, in Chapter 16, revisits Bever's parsing model, in which an initial hypothesis about meaning is the result of a 'quick and dirty parsing', to be later on replaced by the definitive representation of the linguistic expression's meaning. The article critically reviews the relation between mental grammars and parsing, as well as the psychological reality of grammatical representations, and the historical relations between generative transformational grammars and psycholinguistics, focusing on the problems the transformational component brought about for psycholinguistics given the uncertain psychological status of transformational rules. Early empirical research seemed to support the general view about a generative component, but the conclusions were at best elusive when considering transformations (e.g., the impossibility of 'reversing' a transformation). The historical review is helpful for the reader, insofar as it presents hypotheses and experiments carried out in the early days of generative linguistics in a concise and clear manner, as well as the theoretical and empirical challenges the data imposed to transformational models. The author also addresses the difficulties presented by the Derivational Theory of Complexity, and the necessity to critically revisit the basic ideas the DTC presented and are nowadays still in use. Phillips considers, as empirical points the incremental character of linguistic parsing, which is incompatible with Standard Theory's rewriting rules (based on L-grammars), and the problems posited by sentences which had apparently undergone a transformation, whose interpretation required additional stipulations in the psycholinguistic side in order to comply with the model of the grammar. Section 16.4 is focused on discussing Bever's 'double interpretation' model and plausibility-based strategies, suggesting that comprehenders build fully-fledged representations for sentences (p. 306), a claim that is not incompatible with probabilistic heuristics. Phillips' discussion includes numerous references to experiments succinctly described, as well as bibliographical references which are of much use for the reader to have direct access to primary sources of the cases reported. It is to be highlighted that potential objections and counter-arguments (often related to Bever's account and similar purposes) are considered and properly addressed by Phillips, which makes the point stronger and also leads the reader gently into the conclusions. The article makes a point of the necessity of looking for more than one way to account for processing phenomena, suggesting alternatives and considering (within reasonable space limits) the theoretical and empirical implications of each possibility (section 16.7 and the revision of Townsend & Bever's 'analysis by synthesis' is a fine example of this tendency). While acknowledging the importance of Bever's research for the interaction between grammar and psycholinguistics, Phillips presents a critical panorama and points towards several possibilities for future prospects.

Edward Stabler analyzes the relation between language and cognition from a computational perspective in Chapter 17. The author problematizes accounting for linguistic variation, the lack of consensus upon basic theoretical notions (both major issues in current formal linguistics), and the identification of psychological processes involved in linguistic parsing in section 17.1, which helps situating the problematic interplay of computational linguistics and psycholinguistics in context. Section 17.2 presents some basic notions about which there is, apparently, more consensus than often acknowledged, although oversimplifying some issues: While Joshi's (1985) claim that natural languages are mildly context-sensitive is indeed widely accepted (p. 318), it is also to be noticed that there has been research on the Markov nature of morphophonology from the 80's onwards, as well as higher-level, Turing computable models for linguistic theory (see, e.g., Watumull 2012). Therefore, his subset relations between formal grammars, centered on Minimalist Grammars as defined by Stabler (1997) and Michaelis (2001) leaves aside many important and relevant issues both computational (e.g., the alternative models that have been developed beyond Minimalist Grammars, including Unification-based grammars, to give but one example) and empirical (their descriptive/explanatory adequacy). The author centers his attention on Context-Sensitive Minimalist Grammars (CSMG), including a transformational component (as he formalizes the notion of movement in a Minimalist tree, p. 321). All theoretical biases notwithstanding, Stabler attempts to unify formalisms in order to address fundamental issues arising in various versions of the Minimalist Program (like the existence of traces and multidominance alternatives, and the computational nature of Merge, as well as learning methods). Section 17.3 proceeds to briefly discuss the relation between CSMG and psycholinguistic research, departing from Bever's work, and including automata theory. Stabler argues in favor of the existence of computational universals, which he opposes to 'concrete universals', an interesting distinction particularly considering the content of chapters 11 and 13; it might be interesting for the reader to see if the 'conservativity' property Piattelli-Palmarini proposes as a universal holds, and how if so, in a CSMG. Unfortunately, the distinction is not developed to its full extent, no examples are provided (this holds all throughout the article), and the chapter ends quite abruptly. It does, however, provide some future challenges to be addressed from a computational perspective for linguistics and psycholinguistics.

Chapter 18, by Luciano Fadiga & Alessandro D'Ausilio, focuses on the relation between 'action' (so-called 'motor system', although it is clarified that more than a single area of the brain) and language, digging into the problem of how several processes are temporary organized, and how 'actions' obey an end, related to the issue of problem-solving. The issue is relevant for language insofar as there are common characteristics found in problem-solving and language structure, like recursivity (with due distinctions between nested structures and strict sequentiality, the latter of which is problematic), which is not technically defined, but in this paper seems to be synonymous with 'hierarchy' (even though it is not the case that all definitions of recursive functions would be compatible with this approach, particularly given the fact that we can have hierarchy without recursion, in a non-trivial sense, if recursion comprehends [X…[X…]] structures and we operate only with hierarchy without center embedding, appealing to monotonic applications of a generative function). This relevance is somehow difficult to see in the first sections of the chapter, devoted almost entirely to the motor systems, their functioning, and comparison with other systems (like the visual system). The properties of neural networks on which the motor systems depend are also explained, with references where applicable. Only in section 18.3, devoted to Broca's area, can the reader begin to establish a relation with human language, given the key importance of this area in language processing, and in section 18.4 the relevance of the previous discussion is spelled out. It would have helped the reader if this relevance had at least been hinted at in the introduction, or a small summary of the content of each subsection had been provided, so that the reader can prepare to grasp the major points of each section and connect them all in 18.4. Discussion in this section is mainly speculative (e.g., regarding the common origin of language semantics and syntax and the motor systems, a claim that the authors themselves recognize not verifiable), and to some extent also vague. After what seems too long an introduction, more concrete connections with observable aspects of language behavior or careful consideration of language structure would have been expected. However, the only key concept linking both domains, hierarchy, is not technically and unambiguously defined, which seriously undermines the discussion, particularly for formal and computational linguists.

The issue of modular vs. holistic models of the mind is the object of Chapter 19, by Josef Grodzinsky, who makes an introductory history of the debate (going back to Broca's 1861 influential paper in favor of localization, and the holistic reactions it generated), easy to follow and full of relevant references for the interested reader. Within the context of this debate, Bever's position is identified with a form of holism, as he attempted to derive linguistic generalizations from more general cognitive principles. Grodzinsky, on the other hand, stems from Fodor's work on modularity in order to establish four clear delimitation criteria, which are used to discuss literature holding the claim that language is *not* modular, exemplified by the so-called 'mirror neuron theory', and exactly why and how it is insufficient to account for a number of theoretical and empirical problems. As it is essential for the following argument, the case for modular models of the mind is presented in a very neat and clear way, and all throughout the chapter, several perspectives are discussed (neurocognitive, computational, psycholinguistic, and syntactic) and illustrated by means of reported experiments (focusing on different interpretations of the set of data obtained by Fazio et. al.) and, where relevant, concrete linguistic examples (involving relative clauses and quantifier scope, which is an interesting link with the content of chapters 3, 4, and 5). The author considers several possible counterexamples to the modular theory seriously and in detail, which is a feature to highlight in this article. Neurocognitive evidence is focused on the specific—modular—role of Broca's area in language, and whereas its role can be subsumed to more general cognitive principles or not: Experiment report is once again crucial and it is carried out with the utmost care. The author concludes that Broca aphasia is not directly connected to deficits in sequencing, embedding, or action theory (cf. chapter 18), and even considers (although very firefly) cross-linguistic evidence. A clear perspective of what the Broca's area does and does not do emerges clearly by the end of the chapter, and, despite the reader's own position, the discussion is logically consistent and carefully presented, deserving close scrutiny.

Chapter 20 addresses language acquisition, reporting studies with neonates and very young infants. Jacques Mehler builds on Bever's work on cognitive strategies, and suggests that language learning (along with other human-specific cognitive abilities) should be studied before comparing capacities shared with other species. The paper is devoted to discussing breakthroughs in language acquisition research, from Bever's seminal work to current studies on the relation between frequency and abstraction of word order patterns in pre-lexical infants (thus making it a useful reference piece). Discussion in this chapter is centered on phonology, both segmental and suprasegmental, and how it affected theoretical accounts of acquisition from a psycholinguistic point of view; limited space is devoted to lexical learning, but it is taken advantage of: Several proposals are discussed, including pros and cons of each. Other crucial aspects of language, like hierarchy, are mentioned but only in connection to the phonology/prosody aspects, for instance, a neurocognitive differentiation between audios of regular sentences and backward sentences which suggests structure sensitivity. All aspects are related when necessary, for instance, when considering prosody a cue for lexical comprehension or rhythm as a determining factor for distinguishing different languages even at pre-lexical stages; such relations are clarifying and provide unity to the piece. The last part of the article (section 20.5) is devoted to the role of memory (recalling events) and its relation to language acquisition timing, which apparently makes the former more articulate: Language, it is hypothesized, structures event recalling. The section soon enters again the realm of phonology and the status of phonological representations (including syllabic sensitivity) in the brain of neonates. Relevant references are provided when necessary, covering four decades of research, and the discussion is neat and reader-friendly.

Phonology and the syntax-phonology interface is (also) the subject of Chapter 21, by Ewan Dunbar, Brian Dillon & William Idsardi. They focus on the phonological points made in Bever's contribution and Bayesian probabilistic models of parsimony in phonological description. Bayesian approaches are particularly favorable to Bever's 'analysis by synthesis', insofar as they determine that in order to assess the probability of a hypothesis, some prior probability is specified, which is then updated given new data, while allowing reasoning under uncertainly (as they are not limited by Boolean binary operators). For non-specialists, it would have been useful that the authors explained what a Bayesian probabilistic model is, and, at least briefly, summarize its major points, particularly given the fact that the whole chapter revolts around mathematical tools offered by those models and their use in phonological analyses. In my opinion, a couple of sentences in p. 361 are just not enough to fully understand the forthcoming arguments (even though section 2.3.1 is devoted to probability in linguistics, the basic notions of Bayesian probability are not clearly spelled out). Section 21.2 is devoted to concrete problems of Kalaallisut phonology, addressed from a Bayesian perspective on section 21.3, which also includes more general considerations about the pertinence of Bayesian reasoning in linguistics. The authors tackle the issue of acquisition research as one of looking for an optimal grammatical model with respect to primary linguistic data, but the notion of 'optimality' is not defined or formalized. The article assumes a good deal of mathematical knowledge from the reader, but it does provide clarifications for the formulae employed, even though some notions (e.g., 'stochastic model'), with which some readers might not be familiar, are not explained. This, nevertheless, does not undermine the article's overall intelligibility. When discussing learning algorithms and Bayesian inference (e.g. pp. 368ff.), the authors provide the formula(e) in question, which is a very positive feature of this piece, as it entails complete explication of the mathematical tools assumed in this particular framework. Probabilistic models are here compared to their possibility to account for a certain corpus given so-called *Bayes decision rule*, in tone with early generative conceptions of linguistic theory as providing a *decision procedure for grammars* (Chomsky 1957: 52). Despite the highly theoretical character of the piece, and the proliferation of mathematical formulae (sometimes in detriment of concrete examples), the connection to the main topic is never lost, and the application of the mathematical framework to phonetics and phonology (and procedures to decide between models of acquisition, understood as inference from primary linguistic data) is always stressed.

The final research piece is, quite appropriately, an article by Thomas Bever, who updates and advances considerations made in the initial piece, 43 years later. Bever relates his early claims with the subsequent development of the biolinguistic program (said to have emerged in 1974), and reviews a series of later researched points that stress the relation between language and general cognitive and neurological systems. It is worth mentioning those points here, as the reader will see they are recurrent topics throughout the book (once again, giving it rarely found internal coherence and unity):

(A) Statistical and categorial processes interact, so that initial 'draft' representations of meaning are built, to be replaced with definitive representations (cf. chapter 16).
(B) Sentences mix serial and hierarchical processes (i.e., linearity vs. embedding), and derivations include null terminals (i.e., empty categories).
(C) Language has modular basis (something apparently 'logically necessary', p. 389), and syntax is computationally unique.
(D) The neurological basis of language also differ from other skills, perhaps being related to lateralization.
(E) There are some skills involved in language that have parallels in non-humans, but there remains a core linguistic uniqueness that is only human.
(F) There are no external sources of linguistic universals (e.g., physical laws).

The reader might agree or not (I consider, for instance, that the presentation of arguments against the view that universals might be given by physical laws contains *non sequitur* arguments, and that the possibility is not seriously considered), but it is true that the article advances on the introductory chapter, and in doing so also provides a historical account of what has happened in between. Quite appropriately, Bever presents in section 22.8 two challenges for the future, related to the problems of the 'poverty of stimulus' (in both acquisition and real-time comprehension, very much following the line of the 'analysis by synthesis' approach), and the role of genetic variation (including familiar antecedents) in the neurological representation of language. The article follows a neat *past–present–future* pattern which makes it a structurally coherent piece, and includes a good number of references for each period, including brief discussion of current experiments which will surely be reported in forthcoming publications.

The book closes with an Afterword, by Michael Tanenhaus, focused on the impact of Bever's 1970 article, and its offspring: all the debate it generated and all the alternative visions of language and cognition it stimulated, some compatible with Bever's formulation, some contrary to it. The Afterword summarizes in a very clear manner key points of Bever's legacy, including its impact on Chomskyan generative grammar (unfortunately, without mentioning any alternative framework, e.g., Sag & Wasow 2011) and wider aspects of linguistic inquiry, often not directly connected with Bever's ideas, but stemming from the 1970 paper in one way or another. As a way to project the volume to the future, the Afterword features some of the contributors' impressions about the research paths that could arise, flourish, revive, or fall in the future.

The volume is overall a very valuable contribution to the fields of neurolinguistics, psycholinguistics, theoretical and experimental linguistics, and the broader scientific inquiry about the mutual relations between language, cognition, and brain, and has the potential to become a classic on the topics. The variety of views there expressed, and the focus on interdisciplinary work make this book a very important tool for scholars related to any of the aforementioned disciplines, or curious about how we got here with respect to learning, processing, using, and analyzing language. The structure of the book, in terms of internal coherence, dynamic organization, and multiple recurrent *motifs*, is in itself a welcomed change with respect to other volumes on the topic. Structure and content thus combine to make an excellent state-of-the-art volume, featuring some of the most prominent figures on their respective fields, and trying to advance the field with cutting-edge research as well as valuable and useful historical accounts of the development of psycho and neurolinguistics.

References

Altmann, Gerry & Jelena Mirkovic. 2009. Incrementality and prediction in human sentence processing. *Cognitive Science* 33, 583–609.

Behme, Christina. 2014. Biolinguistic Platonism remains an oxymoron. Ms., http://ling.auf.net/lingbuzz/001765.

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

Broca, Paul. 1861. Sur le principe des localisations cérébrales. *Bulletin de la Société d' Anthropologie* 2, 190–204.

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

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

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

Culicover, Peter & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Gödel, Kurt. 1931 [1986]. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme. In Solomon Feferman, John W. Dawson, Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay & Jean van Heijenoort (eds.), *Kurt Gödel: Collected Works*, vol. I: *Publications 1929–1936*, 144–195. Oxford: Oxford University Press.

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

Hornstein, Norbert. 2003. On control. In Randall Hendrick (ed.), *Minimalist Syntax*, 6-81. Oxford: Blackwell.

Joshi, Aravind K. 1985. How much context-sensitivity is required to provide reasonable structural descriptions: Tree adjoining grammars. In David Dowty, Lauri Karttunen & Arnold Zwicky (eds.) *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, 206–250. Cambridge: Cambridge University Press.

Krivochen, Diego. 2013. The emptiness of the realism–conceptualism debate: Its consequences for (bio)linguistics. Ms., http://ling.auf.net/lingbuzz/001757.

Krivochen, Diego. 2014. On phrase structure building: How your theory of labeling gives away your theory of mind. Ms., https://www.academia.edu/6062796/2014_-_On_Phrase_Structure_building_How_your_theory_of_labeling_gives_away_your_theory_of_mind.

Michaelis, Jens. 2001. On Formal Properties of Minimalist Grammars. Potsdam: Doctoral dissertation, Universität Potsdam.

Russell, Bertrand. 1905. On denoting. *Mind, New Series* 14, 479–493.

Sag, Ivan & Tom Wasow. 2011. Performance–compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.) *Non-Transformational Syntax: Formal and Explicit Models of Grammar*, 359–377. Oxford: Wiley-Blackwell.

Shiebert, Stuart. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Brookline, MA: Microtome Publishing.

Stabler, Edward. 1997. Derivational minimalism. In Christian Retoré (ed.) *Logical Aspects of Computational Linguistics*, 68–95. Berlin: Springer.

Watumull, Jeffrey. 2012. A Turing program for linguistic theory. *Biolinguistics* 6, 222–245.

Watumull, Jeffrey. 2013. Biolinguistics and Platonism: Contradictory or consilient? *Biolinguistics* 7, 301–315.

Wurmbrand, Susi. 2014. The Merge Condition: A syntactic approach to selection. In Peter Kosta, Steven L. Franks, Teodora Radeva-Bork & Lilia Schürcks (eds.), *Minimalism and Beyond: Radicalizing the Interfaces*, 139–177. Amsterdam: John Benjamins.

Diego Gabriel Krivochen
University of Reading
School of Psychology and Clinical Language Sciences
Earley Gate, Whiteknights Road
Reading RG6 6AL
United Kingdom
diegokrivochen@hotmail.com

The Non-Hierarchical Nature of the Chomsky Hierarchy-Driven Artificial-Grammar Learning

Shiro Ojima & Kazuo Okanoya

Recent artificial-grammar learning (AGL) paradigms driven by the Chomsky hierarchy paved the way for direct comparisons between humans and animals in the learning of center embedding ([A[AB]B]). The A^nB^n grammars used by the first generation of such research lacked a crucial property of center embedding, where the pairs of elements are explicitly matched ([A1 [A2 B2] B1]). This type of indexing is implemented in the second-generation A^nB^n grammars. This paper reviews recent studies using such grammars. Against the premises of these studies, we argue that even those newer A^nB^n grammars cannot test the learning of syntactic hierarchy. These studies nonetheless provide detailed information about the conditions under which human adults can learn an A^nB^n grammar with indexing. This knowledge serves to interpret recent animal studies, which make surprising claims about animals' ability to handle center embedding.

Keywords: language evolution; animal cognition; syntactic hierarchy; artificial grammar; center embedding

1. Center Embedding and A^nB^n Grammars

One of the properties that make humans unique among animals is language, which has several components including phonology, lexicon, and syntax. It has been debated how much of each of these components is shared between humans and non-human animals (Markman & Abelev 2004, Yip 2006). The component of syntax, which has been receiving much attention in the field of comparative cognition, instantiates linguistic knowledge describable in terms of a finite set of rules. That set of rules is called a grammar. Fitch & Hauser's (2004) seminal work tried to test which type of grammar non-human primates can learn. In doing so, they resorted to the distinction between a *finite-state grammar* and a *context-free grammar*, based on *the Chomsky hierarchy* (Chomsky 1957). Both these grammars can generate sets of surface strings such as "flying airplanes", but only the latter can generate phrase markers associated with surface strings, being able to differentiate between [$_{VP}$ flying [airplanes]] and [$_{NP}$ [flying] airplanes]. As in these

This study was supported by MEXT Grants-in-Aid for the Scientific Research (No. 23240033 to K.O. and No. 23520757 to S.O.) and ERATO, Japan Science and Technology Agency. We thank Cedric Boeckx, Shigeru Miyagawa, and the anonymous reviewers for comments.

ISSN 1450–3417
Biolinguistics 8: 163–180, 2014
http://www.biolinguistics.eu

examples, natural-language sentences in the mind of a native speaker are hierarchically organized into units of phrases. The inadequacy of a finite-state grammar as a model of human grammar can also be illustrated by sentences with *center embedding* (e.g., The boy [the girl liked] smiled), which can be generated only by a context-free grammar (or more powerful ones) (Chomsky 1957). The notion of center embedding played a major role in the studies discussed below.

To compare humans and animals directly in a semantics-free fashion, Fitch & Hauser (2004) expressed finite-state and context-free grammars as simple, meaningless artificial grammars: a finite-state *$(AB)^n$* grammar, which generated sequences such as ABAB through local transitions (Figure 1a), and a context-free *A^nB^n* grammar, which generated center-embedded, "hierarchical" structures such as A[AB]B (Figure 1b). Because finite-state grammars had been observed in non-human animals (Berwick *et al*. 2011, Fitch & Hauser 2004), the crucial question is whether we can artificially induce, in animals, the learning of a "context-free" A^nB^n grammar equipped with center embedding. This question bears direct relevance to the evolutionary uniqueness of human language and generated a series of artificial-grammar learning (AGL) studies driven by the Chomsky hierarchy.

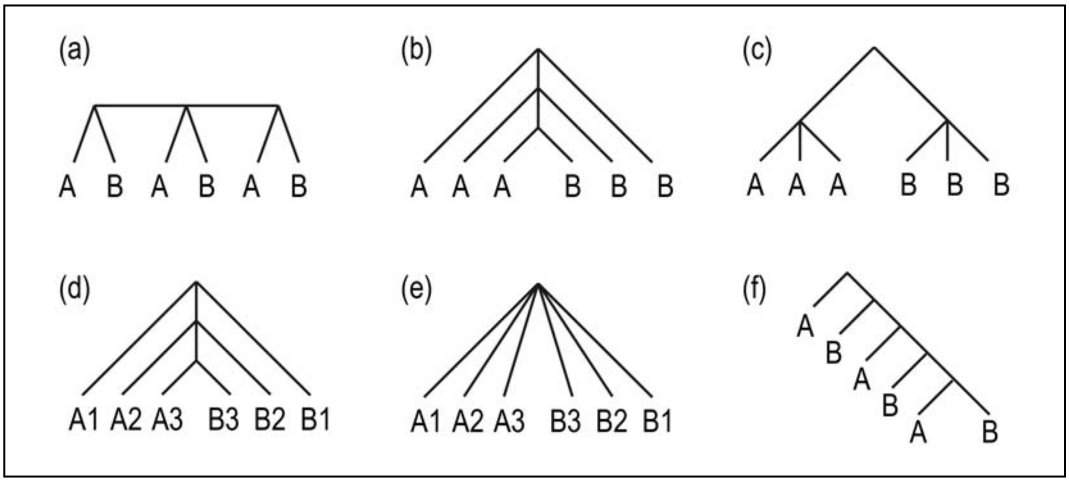


Figure 1: Tree Diagrams Representing Artificial Grammars. (a, b) Original tree diagrams for (a) the $(AB)^n$ "finite-state" grammar and (b) the A^nB^n "context-free" grammar (n = 3 here), used in the first-generation AGL studies. (c) An alternative tree diagram for the first-generation A^nB^n grammar. (d) A tree diagram for the indexed A^nB^n grammar used in the second-generation studies. In this grammar, the pairs of As and Bs are matched from the outer pairs inwards. Numbers (1, 2, 3) attached to As and Bs indicate which A is paired with which B; for example, A1 is paired with B1, not with B2 or B3. (e) An alternative representation of the indexed A^nB^n grammar, in which As and Bs are explicitly paired but no hierarchical information is contained. As in (d), the numbers attached (1, 2, 3) show the unique mapping relations between As and Bs, but nothing is hierarchically higher than anything. (f) The "finite-state" $(AB)^n$ grammar represented as a tail-embedding, hierarchical structure.

The first generation of studies employing $(AB)^n$ and A^nB^n grammars tested a variety of experimental subjects including humans (Bahlmann *et al*. 2006, Fitch & Hauser 2004, Friederici *et al*. 2006), non-human primates (cotton-top tamarins) (Fitch & Hauser 2004), and songbirds (European starlings and zebra finches)

(Gentner *et al*. 2006, van Heijningen *et al*. 2009), and reported striking evidence both for and against the human specificity of center embedding. Neuroimaging studies (Bahlmann *et al*. 2006, Friederici *et al*. 2006) claimed to have dissociated neural correlates of the processing of hierarchical structures (in an A^nB^n grammar) from those related to local transitions (in an $(AB)^n$ grammar).

However, "center-embedded" sequences such as AABB can be interpreted just as As followed by the same number of Bs (Corballis 2007a, 2007b, Perruchet & Rey 2005) (Figure 1c). A violation of this structure can be detected by simply counting the numbers of As and Bs (unequal numbers of As and Bs in an ungrammatical AABA string). Discrimination between "context-free" A^nB^n grammars and "finite-state" $(AB)^n$ grammars can be achieved in similar manners, for example, by counting the transitions between As and Bs (only one A-to-B transition in A^nB^n but multiple transitions in $(AB)^n$). Hence the task assigned to the subject in the first-generation A^nB^n studies could be performed independently of the way the string had been generated by the underlying grammar. The data reported in these studies do not count as evidence either for or against the human specificity of a context-free grammar. More recent, second-generation A^nB^n studies followed a proposal (Corballis 2007a, 2007b, Perruchet & Rey 2005) that the As and Bs in strings generated by an A^nB^n grammar be explicitly matched from the outside pairs inwards (not just [A[AB]B], but [A1[A2 B2]B1]). In the literature (see any of the second-generation A^nB^n studies introduced below), such a relationship is usually represented as in Figure 1d, where elements with the same number are intended to be paired (e.g., A*1* is paired with B*1*, not with B2 or B3). Center-embedded sentences in natural language show this type of pairwise dependencies. For example, native speakers of English would interpret "the boy the girl liked smiled" as having two subject-verb pairs, one (the girl-liked) embedded in the other (the boy-smiled). Hence this sentence not only is in the form of Subject Subject Verb Verb (SSVV) but also contains pairwise dependencies between subjects and verbs (S1 S2 V2 V1), in the minds of those who know the English syntax. An A^nB^n grammar explicitly indexed has been extensively used in the second-generation A^nB^n studies (Abe & Watanabe 2011, Bahlmann *et al*. 2008, de Vries *et al*. 2008, Fedor *et al*. 2012, Lai & Poletiek 2011, Mueller *et al*. 2010). Below we will call this new A^nB^n grammar with indexing an *indexed* A^nB^n grammar for short (but this should not be confused with the index of a context-free grammar (e.g., Salomaa 1969), which has been used in a totally different context).

Unlike the first-generation studies, these new experiments test whether the specific dependencies in the indexed AB pairs have been actually learned. After learning, the subject's sensitivity to grammatical strings having proper AB dependencies (e.g., A1 A2 A3 B3 B2 B1) and ungrammatical ones violating such dependencies (e.g., A1 A2 A3 B3 *B1 B2*) has been tested. Here the strategy of just counting the numbers of As and Bs does not help, because both grammatical and ungrammatical strings have the same number of As and Bs. The implementation of explicit indexing in the A^nB^n grammar has led many authors to assume that the second-generation studies have tested the learning and processing of *syntactic hierarchy* (Bahlmann *et al*. 2008, de Vries *et al*. 2008, Fedor *et al*. 2012, Fitch & Friederici 2012, Friederici *et al*. 2011, Lai & Poletiek 2011, Mueller *et al*. 2010).

2. Hierarchy Is *Not* Involved

Despite the premises of these newer studies, syntactic hierarchy in a strict sense, we argue, has not been learned even in studies using indexed A^nB^n grammars. It is true that an indexed A^nB^n grammar introduces nested pairs and that participants are required to learn and process the dependencies between specific As and Bs. It is a different matter, however, whether humans interpret the strings generated by an indexed A^nB^n grammar as containing syntactic hierarchy. Most of the second-generation A^nB^n studies and a few review articles associate an indexed A^nB^n grammar with the hierarchical structure building of natural language (Abe & Watanabe 2011, Bahlmann *et al*. 2008, de Vries *et al*. 2008, Fitch & Friederici 2012, Friederici *et al*. 2011, Lai & Poletiek 2011, Mueller *et al*. 2010). These papers graph-ically represent the indexed A^nB^n grammar as in Figure 1d. This representation is misleading, in that it gives us an impression that the outer pairs are hierarchically higher than the inner pairs, but such information is not provided during learning as part of familiarization strings and thus cannot be learned. A more accurate re-presentation of the second-generation A^nB^n grammar is in Figure 1e, where infor-mation about pairs is present but information about hierarchy is not. Here, ele-ments with the same number are in a pair (e.g., A*1* is paired with B*1*, not with B2 or B3) as in Figure 1d, but no hierarchy is contained. As long as no hierarchical information is conveyed, the learning of an indexed A^nB^n grammar in the second-generation studies is the learning of center-embedded or nested pairs, but not the learning of hierarchy.

More generally, we cannot make distinctions in hierarchy between "finite-state" $(AB)^n$ grammars and "context-free" A^nB^n grammars, based solely on famili-arization strings. The English sentence "Bob believes Mary came" can be described as noun verb noun verb, or ABAB if A is a noun and B is a verb, but is fully hierarchical in the mind of a native English speaker, who will interpret this sentence as consisting of a higher main clause and a lower embedded clause, as in [Bob believes [Mary came]]. In terms of hierarchy, $(AB)^n$ strings and A^nB^n strings in the studies reviewed here are no different; neither have inherent hierarchical structure and can thus be interpreted either as flat or as hierarchical, depending on lexical items which one imagines inserting. If an AAABBB string is interpreted as having a center-embedded, hierarchical [A[A[AB]B]B] structure, then an ABABAB string can also be interpreted as having a tail-embedded, hierarchical [AB[AB[AB]]] structure (Figure 1f).

In our view, both the first- and the second-generation studies have made the same mistake. The artificial grammar which generated a string is equated with the psychological process involved in the processing of that string, but these two are not the same (Lobina 2011). It is certainly true that hierarchy has been necessary to describe the knowledge of language (language competence); concepts such as *c-command* (Figure 2a) based on syntactic hierarchy have been indispensable for the accounts of many grammatical constructions (Carnie 2006, Chomsky 1981). However, there are examples of non-hierarchical, flat sequences in natural language. Here we take *negation* as an example. Syntactic hierarchy can be easily seen in constructions involving negation. In the sentence "Never dis-agree!", the first negator "never" is hierarchically higher than the second negator "dis–", which is contained in the word "disagree" (Figure 2b). This double negation leads to a (hesitated) affirmative interpretation of the verb "agree" (negative of negative → affirmative). However, we cannot always assume hierarchy of this sort for each negation. For example, it is wrong to do so for the sentence "Never, never say that!". If this sentence had a fully hierarchical representation as in Figure 2c, the first "never" would erroneously negate the rest of the sentence involving the second "never" and would thus lead to an affirmative interpretation of "say that", which is obviously the wrong interpretation. The sentence is singly, not doubly, negative in meaning, and we should postulate a flat representation for the part "never, never" as in Figure 2d.

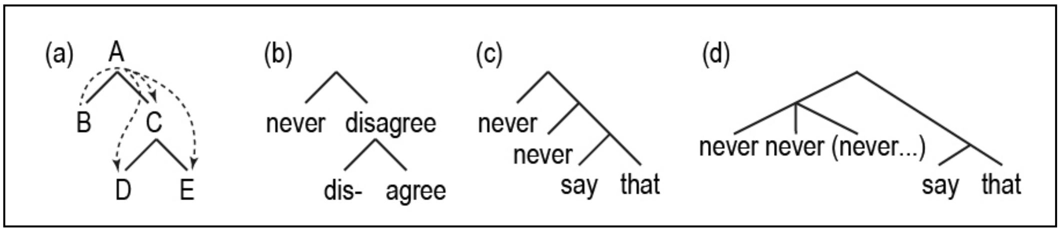


Figure 2: Tree Diagrams Representing Some Grammatical Relations. (a) C-command. B and C c-command each other, while B asymmetrically c-commands D and E (which do not c-command B). (b-d) Interpretation of negation depends on syntactic hierarchy. Double negation (negative of negative) leads to a (hesitated) affirmative meaning, as in (b). However, we cannot always assume hierarchy between two negators; the representation in (c) cannot be correct. We should give "never, never" a non-hierarchical, flat representation as in (d).

For some other constructions, even theoretical linguists did not know for sure (and thus had to debate) whether hierarchy should be assumed. Japanese, which has relatively flexible word order, had once been thought to have non-hierarchical, flat structure in a clause (Hale, 1980, 1982) (Figure 3a).

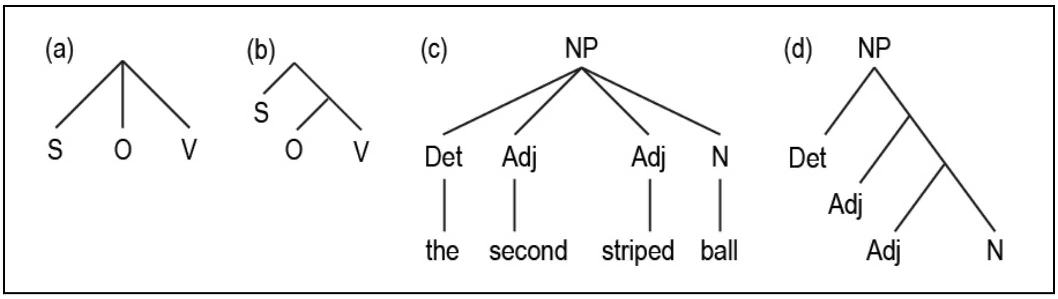


Figure 3: Examples of natural-language constructions, for which theoretical linguists were once divided between hierarchical and non-hierarchical representations. (a) Proposed flat, non-hierarchical representation for subject (S) – object (O) – verb (V) sentences in Japanese. (b) Hierarchical representation for Japanese SOV sentences. (c) Flat noun phrase (NP) proposed for child English speakers. (d) Hierarchical noun phrase.

Later research denied this view and showed that Japanese was as hierarchical as English (Saito & Hoji 1983) (Figure 3b). Also, English-speaking children were once thought to have a non-hierarchical, flat noun phrase (NP) as in Figure 3c. Only a more careful analysis of children's language comprehension revealed that their noun phrase was hierarchical like adults' (Figure 3d) (Crain & Lillo-Martin 1999). English-speaking children form complex interrogatives involving a relative clause in a structure-preserving way as predicted by theoretical accounts of English phrase structure, but this was revealed only through clever experiments (Crain & Nakayama 1987). By analogy, it should not be taken for granted that the subject's mental representations of artificial-grammar sentences are hierarchical; it needs to be demonstrated.

Even if this can be achieved, the actual use of such knowledge required by specific task demands may not depend on the processing of hierarchy. A recent hypothesis questions the involvement of hierarchy in real-time use of natural language, even if its mental representation may still be hierarchical (Frank *et al.* 2012). According to this hypothesis, the involvement of hierarchy must be shown at both the level of mental representations (competence or language knowledge) and the level of real-time processing (performance or knowledge use). The competence/performance distinction is one of the most fundamental concepts in generative linguistics (Chomsky 1965), which is almost exclusively concerned with competence, or the speaker/hearer's internal representation of finite rule sets that generate sentences. Keeping this distinction in mind is not just useful but sometimes necessary, especially where linguists and non-linguists discuss things on a common ground, a primary example of which is AGL studies. The importance of this distinction in experimental studies has been recently reiterated elsewhere (Petersson & Hagoort 2012). In A^nB^n studies, evidence for the involvement of hierarchy in the learning of an A^nB^n grammar has not been provided either at the level of performance or at the level of competence. After all, it has not been studied how the subject processes input strings internally, and we simply cannot know the nature of the internal representations used by the subject in the processing of those strings.

Perhaps those who claim to have studied syntactic hierarchy by using the indexed A^nB^n grammar assume that this grammar automatically introduces the types of hierarchy shown by natural-language sentences conforming to the general pattern of A^nB^n. There are many such sentences, and we can easily see what kind of hierarchy is present in each of them. A typical example would be "John, who Mary liked, smiled.", whose (simplified) tree diagram is shown in Figure 4a. Here the inner sentence "who Mary liked" is attached to the left (to the side of "John"), giving additional information about the subject "John". In a similar sentence, "John, when Mary came, smiled.", that is not the case. As shown in Figure 4b, the inner sentence "when Mary came" is attached to the right (to the side of "smiled") and does not modify the subject "John". If we add another sentence "Bill did *so* too" at the end, it will mean "Bill *smiled when Mary came*, too". This suggests that "when Mary came" is tied to the verb "smiled". One thing we can say here is that in certain natural-language A^2B^2 (more generally A^nB^n) sentences, the inner pair, A2 and B2, is attached either to the left (Figure 4c) or to the right (Figure 4d), but not to the center (Figure 4e) or to nowhere (Figure 4f). The *center*-embedded part is only superficially in the center and is actually attached to just one side, and bears no direct relation to the other.

Even if the direction of attachment is the same, how attachment is done may differ among natural-language sentences. Another typical example of the A^nB^n pattern is found in sentences such as "If either S or S, then S", where S is for Sentence (Chomsky 1957). This sentence has syntactic hierarchy (Figure 4g) that is similar to the one of "John, who Mary liked, smiled". In both these sentences, the inner sentence is attached to the left (to the side of "John" or "if"). However, in the "if" sentence, the inner sentence "either S or S" cannot be deleted ("If then …" is ungrammatical), while in the "John" sentence, the inner sentence "who Mary liked" can be deleted ("John smiled" is grammatical). Here we have the distinction between *complements* and *adjuncts*. "If" must have a complement to stand alone as a syntactic unit and thus requires a sentence. "John" itself is a proper syntactic unit, and we can adjoin something to it but do not have to. Syntactically, complements and adjuncts are in different hierarchical positions (Figure 4h) and are known to behave differently (see Radford 1988 for examples).

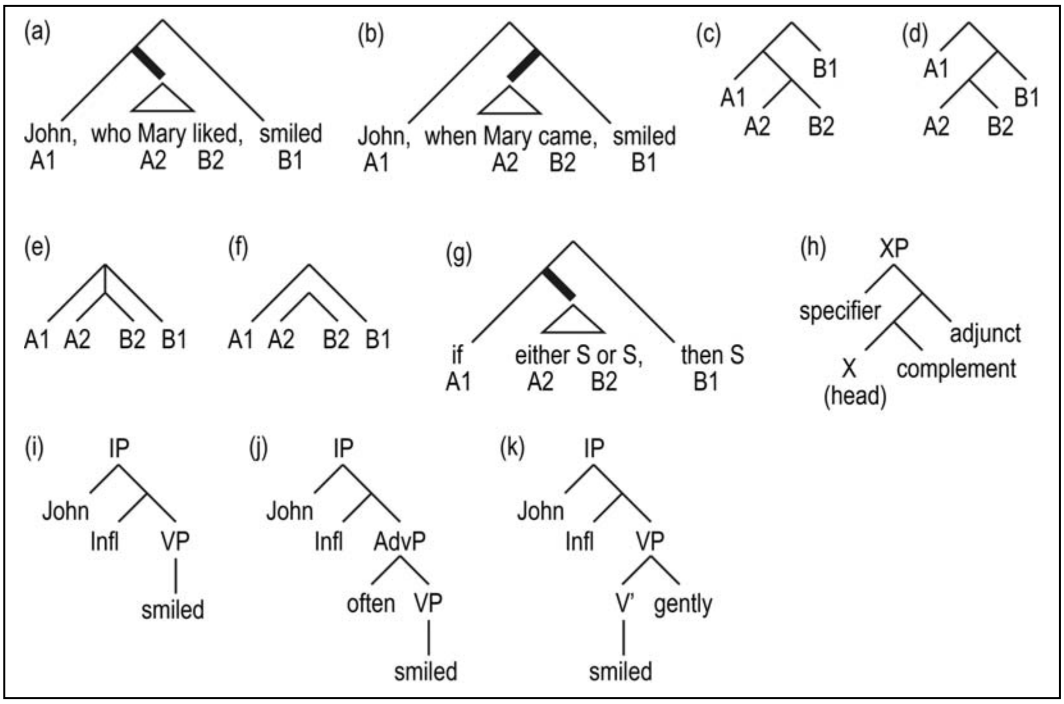


Figure 4: Compatibility between strings and hierarchical representations. In natural-language sentences that conform to the A^nB^n pattern, what is inserted between A and B may actually be attached to the left as in (a) or to the right as in (b). In natural language, the inner AB is attached to the outer A (c) or to the outer B (d), but is not hanging from the center (e) or hanging from nowhere (f). An English sentence containing pairs of "if-then" and "either-or" (g) looks similar in hierarchy to the (a) sentence containing a relative clause, but in (g), the part containing "either-or" is a complement whereas in (a), the relative clause is an adjunct. (h) Complements and adjuncts occupy different hierarchical positions in a phrase. (i) Natural-language sentences conforming to the simple AB pattern may still have hierarchy. Inserting something between A and B (ACB) may create more hierarchy (j), but adding something after B (ABC) may do so, too (k).

It should be clear that the A^nB^n pattern in artificial grammars is compatible with many kinds of hierarchical representations found in natural-language sentences. Just by inserting a pair of AB inside another does not specifically select one of these. In fact, whether and what syntactic hierarchy is created by doing so cannot be known. One may be pleased that at least some hierarchy is created, but just to have syntactic hierarchy of some sort, we do not have to have nested pairs. Just a simple AB pair may be hierarchical in natural language; hierarchy is present even in "John smiled" (Figure 4i). If we put something between A and B ("John *often* smiled"), the sentence may (or may not) have more hierarchy (Figure 4j), but if we put something after the AB pair ("John smiled *gently*"), we may achieve the same thing (Figure 4k). Hence the nesting of AB pairs in an artificial grammar is not special in its compatibility with hierarchical representations of natural-language sentences. Strings without nesting are also compatible with hierarchy present in natural-language sentences.

Some previous AGL research has tested AXB grammars. If, as assumed in the A^nB^n studies, putting a pair of elements (A2 B2) between the two elements of another pair (A1 B1) automatically introduces syntactic hierarchy, then we should equally assume that syntactic hierarchy is present in (and thus can be studied by) a string such as A1 X B1, where X can vary freely while A and B are in a non-adjacent dependency. The A1 A2 B2 B1 string is a special case of this, if the inner pair (A2 B2) is regarded as a unit (X). Artificial AXB grammars have been frequently used to study the learning of non-adjacent dependencies (Newport & Aslin 2004, Newport *et al*. 2004), but have never been claimed to tap syntactic hierarchy. Obviously, X in AXB is not hierarchically lower than A and B, in the absence of explicit evidence that it is. Likewise, the inner pair A2 B2 in A1 A2 B2 B1 is not hierarchically lower than the outer pair A1 B1. In effect, syntactic hierarchy has not been studied in either the first- or the second-generation A^nB^n studies. Some also argue that the Chomsky hierarchy on which the A^nB^n studies are based is not relevant to the neurobiological studies of language at all (Petersson *et al*. 2012).

To sum up, there is no strong evidence that syntactic hierarchy is involved in the learning and processing of either the first-generation un-indexed A^nB^n grammars or the second-generation indexed A^nB^n grammars.

3. The Learnability of Indexed A^nB^n Grammars

The second-generation A^nB^n studies mainly addressed the issue of under what conditions human adults can learn an indexed A^nB^n grammar. The learnability of this grammar revealed recently will inform comparisons between humans and animals in center-embedding learning, but without reference to syntactic hierarchy. The original study which first introduced the second-generation indexed A^nB^n grammar (Perruchet & Rey 2005) reported that the dependencies implemented were impossible even for human adults to learn, if the learning procedure was the same as in the original A^nB^n study (Fitch & Hauser 2004). Inspired by this finding, most of the second-generation A^nB^n studies (Bahlmann *et al*. 2008, de Vries *et al*. 2008, Fedor *et al*. 2012, Lai & Poletiek 2011, Mueller *et al*. 2010) addressed the issue of the learnability of indexed A^nB^n grammars in human adults, and have now managed to describe under which conditions humans succeed or fail in A^nB^n learning. Table 1 (next page) summarizes the key characteristics of 18 experiments which employed an indexed A^nB^n grammar. The results of A^nB^n learning in these experiments are of three types: failure, success, and success possibly aided by a task-taking strategy such as "repetition detection" (de Vries *et al*. 2008). The discussions below exclude the cases where a strategy might have been used.

3.1. *Explicit A^nB^n learning in the visual modality*

As of now, most is known about the *explicit* (as opposed to *implicit*) learning of an indexed A^nB^n grammar in the *visual* (as opposed to *auditory*) modality (Bahlmann *et al*. 2008, de Vries *et al*. 2008, Fedor *et al*. 2012, Lai & Poletiek 2011). The conditions under which human adults' explicit A^nB^n learning in the visual modality tends to be successful include the following: (1) the subject actively searches for rules during familiarization, (2) negative feedback is given about the correctness of the rules the subject found, (3) familiarization strings contain "0-LoE" items and are presented in a "staged" manner, (4) inherent phonological or semantic cues exist between the dependent elements of As and Bs, (5) the level of embedding is one ([A[AB]B]) or two ([A[A[AB]B]B]) (but not three or more), and (6) learning continues for at least 20–30 minutes, and the subject is given 200–300 sentences. Each of these conditions will be discussed in more detail below.

In successful explicit A^nB^n learning in the visual modality, the subject *actively searched for rules* during familiarization phases. Typically, the subject was told that familiarization strings (all grammatical) had been generated by rules, and while those strings were being presented, the subject tried to find those underlying rules.

Negative feedback is provided during rule-testing phases, which are part of learning. During rule-testing phases, the subject can test the correctness of the rules they found during familiarization. Both grammatical and ungrammatical strings are presented, and the subject has to judge each of them for grammaticality. Based on feedback on each judgment, the subject has chances to modify their own rules.

Zero-LoE items (0 level of embedding items) are strings that do not have embedding, that is, simple AB strings (Lai & Poletiek 2011). Zero-LoE items help the subject quickly find out which A is paired with which B. However, for this knowledge to be effective in the induction of the embedding structure of 1-LoE and 2-LoE items, 0-LoE items must be learned first, that is, before 1-LoE and 2-LoE items (Lai & Poletiek 2011). Input that is presented according to the level of embedding (0-LoE → 1-LoE → 2-LoE) is called *staged*. In A^nB^n learning, staged input greatly helps the subject induce the internal structure of complex strings. However, for facilitation to occur, 0-LoE items and staged input must be

Study & Condition	Result	N of A^nB^n	Active rule search	Feed-back	Staged input	0-LoE items	Cues	Exposure length
VISUAL								
Bahlmann 2008	Success*	3	√	√	√	√	√	35 min
de Vries 2008								
Exp. 1, Hier-Scram	Failure	4	√	√				50 min
Exp. 2, Hier-Scram+Rep	Success*	3	√	√			√	50 min
Exp. 2, Hier-Scram	Failure	3	√	√			√	50 min
Lai 2011								
Exp. 1, SS	Success	3	√	√	√	√	√	30 min, 12 blocks
Exp. 1, random	Failure	3	√	√		√	√	30 min, 12 blocks
Exp. 2, SS	Failure	3	√	√	√		√	8 blocks (20 min#)
Exp. 2, random	Failure	3	√	√			√	8 blocks (20 min#)
Fedor 2012								
WS	Success	3	√	√	√	√	√	7.28 blocks
WR	Success	3	√	√	√	√		12.27 blocks
NR1	Success	3	√	√	√	√		16.94 blocks
NR2	Success	3	√	√	√	√		20.25 blocks
Udden 2012, Exp. 2	Success	3						9 sessions (each < 30 min)
AUDITORY								
Perruchet 2005	Failure	3						3 min
Mueller 2010								
C1, no boundary	Failure	2	√				√	12 min
C2, prosody	Success	2	√				√	13 min#
C3, prosody + pause	Success	2	√				√	18 min#
C4, prosody + pause	Success	2	√				√	21 min

Table 1. Characteristics of center-embedding learning in the second-generation A^nB^n studies. Success here is defined as "at least above chance". Success (with an asterisk) means success possibly aided by a strategy. One "block" in column "exposure length" consists of about 10 familiarization strings and about 10 test strings. Abbreviations: 0-LoE = zero level of embedding. Exp. = experiment. Hier-viol = hierarchical-violation. Scram-viol = scrambled-violation. Hier-Viol + Rep = hierarchical-violation + repetition. SS = starting small. WS = words with semantic association paired. WR = words randomly paired. NR1 & 2 = non-words randomly paired. C1 – 4 = condition 1 - 4. # Our estimates.*

combined. If 0-LoE items are presented together with 1- and 2-LoE items from the beginning, facilitation does not occur. Similarly, if input is staged but 0-LoE items are not used (just 1-LoE → 2-LoE), facilitation does not occur, either (Lai & Poletiek 2011). In natural language, input that is staged according to complexity is considered to facilitate the learning of complex structures (Elman 1993). *Starting small*, in the form of staged input or others, may be a natural property of children's first language acquisition (Newport 1990). A theoretical account (Poletiek 2011, Poletiek & Lai 2012) considers the effect of staged input in terms of how much grammatical information is contained in the input strings.

Inherent cues about pairings have been shown to facilitate A^nB^n learning. In many of the second-generation studies, *phonological* cues are provided as to the pairings of elements (Bahlmann *et al*. 2008. de Vries *et al*. 2008, Lai & Poletiek 2011, Mueller *et al*. 2010). An example string would be "de gi ko tu", where "de" and "tu" are paired (outer pair), and "gi" and "ko" are paired (inner pair). The two elements in each pair agree in a phonological feature such as place of articulation (/d/ & /t/, /g/ & /k/). Semantic cues (e.g., semantically related real words such as "you" and "me" paired) greatly facilitate A^nB^n learning (Fedor *et al*. 2012). Facilitation also occurs, to a lesser extent, when real words are randomly paired (e.g., "me" and "lake" for A and B). In the absence of any useful cues, learning occurred to some extent. It is notable, however, that under this condition, 25% of the subjects (normal adults) could not learn pairings, given as many as 400 training sentences (Fedor et al., 2012). Hence to ensure 100% success, some kind of inherent cues about pairings seem to be necessary.

The learning of an A^3B^3 (2-LoE) grammar has been demonstrated, but there is no report on the learning of A^4B^4 (3-LoE), which had been studied in the first-generation studies (Bahlmann *et al*. 2006, Friederici *et al*. 2006). These tendencies may correspond to the limitations on multiple uses of embedding observed in natural language corpora (Karlsson 2007).

Learning continued for at least 20–30 minutes, and 200–300 sentences were presented to the subject. In the earliest A^nB^n studies (Fitch & Hauser 2004, Perruchet & Rey 2005), exposure to the grammar was as short as a few minutes. The learning of an indexed A^nB^n grammar may not be possible in such a short time, even if the other conditions are met.

3.2. *Implicit A^nB^n learning in the visual modality*

At least one study (Udden *et al*. 2012) reports that an indexed A^nB^n grammar presented in the visual modality can be *implicitly* learned. In this experiment (Experiment 2 in the article), most factors that have been reported to facilitate explicit A^nB^n learning in the visual modality are not used. The subject was not engaged in active rule search and was not given negative feedback as to the correctness of their grammaticality judgments (judgments were not done as part of learning). Zero-LoE items were not provided, and input was not presented in a staged manner. Inherent phonological or semantic cues were not present for the AB dependencies. Despite these seemingly disadvantageous features, effects of learning were observed. The secret may lie in the length of learning. The subject went through nine sessions in a period of two weeks. During one session (max.

30 minutes), the subject was shown 100 grammatical strings, which they had to type using a keyboard. In total, 900 strings were presented. This is several times as many as the number of familiarization strings used in the explicit-learning studies. Hence, the implicit learning of an indexed A^nB^n grammar in the visual modality seems to be possible in human adults, given a far larger number of familiarization strings than in explicit learning, even if the facilitative factors already known are not used. We should also note that only this study (Udden *et al*. 2012) used the whole-sentence presentation, where the subject could see the entire sentence on the display, as opposed to successive presentation, employed by the other visual studies, where the sentence was presented in an element-by-element manner.

3.3. *A^nB^n learning in the auditory modality*

The second-generation studies conducted in the auditory modality are a minority, and it is difficult to make a generalization. There may be special effects of sensory modalities (i.e., visual vs. auditory), but this needs to be confirmed by future research. In the auditory modality, only the learning of 1-LoE (i.e., A^2B^2) has been shown (Mueller *et al*. 2010), although in the visual studies, the learning of 2-LoE (i.e., A^3B^3) is reported to be possible (Fedor *et al*. 2012, Lai & Poletiek 2011). This may reflect general difficulty with comprehending embedding in speech streams (Karlsson 2007). Alternatively, methodological differences may be at issue here. In successful A^2B^2 learning in the auditory modality (Mueller *et al*. 2010), input was not staged, and 0-LoE items were not presented. Negative feedback was not used, either, although the subject actively searched for rules in the input. Above-chance learning occurred in conditions where the boundaries of strings (sentences) are marked by prosody or by both prosody and pauses. The artificial grammar in this study utilized phonological cues about pairings. In the auditory modality, center-embedding learning without such cues has not been demonstrated.

As we saw above, the learning of an indexed A^nB^n grammar is possible only under highly specific conditions, even in human adults. When one or more of those conditions are not met, learning becomes difficult or impossible. The findings of the second-generation A^nB^n studies on humans constitute a baseline against which the behavior of non-human animals should be judged.

3.4. *Songbirds*

Currently there are few animal studies on the learning of center embedding in the framework of the second-generation, indexed A^nB^n grammar. To make reliable comparisons between humans and animals, we simply need more research on animals. As we have seen above, much research has already been conducted on humans, and much knowledge about the learnability of an indexed A^nB^n grammar in humans has accumulated. Future research should build on such human research and test animals. That said, we now turn to the few exceptional animal studies that have been published recently.

A songbird species (Bengalese finch) has been claimed to have learned an indexed A^nB^n grammar implicitly and spontaneously (without training or reinforcement), to the level of A^3B^3, via completely passive exposure (Abe & Watanabe 2011). Birds were not trained on ungrammatical strings and were not given positive or negative feedback. Familiarization strings were not given in a staged-input manner. No inherent cues were present in the AB dependencies. Birds were familiarized to grammatical strings only during one session of 60 minutes. In this type of short-exposure paradigm, the learning of A^3B^3 via passive exposure, without negative feedback, without staged input, without inherent cues, has not been demonstrated even in human adults. In fact, humans' learning of A^3B^3 in a meaningless artificial grammar in the auditory modality has not been shown with any learning procedure. Only in a long-exposure paradigm, involving nine sessions of exposure spreading over a period of two weeks, have humans been shown to learn a visual A^3B^3 grammar implicitly (Udden *et al*. 2012).

The claim made in first-generation research that songbirds (starlings) can learn to discriminate grammars with or without center embedding (Gentner *et al*. 2006) merely meant that songbirds, after intensive training, could do something that humans could easily do (without any training, in this particular case). If Bengalese finches can really learn A^3B^3 implicitly and spontaneously in such a short time, this finding can be interpreted as having gone a step further; without any training, birds can do something that humans cannot, or at least have not been proven to be capable of. A close inspection of the test strings used in the Bengalese finch study (Abe & Watanabe 2011) suggested the possibility that the finches behaved according to acoustic similarity among stimuli, rather than grammar (Beckers *et al*. 2012). Methodologically more rigorous research is necessary to precisely describe Bengalese finches' ability to learn center embedding (ten Cate & Okanoya 2012).

3.5. Non-human primates

A recent study reports that non-human primates have a spontaneous tendency to produce center embedding (Rey *et al*. 2102). In contrast to all the other studies above, the subjects in this study, baboons, were not exposed to center-embedding strings at all, and hence did not learn center embedding from external input. They learned pairs of meaningless visual shapes displayed on the monitor. The shapes appeared at random locations, and the baboons were conditioned to touch the correct combinations of shapes in the correct orders (e.g., touch A1 then touch B1). During training, they were required to sequentially touch two shapes at a trial. During test sessions, they were prompted, for the first time, to touch four shapes. For testing, they were shown, for example, A1 first, A2 second, and later, B1 and B2 simultaneously. In this case, they had to touch A1 first and A2 second, but for the latter part they had choices as to which of the stimuli to touch in what order. Specifically, they could choose to touch B2 and then B1 (A1→A2→B2→B1, consistent with center embedding) or B1 and then B2 (A1→A2→B1→B2, not consistent with center embedding). Results show that baboons have a spontaneous tendency to produce more responses which are consistent with center embedding (A1→A2→B2→B1), than those which are not (A1→A2→B1→B2).

This study is special in the second-generation A^nB^n studies, in the sense that it tested whether responses consistent with center embedding are *produced* spontaneously, without conditioning. One might argue that non-human primates' preference to produce center embedding is the evolutionary origin of humans' center embedding, but as of now, no data are available on whether humans show the same preference when put in the same situation. Previous studies on humans have not looked at the issue of center embedding from this perspective. Moreover, it is possible that the preference to put visual shapes in an order consistent with center embedding is not related to center embedding seen in the human grammar at all. In the human grammar, center embedding appears through the interaction of the so-called *head directions* of phrases. Different languages may have different directions for heads (Chomsky 1981, 1986). For example, English is a head-initial language and Japanese is a head-final language. A center-embedding sentence in English like "The boy the girl liked smiled", if directly translated, will not have a center-embedding structure in Japanese (the words will be ordered as in "the girl liked (whom) the boy smiled", to produce the same relative-clause structure with the same meaning). Hence it is one's grammar that determines whether center embedding must be used or cannot be used. Whether the appearance of center embedding in human language has its evolutionary origin in the reported preference of non-human primates to produce center embedding should be supported by further research.

3.6. Comparisons between humans and animals

As we saw above, the second-generation A^nB^n studies on humans as a whole show that the learning of an indexed A^nB^n grammar is very difficult even for human adults and is possible only under specific conditions. It is particularly important to note that this learning is difficult even if humans are *required* to do it. However, both two studies on non-human animals (songbirds and non-human primates) we have just discussed above are in favor of the view that animals also have an ability to handle center embedding. Both those studies provide evidence for this view from animals' spontaneous behavior, without using conditioning or reinforcement. On the one hand, humans' learning of center embedding is difficult even if required. On the other, animals are claimed to have demonstrated center embedding even without being required. This would make more sense if it was exactly the other way around. We have to say that we are in a somewhat odd situation where non-human animals without natural language are claimed to be able to handle a linguistic operation that is difficult even for humans who have natural language. As things currently stand, we are yet to see convergence between animal studies and human studies on the issue of the learnability of an indexed A^nB^n grammar. To move the field forward, each of the two lines of studies should respect the methodological details of the other. Fine methodological details can influence the outcome of AGL (Pena *et al*. 2002). Although the second-generation A^nB^n studies on humans and animals reviewed here all implemented dependencies between As and Bs in a broad sense, humans and animals have not been compared using exactly the same methodologies. Only carefully designed studies can resolve the discrepancy that currently exists between the results of human studies and those of animal studies. The shortage of evidence on animals is also a notable feature of the second-generation studies. Many more studies on animals will be appreciated.

4. Conclusion

The recent, second-generation A^nB^n studies have tested for dependencies in As and Bs, which the older, first-generation A^nB^n studies had not. This led to a currently standard view that the second-generation indexed A^nB^n grammars can be used to test syntactic hierarchy. We argue against this view and claim that syntactic hierarchy cannot be tested with the current experimental setups employed by the A^nB^n studies. These studies offer opportunities to compare humans and animals, within this limitation. The second-generation studies show that the learning of an indexed A^nB^n grammar is fairly difficult even for human adults and is possible under highly specific conditions. This observation is difficult to reconcile with the recent claims that center embedding is observed in non-human animals' spontaneous behavior. Carefully designed comparisons between humans and animals are awaited.

References

Abe, Kentaro & Dai Watanabe. 2011. Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience* 14(8), 1067–1074.

Bahlmann, Jörg, Thomas C. Gunter & Angela D. Friederici. 2006. Hierarchical and linear sequence processing: An electrophysiological exploration of two different grammar types. *Journal of Cognitive Neuroscience* 18(11), 1829–1842.

Bahlmann, Jörg, Richard I. Schubotz & Angela D. Friederici. 2008. Hierarchical artificial grammar processing engages Broca's area. *Neuroimage* 42(2), 525–534.

Beckers, Gabriël J. L., Johan J. Bolhuis, Kazuo Okanoya & Robert C. Berwick. 2012. Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *Neuroreport* 23(3), 139–145.

Berwick, Robert C., Gabriël J. Beckers, Kazuo Okanoya & Johan J. Bolhuis. 2012. A bird's eye view of human language evolution. *Frontiers in Evolutionary Neuroscience* 4, 5.

Berwick, Robert C., Kazuo Okanoya, Gabriel J. L. Beckers & Johan J. Bolhuis. 2011. Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences* 15(3), 113–121.

Carnie, Andrew. 2006. *Syntax: A Generative Introduction*. Oxford: Blackwell Publishers.

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

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
Corballis, Michael C. 2007a. On phrase structure and brain responses: A comment on Bahlmann, Gunter, and Friederici (2006). *Journal of Cognitive Neuroscience* 19(10), 1581–1583.
Corballis, Michael C. 2007b. Recursion, language, and starlings. *Cognitive Science* 31, 697–704.
Crain, Stephen & Diane Lillo-Martin. 1999. *An Introduction to Linguistic Theory and Language Acquisition*. Malden, MA: Blackwell Publishers.
Crain, Stephen & Mineharu Nakayama. 1987. Structure dependence in grammar formation. *Language* 63, 522–543.
de Vries, Meinou H., Padraic Monaghan, Stefan Knecht & Pienie Zwitserlood. 2008. Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition* 107(2), 763–774.
Elman, Jeffrey L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48(1), 71–99.
Fedor, Anna, Máté Varga & Eörs Szathmáry. 2012. Semantics boosts syntax in artificial grammar learning tasks with recursion. *Journal of Experimental Psychology-Learning Memory and Cognition* 38(3), 776–782.
Fitch, W. Tecumseh & Angela D. Friederici. 2012. Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B–Biological Sciences* 367(1598), 1933–1955.
Fitch, W. Tecumseh & Marc D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* 303(5656), 377–380.
Frank, Stefan L., Rens Bod & Morten H. Christiansen. 2012. How hierarchical is language use? *Proceedings of the Royal Society B* 279, 4522–4531.
Friederici, Angela D., Jörg Bahlmann, Roland Friedrich & Michiru Makuuchi. 2011. The neural basis of recursion and complex syntactic hierarchy. *Biolinguistics* 5(1–2), 87–104.
Friederici, Angela D., Jörg Bahlmann, Stefan Heim, Ricarda I. Schubotz & Alfred Anwander. 2006. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America* 103(7), 2458–2463.
Gentner, Timothy Q., Kimberly M. Fenn, Daniel Margoliash & Howard C. Nusbaum. 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440(7088), 1204–1207.
Hale, Ken. 1980. Remarks on Japanese phrase structure: Comments on the papers on Japanese syntax. *MIT Working Papers in Linguistics* 2, 185–203.
Hale, Ken. 1982. Preliminary remarks on configurationality. *Proceedings of NELS [North Eastern Linguistic Society]* 12, 86–96.
Karlsson, Fred. 2007. Constraints on multiple center-embedding of clauses. *Journal of Linguistics* 43(2), 365–392.
Lai, Jun & Fenna H. Poletiek. 2011. The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition* 118(2), 265–273.
Lobina, David J. 2011. Recursion and the competence/performance distinction in

AGL tasks. *Language and Cognitive Processes* 26(10), 1563–1586.
Markman, Ellen M. & Maxim Abelev. 2004. Word learning in dogs? *Trends in Cognitive Sciences* 8(11), 479–481.
Mueller, Jutta L., Jörg Bahlmann & Angela D. Friederici. 2010. Learnability of embedded syntactic structures depends on prosodic cues. *Cognitive Science* 34(2), 338–349.
Newport, Elissa L. 1990. Maturational constraints on language learning. *Cognitive Science* 14, 11–28.
Newport, Elissa L. & Richard N. Aslin. 2004. Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48(2), 127–162.
Newport, Elissa L., Marc D. Hauser, Geertrui Spaepen & Richard N. Aslin. 2004. Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology* 49(2), 85–117.
Pena, Marcela, Luca L. Bonatti, Marina Nespor & Jacques Mehler. 2002. Signal-driven computations in speech processing. *Science* 298, 604–607.
Perruchet, Pierre & Arnaud Rey. 2005. Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin and Review* 12, 307–313.
Petersson, Karl Magnus & Peter Hagoort. 2012. The neurobiology of syntax: beyond string sets. *Philosophical Transactions of the Royal Society B–Biological Sciences* 367, 1971–1983.
Petersson, Karl-Magnus, Vasiliki Folia & Peter Hagoort. 2012. What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language* 120(2), 83–95.
Poletiek, Fenna H. 2011. What in the world makes recursion so easy to learn? A statistical account of the staged input effect on learning a center-embedded structure in artificial grammar learning (AGL). *Biolinguistics* 5(1–2), 36–42.
Poletiek, Fenna H. & Jun Lai. 2012. How semantic biases in simple adjacencies affect learning a complex structure with non-adjacencies in AGL: A statistical account. *Philosophical Transactions of the Royal Society B–Biological Sciences* 367(1598), 2046–2054.
Radford, Andrew. 1988. *Transformational Grammar*. Cambridge: Cambridge University Press.
Rey, Arnaud, Pierre Perruchet & Joël Fagot. 2102. Centre-embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons (*Papio Papio*). *Cognition* 123, 180–184.
Salomaa, Arto. 1969. On the index of a context-free grammar and language. *Information and Control* 14, 474–477.
Saito, Mamoru & Hajime Hoji. 1983. Weak crossover and Move-α in Japanese. *Natural Language and Linguistic Theory* 1, 245–259.
ten Cate, Carel & Kazuo Okanoya. 2012. Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society B-Biological Sciences* 367(1598), 1984–1994.
Udden, Julia, Martin Ingvar, Peter Hagoort & Karl M. Petersson. (2012). Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: Investigating the push-down stack model. *Cognitive Science*

36(6), 1078–1101.

van Heijningen, Caroline A. A., Jos de Visser, Willem Zuidema & Carol ten Cate. 2009. Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences of the United States of America* 106(48), 20538–20543.

Yip, Moira J. 2006. The search for phonology in other species. *Trends in Cognitive Sciences* 10(10), 442–446.

Shiro Ojima
University of Tokyo
Department of Cognitive & Behavioral Sciences
Graduate School of Arts and Sciences
3-8-1 Komaba, Meguro-ku
Tokyo, 1538902
Japan
shiro.ojima@gmail.com

Kazuo Okanoya
University of Tokyo
Department of Cognitive & Behavioral Sciences
Graduate School of Arts and Sciences
3-8-1 Komaba, Meguro-ku
Tokyo, 1538902
Japan
kazuookanoya@gmail.com

Japan Science and Technology Agency
Okanoya Emotional Information Project, ERATO
Nibancho, Chiyoda-ku
Tokyo, 100-0004
Japan

★ FORUM ★

The Interesting Part Is What Is Not Conscious: An Interview with Noam Chomsky

Michael M. Schiffmann

The following is the transcription of an interview held on 23 January 2013 at Chomsky's MIT office in Cambridge, MA, between Michael Schiffmann (MS, in italics) and Noam Chomsky (NC, in regular typeface).

OK, it's January the 23rd, we are at the MIT in Cambridge, Massachusetts. This will be a 60 minute interview with Noam Chomsky on the sixty to sixty-five years of his work, and we will try to cover as many topics as possible. To start off with this, I should put this into a context. I first started to interview Noam Chomsky about this [i.e. the history of generative grammar] two and a half years ago, right here at the MIT, and inadvertently, this grew into a whole series, and today's interview is meant to be the end of the series, but not, hopefully, the end of our talks. [Both laugh.] Well, as I see it, and that's a central part of the research project on your work I'm working on, there are, among many others, several red threads that run through your work, and that would be, first, the quest for simplicity in scientific description, and as we will see, that has several aspects, then the question of abstractness, which we will see in comparison to what went on before and what you started to work with. A closely related question that came to the forefront later was locality, local relations in mental computations. Fourth, the question of biolinguistics, meaning that language can be, and is seen by you, as a biological object in the final analysis, and also, that would be the fifth point, everything you did has always been developed in close collaboration with other people.

So it's not, we are not simply talking about the work of Noam Chomsky, but it's a collaborative effort. Starting in 1946, I remember from my previous interviews that that is actually the period when you got to know who would become your teacher later on, Zellig Harris. And one of the first things you did was to read the galleys for his best-known work, Methods in Structural Linguistics *(Harris 1951). There is another anecdote that I just saw in the morning, when for the first time I saw that Barcelona—I think it was in Spain somewhere in November—talk,*[1] *when you said that another motive, apart from meeting Harris, for going into linguistics, was that you discovered that the Bible, the first words of the Bible had been mistranslated. Can you—maybe that's a good point to start.*

Thanks go to Philadelphia filmmaker Ted Passon, who acted as the sound- and cameraman. Additional thanks go to Chomsky's personal assistant Bev Stohl, who helped to arrange the interview, and to Noam Chomsky himself, who took the trouble to review and correct my transcript of our conversation.

1 Actually, the video, published just one day before the interview, was of a talk Chomsky gave at Princeton University on the invitation of Robert Freidin. See http://www.youtube.com/watch?v=Rgd8BnZ2-iw&list=PL5affof_lboya5kn0tI7YbcQJTE5oX9AS.

ISSN 1450–3417

Biolinguistics 8: 181–203, 2014
http://www.biolinguistics.eu

[Chuckles] Actually, I read the galleys before I had started taking linguistic courses. I had just met Harris, we talked about the field, and he said, you want to learn about the field, why don't you proofread my book.

Nice method!

So I proofread the galleys for him. I think it's probably mentioned in the acknowledgements,[2] but that was my introduction to the field. This was different. I was studying Arabic; that was one of the few fields that I was interested in, and the professor was a very distinguished Arabist and also a wonderful human being. He was an Italian, Giorgio Levi de la Vida, he was an anti-Fascist émigré.[3] We got to know each other pretty well later, but he pointed out to me something, just in conversation, something about Hebrew—I knew Hebrew reasonably well, and knew the Bible, he pointed out to me—I forget the context—that the first few words of the Bible were misvocalized, you know, they were—the original text of the Bible, or the texts, came down without [vowels], it had just consonants. Hebrew, you know, and Arabic, you know, are missing the vowels; they're extra. The vowels were put in about the eighth century by the Masoretes,[4] and they just made a mistake in putting in the vowels. And the phrase that appears is completely ungrammatical. And the translations are wrong.

Is the meaning also—

The meaning, too—I mean, it doesn't mean anything, literally, but it's kind of been reinterpreted so that it means something. And it's not—if you get it correct, it doesn't change the meaning enormously, but I was struck by the fact that for twelve centuries at that point, the first words of something that everybody knows were mistranslated and misvocalized, and nobody had noticed it! And that struck me as meaning: Well, maybe there is something interesting in this field that you want to figure out! [Laughs.]

Right! Well, I guess the morale would be, there is—we think we know so much, but actually, so little is known.

Yeah, we don't pay attention to what we don't know.

Exactly. OK, so you started with Harris, and we covered this a lot, the basic idea really was at the time, in structural linguistics, not only with Harris, to look at things that were supposed to be 'there' in the world and to classify them.

Yes, you study the corpus, you study the corpus of material, when you're a student, let's say in the undergraduate stage, in the late forties when I was a student, you were taught field methods, how to take information from an informant. So it would

2 It indeed is. See Harris (1951: v).

3 Giorgio Levi Della Vida (1886–1967) was a Jewish linguist active in the anti-Fascist movement in his native Italy. After the introduction of anti-Jewish race laws in Italy in 1939, he emigrated to the United States, from where he returned to Italy only in the late 1940s. His main areas were Semitic languages, in particular Hebrew and Arabic. He also worked in the realm of Near-Eastern history and culture.

4 The Masoretes were Jewish scribes between the 5th and 10th century and, among other things, devised a notational system for putting in the vowels missing in the original text of the Bible.

be typically a Native American speaker of Choctaw or something, and you have methods for getting information and data from him about the language. And then you organize that data. And that's a grammar, you know.

You go from phone to phoneme, and on to morpheme, and on to syntax—

You have a rigorous series of steps, you can't mix levels, you can't use syntactic information to analyze the sounds—

Yes, which you later on did in 1956, in that article—[5]

Well, in *1946*, because as soon as I started working myself, I saw this doesn't work, you know.

OK. But you say you immediately saw there is something fishy about that, though that would begin the phase in your life, in your linguistic life, that you have later on called schizophrenic.[6] *On the one hand, you tried these methods—*

Well, I assumed—you know, Harris was an extremely impressive person, did a lot of interesting work and so on, I just assumed it's gotta be right. Thus one part of my life for several years was trying to fix up the procedures and analyses that wouldn't work, because when you looked at them closely, they didn't really work and so on, so I tried to sharpen them up, and overcome difficulties, and I continued with that until about 1953. On the other hand, there was something else, which was kind of like a private hobby. I mean what happened is that – we didn't really study much linguistics as undergraduates, because Harris, Harris's view was that the field was basically over. The *Methods* showed, and a lot of linguists believed this, once we have the methods of analysis, it's all just routine from then on.

He had already begun the level, begun to move beyond the level of sentences into discourse—

The only things we studied—like, there aren't, there weren't—it was a funny department: I mean, there were a few students who were close to Harris; we met separately, and sometimes in his apartment in New York or something like that, and the things that were discussed were what became discourse analysis, he published a couple of articles on it in the early fifties,[7] an effort to apply the methods of linguistics, assumed to be known, to more complex discourses. Essays, articles, chemical papers for the purpose of abstracting, and so on and so

---

5 See Chomsky *et al.* (1956).

6 In conversations with the French linguist Mitsou Ronat in January 1976, Chomsky described his situation between 1947 and 1953 in the following way:

> I did not think myself that I was doing linguistics. In a sense I was completely schizophrenic at that time. I still thought that the approach of American structural linguistics was essentially right. As I told you, I spend a great deal of time trying to improve and to formalize discovery procedures, in order to overcome their obvious defects. But once they were made precise, they led manifestly to the wrong results. Still, for quite a long time I thought that the mistake was mine, due to wrong formulations. In 1953 I published an article in the *Journal of Symbolic Logic* in which I tried to develop a discovery procedure that I hoped might be the basis for something that would really work. That, for me, was real linguistics. Chomsky (1979: 130–131)

7 See Harris (1952a, 1952b).

forth. That's basically all we studied. A lot of it was political. Stuff I could tell you about.

Yeah, I know.

Because that was the real interest that brought a lot of us together, but so, efforts to detect ideological factors by looking at texts from a structural point of view, things like that. There was a—I had to do an undergraduate thesis, an honor's thesis, you know —

And that undergraduate thesis had nothing to do with that, it was different.

No, it was linguistics. And Harris suggested as a topic that I, since I knew it, that I do a structural analysis of Modern Hebrew, which I kind of—I wasn't fluent, but I knew it pretty well, and so, OK, so I went ahead and did what I was supposed to do: I found an informant, a native speaker and went through the procedures, and then I got the phonetics, and on and on. After a couple of weeks of this, I realized that this is totally ridiculous. The only thing that I'm learning is the phonetics, which I don't care about. [For] everything else, I know the answers he is going to give me. Why am I bothering with the questions?

I see.

So I dropped the informant altogether, and just started doing what looked like it made sense, which was in fact a generative grammar. So let's just, let's construct a generative system, which will determine the nature of the expressions of the language. I don't need the informant for this, I can —

Can we briefly move into the syntactic component of this, because there were like twelve pages of syntax in that first work of yours.[8] *And it was highly formal, as you several times mentioned, well, probably nobody ever looked at it because nobody understood it—*

One person looked at it.

Yes, Henry Hoenigswald.

We talked about this.[9]

Yes, and part of that formalization was that it had recursion in it.

This it had to automatically, because it's sort of like a truism that language is unbounded. And it's kind of interesting that later, when I studied the history of the field, I discovered that this had barely ever been mentioned. I can run through cases—

I mean, of course it's kind of obvious that it's recursive, but nobody ever mentioned it—

Nobody mentioned it was infinite! I mean, few people. Scattered cases where it

---

8 This refers to the final, published version of what started as Chomsky's 1949 BA thesis. An extended and revised version of that work served as his 1951 MA thesis. This thesis was once more revised in the fall of 1951 after discussions between Chomsky and his friend Yehoshua Bar-Hillel, revisions that were completed in December 1951. This is the version that was finally published almost three decades later as *Morphophonemics of Modern Hebrew* (Chomsky 1951/1979), from now on, *MMH*.

9 See Chomsky (2010a: 6).

was. They're interesting.

Yes. And here it was right in the first rule: A sentence is something which can be rewritten as something that also contains a sentence, and it goes on from there.[10] *That would be the direct recursive rule, where you have the S on the left side of the arrow and —*

Well, the *system* is recursive; there is no particular rule that's necessarily recursive. The *set* of rules gives a recursive function. It's analogous to — I should say at the same time I was studying recursive function theory and mathematics and so on, and it sort of all fit together. That provided the formal understanding of what a recursive system is.

Exactly, I mean the progress that was made at the time, that was that things were laid out in the open, that had kind of been lingering in the background of people's mind, but never [been] formulated clearly.

Well, you know, Euclidian geometry was kind of a computational system. It didn't work exactly, in fact it wasn't even formalized until the turn of the 19th century. But then the whole concept of algorithms, computations, and so on, that had really been, by the 1940s, that was pretty well understood with, very well understood with Turing, and Church, and Kleene, and other great logicians.

So in Morphophonemics*, you had this expressly recursive system, with the S on one side of the arrow, and also on the other side, and then you had NPs that could contain PPs, and the PPs in turn could contain NPs —* [11]

Yeah.

— that would be another example for recursion, but compared to what you later did, this was not transformational syntax.

No, this was pre-transformational, yeah.

So the transformational work started around here, in Cambridge, right?

In the early 50s.

---

10 The first syntactic rule in *MMH* "S1. Sentence → Elementary sentence <Connective + Sentence>", accompanied by the comment: "S1 must be reapplied until 'sentence' is eliminated" (p. 12). In the conversation between NC and MS, 'S' is used for 'Sentence'.

11 As it stands, this is not really correct, at least for the presentation in *MMH*. Since the syntactic component of Hebrew in this work is only a sketch, and outlined only insofar as it is necessary as input for the morphology, it contains no 'indirect' recursion (i.e. recursion resulting from more than a single rule), as could, for example, be exemplified with the set {PP → P + NP, NP → N + PP, NP →N} of rules whose repeated application could generate strings such as *[for [$_{NP}$ prisoners [$_{PP}$ in [$_{NP}$ prisons [$_{PP}$ within [NP [$_{N}$ prisons]]]]]]]*. Omitting irrelevant detail, rule S5 in *MMH* can be stripped down to PP → P + NP, but the following rule (actually a collection of rules collapsed into one) S6 expanding NPs does not contain the second step necessary for indirect recursion, namely NP → N + PP or some equivalent. As far as I can see, there are also no other syntactic rules in *MMH* that would yield indirect recursion. This system thus only provides for direct recursion, and what is more, the only rule doing so is the one mentioned in the previous note, where the symbol 'Sentence' appears on both sides of the arrow. Also underlining the sketchy character of the presentation is the fact that the only 'Connectives' spelt out in rule S2 are *coordinating* conjunctions, so that even sentence recursion is limited to coordination. The transformations introduced in the *LSLT*/*SS* system (Chomsky 1955, 1957) kicked recursion out of the phrase structure component of the grammar entirely, at least until it was reintroduced in *Aspects* (Chomsky 1965).

When you came to Harvard, as a fellow.

Yeah. Incidentally, the syntax was pretty rudimentary. That work you're talking about was about 95% about the mapping from the syntactic structures to the sounds.

That's pretty clear, and was in fact laid out in the introduction to the booklet: "I'm doing this just to give a base to do the other things." So, one of the things that I remember you telling me that you did was around 1952 when you worked with that famous auxiliary transformation, with the English auxiliary system. Can you give a brief description?

Well, I was struck by the fact—in the descriptive grammars of English, they ran through the paradigms of the auxiliary system, you know: '*eat*', '*was eaten*', '*has eaten*', '*will have eaten*', and so on and so forth. The forms were just listed, and that was it. But if you looked at the forms, there were *clearly* some *patterns* behind them, like when you have '*have*', it was normally followed by something which ended with an '*–en*' element or some other morphological component; '*be*' was going to be followed by '*–ing*', and so on, and they appeared in certain ways. And there was a very simple way to describe all of this, namely, just to say that, the way it was described at that point was: The real element is not '*be*' and then a separate form '*reading*', but '*be–ing*', and '*read*', and then the '*–ing*' moves over and attaches to the '*read*'—or the other way around! But that's a transformation—

And in that case one element determines the form of the next, in that the affix hops from that element on to the next one…

Yeah, what was later on called 'affix hop', years later, by Haj Ross.[12] But—

He had a gift for those formulations.

He liked those terms, yeah. A lot of his terms were—I didn't actually like them, I thought they were kind of frivolous, frankly, [laughs] but I use them—

Like 'island' and so on…

—but now it's called 'affix hop'.

OK.

But—not my term. But the point is, it wasn't a phrase structure rule. And then, when you begin to think about it, if you take, say, "the sandwich was eaten", not only you have this affix hop, what you also have the subject not in the position where it gets its semantic role. Well, and if you think about that, that should be another transformation. Now, Harris's descriptions had—at that time, he was working on co-occurrence, and he tried to, he gave a descriptive argument to try

12 Haj Ross, born in 1938, studied at and got his Ph.D. from MIT (Ross 1967). Despite its great significance for linguistic theorizing, it was only published 20 years later (Ross 1986). Ross has worked, and continues to work, on a huge variety of linguistic phenomena. He was part of the short-lived linguistic current called 'Generative Semantics', and, using a metaphorically very gifted language, he has coined a number of technical terms such as 'scrambling' (the shuffling around of syntactic constituents from their original position to other positions not canonically designed as end points of such shuffling), 'islands' (syntactic structures that constituents are more or less unable to move out of), 'pied-piping' (constructions in which a question word such as *whose* 'carries along' the phrase which it is a part of), and more.

to show that passives had co-occurrence relations with actives, but it didn't really work quite properly. And I thought then, as I say, for a while, for a long time I tried to work it out, but then I realized that this is not the right way to do it. The right way is that there is just an *abstract* underlying structure, and there are various kinds of rules applied to it, one of them being the passive.

And that's the really important thing, you have something abstract from which what you see, finally, or what you hear, is derived, and that's basically the difference between transformations in your work and Harris's.

Yes, and that means that you also abandon the separation of levels. And there was a lot of artificiality that came along with the separation of levels, I mean simple things, like take the words '*writer*' and '*rider*' in English.[13]

The famous example, yes.

Famous example. For structural linguists, who *did work* on this, phonetically speaking, say, in Standard American English, the phonetic difference between '*writer*' and '*rider*' is in the vowel.[14] The medial consonant is identical for the two. And since you're not allowed to mix levels, the description—that's the way the *description* was—intuitively that's obviously crazy. There is a word '*write*', and there is a word '*ride*', and the vowel change depends on the consonant, and then the consonant is neutralized.[15]

And once more the truth of the matter comes from things that you cannot see, because what you see and hear is derived from some computation that's—

Well, the point is you weren't allowed to give this description, because it violated the separation of levels. You had to know the morpheme—

Yes.

Harris had various gimmicks just to try to get around this, but they didn't work, and my feeling was: Let's do it a way that works, that's the way I *had* done it with Hebrew with just ordered rules.

And maybe the most extreme example maybe of mixing of levels would be, on the one hand, you have syntactic structure, and then you have things like contour and stress assignment—

Yeah.

So you have the highest structure and the lowest structure, and then [gestures up and down to connect the two].

That's something that Morris Halle and I worked on in, well, around 1953, '54, and we published an article in '55. One of the major topics at that time in American structural linguistics was pitch and stress.

---

13 See, e.g., Chomsky (2010a: 4), Halle (2010: 18, fn. 1), and sources quoted there.

14 In (one) phonetic notation, the pronunciation is [rʌɪDər] and [ra:ɪDər], respectively. The difference lies in the first vowel, not, as intuition and the spelling of the items might lead one to assume, in the medial consonant.

15 See previous note.

Yes, Trager and Smith.[16]

And there was a system, which was supposed to be an overall system for English, which had four stresses, four pitches, and then you just describe things within that system. But Morris made an interesting observation to me when we were working together. He pointed out that Roman Jakobson makes every imaginable mistake when, in the phonetics of English, but he never gets the pitch and stress contours wrong.

Yes, that's the one thing he gets right.

So we figured: Look, there's gotta be some rule system behind this, and when we took a look at it, it turned out that you can predict pretty complex stress and pitch contours just by regarding them as the reflexes of the syntactic structure with cyclic rules that generate bigger things from smaller things. Of course all of that is totally outside the framework of procedural analysis, because you're using abstract syntactic structure to derive phonetic facts, so it was going to be kind of like incomprehensible in that structuralist framework. But it looked right, and we, Morris and I, then pursued this for years, and finally, in *Sound Pattern of English*, so it's a long discussion, and it goes on from there. But that was basically the idea, that it's in the morpho[pho]nemics—incidentally, I later learned, only much later, that there is a historical tradition, which goes back to classical India, the Paninian tradition, I may have mentioned this before—[17]

Yes, you did.

Yeah, OK.

I have a very little question about the auxiliary system as you described it, and it's really a super-technical question, but, when I with a group of students read Syntactic Structures,[18] *we had this rule for the auxiliary system: C, M, then* 'have + en', *then* 'be + ing'*—what does the 'C' stand for? Where does that 'C' come from? Is it 'Concord' or something like that, do you remember that?*

[Laughs] Oh, that's sixty years ago…

[Laughs] I asked people at home to give me questions along, and I thought this was a good opportunity to nail this one down.

Absolutely. I'm really sorry, I would have to recover pretty old memories to find out what was 'C'. It's not the 'C' that later became standard, the 'Complementizer'.

Of course, absolutely not, that's pretty clear.

It had to do somehow with tense, but I forgot why I picked 'C'…

---

16 Henry Lee Smith and George L. Trager were important structuralist linguists and authors (jointly and individually) of works on the structure of English. In the 1950s, their theory about intonation, and stress assignment was dominant in the US. Chomsky discussed this particular issue as well as phonology in general and even more general issues with Smith at the *Third Texas Conference on Problems of Linguistic Analysis*, May 9–12, 1958 (Hill 1962). For some additional remarks on this, see Chomsky (2010b: 2).

17 See Chomsky (2010b: 19-20) and Chomsky 2010a).

18 See Chomsky (1957).

The only thing that sprang to my mind was 'Concord'—

Well, maybe it was 'Concord', but it was basically the tense morpheme.[19]

Yes, I see. Let me get to the biolinguistic question at the time. The Harrisian tradition was not realist; the idea was not what we are describing here as truth about the world or the language or whatever—when did you start, when did you start and when did Morris start and when did your friend Eric Lenneberg start to develop the different conception? Would that be around that time, about 1951—

Well, the three of us met in 1951. We were all—Morris was here, but Eric and I, we were grad students together, and we met pretty quickly and kind of hit it off, we were all friends and we started thinking, talking about these topics. None of us—one thing we had in common was, none of us believed in structural linguistics, or anything that was going on. But—you have to remember this was a period when strict behaviorism was absolute dogma, in philosophy, in psychology, and we didn't believe in that either.

So right from the start.

Right from the start. So we started reading, you know, together, a lot of the European ethological literature. Morris and Eric could read it in the original, it's German, and I read it in translation, but it was just excluded from American psychology. Almost, I mean there was—you could find pieces in *Comparative Zoology*,[20] in *Comparative Psychology*,[21] but not mainstream psychology. I mean, there are some obvious points about language, I mean, it's just like about as close to a truism as you can get: A language, say, my language, is a property of *me* [points to himself], an individual! It's not some abstract thing out in the universe somewhere.

Well, there are still people who claim that.

I mean, people claim that, it's so crazy I won't talk about it, I mean, this is kind of like as obvious as you can get. Your language is a property of *you*, you're a biological creature, so it's a property of your biology.

If somebody knocks me on the head hard enough, it's gone.

Yes, and in fact, well, furthermore, it's a property mostly of your brain, because when your foot's cut off, you don't lose your language, if your head is cut off, you do lose your language. But that's something that's too obvious to discuss. So it's a biological property of an individual—now, you *can* study language in a different way if you chose.

Right.

---

19 In his final editing, NC adds: "Actually, I suspect it might have been 'component', borrowed from the Harrisian concept *long component*—an inflectional element that spreads over a sequence of elements" (May 2014).

20 The *Zoologischer Anzeiger: A Journal of Comparative Zoology*, one of the oldest German zoological journals, is a scientific quarterly that has appeared since 1878; it is now published by Elsevier.

21 The *Journal of Comparative Psychology* has been published by the American Psychological Association since 1921.

But at least one way of studying it, and in fact a central way, is as a biological property of an individual. Well, and let's think of it as a biological property, what do biological systems have, like the visual system? They have a genetic basis, they grow in particular ways, they're used in particular ways, they evolve, but that was way too far off to even think about… And that's *biolinguistics*!

And that would mean that something that you mentioned often, a quote from Martin Joos, that languages vary without any limits,[22] *can't be right.*

It can't be right, because there would be no way to *acquire* any of them. And I quote him, because he said it clearly, but it was everyone's view. Or essentially without limits, let us assume there are some constraints—

What occurred to me recently was that behaviorism in other realms said exactly the same thing! For example when, I believe it was Watson, one of the co-founders, who said, "Give me some child, and I have my behaviorist methods, and I'll make a criminal or a saint, I can do what[ever]"—

Yes, but it was the same actually in theoretical biology. It was, not really by everyone, but it was generally assumed, I've quoted some of this, too, that organisms can vary—

—without limits—

Almost without limits. You know, when people say, virtually infinitely, which means, just about anything. It's interesting that in biology over the years it's been recognized that in fact, the limits are very narrow.

Yeah. As you recently wrote or quoted, it's just variations of one organism since the Cambrian explosion.[23]

Well that's actually been suggested seriously, as a proposal—

Right. There are other things that you can clearly see, and you mentioned that in another article, for example, the kind of surprising fact that animals rarely use metals in any significant sense, and animals don't have wheels, even though it might be helpful to have ones![24]

[Both chuckle.] Yeah. Somebody pointed it out to me once, a biologist, but [there's] apparently a certain kind of bacteria that have something like wheels —

OK! [Again, laughter.]

Not to get around, just—it's things that spin, you know…

But the basic point really is, it can't be limitless.

---

22 See Joos (1966: 96), citing the 'Boas Tradition', quoted (from an earlier edition) in Chomsky (1968/2005: 68, n. 12) and many of Chomsky's speeches and writings subsequent to the first edition of this work.

23 See Berwick & Chomsky (2011), where the biochemist Michael Sherman is quoted with the suggestion that a "Universal Genome that encodes all major developmental programs essential for various phyla of Metazoa emerged in a unicellular or a primitive multicellular organism shortly before the Cambrian period" (p. 23).

24 Actually, two different sources, namely Chomsky (2002: 103) for the use of metals and Berwick & Chomsky (2011: 22) for the impossibility of using wheels.

You'll do with, you'll use the tools that are available to you. Actually, the most striking—things like this affix hop that you mentioned, it was only really one example. I mean, as soon as you started looking at, trying to actually generate the properties of the language, the actual properties, you quickly found that there are some very puzzling phenomena, things were one way, and not some other way, in which they *could* work. And this was never even noticed, you know. One of the simplest cases, which is still being debated, hotly debated—

I know, yes—

—is structure dependence!

Sure, I know what you are getting at.

Well, that's the kind—

*Structure dependence by the way also gives in my view, structure dependence in my view also gives a very interesting example for locality. Because if you have something like the sentences you quoted in your work in the fifties and also together with George Miller, with multiple dependencies, you have an '*if*' here in the sentence, and in principle, the '*then*' that comes after it might be a million words after it.*[25] *And once you move from linear order to hierarchical structure, it's immediately adjacent, so the relationship becomes local, pure and simple, like in these examples, "Is the man who is in the garden happy?".*

It's a little bit like affix hop, really if they're in a single unit, and then they separate.[26]

Right.

But the structure dependence is much deeper than that. The examples that have usually been used in the literature involve movement, like auxiliary inversion, and that has misled a lot of people. There are all kinds of pointless efforts to find other ways of describing auxiliary inversion, but the very simple examples that don't have any of these properties, like just take adverb interpretation, so suppose you have some simple sentence, one that I've used is "Eagles that fly swim."[27] And suppose you put '*instinctively*' in front of it. Which verb does '*instinctively*' go with? "Eagles that fly swim."

Yeah, that's absolutely clear from the start.

It's perfectly obvious that it's '*swim*', but *why*? I mean, the relation to '*fly*', first of

25 One example would be (Chomsky & Miller 1963: 286):

(i) Anyone$_1$ who feels that if$_2$ so-many$_3$ more$_4$ students$_5$ whom we$_6$ haven't$_6$ actually admitted are$_5$ sitting in on the course than$_4$ ones we have that$_3$ the room had to be changed, then$_2$ probably auditors will have to be excluded, is$_1$ likely to agree that the curriculum needs revision.

26 On the abstract level of syntactic hierarchy, the elements of the previous note (*anyone* – *is*, *if* – *then*, etc.) can be regarded as nothing but (the core of) branches of single units, and the same is true for the subject NP *the man who is in the garden* and the predicate *is happy* with which it agrees. Similar to affixes in affix hopping, the copula verb *be* then separates, in this case by moving to the front of the sentence.

27 For one of the by now many elaborations on this, see Katz (2012), available on the internet, where Chomsky also uses the '*instinctively*'-example discussed immediately below.

all, is more natural; they do fly instinctively, but also, it's local, it's a local—

In the interpretation you even go against the pragmatics of the expression.

Against the pragmatics, and also against *ease of computation*!

Yes.

Because the local relation, the one that's easy to compute, is with the first verb! But you don't, you find—it's also a locality relation, but it's *structural* locality. It's the structurally closest one, not the linearly closest one. And that's much harder to compute!

And you go against this, because language happens to be set up this way—

Well, the real answer to this, I think, which is just becoming clearer now, is that *order* doesn't *exist* in the internal computation. So, like when a child—it's just not available, which has a lot of consequences, we could talk about it. But if order is part of the externalization, it's kind of a reflex of the sensorimotor system. It's outside, which implies that other semantic and syntactic properties also shouldn't depend on order. And there's pretty good evidence [that] that's true; they depend on hierarchy, not on order. Order seems to be extrinsic to language. And from that a lot of other things follow.

Right, we talked about that in Cologne—[28]

We did.

—for example, that islands and a lot of other, similar stuff, that you can, that in principle certain thoughts are fine, but you cannot express them in the externalized language—that that follows from that.

Well another thing—yes. All of this converges on—has a lot of syntactic and semantic consequences, but also much more general ones. Because if order is extrinsic to language, it then follows that every *use* of external language is also extrinsic, in particular the use of language for communication, which has got to be secondary, contrary to dogma, another dogma. Anyhow a lot of these things we were kind of talking about—I'm going back to the fifties—we were kind of playing with, but Eric, he finished his linguistics and went off to medical school and ended up pretty much founding the field of biology of language, and we continued to work with each other for some years.

I was wondering about a point we talked about in Cologne, because externalization prevents you from expressing certain thoughts, and as long as you talk to yourself internally, you are still using the externalized language, and that would mean that in that context it would be the same.[29] *What about unconscious thoughts? Would that remove the barriers to these thoughts?*

Well, when we do what we call 'talking to ourselves'—this is not investigated, but it could be investigated and it should be, but if you just introspect about it

---

[28] To be precise, in the Cologne interview the issue was not order as such, but the island phenomena mentioned by MS in the next few sentences. See Chomsky (2011).

[29] *Ibid.*

carefully, when you think you're talking to yourself, what's actually happening is [that] fragments are passing through your consciousness, not sentences.

Yes.

What's hitting consciousness is bits and pieces, and then you can formulate an expression, using them, and that's talking to yourself, but the real thinking is beyond the level of consciousness. It's something that's going on in there using this whole system; occasionally, bits and pieces of it hit the level of consciousness, and that's what people study, what's conscious, but the interesting part is what's *not* conscious. In fact, there's other information—there are other studies now showing something similar. For example, it's been demonstrated that if you carry out an action, you know, willed action, a tiny time before the action is—you think you're deciding to do it, there is already neural activity in the motor areas, which means that the actual decision is unconscious.

Yes, we talked about it in Cologne, and the absurd fact that people used that as an argument against free will.[30]

Free will, well, no, it's not—

It makes no sense.

It makes no sense, and it's kind of interesting [that] in the history of thought, the concept of unconsciousness has barely entered. I mean, even in Freud, he talks about the unconscious, but you can bring it to consciousness—

Yes, that was the basic idea.

That's the whole point of psychotherapy.

Yes, of course, I know, and it's a big philosophical point which we have no time for… Let's hop back into the 1950s.

OK.

I guess when you wrote that big work of yours, Logical Structure of Linguistic Theory,[31] *what was clearly in the background was that languages can't vary in limitless ways, otherwise, you wouldn't have bothered to write about this logical structure and set up this system—*

Well, the obvious reason why languages can't vary in limitless ways is that there would be no way to learn any of them. I mean, this had to do with—it has to do with the impossibility of induction without a framework, and incidentally, that was one of the topics that Nelson Goodman—

I mean, that sentence alone is sufficient; you have to have a framework, otherwise you can induce nothing…

Which is—it was quite interesting at the time, because the other person[32] I was

30 See Chomsky (2011: 19).

31 See Chomsky (1955/1975), a work that also exists in various unpublished forms dating back as far as the original mimeograph version from 1955.

32 Apart from Willard van Orman Quine. Interestingly, Quine co-authored an article with Nelson Goodman called 'Steps toward a constructive nominalism' (Goodman & Quine 1947)

studying with was Nelson Goodman, who really argued that you *didn't* need a framework! He tried to show in his new literature on induction, in *Fact, Fiction, and Forecast*,[33] he posed fundamental problems for induction, and then argued that there were ways out of them by looking at the way things were used, but it didn't work at all, you know.

Yeah. And looking at the system as you set it up in LSLT *and comparing it to* Morphophonemics, *you have the transformations, so it's a two-part system; on the one hand, you have the phrase structure component out of which the kernel sentences grow,*[34] *and then you have a transformational component, which serves two purposes, namely doing transformations within the kernel, and also, sticking the structures that underlie the kernel sentences, sticking them together. So that would be a different set-up from what you had in* Morphophonemics.

Well, yes, *Morphophonemics* had only some, a few phrase structure rules, but in *LSLT*, it was assumed that, yes, there is a phrase structure component, but it generates abstract structures.

Of course!

And they have to be mapped onto what turns out to be a sequence of morphemes.

The phrase structure component doesn't create the kernel sentences, but the structures underlying *the kernel sentences…*

It generates *only* the structures underlying them. And at that point, I also assumed that there was a finite number of them. I assumed that the recursive character of the language was in the transformational component, in fact in generalized transformations.

That is the point I was wanting to make. Yeah, and I see a very interesting parallel here in having sentences that are stuck together by generalized transformations, in terms of things that now crop up again and again through 'barriers',[35] *and now 'phases'.*[36] *The whole thing seems very similar to me. Because you have some localized things, within that very, very restricted phrase-structural component, and of course you have very local*

---

which appeared in the same journal that was also the venue for Chomsky's first published article (Chomsky 1953).

33 See Goodman (1955).

34 The concept of 'kernel sentences' goes back to Harris; they are, roughly, simple declarative active sentences.

35 This concept was developed in Chomsky's (1986) monograph carrying the same title. At their core, the concept of syntactic barriers deals with the question of how far, and in what steps, syntactic elements can move. With it, Chomsky tried to unify two independent constraints on movement, namely Subjacency and the Empty Category Principle (ECP). For some informal remarks that might elucidate both the notions of subjacency and ECP and the unifying idea of barriers, see Rizzi (2012) and footnotes there.

36 'Phases' are syntactic units roughly corresponding to verb phrases accompanied by the verb's assignment of semantic role to its subject and object, full finite sentences, and probably determiner phrases in one form or another. The idea here is again that the rules of human grammar are first applied in the smallest phase, which can then be 'forgotten', then the next larger one etc. A simple and non-technical presentation of the idea can be found in Larson (2010: chaps. 26–27).

operations now in phase theory. So—

Well, it goes beyond that. Take a look at the work I and others have been doing for roughly twenty years, called the Minimalist Program. It bars, it goes back in a lot of ways to something like generalized transformations. For one thing, there's no phrase structure grammar. Phrase structure grammar was basically eliminated in the 1960s.[37] Phrase structure grammar involved a tremendous number of stipulations, like, why do we say, "V → **V",** [I mean,] "VP → V + NP", why don't we say "NP → V + NP", you know? It's just stipulated, so endocentricity is stipulated, where the phrase is compared against the head.[38] There is massive stipulation we want to get rid of. It was more or less gotten rid of by X-bar-theory, and that has its own stipulations—[39]

I think at the time it was not formulated that way, but for example, if one looks at the work that comes after Aspects of the Theory of Syntax, *and one scans the work, your work for example, for phrase structure rules, you will find them as an expository device, for example, the two-part phrase structure of X-bar-theory;[40] it's nothing more than that, and the meaning is completely different, in that things are flipped around—phrase structure is top down, and X-bar is also [makes a gesture from below]*

It's projected from the lexical items, the smallest atoms. Actually, there is new work I have on this, which suggests a slightly different way of looking at it, but I think you can get around, eliminate a lot of stipulative character of X-bar theory as well, but when you finish this, you will really do a lot of parallel computation, so if you construct not even—I mean any sentence, like "The man saw the boy", let's say. '*The man*' and '*the boy*' are being constructed in parallel, '*see the boy*' is being constructed out of '*see*' and '*the boy*', and then '*the man*' is being added to that, that's all parallel computation, which kind of has the spirit of generalized transformations in a way, so if the only combinatorial operation is the absolutely simplest one, namely, just Merge two things to make a new thing, then you're going to have parallel computation, extensively, richer even than generalized transformations.

Yes, I see. Let me, since our time is already compressing very rapidly, let me get back to the 1960s about things that we haven't talked about yet so much. One thing that really

---

37 Reference here is to the introduction of X-bar theory in Chomsky (1972). From that time on, X-bar theory in its various formulations, started to replace the previous multitude of phrase structure rules that had characterized generative grammar from its beginnings and even more so after the recursive component, the part of the grammar that enabled infinitely long sentences, had been assigned to phrase structure rules instead of generalized transformations in Chomsky's (1965) classical *Aspects of the Theory of Syntax* in what henceforth was often called the 'Standard Theory'.

38 In his final 2014 corrections, NC says: "Might be better to delete this [namely, 'so endocentricity is stipulated, where the phrase is compared against the head']. Endocentricity really comes in with X-bar theory, not phrase structure grammar." I have kept the passage in order to illustrate how the attempt to get rid of one set of stipulations can inadvertently lead to new stipulations that may come under scrutiny only much later.

39 See previous note.

40 Simplified to the core structures, the two residual phrase structure 'rules' in Chomsky's (1986: 3) *Barriers* are X' = X + XP and XP = XP + X'. Presumably, it is hardly an accident that Chomsky doesn't use the arrow '→', meaning 'consists of', to present the relation, but rather the more neutral symbol '=', which suggests a two-way relation at the very least.

became foregrounded in the 1960s was what you just mentioned, namely the question of language acquisition. It was always there, in that you had to have an inductive system, but starting from the 1960s it was also applied in research, including the research of your wife,[41] including the research of Lila Gleitman[42] and so on and so forth. What do you think are some of the important things that happened in terms of this, of language acquisition research in the 1960s?

Well, actually the first—before that, language acquisition studies of course had taken place, but they were very descriptive. Experimental work began, and in order to study how a system is acquired, any system, you have to know what the system is! Now the earlier work didn't really ask: "What's the system that's being acquired?" They just looked at, you know you have to, children learn simple words, and then they learn two words, and then they learn three words, and so on. But as soon as you ask, "What system is it that's being acquired?", you ask different questions. So for example take Lila's work, Lila Gleitman. She was asking questions like, what does a child *understand* when it's producing only two words. Well, it turns out it understands much more complex structures. With the right kind of experimentation, you can show that—

In that talk in Spain in November last year,[43] you said, well, research tends to show that pretty much everything about the language is in place maybe at the age of two.

That's — there is work that *suggests* that. Actually, it's my wife's work.[44] But for example, if you take cases like, there is the famous Helen Keller type case. Helen Keller was, she lost speech and hearing after around twenty months, and then she developed —

An extreme case of the poverty of the stimulus!

Extreme case. It *looks* like that. However, I think that's a little misleading, because the evidence at least suggests that by twenty months she already knew the language! And everything else was extracted—

I see the point. So which kind of shift the PoS argument because that, doing this in that period as a small baby is also extremely impressive, isn't it?

That just makes it even more dramatic. I mean the amount that's known—if anything like this is correct, and my guess is that some of it will be shown to be correct, it's hard to do—, that would mean that basically, it is pretty much like the growth of vision. I mean, the structure of the system is already fixed, and you just tinker some little bits and pieces of it.

Another thing that it took people a long time to even think about, because it came so natural, of course, you grow up, you see, it's normal, it's there, but there is a lot to explain

41 See e.g. C. Chomsky (1969). This was Carol Chomsky's doctoral dissertation written under the supervision of Roman Jakobson.

42 For an incomplete account of Lila Gleitman's many contributions, see her website at http://www.psych.upenn.edu/~gleitman, her CV posted there, and also Gleitman (2013).

43 Should read Princeton, actually—see footnote 1 above.

44 Part of Carol Chomsky's far too little known work beyond her seminal contribution (C. Chomsky 1969) is listed in the bibliography of a memorial volume (Piattelli-Palmarini & Berwick 2013).

here.

Oh yeah!

The other thing that became prominent in your work in the 1960s, but I guess you've been very interested in that before, is the philosophy of language and the history of linguistics also. Can you tell me when you got into this?

Well, I mean, in the 40s and early 50s, I was mostly studying philosophy! With Quine, Goodman, and others, I didn't really agree with it, but I started reading earlier philosophy, mostly on my own. And of course I met, I was in contact with people like John Austin, close contact, I knew the British analytic, contemporary analytic philosophers of the time, and by the late 50s, I was reading and thinking about earlier philosophical traditions.

Did you also occupy yourself — well, that's a rhetorical question because I know you did — with Wittgenstein?

I read Wittgenstein, I didn't meet him, everybody read Wittgenstein, and so, yes, I read Wittgenstein, and well, some things were there which I thought were interesting. So I had— but the actual course work I was taking, such as it was, was mostly philosophy. And though I didn't really agree with Goodman, I was very close to him, and studied what he was doing. He was very, kind of like Harris, very impressive person, and interesting ideas, I thought they were wrong but very interesting, and stimulating. And then I got, I try to figure out how, but I think it was probably through, maybe Thomas Huxley, that I started finding out something about earlier ideas about language. They were in the philosophical tradition, I started reading Descartes, and Locke, and others, and —

Yeah, and you wrote that famous book Cartesian Linguistics.[45] *What made you pick that title?*

Well, actually the title, I probably should have picked a different title, because most people who read it didn't get beyond the title. So if you get to the first page, it points out that there is no such thing as Cartesian linguistics, there is just linguistics which turns out had borrowed and adapted certain, and developed certain ideas, which were also developed in Descartes, and used [in] and crucial for *his* philosophy, but that were developed in different ways —

What were these crucial points?

Well, the crucial point in Descartes, which was actually very significant in his own philosophy — he was basically a scientist, we call it philosophy —

Yes, that distinction didn't exist.

His thought, his core research was to try to show that you could give a mechanistic description of everything in the world. That's the 'mechanical philosophy' as it was called, what was assumed by Galileo, by Descartes, by Leibniz and Huygens, and Newton, you know, and this was, this is the core of early modern science. Try to show that the world is basically a machine. Of the kind that could be

45 See Chomsky (1966/2009), originally published in 1966 by Mouton, The Hague.

designed by a skilled artisan. In fact, it *was* designed by a super-skilled artisan, you know,[46] but meaning levers, and things pushing and pulling each other, and so on and so forth.

So the spontaneous kind of physics that you have, as an ordinary human being.

It's so intuitive, isn't it, and it was stimulated at the time by the fact that skilled artisans were *doing* things like that. Like if you walked in the gardens of Versailles, there were all kinds of complicated machines—they were stimulating in much the way computers stimulate the imagination of today. So the basic idea was, look, that's what the world is, and we will show it. Galileo for example was frustrated to the end of his life because as he said he just couldn't figure out how to give a mechanical explanation of the tides, or the motion of planets and so on, and Descartes thought he had a system which *would* do it, and that was his basic contribution. Nobody studies it these days, because it is known to be wrong, but then he did notice that there are things that don't seem to have a mechanical explanation. And crucial for him was *language*. He pointed out in the *Discourse on Method*[47] [that] every normal human being can carry out what I have sometimes called the 'creative use of language', you speak indefinitely, you can use indefinitely many sentences, new ones all the time, never heard before, other people understand them, they're appropriate, you use them as appropriate to situations, but it's not caused by situations, which is a crucial difference, and it's not caused by internal mechanisms of your particular physical state: You could produce many different things, and it's coherent and intelligible, and so on. These properties—

Descartes did not talk about rule systems or anything like that.

But that happened very quickly. That's when you get what I was calling 'Cartesian linguistics'. For Descartes, this[48] was fundamental.

This is exactly what I wanted to know.

For Descartes, this is the core of his mind/body distinction. He says, well, we have body, you know, the physical thing, which is going to be a machine, but then there are these properties that you cannot account for mechanically, so we need—he's a scientist—we need a new principle. And in the substance metaphysics of his day, that meant a new substance, so that's *res cogitans*, and then the main scientific problems are, find out its properties and show how it's linked to the other substance. You know, that's where you get the speculations about the pineal gland and so on.[49] But this is fundamental to Cartesian philosophy. Now if you look at the work of the period where the people were maybe influenced by

---

46 Namely, God.

47 See Descartes (1637/1998), one of the many accessible editions.

48 That is, the creative and non-deterministic use of language, a use that for Descartes was at the same time neither arbitrary nor random.

49 Descartes's postulation of two substances, one thinking and free—*res cogitans*—, the other material and determined—the body—, posed the problem of how the two interacted with each other and what, if any, the particular locus of this interaction was. Descartes suspected that the pineal gland, located as it is right in the midst of that mysterious organ, the brain, was that locus and served as the means of that interaction.

Descartes, maybe not, some of the people were *directly* influenced by Descartes, like his kind of like junior associates, people like Cordemoy,[50] he also pointed out that whatever this is, it is unique to humans, animals can't do it, machines cannot do it…

He made that very explicit.

Very explicit. So you have—

And even human idiots have it and animals don't.

Even idiots have it, but that led to immediate experimental questions, can we design experiments to determine if another entity who looks like us has this property or not.

Yeah, that would be the test. So that creature has a mind like us.

That's where the basic work was done by people like Cordemoy and others. And that's kind of like a litmus test for acidity: Do you have this property? It's kind of similar to what people now call the Turing test, but much more significant, because this is a real scientific issue: We want to find an absolute—well, is there a test for this property? Then, when you get to the logicians and the linguists, especially at Port Royal, you know, the Port Royal grammarians where you have — they produced two great works, the *Grammar*[51] and the *Logic*,[52] they began to provide linguistic mechanisms, they were interested—they weren't Cartesians, like they were anti-Cartesian in many ways, but the basic idea that permeated the intellectual atmosphere, they worked out methods of—first of all, they studied the vernacular, which was unusual at the time. Descartes did too, he wrote it in French, which was unusual, but the Port Royal logicians and linguists also studied the vernacular, and they noticed puzzles, in fact there were, there was a famous puzzle, called the Rule of Vaugelas,[53] which was a descriptive observation about French, and it was very puzzling at the time, and they worked out an analysis of it, which is based on principles that are very similar to 'extension' and 'intension' in the Fregean modern sense.

Yes, I see.

And they also put a lot of this in the framework of rules of, sort of like generative rules, it's a little bit anachronistic, but if you look at it, they're kind of precursors to generative rules.

OK! Right. I think we are rapidly approaching—

Incidentally, as far as Cartesian—in the same book, I went on to carry this for-

50 Géraud de Cordemoy (1626–1684) was a 'second generation' Cartesian who, different from Descartes himself, had a keen interest in the concrete workings of language. A glimpse into his thinking is available in de Cordemoy (1668/2003).

51 See Arnauld & Lancelot (1660/1975).

52 See Arnauld & Nicole (1662/2012).

53 See, for example, the 3rd edition of Chomsky (1966/2009: 96), as well as the passages before: "In Chapter IX, the Port-Royal *Grammar* first notes a variety of exceptions to this rule and then proposes a general explanatory principle to account both for the examples that fall under the rule of Vaugelas and for the exceptions to his rule."

ward to —

Let me get to a question quickly which I think is very often misunderstood, namely, when you wrote in the 1950s, we can't found syntactic analysis on semantics. I think what's at issue here is simply the use of the word 'semantics', which when by semantics you understand how language is used*, which was, –*

Yeah, but there was already a straightforward point at the time, which has been totally misunderstood.

Yeah!

The standard view at the time, expressed very clearly by Quine, but also in the linguistics and so on, was that you could do an analysis of language just as a system of form, but you had to appeal to meaning at two points: One point was for phonemic distinctness. In order to show that, you know, let's say, '*writer*' and '*rider*' are distinct, you have to appeal to the meanings. So the—Harris actually had a way around this, but the general view was you had to appeal to the meaning. Also to decide whether a sentence was grammatical, you had to appeal to the meaning. And the specific technical argument having to do with this was to show that neither of these would work. If you actually paid attention to the meaning, you couldn't determine grammatical status or phonemic distinctness on the basis of the meaning, and in fact, there was really no point at which *use* of meaning, the facts of meaning, entered into determining what the, say the way a transformational rule works—that has nothing to do with the study of meaning! In fact, *LSLT*, a lot of which is structuralist, it's basically a study of semantics! It's a study of, you know, why the sentences mean what they do. Like, why does "John is eager to please" mean something different from "John is easy to please". That's essentially a study of semantics. It's not *based* on *semantic intuitions*; the idea is that we have a formal system in our heads, we *use* it in particular ways, to express thoughts, to talk about the world, in all kinds of ways. And that's semantics. Now it's not even very clear if you look closely, that language *has* a semantics in the technical sense.

Yes, yeah, absolutely. I think we will tease this very interesting and I think decisive question, we'll tease out the finer points simply in writing –

OK.

Because –

Too late.

– we are obedient human beings and –

[Points to Ted Passon, sound- & cameraman] He's tired. He wants to go home.

And we don't destroy the schedule that Bev wisely set up for you.[54]

No, OK!

54 Bev Stohl, the ever advertent and sometimes stern guardian of Chomsky's well-being, who also successfully saw to it that the interview would not last longer than the allotted 60 minutes. With 59 minutes and 36 seconds, Noam and I should not rank badly in a list of well-behaved citizens ready to follow reasonable advice and orders.

Thank you very much, Noam, for doing this.

OK.

And I hope to continue the conversation.

What are you going to do with all this stuff?[55]

Well, at the moment, it looks like it might be rolled into a book. […] But first of all, I'll have a lot of work to do on this.

References

Arnauld, Antoine & Claude Lancelot. 1660/1975. *General and Rational Grammar*. The Hague: Mouton.

Arnauld, Antoine & Pierre Nicole. 1662/2012. *Logic, or the Art of Thinking*. London: Forgotten Books.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its development. In Anna Maria di Sciullo & Cedric Boeckx (eds.). *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 19–41. Oxford: Oxford University Press.

Chomsky, Carol. 1969. *The Acquisition of Syntax in Children from 5 to 10*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1951/1979. *Morphophonemics of Modern Hebrew*. New York: Garland.

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

Chomsky, Noam. 1955/1975. *The Logical Structure of Linguistic Theory*. New York: Plenum.

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

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

Chomsky, Noam. 1966/2009. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*, 3rd edn. Cambridge: Cambridge University Press.

Chomsky, Noam. 1968/2009. *Language and Mind*, 3rd edn. Cambridge: Cambridge University Press.

Chomsky, Noam. 1972. Remarks on nominalization. In Noam Chomsky, *Studies on Semantics in Generative Grammar*, 11–61. The Hague: Mouton.

Chomsky, Noam. 1979. *Language and Responsibility: Based on Conversations with Mitsou Ronat*. Sussex: Harvester Press.

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

Chomsky, Noam. 2002. *On Nature and Language*, ed. by Adriana Belletti & Luigi Rizzi. Cambridge: Cambridge University Press.

Chomsky, Noam. 2010a. On the failures of structuralism and the necessity of abstract analysis. Interview with Michael Schiffmann. Cambridge, MA, July 29.

[55] That is, the previous interviews with Chomsky himself, but also with all the other linguists involved in this series.

Chomsky, Noam. 2010b. The place of displacement and other questions demanding an answer. Interview with Michael Schiffmann, Cambridge, MA, August 5.

Chomsky, Noam. 2011. Poor God. Must be bored to death! Interview with Michael Schiffmann, Cologne, June 7.

Chomsky, Noam, Morris Halle & Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle, Horace Lunt, Hugh McLean et al. (eds.), *For Roman Jakobson: Essays on the Occasion of his Sixtieth Birthday*, 65–80. The Hague: Mouton.

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In Duncan R. Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of Mathematical Psychology*, vol. *II*, 269–321. New York: John Wiley & Son.

de Cordemoy, Géraud. 1668/2003. *Philosophical Discourse Concerning Speech Conformable to the Artesian Principles*. Whitefish, MT: Kessinger Publishing.

Descartes, René. 1637/1998. *Discourse on Method*. Indianapolis, IN: Gardner's Books.

Gleitman, Lila. 2012. When the duck *gorps* the bunny: Generative grammar and language acquisition. Interview with Michael Schiffmann. Philadelphia, January 21.

Goodman, Nelson. 1955. *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Goodman, Nelson & Willard van Orman Quine. 1947. Steps toward a constructive nominalism. *Journal of Symbolic Logic* 12, 105–122.

Halle, Morris. 2010. Jakobson, Chomsky, and the quest for linguistic rules. Interview with Michael Schiffmann, Cambridge, MA, August 4.

Harris, Zellig. 1951. *Methods in Structural Linguistics*. Chicago: University of Chicago Press.

Harris, Zellig. 1952a. Discourse Analysis. *Language* 28(1), 1–30.

Harris, Zellig. 1952b. Discourse Analysis: A Sample Text. *Language* 28(4), 474–494.

Joos, Martin (ed.). 1966. *Readings in Linguistics*, 4th edn. Chicago: University of Chi-cago Press.

Hill, Archibald (ed.). 1962. *Third Texas Conference on Problems of Linguistic Analysis, May 9–12, 1958*. Austin, TX: The University of Texas.

Katz, Yarden. 2012. Noam Chomsky on where artificial intelligence went wrong. *The Atlantic*, Nov. 1.

Larson, Richard. 2010. *Grammar as Science*. Cambridge, MA: MIT Press.

Piattelli-Palmarini, Massimo & Robert C. Berwick (eds.). 2013. *Rich Languages from Poor Inputs*. Oxford: Oxford University Press.

Rizzi, Luigi. 2012. From Pisa to Siena: Parametric syntax and the Minimalist Program. Interview with Michael Schiffmann. Frankfurt, March 9.

Ross, John Robert. 1967. Constraints on variables in syntax. Cambridge, MA: MIT dissertation.

Ross, John Robert. 1986. *Infinite Syntax!* New York: Ablex Publishing.

Noam Chomsky
Massachusetts Institute of Technology
Department of Linguistics and Philosophy
77 Massachusetts Avenue, Bldg. 32-D808
Cambridge, MA 02139
USA
[*chomsky@mit.edu*](mailto:chomsky@mit.edu)

Michael Schiffmann
University of Mannheim
English Department
Schloss, Ehrenhof West 266
68131 Mannheim
Germany
[*mikschiff@t-online.de*](mailto:mikschiff@t-online.de)

★ FORUM ★

Structural and Functional Organizing Principles of Language: Evolving Theories

Ádám Szalontai & Katalin Csiszár

1. Introduction

The overall goal of this paper is to evaluate theories that attempt to address the organizing principles of language and review the development of these theories toward the integration of language within an interactive network of higher-level cognitive functions. Commencing with an overview of traditional concepts of language as modular, distinct, and innate, we focus firstly on areas that highlight the foundation of modularity theory including various module definitions and criteria, and applications of modularity in information processing and biological systems. We also discuss challenges to the overall applicability of a modular system and limitations of modular models in dealing with adaptation, novelty, innate versus learned, domain-general and domain-specific features, and developmental and age-related changes of cognitive organization.

Prompted by the rapidly increasing amount of empirical data on the functional elements of the human brain, we then evaluate several major theories of cognition, including views that oppose modular organization and those that integrate modular and semi-modular views with topological modularity in simpler, and dynamic integration in higher-level cognitive functions.Within this framework, modular and non-modular components of linguistic knowledge, organizing principles of language viewed either as specific or derived from other systems, and concepts of language as one of the cognitive functions or the outcome of unique interactions among cognitive components are discussed.

Emerging theories that integrate interactive network models support a cognitive architecture as a mosaic of domain-specific and domain-general processes involving both functional segregation and integration within a global neuronal workspace. Within this anatomically distributed workspace, the language function represents unique interactions among cognitive components consistent with an organization that is task-dependent with a continuum between degrees of modular and shared processing. As a higher-level, learning-based, and effortful cognitive process language transiently enlists a less modular organization for an efficient network configuration in interaction with several cognitive systems and the domain-general cognitive control/multiple-demand network.

We are very grateful to Dr. Ágnes Lukács, Department of Cognitive Sciences, Budapest University of Technology and Economics, and Dr. Balázs Surányi, Research Institute for Linguistics, Hungarian Academy of Sciences, for their helpful comments during the development of this paper.

ISSN 1450–3417

Biolinguistics 4.8: 204-225, 2014
http://www.biolinguistics.eu

2. Traditional Views of Language and Cognitive Architecture: Limitations and Challenges

Traditional explanatory models considered the language faculty (FL) as an organ of the body determined by genetic endowment (Universal Grammar, UG), experience, and possible design features, mainly centered around the concept of computational efficiency and conceptualized as third factor principles in subsequent theories. Some of these features, based on the initially assumed lack of difference in basic language capacity among contemporary humans, have also been considered language (UG) and/or even organism (third factor principles) independent (Chomsky 2005, 2011). Earlier concepts also held language as different and distinct from other cognitive functions with considerations of UG as a specific language module or a modularized knowledge or element of the FL (Hauser et al. 2002).

Modularity, originally an engineering notion of near-decomposability, refers to a system that is made of components whose workings are independent of each other and in which the modular structure allows parts to be modified without the whole system ceasing to function (Simon 1969). The concept of modularity including certain characteristic and/or necessary model features (Fodor 1983), and a modular design was considered central in biology (Marr 1982) with a general modular theory also proposed for perception and cognition (Coltheart 1999). The modular organization hypothesis has been also linked historically to the claim that aspects of the human mind are innately specified (Fodor 1983, reviewed in Twyman & Newcombe 2010).

The integration of modular organization in certain biological areas such as development has high explanatory value. However, a strictly modular view is controversial in the cognitive sciences as it precludes complex processes such as associative learning, attention, working memory, or general intelligence that cut across domains (Shettleworth 2012). The assumption that much of the cognitive processing is modular was challenged early on by fMRI data of brain activation (Wojciulik et al. 1998), and studies of neurodevelopmental language disorders that proposed relative modularity that is only achieved after an extensive period of developmental time (Bishop 1997, Karmiloff-Smith 2007). The widely different definitions of what constitutes a module remain controversial and the minimum criteria that are required for a system to be considered modular are unsettled (Twyman & Newcombe 2010, Kaltenbach & Stelling 2012). Important theoretical challenges also remain regarding function-centered decomposition of dynamic biological networks (Kaltenbach & Stelling 2012). The controversial issues include the integration of modular clusters in a larger-scale as modalities in the brain need to be both isolated and sufficiently connected for coherent functions (Gallos et al. 2012), questions as to how does a modular system deal with novelty (Anselme 2012), the innate and/or developmental nature of modularity (Thomas & Karmiloff-Smith 2002), the lack of one-to-one correspondence in function to structure mapping (Petersson et al. 2012), the relationship between domain-general and domain-specific cognitive processes (Meunier et al. 2010, Kitzbichler et al. 2011), the continuum between strictly modular processing and degrees of task dependent shared processing (Borowsky et al. 2007), and functional specialization as a matter of discrete units or as a matter of degree (Kanwisher 2010).

2.1. Universal Grammar: A Disputed Language Module

One of the most influential traditional theories of language assumed an innate faculty with a defining influence of neural circuitry shaped by biological determinants (Chomsky 1965, 1968, 1995, 2005). A few genetic events (summarized by Chomsky 2005), subsequently also interpreted as a single mutation scenario (reviewed in Jackendoff 2011), had been proposed to rewire the human brain and create an abstract cognitive mechanism responsible for the development of language. Within this framework a modular organization of cognitive functions was also assumed in which language constitutes its own module (Hauser et al. 2002). The traditional linguistics term of language refers to an internal component of the mind and/or brain (internal or I-language). The the FL in the broad sense (FLB) was proposed to include a narrower internal computational system (FLN) and at least two other internal systems (sensorimotor and conceptual-intentional) and cognitive resources or abilities necessary for the acquisition and use of language that together constitute UG (Hauser et al. 2002). FLN was considered to only include recursion and to be the uniquely human component of the FL and a biological capacity of humans that allows mastering of language without explicit instructions (Hauser et al. 2002). FLN by this definition cannot be compared to anything existing in the mind of other species, not even in other domains of the human mind (revisited in Boeckx & Longa 2011 and Traxler et al. 2012).

Alternative theories include those that argue that the language-specifically adapted sensorimotor systems should be part of FLN, traditions that consider the conceptual-intentional system as intrinsic part of FLN (Jackendoff 2011), and views that language is built on biological and cognitive foundations that pre-date the emergence of language (Chater et al. 2009, Christiansen et al. 2009). Although language is used to express recursive thoughts, their recursive nature is independent of language and likely preceded its evolution (Corballis 2011). Recursion may have evolved in order to solve other computational problems such as navigation, quantification, or social relationships (Hauser et al. 2002). Furthermore, as recursion is not unique to language, it cannot constitute FLN (Jackendoff 2011).

Theories that held FL as a distinct module among other cognitive modules, also considered UG as a specific and modularized knowledge/element of the language faculty. UG as a genetically determined language acquisition device that constrains the parametric options available for natural languages has been viewed as the key component that explains both the linguistic universals and the assumed quick and uniform path to language acquisition (Hauser et al. 2002). Based on the shared common core of human languages (Berwick et al. 2013), the poverty of stimulus for UG argues that invariant properties of the human mind, including the structure dependence of grammatical rules and certain constraints on question formation, reflect an innate human endowment (Berwick et al. 2011). The principles and neural mechanisms of UG have also been regarded not only as innate, but distinctly modularized and independent of (though connected to) other parts of cognition (Grodzinsky 2006).

UG however is a disputed notion (Elman et al. 1996, Boden 2006, Clark & Lappin 2011). Some argue that the simplest idea is that there is a universal set of cognitive capacities underlying human linguistic competence. Others question the reality of UG given that there is no consensus on the very notion of UG, and also question the species specificity, poverty of stimulus, ease of acquisition, and uniformity of the knowledge of language across a population (Dąbrowska 2004, Evans & Levinson 2009 reviewed in Irurtzun 2012). During language acquisition there is no need for UG to figure out what grammar/syntax underlies a particular language, the ability to learn signs also enables the learning of combinatorial signs or dedicated order of signs (Bouchard 2012). There are factors other than language-specific UG conditions that can canalize grammar very stringently: Properties of the perceptual and conceptual systems necessarily impose boundaries within a highly circumscribed course of language development and these properties are considered sufficient elements of language. Further challenging the existence or need for innately specified knowledge of language, a Bayesian framework for grammar induction showed that given certain innate domain-general capacities, a learner can recognize the hierarchical phrase structure of language without having this knowledge innately specified (Perfors et al. 2011).

2.2. The Computational View of Language

The basic design of language as a biological subsystem reveals a system of discrete infinity, that is unbound number of expressions. Language (the unified nature of language) has been proposed to arise from a shared species-specific computational ability that is grounded in a neuronally realized computational mechanism that yields an infinite array of structured expressions. At minimum, this computational mechanism is able to combine one linguistic representation with others, yielding new and larger linguistic objects. The computational mechanism includes some operation (Merge) that constructs new representational elements from already constructed elements which must then be transposed to linear representations, a constraint imposed on the sensorimotor systems input–output channel (words must be pronounced sequentially) and on language perception (listeners analyze sequentially ordered acoustic sequences) (Chomsky 1995). In this view, the FL as a computational device is capable of processing symbolic elements and externalizing and internalizing the output of such computations. The computational system is viewed as the outcome of interactions between a sequencer (activity performed by the basal ganglia) and working memory (the activity of diverse cortical structures) (Benítez-Burraco 2012). Accordingly, language acquisition depends on the interplay of the shared initial genetic endowment (UG, the language- and human-specific module), conditions imposed by the structure of the brain, cognitive preconditions (statistical analytical capacity), external influences (environmental stimulants), and certain general principles such as external laws of growth and form, and minimalization of computational complexity (Berwick et al. 2013).

2.3. Modularity in Information Processing

The ubiquity of modularity and hierarchical modularity across technological and biological systems prompted a search for dynamic, adaptive, or anatomical constraints that may drive the evolution of networks towards a modular architecture (Meunier et at. 2010). A system built of multiple and sparsely interconnected modules allows efficient adaptation. Evolution of such a modular system can take place by change in one module at a time or by duplication or mutation of modules without loss of function of well-adapted modules. The evolution of individual modules does not jeopardize the function of the entire system and results in robustness, a major advantage for any system evolving under changing selection criteria. High clustering of connections favor locally segregated processing of specialized functions (as in visual motion detection), while short path length supports globally integrated processing of generic functions (as in working memory) (Meunier et al. 2010, Sporns et al. 2010, 2013). Modular topology is associated with rich non-linear dynamic behavior including time-scale separation (fast intra-modular and slow inter-modular processes) and high dynamical complexity due to the coexistence of both segregated and integrated activity. The feedback between structure and function including reinforcement of links between synchronized units and pruning of links between asynchronized ones, naturally drives the emergence of inhomogeneities and a modular network. Optimality at performing tasks in a changing environment, where different goals share basic sub-problems and where rapid adaptation to each of the different goals is enhanced, produces networks, modules, and modular units that specialize in these sub-problems. Thus a modular network is a topologically modular and nearly decomposable system made of component modules each of which comprises a number of nodes that are densely intra-connected to each other but sparsely connected to nodes in other modules (Meunier et al. 2010).

2.4. Modularity in Biological Systems

In biomedical research, mechanistic explanations dominate by which a phenomenon is explained by revealing the set of entities and activities that are spatially, temporally, and causally organized. The modular partitioning and hierarchical structure of the biological space emerged as a symmetry-breaking phase transition exemplified by metabolic networks, gene networks, protein interaction networks, or social networks (Lorenz et al. 2011). The biological norm has been considered to be a set of specialized modular systems and this type of organization presumed so ubiquitous that all functional systems were anticipated to be subject to the same organizational principles. In evolutionary and developmental biology, modularity, defined in general terms as a property of being made up of self-contained and independently functioning parts, is regarded as a key principle (reviewed in Shettleworth 2012).

Connotation for modules (Fodor 1983) has been defined as autonomous (operate independently of other systems and are independently disruptable), domain specific (responsive to a distinctive class of stimuli), innately specified, informationally encapsulated (impervious to information outside the modules domain, a feature considered to be at the heart of modularity, Fodor 2001), peripheral (as opposed to central decision making), fast acting (as a reflex), mandatory (not under conscious control), obligatory (acting regardless of circumstances) and hardwired. While innateness and modularity are different concepts, modules are often held innate in the sense that they develop similarly across individuals regardless of environmental input (Barrett 2012).

These definitions have been interpreted not as necessary conditions for the applicability of the term but rather as a general theory of perception and cognition and features that are characteristic of modules (Coltheart 1999). While it has been recognized that the notion of modularity ought to admit to degrees (Fodor

1983), it has also been postulated that if a system has most of the modularity properties, then it is very likely to have all of them, a mechanistic explanation derived from the computational information-processing paradigm. Of the various module-associated concepts, however, only the concept of autonomy corresponds to strict modularity in the sense of independent disruptability, or dissociability (it can be selectively disabled with no effect on other capacities of the same system) (Menzies 2012).

The most basic term of a module is a capacity that is functionally individuated in terms of its input and output conditions with domain-specificity as its most important property. The generalized and widely applied conceptualization of a module involves a set of related elements that maintain a strong connectivity within but a weak connectivity among other equivalent sets of elements. In this general sense, a module may include sets of functionally interrelated genes, interactions among regulatory elements, interrelated set of neuronal structures, or coordinated actions of biological structures (Benítez-Burraco 2012).

While the traditional analysis of complex biological networks relied on decomposition into smaller, semi-autonomous units (e.g. signaling pathways), with the recently increased scope of systems biology the different definitions of what constitutes a module or a modular structure and the function centered decomposition of dynamic biological networks sparked controversies (Kaltenbach & Stelling 2012). There are considerable debates about the extent to which any modular structure in a mature adult is inborn or emerges through experience (Thomas & Karmiloff-Smith 2002). Other views point out that biological systems are actually not all highly specialized and modularized and are rather varied in how they constrain processes (Gallistel & Gibbon 2000). Moreover, the modularity assumption does not apply universally in all domains, there is more depth to mechanistic explanations than box-and-arrow diagrams, and the spatial and temporal organization of mechanisms are often as significant as the causal organization (Gallistel & Gibbon 2000, Menzies 2012).

3. Organizing Principles of Cognition: From Modules to Global Neuronal Workspace

3.1. Modular Models of Cognitive Functions

Modularity has been traditionally presumed as being essential at both cognitive and neural levels, yet the notions of neural and cognitive modularity remain controversial (Marcus 2006). Of the brain mechanisms, commonly thought of as falling into two categories: specialized and general-purpose, the specialized mechanisms being frequently associated with the idea of modules (Barrett 2012, Barrett & Kurzban 2006). In this model, independent disruptability has been assumed to be a basic meaning of modularity as two sub-processes (mental or neural) can be modules if and only if each can be changed independently of the other (Sternberg 2010). Thus the notion of modularity, applied in a strict sense to the organization of the brain, envisions a system made of components whose workings are independent of each other with parts modifiable (Simon 1969) exemplified in the modular account of the visual system viewed as a computational system made of a collection of small independent sub-processes (Marr 1982).

The general application of the modularity concept to cognitive functions (Marr 1982, Coltheart 1999, Anselme 2012) resulted in the assumption of massive modularity for the human mind while other theories of modularity such as the (i) core knowledge modules position (Spelke & Kinzler 2007) distinguished limited number of core knowledge modules (object, action, number, geometry, and social partner representation), or in a version of core modules, six big traits (intelligence, openness, contentiousness, agreeableness, emotional stability, and extraversion) modules presumed to be domain specific, innate, and even shared across species (reviewed in Twyman & Newcombe 2010). Modules defined by these theories, however, do not conform to the module definition (Fodor 1983), few are encapsulated, and most are involved in various cognitive functions.

Among various additional theories, the (ii) functional modularity approach argues that the key property of cognitive modularity is functional specificity given that distinct domains of information require specific processes to operate on them (Barrett & Kurzban 2006). In this view, functional modules constitute a subsystem of the cognitive system each dedicated to specialized functions. (iii) Anatomical modularity is an additional thesis by which each functional module is implemented in a dedicated, relatively small, circumscribed neural hardware (Anderson 2010). In this paradigm, various kinds of cognitive modules have been distinguished: developmental, (neuro)functional, mental, or even virtual (a pattern of dissociability between aspects of the systems that does not correspond to separate neural systems) (Griffiths 2007). These modules have been generally assumed to be localizable within the brain and to be neurally specific. In contrast, it has also been proposed that there is no necessity for a cognitive module to be associated with a localized fixed neural architecture (Coltheart 1999). An alternative hypothesis to the functionally distinct, independent neurocognitive modules (e.g., in an extreme view, language module owes nothing to other cognitive devices) is the (iv) descent-with-modification view by which modules are shaped by evolutionary changes. Thus, from common origin ancestral cognitive or neural modules/capacities, relatively recent modules (such as language) may derive and draw on general cognitive resources, consistent with features of neurodevelopmental disorders and developmental language disorders (Marcus 2006).

Given the many invocations of the term modularity, also referred to as (v) innatist-modularity (inborn modularity without modification), the modularity definition is considered so vague as to be essentially untestable (Twyman & Newcombe 2010).

3.2. Challenges to Cognitive Modularity: Domain-Specific and Domain-General Processes

While cognitive modules clearly function in dynamic environments and have to deal with change-induced novelty and uncertainty, the novelty of stimulus is problematic for the modular concept as it does not satisfy the modules criteria for domain specificity. In order to overcome this problem, a potential for (vi) transient variations in domain-specificity (behavioral transitions from exploratory activity to habit formation) have been proposed (Anselme 2012). Explanations for information processing adaptation, however, remain controversial and the form in which such adaptation may take is disputed.

Brain mechanisms, if evolved through processes of descent with modifications are likely to be heterogeneous rather than limited to two specific kinds (modular and non-modular). As new structures evolve from older structures, adaptations represent a mix of ancestral and derived features of which the older ones are shared more widely across structures, while relatively recent ones (properties of specialized brain regions) are more narrowly distributed in a hierarchically organized fashion (Caroll et al. 2005). The mix of these mechanisms are also likely to be highly plastic, environmental factor dependent for development, and interactive with other systems (Barrett 2012). A structure recaptured in the (vii) dual process theory proposes two kinds of human cognitive domains: simple basic processes shared with other animals and slower developing uniquely human processes, a human cognitive architecture that is a mosaic of modular and domain-general processes (Shettleworth 2012). Based on this model, variations in developmental outcomes across individuals or environments may be standard for brain adaptations. Adaptations for language exemplify this theory producing highly variable outcomes in various languages (Evans & Levinson 2009).

Additionally, domains/modules are not persistent (inborn) and among the modules, for example, language is viewed as a mechanism that moves infants from an innate modular representation to integrated cognition in adulthood (Twyman & Newcombe 2010). Or in contrasting opinions, infant brains start out highly connected and only over developmental time do the networks become increasingly specialized with domain-general and domain-specific processes (Steele et al. 2012). The organization of brain networks overlapping with functional domains (executive and auditory/language processing) has been reported to also demonstrate aging related changes, a reduction in functional segregation and intra- and inter-module connectivity (Chen et al. 2011). Function related dynamic changes have also been noted to occur such as an increase in connectivity between working memory regions and language regions concomitant with processing load increase (Makuuchi & Friderici 2013).

While specialized (modular) mechanisms appear innate, domain specific, and/or isolated from other brain systems, the generalized mechanisms (non-modular) are considered developmentally plastic instead of innate, domain-general and interactive. Along these lines, the (viii) dual system view equates specialization with highly local, narrow, and stereotyped (modular) processes, while the general-purpose processes are defined as those outside modularity. Thus, features of developmental plasticity, interactivity, or the ability to respond to novel stimuli are taken as evidence that a brain process or region is not evolutionarily specialized (Barrett 2012).

3.3. Brain Networks: Structurally Distributed and Functionally Diverse Connectivity

The human brain has been subjected to extensive multiple scale studies from neurons, circuits, anatomically defined areas to functional networks. Anatomically localized and functionally specific brain regions (developed through the maturation of specialized groups of similar fate cells sparsely connected to groups of cells of different fates) and their connecting networks, considered as information processing systems, share some of the organizational principles of modular systems. As other complex systems, brain networks also demonstrate hierarchical structure, or modularity on several topological scales (submodules and sub-submodules) that ensure robustness, adaptivity, and ability for evolution of network function (Meunier et al. 2010). Brain organization studies based on data obtained in cat and macaque brain connectivity using fiber tracking have identified four hierarchically organized major sub-networks classified as visual, auditory, somatosensorimotor, and frontolimbic. Human anatomical network analyses of cortical regions reproduced some of these functionally localized areas such as auditory/language, strategic/executive, sensorimotor, visual, and mnemonic processing (Chen et al. 2008) and interpreted the data as a modular — though with modules loosely defined as groups of connected cortical regions — brain structural networks (Hagman et al. 2008).

In more complex cognitive functions, to establish the association of anatomical brain structures with specific function ranging from synapses to entire brain regions, remains a challenging task (Fotopoulou 2013). The integration of modular clusters in a larger-scale has also been problematic as modalities in the brain that process different characteristics need to act in an isolated fashion for efficient computations, yet need to be sufficiently connected to perform coherent functions.

In order to overcome the limitations of the modular cognitive model (with modules variably defined as network areas based on connectivity features: high intra- and sparse interconnectedness), several types of alternatives have been proposed. The network view emphasizes that a complex system is shaped by the interactions among its constituents driven by universal selection criteria, such as high efficiency of information transfer at low physical connection cost (Sporns 2010). The functional integration-convergence model further emphasizes the significance of connectivity patterns among various interconnected, functionally diverse, and structurally distributed components of the nervous system (Fotopoulou 2013). Another alternative is the adaptive combination model supported by results of a study that challenge the existence of a geometric/reorientation module and hypothesizes that information (geometric and featural in this case) is utilized to varying degrees dependent on the certainty and variance that these two kinds of information represent (Ratliff & Newcombe 2008, Twyman & Newcombe 2010). Additionally, a recent combination of high-temporal resolution fMRI and network analysis tools have revealed both functional and topological fractal properties of brain networks, described as a two-layer structure (strong ties in a sea of weak ties) that fulfill the need for information flow within complex structures (Gallos et al. 2012).

3.4. High-Level Cognitive Functions Enlist a Global Neuronal Workspace

Adult human cognition shares simple basic processes with the cognition of other animals while additionally includes unique, slower-developing, usually slower acting, more explicit, and consciously accessible processes, among which the kind versus the degree of cognitive difference has not always been clearly defined. The relatively low level cognitive or perceptual processes based on features such as domain-specific, informationally encapsulated, fast, automatic, and anatomically localized, can be characterized as physiologically modular. In contrast, higher-level integrated, effortful, and conscious cognitive processes have been linked to anatomically distributed neuronal workspace architecture that may have emerged by breaking modularity of the background modular system (Dehaene et al. 1998), suggesting that modularity and/or non-modularity of brain network organization may be related to the type of cognitive processing that it can support (Meunier et al. 2010), and by this definition certain elements may be modular, while others are not (Barrett 2012).

Consistent with this theory is the basic modular architecture of the visual system (Magen & Cohen 2007) and evidence in favor of modular processing of verbal and spatial information in short-term memory (Guerard & Tremblay 2008). A contrasting non-modularity characterizes the central auditory function involving higher order performance (Musiek et al. 2005), the task-dependent activation of multiple/alternative pathways in prelexical and semantic processing, and the dynamically determined cortical network supporting language comprehension (Price 2010). Additionally, there is a continuum between strictly modular processing (in perceptual tasks) and varying degrees of modular and shared processing (in analytical tasks) that depends on the nature of the task (Borowsky et al. 2007).

Effortful cognitive performance that depends on the formation of a global neuronal workspace enlists — with increasing demand and faster performance — a more global, less clustered, and less modular networks with more long-distance synchronizations to allow the transient adoption of functional networks for less economical but more efficient configuration (Kitzbichler et al. 2011).

3.5. Functional Segregation and Integration: Continuum and a Matter of Degree

Neuropsychological theories that infer the functional role of certain brain areas on the basis of the consequence of damage to these areas (localizationist and anti-localizationist theories) served as the bases for two central principles of structure-function relations: functional specialization/segregation (specialized neurons form segregated regions responsible for discrete mental function) and functional integration/convergence (mental functions are based on connectivity patterns among various functionally diverse and structurally distributed components of the nervous system) (Friston & Price 2011, Fotopoulou 2013).

The segregation model derives from the long tradition of concepts for specialized organs and specialized brain modules and the logic of information processing systems that perform a series of formal operations (reviewed in Kanwisher 2010). Based on the theory that the mind is modular in its core conception, organized in computationally autonomous serially organized domains of function, brain damage was anticipated to result in a selective, encapsulated impairment of a component of cognitive processing without affecting other components. Earlier studies based on these assumptions aimed to identify behavioral dissociations to suggest new modular division in which cognitive information followed paths along serially organized modules each serving a different core cognitive function. In structure-function mapping studies (neuroimaging based mapping between neuronal activity and cognitive function), modularity of processing and processes was a key reference for establishing functional segregation as a principle of brain organization. The modularity theory also contributed to characterization of distributed brain responses in terms of functional integration or coupling among different brain areas (modular but coupled) (Friston & Price 2011).

The contrary concepts of distributed neural and cognitive processing have an equally rich tradition (Kanwisher 2010). It has been increasingly recognized that correlations between mental tasks and surrogate brain signals in functional neuroimaging studies have provided only indirect evidence (inferences about cognition based on neural activation). The initial and relatively simple imaging and statistical analyses resulted in simplistic localizationist and modular arguments about the role of certain brain areas in complex mental functions. From mapping of sensory functions into functionally specialized areas in the human cortex (spatial segregation) does not follow that similar kind of mapping would apply to complex cognitive and emotional functions.

The use of more refined methods to investigate the neural basis of the mind *in vivo* allowed insight into functions such as semantic processing and memory, and beyond these, into emotion and empathy. The concept that the human mind can be understood by examining exclusively cognitive functions has also undergone considerable criticism with increasing support for the view that mental abilities are defined also by emotions and motivations and are subject to intricate interactions with interpersonal, social, and technological environments. As a result, there is a recent change in emphasis from functional segregation to considerations of functional integration and to methods that allow the capture of dynamic large-scale operations in the brain. The possibility to observe structural connectivity, such as non-tasks specific large-scale distributed networks and (non-stimulus driven) self-organizing endogenous brain activity, reveal a neurocognitive organization that surpasses the classical modular and computational centered view of the mind (Fotopoulou 2013). There are still debated questions as how specialized regions of the brain are, how much of the mind is made up of specialized components, and importantly, whether the functional specialization is all or none or it is a matter of degree (Kanwisher 2010).

4. The Organizing Principles of Language

4.1. A Distributed and Hierarchical Language-Serving Network Structure

The brain regions that serve the capacity for language, collectively provide the semantic, syntactic, phonological, and pragmatic operations required for language comprehension and production. Previously developed language models such as the Wernicke-Lichtheim-Geschwind model that describe left-lateralized language functions give only limited view and interpretation of language processing in the brain. The language networks proved more extended than those defined (based on the earliest brain imaging research) as the classical language regions/modules (Broca and Wernicke areas) and include, as part of a prominent network-forming region, the lateral surface of the left frontal, temporal and parietal cortices and a number of other cortical, subcortical and cerebellar regions (reviewed in Fedorenko 2014). The division of labor between Broca's region (frontal cortex) and Wernicke's region (temporal cortex) does not correspond to language production versus language comprehension. Contrary to earlier reported functional distinction between language and other cognitive processes (Fedorenko et al. 2011), none of the language-relevant regions and none of the language-specific neurophyschological effects have proved language-specific as these are also triggered by other input

(Koelsch et al. 2002). For language as for other cognitive functions, the function-to-structure mapping as one-to-one correspondence is almost certainly incorrect (Petersson et al. 2012).

An analysis of the network structure for associative-semantic processing that also sub-serves many important cognitive functions, identified networks, sub-networks, and hub-status nodes with local clustering and discerned four major communities or sub-systems (Vanderberghe et al. 2012). In an attempt to identify cognitive elements involved in semantic circuitry and to capture the entire network (as opposed to individual functional components localized by methodologically limited neuroimaging studies and linear modeling) a group independent component analysis (ICA), providing both spatial and temporal information, identified a more complicated language distribution pattern, an elaborate network involving several additional spatially independent brain regions (eight task-related group ICA maps) sub-serving semantic decision. The authors of this work recognized the importance of information gained when analyzing cognitive functions in terms of underlying network structures, demonstrated that the semantic network comprises left, right, and bilateral sub-networks, concluded functional connectivity, and proposed a hierarchical cognitive model for semantic decision tasks, yet summarized this complex structure as modular with such broad functional (module) categories as verbal encoding and mental imagery and semantic decision making as sub-modules (Kim et al. 2011). In spite of revoking modularity in this study, the authors also propose each brain regions activation/function not as that of an isolated module(s), but rather as part of a network.

In these reports as in many others, the use of the terms module and/or modular have many connotations but few, if any, defined characteristics or specific criteria and critical attributes of modules (automatic, encapsulated, or neuronally specialized). Moreover, the terms modular and modules are often used as synonymous with various functional unit(s), however, with diversely defined content, and interchangeable with circuitry, networks, sub-networks and/or nodes, overlapping with the non-modular, dynamic network system concept.

4.2. Interactive Language-Related Abilities: The Role of Developmental Time

Interpretations of some the neurodevelopmental disorders have been viewed both as evidence for modular preservation of language or evidence for non-modular cognitive development (Brock 2007). Some of these disorders have been considered as conditions in which selective modules are impaired while others (language) appear normal (reviewed in Szalontai & Csiszár 2013). Williams syndrome with selective cognitive deficits but relatively preserved language, had been initially proposed as an example for modularity of language. Abundant subsequent evidence on alterations in brain development, language features, and interactions among cognitive capacities in Williams syndrome, however, point to contrary hypotheses. Genetic (Vanderweyer et al. 2012), and extensive developmental (Karmiloff-Smith 2007, 2012) studies revealed that in Williams syndrome, deficits profoundly affect synaptic activity, neuronal density, brain size and morphology, and functional connectivity. Among Williams-specific language features, spatial language deficits mirror deficits in nonverbal spatial cognition (Brock 2007). There is dissociation between grammatical rules and the mental lexicon in the production of inflected form for irregulars, as well as a correlation between performance on morphological tasks and phonological short-term memory (Pléh et al. 2003), and between semantic organization and reading levels (Lee & Binder 2014). Furthermore, the common basic auditory processing shared by prosody and music, is also affected (Don et al. 1999). While patients do process music and prosody through shared mechanisms, these are different from those in non-affected individuals (Martinez-Castilla & Sotillo 2014). The language phenotype in Williams syndrome, therefore, is not an indicator for a selectively spared (language) module, but the interactive result of multiple altered neural and cognitive processes during development (D'Souza & Karmiloff-Smith 2011).

In specific language impairment (SLI), traditionally considered as a single impaired function within a normally functioning brain with intact cognition, affected children proved to have lower performance IQ (Botting 2005), an overall increased radiate white matter, altered intrahemispheric and corticocortical connections (Herbert et al. 2004), asymmetry in their language-association cortex (De Fosse et al. 2004), and abnormal development of brain structures that constitute the procedural memory system (Ullman & Pierpont, 2005). This complex phenotype in SLI does not support the involvement of a putative single language module. A developmental model of SLI proposes a higher order of complexity: As language emerges from multiple abilities (attention sharing, speech pattern detection, phonetic and phonemic discriminations, speech processing speed), contribution from lower level deficits in any or several of these abilities during development can contribute to the phenotype (D'Souza & Karmiloff-Smith 2011). Comparisons of adult and developing cognition including language revealed a strong role for developmental time in both typical and atypical development as infant brains start out highly connected and only during development do the networks become increasingly specialized. Understanding the concurrent and longitudinal constraints can cast a broader light on the role of development and relationships between domain-general (attention) and domain-specific (vocabulary, letter knowledge, phonological skills) processes (Steele et al. 2012).

4.3. Task-Dependent Recruitment of Perceptual and Cognitive Processes

Language used to be widely considered as different and distinct from other cognitive functions with its own specific organizational principles. Subsequent views, while considered some elements or principles specific to language (basic primitives, features, syllables that allow to begin to distinguish different types of patterned stimuli), recognized that some characteristics may become grammaticised over time. Furthermore, it was also recognized that interactive constrains on linguistic performance and structure arise from cognitive constrains on learning and real-time processing (Christiansen & Chater 2008, Newport 2010).

These organizing principles, while characteristic of language, are not unique to language and also include organizing principles for other functional domains (motor behavior). The basic language organizing principles recognized as shared with other cognitive domains include computation of mutual information, entropy, conditional probability, contingency or predictiveness between elements and computed over hierarchical rather than linear distance in a recursive fashion (Newport

2011). Consequently, a unique combination of cognitive functions constrains language and the localization of these cognitive functions arises not from the inherent localization of cognitive modules but from the interactions of multiple cognitive and perceptual processes involved in a particular function (Newport 2010).

The cortical network thought to be domain-specific for language processing has been shown to also process musical information suggesting that this network is less domain-specific than previously believed (Koelsch et al. 2002). While considerable research supported the view that faces and words are subserved by independent neural mechanisms located in the ventral visual cortex in opposite hemispheres, a current study demonstrated a co-mingling of face and word recognition mechanisms. This co-mingling is unexpected from a domain-specific perspective, but follows as a consequence of an interactive, learning-based account in which neural processes for both faces and words are the results of an optimization procedure with specific principles and constraints. A comparison of pseudoword and face identification revealed that both stimulus types exploit common neural resources within the ventral cortical network (sublexical orthographic representations within the left ventral cortex and continuity of reading with other visual recognition skills) (Nestor et al. 2012). Thus cognitive functions appear to arise not from localized cognitive modules (language or face perception) but from the interaction of multiple perceptual and cognitive processes that underlie a particular task (Behrmann & Plaut 2013).

A meta-analysis of comparative functional anatomy data for speech comprehension and production in healthy adult brain including activation patterns for prelexical speech perception, meaningful speech, semantic retrieval, sentence comprehension, and incomprehensible sentences, identified association with the use of prior knowledge of semantic associations, world sequences, and articulation that predict the content of the sentence. Speech production activated the same regions as speech comprehension and additional areas for word retrieval, articulatory planning, the initiation and execution of speech, and suppression of unintended responses (Price 2010). The observation that prelexical and semantic processing of spoken words extend into anterior, ventral, and posterior directions suggested that the same speech input can follow multiple different pathways in which the location of activation is determined by the task demands similar to alternative strategies and dual routes, featuring both a direct and an indirect route (noted earlier for phonological processing (Heim 2005)).

Collectively, results from these studies (without even addressing functional connectivity of the activated regions), strongly promote the view that the cortical networks supporting language comprehension are dynamically determined by the task and context.

4.4. Domain-General Cognitive Control and Functionally Specialized Language Regions: Division of Labor

Abundant evidence shows that the network-forming language system with somewhat varying functional definitions (the lateral surface of the left frontal, temporal and parietal cortices and a number of other cortical, subcortical and cereballar regions) (Fedorenko & Thompson-Schill 2014), interacts with several cognitive systems including the visual system, social cognition supporting system, and, importantly, the working memory/cognitive control network/mechanisms (Fedorenko 2014). Some of these cognitive mechanisms are also known to be shared between language and other functions such as musical ability with highly similar structural and expressive features (Perrachione at al. 2013). The cognitive control network/multiple-demand system, also referred to as task-positive network, or fronto-parietal attention network (including parts of the dorsolateral prefrontal cortex, parts of the insular cortex, regions along the precentral gyrus, presupplementary and supplementary motor area, parts of the anterior cingulated, and regions in and around the intraparietal sulcus), is domain-general and flexible according to task demands and is implicated in a broad range of goal-directed behaviors. While its role in complex behaviors is not fully understood, the cognitive control system has been implicated in attention, working memory, cognitive control, structure building, timing/sequencing, attentional episodes, and conscious awareness (Fedorenko 2014). The cognitive control network is spatially and functionally distinct from the language system, however, it responds to linguistic input (both to pseudowords or processing of natural sentences) as much as the language system. While domain-general regions are engaged during language comprehension, dissociations from the language network also exist, indicating that the cognitive control regions may not be essential for language comprehension. Yet these interactions may still function in facilitating efficiency or speed of comprehension, providing workspace and alternative routes, or support predictive processing (Fedorenko 2014). The involvement of domain-general processes not only in language comprehension, but also in language acquisition is supported by the impairment of implicit sequence learning in SLI (Lukács & Kemény 2014).

5. Conclusion

The theory of modularity as a general principle with traditions in informatics has been historically applied to aspects of human cognition. Modularity views have been specifically influential in characterizing the organizing principles and structural and functional elements of language. Application of strict modularity, however, has been controversial as it precludes complex cognitive processes. With recent advances in brain activation analysis and systems biology interpretation of these results, the various and controversial definitions of what constitutes a module or a modular organization have sparked profound theoretical debates. Criteria for a cognitive modular system remain inconsistent and range from the definition that emphasizes topological modularity (dense intra- and sparse inter-connectedness) (Meunier et al. 2010), to a system with independently disruptable components (Menzies 2012, Sternberg 2010), to a property of being made up of self-contained and independently functioning parts (Shettleworth 2012), to the most basic use of the term as a capacity with functionally individuated input and output conditions (Menzies 2012).

Alternative models of modularity and alternatives to modular organization have been suggested in order to resolve some of these controversies and address the limitations inherent in modular system organization. Examples of these models include the functional and topological two-layer structure for information flow integrating modular clusters in a large scale (Gallos et al. 2012), introduction of the feature of transient variations in module domain-specificity (Anselme 2012), the modular-but-coupled theory (Friston & Price 2011), and the descent-with-modification modularity involving evolutionary changes (Marcus 2006).

Specialized modular mechanisms are considered innate, domain specific and isolated, in contrast, non-modular mechanisms are generalized, developmentally plastic, domain-general, and interactive, overall, more suitable to capture the working principles of the cognitive system. Theories that embrace generalized mechanisms take into consideration the effect of the variance of types of information (Ratliff and Newcombe 2008, Twyman & Newcombe 2010), the dynamic interaction of multiple neural and cognitive processes and the role of developmental time (D'Souza & Karmiloff-Smith 2011, Karmiloff-Smith 2012), the significance of connectivity patterns among functionally diverse and structurally distributed components of the central nervous system (Fotopoulou 2013), and equate specialization with local, narrow, and stereotyped functions, and general processes as those outside modularity (Barrett 2012) with a view of the cognitive architecture as a mosaic of modular and domain-general processes (Shettleworth 2012).

The large-scale operations in the brain have been interpreted with a recent emphasis on dynamic functional integration rather than segregation. Structural connectivity data, including distributed networks and endogenous brain activity, have revealed a neurocognitive organization that surpasses the classical modular and/or computational centered view (Fotopoulou 2013) and suggest a brain network organization that is determined by the type of actual cognitive processing (Meunier et al. 2010). A task-dependent continuum has been also noted between modular processing and varying degrees of modular and shared processing (Borowsky et al. 2007), and functional specialization has been formulated as a matter of degree (Kanwisher 2010). Higher-level and effortful cognitive processes proved to be linked to an anatomically distributed neuronal workspace architecture (Dehaene et al. 1998) that and enlist with increasing demand, a global network configuration with long-distance synchronizations and transient adoption of functional networks (Kitzbichler et al. 2011).

Theories of language organizing principles are profoundly shaped by the facts that none of the traditionally defined language-relevant regions and none of the neurophyschological effects proved language-specific, there is no one-to-one correspondence for cognitive function and structure (Petersson et al. 2012), the principles of language organization are not unique but shared with other cognitive domains (Newport 2011), and are the developmental time-dependent result of interactions of neural and cognitive processes (Karmiloff-Smith 2012), the language-serving network is more elaborate than previously anticipated and involves connectivity of several newly recognized and spatially independent brain regions (Kim et al. 2011), language associated cognitive functions arise from the interaction of multiple perceptual and cognitive mechanisms, cortical networks while not domain-specific for language (Koelsch et al. 2002) reflect a learning-based mechanism in which neural processes are the results of an optimization procedure (Behrmann & Plaut 2013), with a multiple pathways activation pattern determined by the task demands (Price 2010).

Together with the activation of the language network, the domain-general multiple-demand system also engages (Fedorenko 2014), and while not essential, it may facilitate efficiency or speed by providing workspace and alternative processing routes. As cognitive function of any region depends on the areas that it interacts with (Price 2010), the functional association of language regions can only be revealed in the context of their interactions with other brain regions and with the understanding of the task-dependent modulation of these regional interactions.

References

Anderson, Michael L. 2010. Neural reuse: A fundamental organizational principle of the brain. *The Behavioral and Brain Sciences* 33(4), 245–266; discussion 266–313.

Anselme, Patrick. 2012. Modularity of mind and the role of incentive motivation in representing novelty. *Animal Cognition* 15(4), 443–459.

Barrett, H. Clark. 2012. A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences of the United States of America* 109 Suppl 1, 10733–740.

Barrett, H. Clark & Robert Kurzban. 2006. Modularity in cognition: framing the debate. *Psychological Review* 113(3), 628–647.

Behrmann, Marlene & David C. Plaut. 2014. Bilateral hemispheric processing of words and faces: evidence from word impairments in prosopagnosia and face impairments in pure alexia. *Cerebral Cortex (New York, N.Y.: 1991)* 24(4), 1102–1118.

Benítez-Burraco, Antonio. 2012. The 'language genes'. In Cedric Boeckx, María del Carmen Horno-Chéliz & José-Luis Mendívil-Giró (eds.), *Language, from a Biological Point of View: Current Issues in Biolinguistics*, 215–262. Newcastle upon Tyne: Cambridge Scholars Publishing.

Berwick, Robert C., Angela D. Friederici, Noam Chomsky & Johan J. Bolhuis. 2013. Evolution, brain, and the nature of language. *Trends in Cognitive Sciences* 17(2), 89–98.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the stimulus revisited. *Cognitive Science* 35(7), 1207–1242.

Bishop, Dorothy V. M. 1997. Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology* 50(4), 899–923.

Boden, Margaret. 2006. *Mind as Machine: A History of Cognitive science*. London: Clarendon Press.

Boeckx, Cedric & Víctor M. Longa. 2011. Lenneberg's views on language development and evolution and their relevance for modern. *Biolinguistics* 5(3), 254–273.

Borowsky, Ron, Carrie Esopenko, Jacqueline Cummine & Gordon E. Sarty. 2007. Neural representations of visual words and objects: A functional MRI study on the modularity of reading and object processing. *Brain Topography* 20(2), 89–96.

Bouchard, Denis. 2012. Solving the UG problem. *Biolinguistics* 6(1), 1–31.

Brock, Jon. 2007. Language abilities in Williams syndrome: A critical review. *Development and Psychopathology* 19(1), 97–127.

Carroll, Sean B., Jennifer K. Grenier & Scott D. Weatherbee. 2004. *From DNA to*

Diversity: Molecular Genetics and the Evolution of Animal Design. Malden, MA: Wiley-Blackwell.
Chater, Nick, Florencia Reali & Morten H. Christiansen. 2009. Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences* 106(4), 1015–20.
Chen, Zhang J., Yong He, Pedro Rosa-Neto, Jurgen Germann & Alan C. Evans. 2008. Revealing modular architecture of human brain structural networks by using cortical thickness from MRI. *Cerebral Cortex* 18(10), 2374–2381.
Chen, Zhang J., Yong He, Pedro Rosa-Neto, Gaolang Gong & Alan C. Evans. 2011. Age-related alterations in the modular organization of structural cortical network by using cortical thickness from MRI. *NeuroImage* 56(1), 235–245.
Chomsky, Noam. 1968. *Language and Mind*. New York: Harcourt, Brace and World.
Chomsky, Noam. 1969. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1), 1–22.
Chomsky, Noam. 2011. Language and other cognitive systems: What is special about language? *Language Learning and Development* 7(4), 263–278.
Christiansen, Morten H. & Nick Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31(5), 489–509.
Christiansen, Morten H., Nick Chater & Florencia Reali. 2009. The biological and cultural foundations of language. *Communicative & Integrative Biology* 2(3), 221–222.
Clark, Alexander & Shalom Lappin. 2011. *Linguistic Nativism and the Poverty of the Stimulus*. Malden, MA: Wiley-Blackwell.
Coltheart, Max. 1999. Modularity and cognition. *Trends in Cognitive Sciences* 3(3), 115–120.
Corballis, Michael C. 2011. *The Recursive Mind: The Origins of Human Language, Thought, and Civilization*. Princeton, NJ: Princeton University Press.
Dąbrowska, Ewa. 2004. *Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh: Edinburgh University Press.
De Foss, Lies, Steven M. Hodge, Nikos Makris, David N. Kennedy, Verne S. Caviness, Lauren McGrath, Shelley Steele, David A. Ziegler, Martha R. Herbert, Jean A. Frazier, Helen Tager-Flusberg & Gordon J. Harris. 2004. Language-association cortex asymmetry in autism and specific language impairment. *Annals of Neurology* 56(6), 757–766.
Dehaene, Stanislas, Michel Kerszberg & Jean-Pierre Changeux. 1998. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences* 95(24), 14529–14534.
Don, Audrey J., Glenn E. Schellenberg & Byron P. Rourke. 1999. Music and language skills of children with Williams syndrome. *Child Neuropsychology* 5(3), 154–170.
D'Souza, Dean & Annette Karmiloff-Smith. 2011. When modularization fails to occur: A developmental perspective. *Cognitive Neuropsychology* 28(3-4), 276–287.
Elman, Jeffrey L., Eliyabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith,

Domenico Parisi & Kim Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *The Behavioral and Brain Sciences* 32(5), 429–448; discussion 448–494.
Fedorenko, Evelina. 2014. The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology* 5, 335.
Fedorenko, Evelina, Michael K. Behr & Nancy Kanwisher. 2011. Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences of the United States of America* 108(39), 16428–16433.
Fedorenko, Evelina & Sharon L. Thompson-Schill. 2014. Reworking the language network. *Trends in Cognitive Sciences* 18(3), 120–126.
Fodor, Jerry A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass: A Bradford Book / MIT Press.
Fodor, Jerry A. 2001. Doing without what's within: Fiona cowie's critique of nativism. *Mind* 110(437), 99148.
Fotopoulou, Aikaterini. 2014. Time to get rid of the 'modular' in neuropsychology: A unified theory of anosognosia as aberrant predictive coding. *Journal of Neuropsychology* 8(1), 1–19.
Friston, Karl J. & Cathy J. Price. 2011. Modules and brain mapping. *Cognitive Neuropsychology* 28(3-4), 241–250.
Gallistel, C. Randy & John Gibbon. 2000. Time, rate, and conditioning. *Psychological Review* 289344.
Gallos, Lazaros K., Mariano Sigman & Hernn A. Makse. 2012. The conundrum of functional brain networks: Small-world efficiency or fractal modularity. *Frontiers in Physiology* 3, 123.
Griffith, Paul E. 2007. Evo-devo meets the mind: Towards a developmental evolutionary psychology. In R. Brandon & R. Sansom (eds.), *Integrating Evolution and Development: Form Theory to Practice*, 195–225. Cambridge, MA: MIT Press.
Grodzinsky, Yosef. 2006. The language faculty, broca's region, and the mirror system. *Cortex* 42(4), 464–468.
Guérard, Katherine & Sébastien Tremblay. 2008. Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 34(3), 556–569.
Hagmann, Patric, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J. Honey, Van J. Wedeen & Olaf Sporns. 2008. Mapping the structural core of human cerebral cortex. *PLoS Biology* 6(7), e159.
Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598), 1569–1579.
Heim, Stefan. 2005. The structure and dynamics of normal language processing: Insights from neuroimaging. *Acta Neurobiologiae Experimentalis* 65(1), 95–116.
Herbert, Martha R., David A. Ziegler, Nikos Makris, Pauline A. Filipek, Thomas L. Kemper, Joseph J. Normandin, Heather A. Sanders, David N. Kennedy & Verne S. Caviness. 2004. Localization of white matter volume increase in

autism and developmental language disorder. *Annals of Neurology* 55(4), 530–540.

Irurtzun, Aritz. 2012. The present of UG. *Biolinguistics* 6(1), 112–123.

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

Kaltenbach, Hans-Michael & Jörg Stelling. 2012. Modular analysis of biological networks. *Advances in Experimental Medicine and Biology* 736, 3–17.

Kanwisher, Nancy. 2010. Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences of the United States of America* 107(25), 11163–11170.

Karmiloff-Smith, Annette. 2007. Atypical epigenesis. *Developmental Science* 10(1), 84–88.

Karmiloff-Smith, Annette. 2013. Challenging the use of adult neuropsychological models for explaining neurodevelopmental disorders: developed versus developing brains. *Quarterly Journal of Experimental Psychology (2006)* 66(1), 1–14.

Kim, Kwang Ki, Prasanna Karunanayaka, Michael D. Privitera, Scott K. Holland & Jerzy P. Szaflarski. 2011. Semantic association investigated with functional MRI and independent component analysis. *Epilepsy & Behavior* 20(4), 613–622.

Kitzbichler, Manfred G., Richard N. A. Henson, Marie L. Smith, Pradeep J. Nathan & Edward T. Bullmore. 2011. Cognitive effort drives workspace configuration of human brain functional networks. *The Journal of Neuroscience* 31(22), 8259–8270.

Koelsch, Stefan, Thomas C. Gunter, D. Yves v Cramon, Stefan Zysset, Gabriele Lohmann & Angela D. Friederici. 2002. Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage* 17(2), 956–966.

Lee, Cheryl S. & Katherine S. Binder. 2014. An investigation into semantic and phonological processing in individuals with williams syndrome. *Journal of Speech, Language, and Hearing Research* 57(1), 227–235.

Lorenz, Dirk M., Alice Jeng & Michael W. Deem. 2011. The emergence of modularity in biological systems. *Physics of Life Reviews* 8(2), 129–160.

Lukács, Ágnes & Ferenc Kemény. 2014. Domain-general sequence learning deficit in specific language impairment. *Neuropsychology* 28(3), 472–483.

Magen, Hagit & Asher Cohen. 2007. Modularity beyond perception: Evidence from single task interference paradigms. *Cognitive Psychology* 55(1), 1–36.

Makuuchi, Michiru & Angela D. Friederici. 2013. Hierarchical functional connectivity between the core language system and the working memory system. *Cortex* 49(9), 2416–2423.

Marcus, Gary F. 2006. Cognitive architecture and descent with modification. *Cognition* 101(2), 443–465.

Marr, David. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H. Freeman.

Martínez-Castilla, Pastora & María Sotillo. 2014. Pitch processing in children with Williams syndrome: Relationships between music and prosody skills. *Brain Sciences* 4(2), 376–395.

Menzies, Peter. 2012. The causal structure of mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43(4), 796–805.

Meunier, David, Renaud Lambiotte & Edward T. Bullmore. 2010. Modular and hierarchically modular organization of brain networks. *Frontiers in Neuroscience* 4(200).

Musiek, Frank E., Teri James Bellis & Gail D. Chermak. 2005. Nonmodularity of the central auditory nervous system: implications for (central) auditory processing disorder. *American Journal of Audiology* 14(2), 128–138; discussion 143–150.

Nestor, Adrian, Marlene Behrmann & David C. Plaut. 2013. The neural basis of visual word form processing: A multivariate investigation. *Cerebral Cortex* 23(7), 1673–1684.

Newport, Elissa L. 2010. Plus or minus 30 years in the language sciences. *Topics in Cognitive Science* 2(3), 367–373.

Newport, Elissa L. 2011. The modularity issue in language acquisition: A rapprochement? comments on gallistel and chomsky. *Language Learning and Development: The Official Journal of the Society for Language Development* 7(4), 279–286.

Perfors, Amy, Joshua B. Tenenbaum & Terry Regier. 2011. The learnability of abstract syntactic principles. *Cognition* 118(3), 306–338.

Perrachione, Tyler K., Evelina G. Fedorenko, Louis Vinke, Edward Gibson & Laura C. Dilley. 2013. Evidence for shared cognitive processing of pitch in music and language. *PLoS ONE* 8(8).

Petersson, Karl-Magnus, Vasiliki Folia & Peter Hagoort. 2012. What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language* 120(2), 83–95.

Pléh, Csaba, Ágnes Lukács & Mihály Racsmány. 2003. Morphological patterns in hungarian children with Williams syndrome and the rule debates. *Brain and Language* 86(3), 377–383.

Price, Cathy J. 2010. The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences* 1191, 62–88.

Ratliff, Kristin R. & Nora S. Newcombe. 2008. Is language necessary for human spatial reorientation? Reconsidering evidence from dual task paradigms. *Cognitive Psychology* 56(2), 142–163.

Shettleworth, Sara J. 2012. Modularity, comparative cognition and human uniqueness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367(1603), 2794–2802.

Simon, Herbert A. 1969. *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Spelke, Elizabeth S. & Katherine D. Kinzler. 2007. Core knowledge. *Developmental Science* 10(1), 89–96.

Sporns, Olaf. 2010. *Networks of the Brain*. Cambridge MA: MIT Press.

Sporns, Olaf. 2013. Network attributes for segregation and integration in the human brain. *Current Opinion in Neurobiology* 23(2), 162–171.

Sternberg, Saul. 2011. Modular processes in mind and brain. *Cognitive Neuropsychology* 28(3-4), 156–208.

Szalontai, Ádám & Katalin Csiszár. 2013. Genetic insights into the functional elements of language. *Human Genetics* 132(9), 959–986.

Thomas, Michael & Annette Karmiloff-Smith. 2002. Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *The Behavioral and Brain Sciences* 25(6), 727–750; discussion 750–787.

Traxler, Matthew J., Megan Boudewyn & Jessica Loudermilk. 2012. What's special about human language? The contents of the "narrow language faculty" revisited. *Language and Linguistics Compass* 6(10), 611–621.

Vandenberghe, Rik, Yu Wang, Natalie Nelissen, Mathieu Vandenbulcke, Thijs Dhollander, Stefan Sunaert & Patrick Dupont. 2013. The associative-semantic network for words and pictures: effective connectivity and graph analysis. *Brain and Language* 127(2), 264–272.

Vandeweyer, Geert, Nathalie Van der Aa, Edwin Reyniers & R. Frank Kooy. 2012. The contribution of CLIP2 haploinsufficiency to the clinical manifestations of the Williams-Beuren syndrome. *American Journal of Human Genetics* 90(6), 1071–1078.

Wojciulik, Ewa, Nancy Kanwisher & Jon Driver. 1998. Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology* 79(3), 1574–1578.

Ádám Szalontai
Hungarian Academy of Sciences
Research Institute for Linguistics
Benczúr utca 33.
1068, Budapest
Hungary
adam.szalontai@gmail.com

Katalin Csiszár
University of Hawai'i
John A. Burns School of Medicine
1960 East-West Road, Biomed T415
Honolulu, HI, 96822
USA
kcsiszar@aol.com

What Is Un-Cartesian Linguistics?

Wolfram Hinzen

Un-Cartesian linguistics is a research program with the aim of rethinking the nature of grammar as a domain of scientific inquiry, raising new questions about the constitutive role of grammar in the organization of our (rational) minds and selves. It reformulates the 'Cartesian' foundations of the modern Universal Grammar project, shifting emphasis away from the study of a domain-specific 'innate' module separate from thought, to the study of a *sapiens*-specific mode of cognition conditioned by both grammatical and lexical organization, and thus a particular cognitive phenotype, which is uniquely also a linguistic one. The purpose of this position paper is to introduce and motivate this new concept in its various dimensions and in accessible terms, and to define the 'Un-Cartesian Hypothesis': that the grammaticalization of the hominin brain in the evolutionary transition to our species uniquely explains why our cognitive mode involves a capacity for thought in a propositional format.

Keywords: (un-)Cartesian linguistics; concepts; meaning; reference; truth; universal grammar

1. Cartesian Linguistics

Non-human primates listening to sound sequences governed by an artificial 'grammar' can extract formal rules by which such sequences are formed (e.g., Wilson *et al*. 2011, Rey *et al*. 2012). Grammar proper, on the other hand, as used by humans to structure and convey propositionally meaningful information, remains a barrier that no other species has crossed (Tomasello 2008). What then is the significance of grammatical organization for our particular cognitive phenotype?

In theory, grammar could be an arbitrary convention, carrying no particular significance for the genesis of our species-specific mode of cognition. Consistent with that, a popular conception identifies different domains by the terms 'thought' and 'language', with the former usually regarded as primary and the latter as an *expressive tool* (though on some Neo-Whorfian views, a particular language may also *influence* thought, which thereby comes to 'depend' on language). On this conception we first *think* (or intend to say something), and then we

I thank two anonymous referees for their extensive and very helpful comments. I also thank Michelle Sheehan for commenting extensively, and to Peter McKenna for much recent discussion.

say what we think. *How* we say it is a contingent cultural convention. Speaking grammatically is then no more or less of a mystery than that we follow social norms. No special explanatory riddle arises, and there could be no such thing as a 'science of grammar' unveiling them.

Another option, however, is that thought of a *sapiens*-specific variety and language are inherently integrated as two sides of one coin. Consistently with this other intuition, language without thought expressed in it would be a parody, and thought not expressible in language would not be thought of the same kind. Moreover, since no one would want to identify language with a system of pronunciation, and it is clear that language *is* (almost continuously in our waking lives) used internally for purposes of thought as well, in addition to being used for communication, it is a natural suggestion that the cognitive mechanism generating human-specific thought and those generating language should be the same. That they are is the Un-Cartesian thesis, and Un-Cartesian linguistics is the research program seeking to determine to what extent it holds.

What is meant by 'language' here? Very crudely, human language exhibits two major principles of organization: the lexicon (the words) and the grammar (relations between the words). Both correlate with the existence of different kinds of meanings, which structure our cognitive space insofar as it is human-specific. The Un-Cartesian suggestion is that language thereby becomes a *principle of cognitive organization* and no *separate* theory of human-specific thought is required. The naturalization of grammar as a scientific domain, on this view, does not proceed by de-semanticizing grammar into a purely formal domain of 'syntax', but by re-describing it as a *cognitive principle* from which the major dimension of human-specific thought—reference, predication, and truth-conditional content—fall out.

On what will be called the contrasting 'Cartesian' view here (Arnauld & Lancelot 1660), thought is rational and universal by definition, while language in the ideal case 'mirrors' thought sufficiently so as for grammar to become a 'science' (i.e. be 'rational' or 'universal' grammar).[1] No explanation is here offered for why thought of this kind exists, and language in particular cannot now be invoked as the mechanism to explain it. An answer to the above question of what cognitive change grammar induces might now just be: none. The exact *same* meanings and thoughts are available, whether or not we express them in language. Put differently, although language is a unique 'window' into the mind, it does not *constitute* what is seen through that window, which is accessible more directly non-linguistically through introspection, turning language into a kind of 'detour'. The mind, in short, while perhaps not *externalizable* without language, would be the same if it didn't cognize grammatical relations: for it has its own principles of organization—perhaps those of logic or perhaps it can engage in processes of 'representation', 'social cognition', or 'mentalizing', regarded as independent of language.

Chomsky's 20th century version of the 'science of language' picked up the banner of 16th century Cartesian linguistics (Chomsky 1966), yet it took a stance

1 A historically interesting question, nonetheless, is how 'Cartesian', in these terms, Descartes himself ultimately was; see in particular Cottingham (1998).

on language and thought different from the early Cartesian one, if only for reasons of methodological caution. If language is to become subject to scientific inquiry as a natural object in its own right, the suggestion is, the proper object of study is to be 'I-language', a formal computational system internal to the brain interfacing with, but different from thought, and in particular seen in abstraction from the use of language for such purposes as reference. A 'formal' and 'internalist' stance was thus adopted, with 'syntax' viewed as separated from 'semantics' and forming the core of the enterprise.[2] Today, the standard view remains that language is divided from 'thought' by an 'interface', though an important idea has gained ground recently according to which thought might be *optimally* represented or expressed by language, making this interface particularly *tight*.[3]

Since the Un-Cartesian research program seeks to make plausible that grammar yields—rather than (optimally) 'expresses'—a form of meaning that would not exist without it, there is *no* 'interface' between language and (a *sapiens*-specific mode of) thought: The organizational principles of grammar *are* the ones that define a thought system unique to our species, explaining its apparent absence in species that cannot fathom grammar. This is a claim about the nature of grammar as a domain of scientific inquiry, and as such it is ipso facto a *universal* claim: a claim about what grammar, *as such* or by its nature, happens to be. Grammar, on the other hand, is *not* said to be universal because it is 'innate', let alone 'modular'. Universal grammar is also not here defined as the study of a genetic endowment for a formal-computational system underlying syntax, but as the study of a cognitive type. Nor does the program exclude cross-linguistic variation, the study of which is central to this research program, insofar as grammar is nowhere manifest except in the languages spoken around the globe. The claim, rather, is that such variation does not *affect* the organizational principles of grammar that are those of universal and *sapiens*-specific thought.[4]

Neither does the program take issue with the desire of some to apply the term 'thought' to pets, insects, or computers, or with the view that we can think in images or feelings, or the view that there is a generic notion of 'mental repre-

---

2 See Chomsky (2000). In a footnote to section 9.3 of *Syntactic Structures*, Chomsky remarks that "much of our discussion can be understood as suggesting a reformulation of parts of the theory of meaning that deal with so-called 'structural meaning' in terms of the completely nonsemantic theory of grammatical structure" (Chomsky 1957/2002: 103). The notion of 'grammatical meaning' developed below could be regarded as a further development of this early notion of 'structural meaning'. I thank an anonymous reviewer for clarifying this paragraph.

3 Chomsky (2007), Berwick & Chomsky (2011). It remains the case in this framework that human thought—our species-specific cognitive type—is not the subject of the inquiry. The subject is a formal computational system, whose functioning as a thought system depends on its association with language-external systems, which in particular account for reference and truth (see also Pietroski to appear). A 'Cartesian' view is also maintained in many discussions of the evolution of language today, where the 'evolution of thought' is barely thematized. A Cartesian assumption is even maintained in a putatively non-Cartesian framework such as Davidson (2004), where thought is said to *depend* on language. For the propositionality of such thought does not, on this view, strictly come *from* language, but from the embedding of language in an interpretative infrastructure in which rational agents try to make sense of the propositional attitudes of other agents.

4 This includes all variation of the kind documented in, say, Evans & Levinson (2009).

sentation' that applies to many species, independently of conscious apperception or intentional modes of reference. The concern is solely with the explanation of the fact that, at a species level, humans think differently from any other animal—differently, indeed, even from other species within the genus *Homo* (Tattersall 2008), judging from their archaeological records, not to mention other primates. Since this difference must fall out from something, the hypothesis that it falls out from grammatical organization competes with others. The claim here is that generic notions such as domain-generality, 'recursion', 'social cognition', or 'theory of mind' will turn out to be too unspecific, whether in conjunction or in isolation. That is, non-linguistic explanatory constructs are either insufficient or turn out to be language-dependent.

2. What *Is* Grammar?

On the one hand, grammar is one of the most immediate aspects of human experience. We cannot as much as open our mouths when uttering a word in context without such words exhibiting grammatical properties. They will come out as nouns or verbs, predicates or subjects, modifiers or arguments, etc., and they never *lack* such properties.[5] Most words, moreover, don't come alone and when they combine with others they combine in grammatical ways, largely so as to form *sentences*. These are structures unique to our species and, somewhat mysteriously, capable of truth and falsehood.[6] Because of this, they are also capable of conveying *knowledge*: If they are true, the world is as the sentences say, independently of whether we believe this, for truth is independent of belief.

Grammar is also present at least in fragments when we think silently, and although we can speak ungrammatically at times as our attention slips, we cannot violate the laws of grammar: It not only takes extraordinary efforts to try to speak ungrammatically, but we could not convince ourselves that 'the with happy no' is grammatical, or that 'John slept' is not, any more than we could convince ourselves that 2+2=5 is an arithmetical truth, while 2+2=4 is not. Grammar is present, in a reduced form, even in interjections and fragments (Merchant 2004), and it is hard to imagine the state of mind in which it would be missing altogether, though states of very severe thought disorder (manifest clinically as 'word salad') and catatonia may approximate the phenomenon (McKenna & Oh 2005).

On the other hand, grammar is also completely invisible. What we primarily see or hear in a language is *words* (or, from a linguistic point of view,

5 To illustrate, Holmberg (2013) interestingly shows that even answers such as 'Yes' or 'No' are syntactically complex expressions.

6 One reviewer points out that this line of reasoning is based on the *assumption* that human languages generate truth-evaluable expressions—which is controversial (see Pietroski to appear). What is not controversial and sufficient for present purposes is that some things are true and others false: e.g., 'Snow is white' vs. 'Snow is red'; and that the former proposition is true if and only if snow is white; and that 'the snow's colour' is not true or false, showing that truth-evaluability correlates with grammatical distinctions. That two systems (grammar and something else) account for the emergence of truth-evaluable structures is a possibility, but uneconomical if and when language is sufficient.

morphemes). The particular grammatical *relations* that hold them together, like the relation of predication holding between 'dog' and 'bite' in the sentence *Dogs bite*, are not there, physically speaking. What we see is 'Dogs' and 'bite'. Nothing in the visual (or auditory) appearance of these words, nor their order, even accounts for one being a *noun*, the other a *verb*, i.e. their status as particular 'parts of speech'; nor for one being syntactically a Noun Phrase (NP), the other a Verb Phrase (VP), i.e. for their phrasal status; nor for one functioning grammatically as a *referential* expression (referring generically to dogs), the other as a *predicate*, which together yield a *truth value*.

Sometimes little words like *do* or *is*, as in *Dogs do bite* or *This is a dog* can betray the existence of grammatical relations at a morpho-lexical level. Yet neither 'do' nor 'is' have much lexical content and no such lexical overtness is required for a grammatical relation to obtain. Thus the word 'is', which is grammatically obligatory in *This is a dog*, disappears in a structure like *I consider [this a dog]*, in which the grammatical relation of predication between 'this' and 'a dog' is otherwise the same (and in many languages, copulas such as 'is' are missing altogether). Grammatical relations are purely *structural* in this sense. While we can sometimes glean grammar from little words that may signal it, grammar is simply not about words, and we will argue later that it cannot be reconstructed from the content or feature specifications of words either. It is an independent structuring principle, with cognitive effects.

3. Words and 'Concepts'

Words *have* been the focus of the analytic philosophy of language, where the theoretical interest is meaning and more narrowly word meaning and where grammar has barely been a topic of inquiry in its own right.[7] Thus a typical introduction (e.g., Lycan 2008) will begin from presenting the 'base theory' of meaning, which simply says that the meaning of a word is its referent, a non-linguistic entity. Complications then ensue, and the more elaborate 'Fregean' model is discussed, whereby a word also has (and again lexically) a 'sense', or the 'Kaplanian' model, where it also has a 'character'. The content of grammatically complex expressions then has to be reconstructed from that of words, which is the origin of the idea of *semantic compositionality*, to which we return. The lexicon, then, exclusively, injects 'content' into grammatical configurations, grammar has no meaning of its own, and meanings are non-linguistic entities.[8]

In psychology, too, words are taken to express putative non-linguistic 'concepts', which are conventionally referenced through capitals (DOG, MAN,

---

7 Historically, the project has been that of a pure 'logic of thought', with language as no more than an analytic tool (typically regarded as a deficient one).

8 In fact there are six major philosophical textbook answers to the question of what 'meaning' is: (i) a non-linguistic mental 'idea', (ii) a mind- and language-independent 'proposition', (iii) 'reference' (in a purely semantic or causal sense), (iv) a convention of use, (v) beliefs, (vi) nothing (meaning does not exist). Strikingly, in every single of these approaches, grammar is systematically irrelevant to the existence of propositional meaning (see Hinzen 2006, 2007, for discussion).

etc.) (e.g., Fodor 1998). The meaning of words ('language') resides in concepts ('mind', 'thought'), and meaning again pre-exists language and is independent of it. However, concepts such as DOG or MAN identify words such as *dog* or *man*, and a word in any given language *is* a lexeme only when viewed *together* with a given meaning or concept that it encodes. If so, DOG is nothing *other* than the English word *dog*, viewed together with its meaning but in abstraction from its *sound* (which unlike meaning is rarely regarded as an inherent property of concepts by concept theorists). What then might we be talking about, if we talk about concepts and *don't* mean words?

Concepts of some kind can be found in infants that do not yet articulate language, and in non-human species (Carey 2009, Gallistel 1998). Yet it is clear and widely accepted that the 'concepts' available to non-linguistic and pre-linguistic beings are not the same.[9] The question therefore again arises whether speaking of concepts in a narrowly human sense, on the one hand, and of words, on the other, really makes an empirical distinction. For every concept that we know or have, there is a word that identifies this concept, and that is identified by it. If there was no word or phrase to identify it, how would we know a concept existed and what that concept was?

Beyond the word level, the theoretical utility of the non-linguistic term 'concept' becomes even less clear. Do sentences 'express concepts', too? 'Complex concepts', perhaps? That we should have simple 'concepts' such as DOG or MAN makes good pre-theoretical sense: We have such 'concepts' in the sense that we know, in a general if vague sense, what kind of things these are, often based on their perceptual features: Thus we know that the former bark, the latter don't; the latter talk, which the former don't, etc. That is, concepts in this sense connect with our *semantic memory* (Tulving 1972). One could not say the same of the (typically phonologically 'light-weight') words and morphemes like 'the', 'to', 'do', 'is', 'has', '–ed' or '–s', which express grammatical relations in sentences: The language-independent notion of 'concept' has no grip on these items. The same applies to the 'silent' words posited by classical generative linguistics, like the implicit PRO subject of the embedded verb *blow* in sentences like *John likes [PRO to blow his nose]*, or the trace *t* of the moved *wh*-expression in *What did you blow t?*. In fact, we only need to go into the domain of lexical verbs, and we will see our intuitions wavering as to whether a word like 'bite' expresses a 'concept' (of an object that is an 'event') or perhaps rather a relation *between* concept pairs, such as DOG and MAN or DOG and SAUSAGE.

If it comes to a proper grammatical relation such as predication, it seems that we have no 'concept' of predication at all. If one was posited, how would it relate to our grasp of the grammatical relation itself that holds between the concept functioning as the subject and the concept functioning as predicate? Why is understanding the grammar not sufficient? In general, how we *count* putative

9 In particular, as Carey stresses, conceptualization in non-linguistic beings is still continuous with perception; it also remains stimulus-controlled, non-combinatorial, and non-propositional, and concepts are not employed for purposes of intentional reference, with a capability to refer to anything at all no matter how remote in space and time (Fitch 2010: 187–194). Arguably, no non-human primate ever *learned* any human concept either (Pettito 2005). See further DeVilliers 2007, in press, Penn *et al*. 2008, Terrace 2005).

non-linguistic concepts, which are largely inaccessible without language, is opaque. We could ask whether *Shem kicked Shaun* and *Shaun was kicked by Shem* express 'the same concept'. Yet it is their grammatical identities and differences that capture precisely in what way they are identical and in what other ways they are different, and no appeal to a realm of 'concepts' appears to be required. We could ask whether *John sleeps* and *John is sleeping* express 'a different concept'. But when asked which difference this is, it is not clear whether we could do better than recapitulating the grammatical distinction between the two kinds of predicate involved, verbal and gerundive, respectively, and the meaning differences that this distinction entails.[10]

Rather than saying we have 'concepts' of grammatical function words, grammatical constructions, or parts of speech distinctions, it may make more sense to say that, in addition to concepts in the sense of a form of semantic memory, we have *grammar*, and that it is for this reason that we understand constructions beyond the level of content words. If concepts are words, and over and above words there is only grammar, what does the term 'thought' even add?

4. Meaning in Language

A typical traditional *semantic network* representation of semantic memory will indeed represent concepts in the form of words, which form the nodes of a network (e.g., BIRD, ANIMAL; see Collins & Quillian 1969, Baddeley 1990). The nodes are viewed as *categories*, each stored along with a number of associated *properties*, which reflect our *general and shared knowledge* about members of the category. Thus BIRD might be associated with HAS WINGS, CAN FLY, and HAS FEATHERS. As one node is activated, activation spreads to associate nodes. Connections can come with different degrees of 'strength', reflecting statistical data about co-occurrence of experiential features in our environment, and there can be typicality effects (e.g., CANARY > OSTRICH in the category BIRD). Such a network can also be structured 'hierarchically' in the sense that the node represented by CANARY includes all the properties typically associated with BIRD, which in turn includes all the typical properties of ANIMAL. Because of its structure, a network of this nature can verify propositions against knowledge maintained in long-term memory, such as 'A canary is a bird' or 'A canary can fly'.[11]

A semantic network of this kind appears as a possible model for how our *lexical knowledge*—meaning at the level of words, without grammar—might be structured. Semantic associations will bind various words together and e.g. entail

---

10 As M. Sheehan notes, the assumption here is that at the level of grammatical meaning, the same meaning differences can be expressed in all languages, though they need not be expressed with the same lexical resources, which are variable. Mandarin can expresses Tense and Aspect distinctions, but lexicalizes only the latter.

11 Neurobiologically, semantic memory is spread out widely across the brain, with reliable activations throughout the left temporal and parietal heteromodal cortex (Binder & Desai 2011, as consistent with temporal atrophy seen in semantic dementia, which specifically affects semantic memory (Hodges & Patterson 2007). It is not, that is, in one particular locus or a modular notion separate from where 'language' processing might take place.

that DOG is highly associated with BARK but not with FEATHERS, while the reverse is true for BIRD. These associations are not *grammatical relations*: Grammar is what a semantic network *lacks*. A semantic network in this sense thus is a model of how much meaning we can obtain in the absence of grammar.

Incipient grammar on the other hand *is* visible in 'property terms' such as 'CAN EAT' or 'HAS WINGS', where 'can' and 'has' effectively express grammatical relations between subjects (identifying categories or instances of them) and predicates (depicting properties attributed to these categories or instances). The notion of 'property' itself appears as a grammatical term in disguise, since whether the concept 'WING', say, functions as a property is something that we can in general only tell from its grammatical position in a sentence. Thus a sentence might be either about a wing (saying that it is broken), or about a bird (and say that it has wings). It is grammar, too, that makes a distinction between *A collie is a dog* and *A dog is a collie*, only one of which is true, and which differ merely with regard to which indefinite noun phrase plays the role of the grammatical predicate and which plays the role of the referential expression (subject). Why could the mechanism by which our brain *connects* concepts productively and propositionally, not just *be* grammar?[12]

Evidence that it must be, at some point, is that grammatical relations between words are *sui generis* and crucially *independent* of the statistical and hierarchical organizational principles of semantic memory, opening up a richer and different semantics. Thus the utterance *This dog has feathers* could be true (though surprising), even if dogs generally lack feathers. *Birds fly* could be true, yet not necessarily because most birds fly (they might not, for some reason). Crucially, *Birds fly*, *Birds often fly*, and *Most birds fly* are three grammatically distinct sentences, and they express three distinct propositions or thoughts. *A bird could have two heads* and *Necessarily, 2+2=4* might both be asserted, yet the former does not mean that I have found or expect to find such a bird, and the latter does not mean that I have never found that 2 and 2 made 5. Truths *are* often asserted on the basis of experience, but truth is not an experiential category. We know that if what we say is true, it is not true because we believe it is: It is true because of what it says and the world is as the sentence says: an objective matter.

If thinking was a matter of semantic associations only, there would be no subjects and predicates, no topics and comments, no presuppositions and assertions, no truth values. It might also be that we would be lacking another memory: episodic memory, which Tulving (1972) classically distinguished from semantic memory. The former refers to *individually and first-personally experienced* memories specific to time and place (e.g., *I was bitten by a dog yesterday*), with a connection to a notion of 'what happened to me' (re-lived experience) and the circumstance of its acquisition. Such memory is necessarily conscious in the sense that it only exists when retrieved. I may know for a fact that a dog bit me yesterday, yet this is episodic memory only when I *re-live* the experience in an episode of thought. Else it is a form of 'personal' memory, which comprises facts

12 A few years after Collins & Quillian (1969), Collins & Loftus (1975) introduced a revised network model exhibiting a range of different 'links' such as 'IS', 'CAN' and 'CANNOT', viewed as independent of the categories themselves. An interesting possibility, however, is that a better name for the exact 'links' required would simply be 'grammar'.

about me of the sort that other persons can know, too (Renoult *et al*. 2012). In contrast, I *know* what 'dog' means, and that dogs are mammals or bark, whether or not I happen to use this word on a particular occasion, or whether or not I am engaged in a particular episode of conscious thought about dogs. Moreover, that knowledge is widely shared.

We might put this difference by saying that episodic memory is by its nature *indexical*, whereas semantic memory remains purely *conceptual*. Since the former involves an event specified for its time and place and in relation to the first-person experiencing self, while semantic memory is crucially impersonal and generic, the two kinds of memory are necessarily expressed in a grammar of a different kind, even if the same lexical concepts are involved. The former will involve indexical reference, referential specificity, reference to an event specified for time and place, and grammatical Person distinctions as relevant for a first Person perspective. The latter will be grammatically much simpler.

This difference in grammar could be an accident – a fact merely at the level of the external 'expression' of the memories in question. Yet it is not clear why it should be. If episodic memory involves a psychological process of 'scene construction' (Hassabis & Maguire 2007), in particular, a generative system is required that provides for the relevant constructional principles, can apply them productively, and capture the right distinctions. Why should grammar *not* be considered in this regard, when key distinctions involved *are* grammatical and referential rather than merely conceptual in nature? If it is not, another system is needed that will provide for the exact same distinctions at the level of non-linguistic 'thought'. It is not clear which system this might be, and whether it would not have to re-state grammatical distinctions.

The considerations applying to memory apply to thought as such. We retrieve our concept DOG for much of our dog-related activities, like planning to buy a dog, avoiding one seen in the distance while running in the park, recalling playing with one yesterday, or regretting that we will never own one. Yet in any such mental episode this concept is retrieved along with others, and the relations between them will be different across such episodes. A system is needed that can specify these distinctions as fine-grained as is required, and construct representations of the relevant scenes, desires, or thoughts. Grammar is such a system, and as such it comes for free. It is clearly one way—and perhaps the only empirically known way—in which concepts *can* be systematically combined creatively and so as to generate an infinity of possible thoughts about either possible or actual worlds and situations. Again, since language is clearly not only there to talk or communicate, it makes sense that we would use such a system for purposes of thought and memory as well.

5. Reference from Grammar

To summarize so far, grammar is invisible yet powerful in establishing relations between words. These relations seem quite unique and are not of a generic 'associative' kind. They go with phenomena such as reference and truth that associative principles do not entail. All this could be an accident, and all the meaning there is could have always existed, independently of language. But this becomes unlikely when non-linguistic substitutes for linguistic notions of meaning—e.g., 'concept', 'memory', 'propositions'—do not capture the right distinctions, short of re-stating the grammatical ones that exactly fit the bill.

I will call the specific kind of meaning that goes with grammatical organization *grammatical meaning* in what follows. With grammar, we can refer and predicate, and the result is propositional truth. Along with truth comes another distinctive human privilege: the making of mistakes.[13] That is, a guarantee for real-world content, which is still there in the case of perception (if we ignore hallucinations), is lifted, and we almost never know for sure whether what we say is true is indeed true: With propositional truth, correctness is not anymore a matter of what we know, nor of what others know, but of how the world is. Adding a grammatical system to a system of concepts does not then result in more concepts, but instead something entirely different: a capacity for using a resource of stored concepts (semantic memory) for a new purpose, namely referentiality, which does not exist in the same form in animals (Fitch 2010). That is:

(1) a. WRONG EQUATION: concepts + grammar = more concepts
 b. RIGHT EQUATION: concept + grammar = reference

As grammatical complexity unfolds, reference gets specified in a more fine-grained way, and a system of *formal-ontological distinctions* arises in terms of which kinds of objects are being referred to: In particular, referents can formally be substances, objects, events, propositions, and facts, depending on which grammatical complexity is involved. So grammar maps concepts onto referents with a formal ontology, and referents are not new concepts. If the world is nothing other than the totality of facts (Wittgenstein 1922), and facts as a particular formal ontology of reference arise with grammar, grammar gives us a sense of 'the world' or of what there is.

That reference is an instance of grammatical meaning appears independently empirically correct: My lexical knowledge of what DOG means cannot distinguish between the dog I saw yesterday, this dog I see as opposed to that one, some dogs walking through the streets, all dogs, dog-meat, or the species 'dog'. Referential distinctions of this nature are a kind of meaning that words can in principle not encode. Reference is an instance of grammatical meaning in this sense. No word, as such, and not even any complex word (compound), can be used to *refer* to particular objects or events in the world as placed in space, time, and discourse, and in relation to our own personal

13 All animals can respond to stimuli *non-adaptively*, which at times will cause death. But death does not prove falsehood, as history documents. Nor does it show that a creature will be capable of making mistakes. It is also not clear whether having a notion of truth can count as adaptive. A notion of mistake has been applied to perception as well in some philosophical accounts, yet perception unlike language is stimulus-controlled: We cannot will ourselves into perceiving the face of the person behind us, without turning our heads; and if we perceive it as a face, we cannot will ourselves to perceive it as a car. By contrast, we make decisions on what to believe and assert.

experience or self.[14] The point can be made in very simple terms: 'dog' is a lexical item; 'the dog' is not. The former cannot refer; the latter can (in the right grammatical context).

In any act of reference, then, grammar gets involved imposing specific constraints. Thus at the nominal level, grammar requires a decision as to whether the reference to an object is abstract/generic, indefinite/quantificational, definite, rigid, deictic, or personal. At the sentence-level, every assertion requires a specification of grammatical Tense (finiteness), which has the effect that referential specificity for an event referred to is obtained. Thus in *John is sleeping*, the present tense marking on 'is' is used to locate an event of John sleeping relative to the act of speech: The event is indexically *co-located* with that act, as and when it takes place, and captured as ongoing, outlasting the speech act.

Every such act is further constrained to provide 'new information', expanding our knowledge beyond memory, and to anchor its content in the context of the utterance and with respect to the epistemic state of the subject. The act is also conscious as and when it takes place, where such consciousness is inherently first-personal: It would not be enough to know it for a fact that *Peter* (which is grammatically 'third Person') is talking, if I am Peter, and talk. What I have to know is that *I* (the grammatical 'first' Person) am talking, here and now, making this or that claim about the world, to this or that other (grammatically 'second') person. It is from the *lack* of such knowledge—the breakdown of the deictic frame—that we predict the occurrence of the 'nuclear symptoms' of schizophrenia: one's thoughts heard as spoken out loud, or inserted by an outside force (Crow 2008, 2012).

The first instance of reference in this sense is index-finger declarative pointing occurring universally in humans around the first birthday, often accompanied by verbalizations of the first words. In the course of the next one and a half years in development, and crucially in both the signed and spoken modalities (Pettito 1987), such units start being fully grammaticalized: The same kind of deictic reference can now proceed in the absence of non-linguistic gestures in the visual modality and any particular stimulus processed online as the word is used, as in an act of reference to 'the dog I saw yesterday'.

In units of referentiality like this, deictic reference comes from four things: (i) the intuitively 'meaningless' word 'the', the *determiner* (D); (ii) a lexical *description* that, in the context of the utterance, will typically involve a (complex) concept uniquely true of one particular relevant dog in the context of speech; (iii) the *grammar* creating a unit consisting of the determiner plus the description, and connecting it with the rest of the structure, and (iv) a *time* and *location* in which the speech act takes place, which sets up a particular space for deictic reference, in which objects fall under descriptions from the speaker's and hearer's perspectives.

(iii) is required, since to answer the question of whether a given phrase is

---

14 Thus a compound like 'dog hater' only applies generically to people hating dogs, but not to a particular episode of a particular person hating a particular dog. We can utter 'dog' and this may pick out a particular dog, but only when the utterance is accompanied by a deictic gesture—of the sort we find in infants in the one-word stage around the first birthday. It is the word together with the gesture that yields a unit of reference in that case.

referential or else a mere predicate we have to look at its grammar, and nothing else will tell. Thus, for example, in *I wished her husband wasn't her husband* (from Lycan 2008), the exact same phrase 'her husband' is referential in its first occurrence, where it is the grammatical subject, and predicative in its second, where it is the predicate. The sentence says that I wish that a certain person I refer to under the description 'her husband', does not fall under this description. Reference in the case of a structure like D+NP thus does not come from the lexical content involved in the description (in the NP), nor even from a definite determiner like 'the', but from the determiner and the NP co-occurring as a single unit of grammatical organization, in the right grammatical relations, in a speech context.[15]

With grammar, then, a human-specific *deictic frame* is set up: a logical space in which we can think and act rationally. In this frame, nominal phrases serve to place a given concept in space, verbs to place it in time, clauses to place it in discourse—all in relation to the speaker's first-personal self, the center of the deictic space. Before grammar, a defined relation to the world is not there. Reference in a human-specific sense is in this way an evolutionary riddle and a profound explanatory *problem*. It is not the *solution* to the problem of meaning, as a term like 'the referential theory of meaning' (Lycan 2008) suggests. Reference is what *poses* the problem, and grammar is its answer.

The point can be made in a different way. The core meaning of the term 'semantic' is 'relations to the world'. But a semantics in this general sense is carried by myriad cognitive or even non-cognitive systems, including mental representations that we find in navigating insects (Gallistel 1998), the perceptual systems of pre-linguistic infants (Carey 2008), the functionally referential alarm calls of monkeys or chicken (Hauser 1996), and the percepts of the material qualities of objects as studied in vision science (Mausfeld 2011). Ipso facto, such a generic notion of 'semantic' will neither predict nor illuminate the specific intentional ways of referring to the world that we find in our species. Nor will any appeal to 'causal' theories of reference (Devitt & Sterelny 1987) help in explaining our species-specific deictic frame: An infant and pet kitten may be exposed to the exact same causal relations, yet only one of them will start to intentionally refer, on a biologically timed course.

6. The Hierarchy of Referentiality

In translations of natural language into the idiom of philosophical logic, nominals are regarded as either referential or not. The 'non-referential' expres-

---

15 This is to make the point in terms of determiners, which exist in English but not in other languages with other lexical resources to indicate grammatical functions. Nonetheless, all languages appear to be able to enact the same forms of grammatical reference, and the absence of lexical determiners in a language like Chinese *supports* the point that the relevant mechanism is grammatical (see Cheng & Sybesma 1999, and subsequent discussion including Wu & Bodomo 2009). Even in English, the case of proper names moving to D when they are used referentially (Longobardi 1994) shows that referentiality is partially independent of lexical resources such as particular kind of determiners. For further typological considerations, see also Longobardi (2001), and see further sections 11–12.

sions are equated with 'general' or 'quantificational' ones, while the 'referential' ones are said to be 'singular'. Empirical complications ensue immediately, however: Linguistic reality does not allow for such a categorical divide. Thus the divide makes it hard to see—as a debate raging for more than century since Russell (1905) illustrates—how to classify a definite description like 'the murderer', which has no quantifier in it grammatically, is nonetheless not used *referentially* in the way that a proper name is, yet clearly *can* be used by speakers to refer to a particular person, as in 'The murderer entered'. In a similar way, there is no proper name in the assertion 'Dogs bite', and yet the speaker seems to be referring generically to dogs; and a person saying 'I saw dogs' is referring to an indefinite number of dogs that he saw. In short, it appears that reference in language can take a number of different forms. More specifically, we observe that:

(i) These forms are ordered in a *hierarchy* to be defined below;
(ii) specific forms of *grammatical complexity* correspond to each layer of the hierarchy;
(iii) grammatical complexity *increases* as we move from the beginnings of referentiality to its maximal forms; and
(iv) none of them are ever found *outside* of a grammatically structured system, or are lexical.

Consider a simple progression like the following, which has nothing to do with a change in the *conceptual content* of the lexical item *dog*, but instead solely with different *ways in which it can be embedded in the deictic space* that the speaker shares with the hearer:

(2) I ate **dog** < I ate **dogs** < I ate **a dog** < I ate **the dog** < I ate **this dog**

In the first example, the lexeme DOG is morphologically maximally simple and the determiner phrase of which it is a part lacks an overt determiner and can only have a *mass* reading involving a quantifier (cf. 'I ate some dog-meat'). With number marking added, as in the second example, we now refer to *individualized* specimens of the dog kind, though only *generically*. With the determiner 'a' added ('weak' in the sense of Milsark 1977), we can descriptively refer to something or other that falls under the description 'dog', but we also obtain the option of indefinite *specificity*. With the ('strong') determiner 'the', we obtain *definiteness*. With a deictic element added, as in 'this' (=the+here), we obtain *indexicality*.

We therefore move from a maximally indefinite to a maximally definite form of reference, as grammatical complexity of the determiner phrase increases. This process finds an end when the deictic stage is reached and referential specificity is maximal: At this stage, the speaker could perhaps add further deictics outside of this phrase, or descriptions that act as adjuncts, but a redundancy effect now arises (*I ate this dog here*; *I ate this dog here and now*; *I ate this brown dog right here and now that you selected*, etc.). So we have reached the end of the process of fixating reference that nominals can subserve. In line with that, the further we go up on the scale of increasing referential strength, the less easily do the nominals in question lend themselves to predicative uses: 'a dog' is a perfect and unmarked sentential predicate, as in *He is a dog*; the same is even more true for empty determiners, as in *I saw dogs*; but already 'the dog' becomes marked as such a predicate, as in *He is the dog*, where we are starting to get the intuition that 'the dog' is referential. If we finally come to *He is this dog*, we see the predicative option essentially disappearing. Insofar as we are moving here from 'cannot be referential' to 'must be referential' we can speak of a 'hierarchy of reference' in the nominal domain.

At the end of this progression, the grammatical process of increasing complexity essentially stops: More than embedding a lexical description in deictic space we cannot do in the case of lexical nominals; we have nailed down our exact object of reference. At no point in the progression did adding grammatical complexity co-vary with any *other* function than, broadly speaking, reference, or the conceptual content of the word 'dog' changed. What changed instead is what we have called the formal ontology of reference: whether DOG comes out as a *mass*, a set of non-specific individual *instances* of a kind, one *particular* individual, a *part* of an individual, etc. This formal ontology is moreover not assorted. Instead, the small number of distinct formal types are ordered by inclusion relations: Individual reference presupposes a substance/mass, specific individuality requires individuality, deixis requires definiteness, etc.

An analogous progression towards increased referential specificity as grammatical complexity is built up can be seen in the domain of clauses. In *She seems* ***to be a man***, there is an embedded, non-finite and non-tensed clause (as seen from the impossibility of specifying its independent tense, as in **She seems to be a man tomorrow*). Next such clauses can become tensed, as in *She wants* ***to become a man (tomorrow)***, and they then can become finite, in which case they also project their own subject in addition to an expletive one at the matrix level, as in *It seems* ***she is a man***, or with two full-blown subjects, as in *We think* ***she is a man***. Yet no finite embedded clauses can ever occur as an assertion, or denote the truth. For an assertion of truth to occur, we require a matrix (non-dependent) clause, as in ***She is a man***. Once such a clause is configured grammatically, truth is denoted, and the grammatical process again essentially stops: More cannot be done in clausal grammar. No more extensional form of reference can be reached. In English we can add a tag, negotiating the truth value assigned with the hearer: *John blew his nose, didn't he?* Or we can add a qualification, as in so-called 'slifting': *John blew his nose, I believe*. But neither process is recursive: **John blew his nose, didn't he did he?* or **John blew his nose, I believe, he claims*. Moreover, neither process changes the propositional meaning, or what was asserted.

Truth is thus the maximally *extensional* form of reference that is possible in grammar. If Superman is Clark Kent, it doesn't matter whether we know it, or how he is described: If it is a fact that one is the other, it is a fact no matter what we know or don't know. No embedded clause ever reaches extensionality in this maximal sense (including factive clauses, which retain an intermediate degree of intensionality; Sheehan & Hinzen 2011). Only (whole) sentences carry truth values, which are reached only at the root of the tree, at the end of the grammatical process.

These observation in the domains of reference (nominals) and truth (sentences) reinforce and differentiate the picture that there are two kinds of meaning: One is lexical/conceptual, presumably ultimately based on perception and sensory-motor processing, though also supra-modal, and as such constituting semantic memory; another is grammatical, which adds nothing to lexical content, is about reference, predication, and truth instead, and an expression of our personal creativity in changing and updating a body of shared knowledge. The latter kind of meaning is the content of grammar and involves a progression indexed by grammatical complexity.

7. Rational Grammar

Could reference be the most *foundational* grammatical concept? Are there core processes of grammar *unrelated* to its apparent involvement in the task of configuring acts of reference, or that do not relate to linguistically specific forms of meaning at all? If yes, this would support the 'autonomy' of grammar and question its inherent role in the constitution of a particular cognitive type. Yet we cannot as much as stick two words together without creating a grammatical relation that has consequences for reference. We can write on a shopping list: 'bread, juice, butter, beer', and each of these will then refer separately, to a different item each. If we change the grammar from that of a list to that of a compound, as in 'dog food', the consequences for reference are different: 'dog food' is used not for referring, in sequence, first to dog and then to food, but uniquely to refer to food, namely food for a dog; it is not used to refer to a dog, perhaps a dog for fetching food, which 'food dog' would refer to. The reason is that compounds are, in grammatical terms, 'headed', unlike lists. But headedness, in the sense just seen, translates into reference: It is not a merely formal notion. Headedness again disregards statistical facts: 'dog art', unlike 'art dog', is a kind of art, not a kind of dog, no matter how rarely or often dogs are involved in art.

With reference, we also get its opposite, predicativity, accounting for the basic structure of a sentence, consisting of a subject, which is referential in one of the above ways, and another term, which functions as the 'predicate'. Two referential terms cannot make a sentence. With predicates, we also obtain adjuncts, which are modifiers to already given referents. Referents combined with predicates yield propositional claims, hence truth values, which are for the clausal case what objects are in the case of referential nominals. On the way there, we obtain reference to events with verbs applying to arguments, as in *kill Bill*, where *Bill* plays the thematic role of 'Patient'. Events necessarily have a participant as an inherent part: *thematic structure*. With grammatical (and finite) tense, as occurring in a full sentence where *Bill* is the sentential object and something else is the grammatical subject, we obtain reference to a *specific* event and a truth value, as in *She has killed Bill*. In that case, there will moreover ipso facto be an event and a state, the latter an inherent part of the former: that she has killed Bill, and that Bill is dead. Without arguments playing thematic roles, and without grammatical relations such as finiteness requiring the further articulation of thematic structure into subjects and objects, there will be no cognizing of facts in this sense, and no reference to events or states. The relations in question are marked by the so-called 'structural' Cases in grammar (Nominative and Accusative), which find an interpretation in (grammatical) semantics for this very reason.[16]

If everything in grammar sub-serves this goal of reaching the truth value (the establishment of facts that, once asserted, can be de-indexicalized and re-enter semantic memory as truths about the world), it makes sense to rationalize grammar in these terms. In turn, we will regard propositional cognition as grammatical in nature: As grammatical structures are built up, a formal ontology falls into place, which begins from objects, proceeds to events and ends with propositions, which, in the case of matrix declaratives, are true or false. This ontology is what rational thought requires. A grammaticalized world is thus a *rational* one, which exhibits a structure and system of formal distinctions that cannot be found in perception.

8. Compositionality?

Grammar, apart from yielding referents rather than new concepts, also never quite *combines* concepts, and it is not obvious that this ever happens outside of grammar, either.[17] Thus, in grammar we may combine 'dog' with 'the', the latter of which is intuitively not a concept (but the expression of a grammatical function, namely reference), and then 'kill' with 'the dog'. But then, 'kill' is not combined grammatically with 'dog', *directly*: a determiner intervenes.[18] Nor does 'she' combine directly with 'kill the dog'. 'She' rather combines with something conceptually meaningless and grammatical first, namely finite Tense, as marked morphologically on the verb or through an auxiliary, but it does not combine with the lexical verb as such. It appears as if concepts have to be lexicalized *and* grammaticalized *first*, before they can combine productively and enter into structures with propositional meanings. What enters this combinatorics as a 'part' is never a pure lexical concept like DOG to start with, but a *part of speech*: a noun, verb, etc.; it is a *noun*, or category N, not DOG, that 'the' combines with.[19]

Even in the case of compounds (like 'dog food'), grammar never combines concepts directly, but a *head* is combined with a *modifier*, which are also particular *parts of speech*. Both are *lexicalized* through a concept, but these concepts only

---

16 See further Hinzen & Sheehan (2013: Ch. 6), and Hinzen (2014). Evidence in favour of the common view, that structural Case *is* strictly uninterpretable, would be evidence against (or limit) the present program, as it would identify a crucial dimension of grammar apparently irrelevant to meaning.

17 That 'concepts compose' is one of the prime axioms of research on 'concepts' in philosophy such as Fodor (1998). On this view, there are simple ('atomic') and complex (structured) concepts. But it is actually not clear what the evidence is for a combinatorics that applies to 'concepts' (rather than parts of speech or phrases) and produces new such 'concepts', when the actual evidence for a productive conceptual combinatorics governed by non-statistical principles comes from *grammatical* relations in which we see concepts appear.

18 A covert one in the case of 'eat dog', as an anonymous reviewer notes.

19 Evidence that even part-of-speech distinctions are not lexical but already reflections of grammatical functions, comes from the fact that grammar can overrule any lexical part-of-speech specifications (as in *manning a flight*, *topping the agenda*, etc.). See Levelt *et al*. (1999) and Vigliocco *et al*. (2011) for psycholinguistic and neurolinguistic evidence.

combine in virtue of their grammatical roles and part-of-speech status. These grammatical meanings are not *determined* by the lexical contents involved, as the fact shows that in 'food dog', the same lexical contents are involved, but the referentiality facts are the reverse. Lexical meanings are powerless to determine their grammatical functioning. One might be tempted to think that adjectives like *bald* will necessarily denote properties, not objects: They *must* be predicative. But they, too, need not; cf. *The bald tend to be sexy* or *Baldness is sexy*.

How then could grammatical meaning be reconstructed by 'composing' lexical meanings in a 'semantic component', regarded as separate from grammar itself? The student of philosophical logic is taught that the meaning of *John sleeps* is 'composed' from John, who is said to be the meaning of 'John', and the property of sleeping, which is said to be what 'sleeps' denotes. But how do we know that these are the right mappings? Plainly, from *understanding the grammatical structure of the sentence*, which turns 'John' into a referential expression (which it need not be in grammatically different contexts, such as *No John is proud of his name*, where 'John' refers to a property, namely being called 'John'), and 'sleeps' into the sentential predicate. Hence the mapping to the semantic values in question does not explain our grammatical understanding: It depends on it.

What, moreover, is a 'property'? Davidson (2005) argues that a characterization of the notion of 'property' that is logically independent of the notion of 'predicate' has never been provided. It does not help to interpret 'sleeps' as a mathematical function either: the function mapping John onto the proposition that John sleeps. For the compositional emergence of this proposition was what we were promised to obtain. It does not cash out this promise to hear that, in order to obtain this proposition, the sentential predicate has to be mapped onto a function that has been defined to yield this very proposition, when applied to John (Davidson 1967). Nor does it help to define 'sleeps' as the set of sleeping things. For the definition of this set will have to exploit our understanding of the predicate 'sleeps'.

With a notion of grammatical meaning missing, the only way for sentential meanings to arise from word meanings *is* for us to 'compose' the word meanings. This is a lexicalist model for solving the problem that sentential meaning poses. But as we just saw, grammatical meanings arise from the grammatical roles that words play: their roles as subjects or predicates, heads or modifiers, arguments or adjuncts. No word plays any of these roles lexically. Hence grammatical meanings cannot be composed from lexical meanings. Not only are lexical meanings never strictly composed, but compositionality of lexical meanings is not the solution to the existence of grammatical meanings. Grammar is this solution. Grammar is *foundational* for the human cognitive mind in this sense.

9. Grammar and Species

As I have presented it, studying language formally in abstraction from its role in thought and use is a methodological choice: It is not to study form as an object of nature. There is no 'form' in nature in addition to 'content', and a theory is needed to connect them. Nonetheless, in practice, generative grammar has not merely posited principles of language described formally, but formal principles: principles that don't in any way *appear* to illuminate the rational structure of thought or of meaning, and in this sense appear arbitrary. At best, they are instances of natural law, including the economy principles that Minimalism has moved to the forefront of inquiry. The present framework suggests a partially different inquiry: to rationalize the principles of grammar by regarding them as the principles of a rational cognitive type that is part of a speciation event in the genus *Homo* (Crow 2002, Stringer 2011, Tattersall 2008, Hinzen & Sheehan 2013: Ch. 7). That perspective cannot be pursued through the study of the 'principles' and 'parameters' of cross-linguistic variation (Chomsky 1981, Newmeyer 2005), since even if these were finally understood, the same question would arise: What is the cognitive function of grammar? It is difficult to address this question by comparing speakers of French and Japanese, or any other language, since these represent the same cognitive type. But it *can* potentially be studied by comparing cognitive phenotypes *within* our species, which may differ in their specifications for UG (Crow 2008).

In this regard the view that the genetics of language is the genetics of *sapiens*-specific thought predicts that different cognitive phenotypes within our species should co-vary with particular, identifiable linguistic profiles, i.e. linguistic phenotypes. Formal thought disorder in schizophrenia, which is diagnosed as a language abnormality (McKenna & Oh 2005), is a case in point. The cognitive changes involved should map onto grammatical changes in a systematic fashion, which should stand in a meaningful relationship to the changes in the cognitive phenotype that we observe, illuminating symptoms. In turn, where our mind remains rational, but language is produced only with difficulty, as in agrammatical aphasia, the profile of the changes should be different (see further Hinzen & Sheehan 2013: Ch. 8).

10. The Rise of the LOT

Nothing prevents us to take a metaphysical stance and think of formal distinctions in the ontology of reference as written into the very fabric of the universe, independently of language. But positing such a metaphysical ontology and formalizing it won't explain its existence in a particular cognitive type. The present framework claims that grammar does explain, for free, why thought takes place in a deictic space that exhibits such an ontology. By hypothesis, grammatical distinctions co-vary with the formal ontological ones and can be rationalized in these terms. By contrast, there is no non-linguistic and specifically no non-grammatical evidence for such a formal ontological structure. A cat does not refer to things in the same way as we do, and it does not distinguish in a systematic fashion between reference to facts and propositions, properties and states, or situations and events. Distinctions in the formal ontology of semantics are not distinctions at the level of the physics of perception of the external world either. *Mary smiles* and *Mary's smile* can be uttered in the exact same external circumstance, yet in the first case I will have said something true or false, and referred to an event as co-located with the act of speech and as ongoing with respect to it. In the second case, I will only have referred to what is formally an object—a difference in formal ontology.

In line with the idea that grammar and a particular kind of content are inseparable, it is impossible to produce a grammatical expression that is meaningless and that does not exhibit the formal ontology in question. *Colourless green ideas sleep furiously*, while often misinterpreted as evidence to the contrary, is a case in point: It has rational uses.[20] Grammar cannot play the role of deriving the formal ontology of thought or explain why there is one, if it is a formal system only. In that case, distinctions at the level of content will lead an independent life. Theories of meaning and content will not invoke grammar, and will tend to look for such factors as belief, causality, social cognition, or mental representation instead. Indeed it is the formal nature of the study of grammar in the 20th century that has led philosophers to conclude the philosophical insignificance of generative grammar.[21] Linguistics was meant to be formal because it was meant to be scientific and naturalistic—yet if it is merely formal, it cannot address philosophical questions about content, suggesting continuing a long tradition of language-neglect in philosophy.[22]

With the idea of 'arbitrary' principles of grammar firmly enshrined since the 1970s, it is also clear why it is natural for philosophers to conclude that 'thought' has to have its independent 'language': the Language of Thought (LOT) (Fodor 2008). Positing a LOT is particularly motivated if we see the kind of thought that it characterizes widely distributed in non-linguistic species (or pre-linguistic humans). In line with that, Fodor & Pylyshyn (1988: 28) wrote: 'that infraverbal cognition is pretty generally systematic seems, in short, to be about as secure as any empirical premise in this area can be'. All cognitive organisms, in short, verbal or not, think in much the same way. They are symbol-users (not externally, but in their LOT) and their 'mental representations' have a systematic and compositional semantics. In fact, though, it is 'about as secure as any empirical premise can be' that chimpanzees, say, do *not* think like us, and that it is merely the tragic lack of an expressive module that prevents them from telling us. If they did think like us, our current practices in treating them would be ethically indefensible, and should be switched for our treatment of agrammatical aphasics, whose thoughts we take to be rational and different from those of a chimpanzee, while the tool is broken that would normally convey them.[23]

---

20 As in this famous poem: "It can only be the thought of verdure to come, which prompts us in the autumn to buy these dormant white lumps of vegetable matter covered by a brown papery skin, and lovingly to plant them and care for them. It is a marvel to me that under this cover they are labouring unseen at such a rate within to give us the sudden awesome beauty of spring flowering bulbs. While winter reigns the earth reposes but these colourless green ideas sleep furiously." (C. M. Street)

21 Davidson (2004: 132–133) in particular rejects the philosophical significance of generative linguistic theory *because* of its stated aim to be a naturalistic science, and such a science is only possible for 'syntax', which he takes to be meaningless by standard definition.

22 From Frege to Russell, to Carnap, and to Quine, language had barely been regarded as more than a poor translation of logical form—that is, logic seen through the distorting lenses of a conventional system of 'signs'.

23 Tomasello (2008) agrees on a principled difference between a modern human and chimpanzee cognitive infrastructure but takes grammar to be an epiphenomenon of a non-linguistically specific adaptation for 'culture'. Yet a *linguistic* culture is intended here. What

With grammar characterized in the present terms, the distinction between a grammatical and a thinking being becomes incoherent: To structure one's thinking grammatically *is* to embed it into a triangular deictic space, with each corner of the triangle corresponding to a grammatical Person. This is what it *means* to think. There is nothing left for a LOT to accomplish and claims about the structure of the LOT, insofar as they are empirical ones, will be ones about grammar.

11. Person, Grammar, and the Self

Lexical items or concepts as units of semantic memory are impersonal—no person can claim ownership of any lexical concept, and they are given to us as infants in the form of an already existing and shared classificatory scheme for human experience. Reference by contrast is only enacted on particular occasions of language use, in a deictic space in which a grammatical first person refers *for* a grammatically second person *to* an object (the grammatical 3rd or non-Person), which is independent of both the pointer and the hearer, as well as the signal.[24] All of human reference is *triangular* in this sense, with the grammatical 1st, 2nd and 3rd persons labeling the corners of the triangle. A creature inhabiting this deictic frame is a *rational* one according to Davidson (2004), and this triangular infrastructure of rationality is the structure of thought itself: Nothing has to be added to it to qualify as such. A form of thought that didn't share its basic structure, moreover, wouldn't be thought in the same sense. We can share a deictic frame with our pet dog, but it is a frame of a different kind, whose baseline is not language.

Grammar, in setting up a novel deictic space, may therefore be well placed to address another problem: Human thought is not explained before we have explained how and why it is *personal*. We consider ourselves and other members of our species persons, but not the members of any other species, with non-language using humans such as fetuses or comatose patients as an unclear intermediate case. How come that we language users are persons and non-language users are not? Once again, the correlation between persons and language could be accidental—but then why should it be?

If we address the question of what persons are metaphysically, discussions of the nature of the self usually begin from a view ascribed to Descartes, according to which selves are simple, immaterial substances. A second possibility

drives the evolution of language is said to be *expressive needs*, communicative intentions of the kind that we have, but chimpanzees lack grammar of the sort that we see in human languages. Evidence for this view would be that the *kind* of communicative intentions that we see expressed in language can exist without language. Apart from that, a drive to communicate and an ability and motivation to share intentions neither predict nor explain a thinking ability in the present sense. Before we communicate propositionally, we need a mind that can think propositionally.

24 Crucially, the second person can be the speaker himself, as in self-talk, in which it is critical that the two persons are grammatically distinct: A person standing in front of a mirror says 'I hate you!', not 'I hate me!' (see Holmberg 2010). This shows that the non-grammatical notion of 'self-reference' does not capture the right distinctions: One can refer to the exact same person in the grammatical 1st, 2nd, or 3rd Person. The grammatical distinctions are needed.

is then taken to be that we are indeed such a substance, but that we *also* have physical properties: Selves are immaterial substances, but they have bodies (Lowe 2008). A third is an alternative famously proposed by Hume, that qua selves we are simply bundles of ideas: Each of us is 'a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement' (Hume 1737/1888: 252). A fourth is that we are our bodies. A fifth is that we are computer programs, hence abstract objects, which can be 'multiply realized'. A sixth is that we simply don't exist: There are no selves (see Olson 2007 for an overview). It is again remarkable that in each and every of these proposals language simply does not seem to matter to our existence as selves. It is as if language was irrelevant to having the kind of self that we do and was merely a way of *talking* about it. In the cognitive neuroscience of the self, too (e.g., Kircher & David 2003), traditional notions from philosophical epistemology and phenomenology ('subject' and 'object', 'self' and 'other', 'subjective experience', 'first person perspective', 'meaning', etc.) prevail and are used in non-linguistic senses. The 'first-person perspective' (Zahavi 2006, Lowe 2008) is stressed in its importance, yet the Cartesian assumption appears to be that such a 'perspective' is somehow available independently of language, through introspection—despite the fact that 'first person' is a *grammatical* distinction: The notion 'first person' cannot here have a *non-grammatical* sense, as in 'the first person entering the room'.

The very fact that we speak of a 'first-person perspective' indicates that the notion of a perspective as such is insufficient to capture the required distinction: Grammar has to be added. Zahavi (2006: 27–29) stresses that self-awareness involved in conscious states cannot be construed along 'subject-object' lines, yet it is the grammar of self-reference that precisely tells us how this is the case: The grammar of 3rd-person object reference ('the/my self', 'my body/brain') is different from that of 1st-person reference ('I'; Bianchi 2006, Martín & Hinzen 2014). The grammatical distinction is essential: We would worry about an English-speaking child who never referred to itself in the grammatical 1st person. As the cases of abnormal cognition in schizophrenia and in autism both indicate (see, e.g., Rochester & Martin 1979, Hobson *et al*. 2010, respectively), pronouns, which paradigmatically incorporate grammatical Person distinctions, are a particular locus of vulnerability in these disorders, indicating a connection between disturbances of selfhood and grammar.

Independently of this connection, pronouns *are* highly grammatical creatures in the sense that person systems that they incorporate interact with inherently grammatical organizational principles, such as structural Case. It is significant, moreover, that the personal pronouns are grammatically the most complex, coming still on top of the hierarchy of reference above (Martín & Hinzen 2014). At the beginning of this hierarchy, we saw nominal arguments without overt determiners fixing reference merely via their descriptive lexical content (e.g., *I'd like to have* ***dog*** or *I like* ***dogs***). Nominal arguments with the strong determiner 'the' will normally be definite-referential, except when special operators are involved that compromise such referentiality, such as 'occasional' in ***The occasional dog*** *passed by*, where no particular dog is referenced. If we expand 'the' into 'this' by adding a deictic element, referentiality becomes an absolute requirement (**this occasional dog** *passed by*). By the time we have reached demonstrative reference (*I like* ***this dog***), reference succeeds with a pronominal on its own without any support from the lexical nominal: *I like* ***this/him/her/it*** vs. **I like* ***the***. Extending this progression into the personal (1st and 2nd person) pronouns, we find these deprived of Gender (and arguably Number) specifications, which makes them even poorer in terms of lexical content: ***I*** *like* ***you***.

While 3rd person pronouns still allow modifying relative clauses, moreover, the personal pronouns don't; cf. ***He*** *who enters through this door will be shot* vs. ****You*** *who enter(s) through this door will be shot*. There is a further progression from the 2nd Person to the 1st Person in that the former, at least in the plural, *can* co-occur with a lexical nominal, as in ***You linguists*** *are crazy people*, and at a stretch even in the singular (***You linguist*** *will never get this*), the 1st Person is completely unmodifiable: ****I linguist*** *like my job*.[25] Finally, while 3rd person pronouns can paradigmatically be bound, losing their referential independence or force, as in *John thinks* ***he*** *is smart*, 1st person pronouns cannot be.[26]

We thus see a progression from readings that are maximally non-specific and descriptive to readings that are maximally specific and non-descriptive, with quantificational/indefinite, definite referential, and deictic forms in the middle, again ordered with respect to one another. This is mirrored in the grammatical complexity of the grammatical argument we see, which goes from obligatorily absent or optional determiners to obligatory ones, until the complement becomes optional, and finally the complement becomes obligatorily absent, as with personal pronouns. Put differently, reference goes from being maximally co-mediated lexically to being maximally mediated grammatically—until all lexical content is lost in the case of 'I' and the referent is not described at all: the case coming closest to what Russell called a 'logically proper name'. We may summarize this hierarchy as follows, where a star within a bracket indicates obligatory deletion and a star outside of a bracket obligatory presence of the material in the bracket:

(3) (*the) *(NP) < *(a) *(NP) < *(the) *(NP) < *(this) (NP) < *(he) (*NP) < you < I

Personal forms of reference, and especially reference in the 1st Person are therefore a maximally *grammatical phenomenon*. It is not a long step from here to argue that, given that, by a wide agreement, selves are identified in the grammatically 1st Person, selves are *also* individuated grammatically. Knowing oneself under a description, including one's own proper name, is not enough to know that one is identical to that person, in the sense that one would know: 'I am that person'. In this sense, descriptive forms of reference in the 3rd person ('that person', 'Rudolf Lingens', etc.), do not substitute for forms of reference in the 1st

---

25 As Michelle Sheehan (p.c.) notes, the Person Case Constraint and systems of inverse agreement also illustrate that the 1st person can be seen to be less 'object-like' than the 2nd person. Even in English, '?She lent me you' is much better than '*She lent you me'.

26 Where they *appear* to be bound, the person features are often not interpreted, as the fact of the substitutability of 'his' for 'my' shows in the following: *I am the only one around here who takes care of my/his children* (example from B. Partee).

person (cf. Perry 1993). If the use of the 1st person in ascribing a property to oneself is both necessary and sufficient for the relevant form of self-knowledge to exist, the argument that there are no selves without the grammar of Person would be complete.[27]

12. Topological Mapping

The above account entails that the forms of reference can be mapped from the grammatical configurations involved: No extra-grammatical principles are required. Longobardi (1994) provided crucial evidence that this applies as well to a case that prima facie seems to contradict the account: that of proper names. At least in languages like English, proper names *are* paradigmatically used definite-referentially, yet a determiner is obligatorily *absent* in such uses. As we noted, where proper names are used referentially, this is never a lexical fact, since all proper names can also be used predicatively, as noted. Yet, it still *prima facie* contradicts the above hierarchy, in that a (strong) determiner should not be absent in a referential interpretation. In this regard Longobardi raised a fundamental question: Where a proper name is used referentially (without a determiner), is it grammatically in the position of D or of N, within the scheme [D[N]]?

Cross-linguistic evidence supports the former answer: The proper name, first generated in N, *moves* to the D position in the uses in question. In line with this proposal, where, in English, the determiner is present and movement is blocked, a *descriptive* reading is derived, as in *the early Russell* or *the Russell of 1905*, which refers not to Russell to directly, but one particular kind (or stage) of Russell. This proposal, further corroborated in Longobardi (2001, 2005, 2008), has two striking implications that are of foundational significance (Hinzen 2007): First, it has long since been noted in the philosophical literature that proper names, while their referents can of course be associated with all sorts of descriptive properties (John is the mayor, he is handsome, etc.), are nonetheless standardly used in such a way that none of these descriptive properties are crucial to the identity of the referent. Thus, if it turns that John isn't the mayor, in fact, or he is ugly, the name 'John' doesn't change its reference; it will refer to the exact same person.

Termed 'rigidity' by Kripke (1972), this important empirical observation finds a natural explanation in Longobardi's account: While all proper names have lexical descriptive content (the least we know about John is that he is called 'John'), they have no other lexical specifications. Even their single lexical-descriptive specification (being called 'John'), however, is not expected to determine reference when the nominal moves from the N to the D position: Properties serving to identify the referent are specified in the N-position, the complement of the determiner. Where they are gone, no identifying conditions are expected to

[27] See Martín & Hinzen (2014) for evidence for the premise: Control constructions, in particular, like *John expects PRO to get a medal*, do not enforce the relevant *de se* readings, which require a 1st Person subject.

co-determine the identity of the referent, and rigidity follows (Hinzen 2007).

The second consequence is that reference is *never* lexically determined—even in the case of proper names it depends on N-to-D movement, and like in all of the other forms of reference we have encountered above, the form of referentiality that we obtain depends, in Longobardi's terms, on the 'topology' of the phrase in which they appear. It depends on facts like whether the determiner position is empty, or whether the N-position is empty, or whether both are filled, and co-determine the act of reference.

Sheehan & Hinzen (2011), modifying Longobardi's mapping principles for nominal slightly, expanding them to 3rd person pronouns and implementing them in an architecture based on the 'phases' of Chomsky (2001, 2007), argue that the exact same consequence holds in the clausal case. Frege, suggesting that clauses are so-called 'derived proper names', argued that clauses, too, can be referential expressions, and matrix declarative clauses in particular refer to truth values, which are their object-like referents. Sheehan & Hinzen show how this intuition can be naturalized on a grammatical path: Truth-asserting clauses involve V-to-C movement (movement of the verb to the clausal edge) either overtly or covertly, the exact clausal equivalent of nominals that have moved to D. Their account is argued to also explain why, in a language like English, if we are to make an assertion (claim a truth value), the complementizer 'that' must be exactly as absent as the determiner 'the' must be in the nominal case when we are to refer to an object:

(4) a. *the Russell
    b. *that Russell is a philosopher[28]

The reason is that for the truth-denoting reading to be derived, the verb must move to the position of the complementizer, which means that the latter cannot also be present. For empirical evidence for overt/covert V-to-C movement in a range of languages, the clausal equivalent of N-to-D movement, see Sheehan & Hinzen 2011, Hinzen & Sheehan 2013: Ch. 4).

This account reinforces Longobardi's point that reference is regulated *topologically* (and hence grammatically) rather than lexically, for the way it comes about is even independent of the lexical category (noun or verb) involved. This suggests a generalisation of the D+NP scheme above to that of a general unit of referentiality in grammar, of the form <edge, interior>, with a broad division of function: <reference, description>; the edge involves reference-regulating elements like determiners in the nominal case and complementizers in the clausal case, while the interior involves the lexical description of the referent that such units can be used to refer to.

As for embedded clauses, they can be either fact or proposition-referring (factive or non-factive). Canonically factive clauses, in particular, represent the case of a referential in an embedded position, where the complementizer is obligatorily present exactly as 'the' is in the case of nominals, if the reading is to

[28] The reading that 'that' is a demonstrative determiner here (as in 'that Russell-chap annoyed me') is not intended and irrelevant.

be definite:

(5) a. *(the) man
b. I resent *(that) John left.

In canonically intensional (proposition-denoting) embedded clauses, by contrast, where the proposition is not yet evaluated for truth or falsehood, the complementizer tends to be optional:

(6) a. I saw (some) men.
b. I believe (that) John left.

The proposal also features intermediate phenomena such as 'slifted' clauses, as in *Russell is a philosopher, I believe*, where the fronted clause is genuinely asserted, and the complementizer is obligatorily absent, as predicted, and where a canonically factive verbs cannot be employed, given that factive clause, on this account, are not asserted (truth-referring), but fact-referential:

(7) a. (*That) Russell is a philosopher, I believe.
b. *Russell is a philosopher, I resent.

Not only object-reference is regulated topologically, then, but truth, fact, and proposition-reference as well. If so, the entire formal ontology of semantics falls into place as grammatical complexity is built up, unit by unit, with each corresponding to a Chomskyan 'phase' of the derivation. The formal ontology of meaning co-varies with grammar, not with non-grammatical factors, and hence this formal ontology is indeed an aspect of grammatical meaning in our sense.

13. Truth

No account of thought can be that: an account of thought, without it being an account of the *content* of thought. This is how the term, as an object of study, has been traditionally defined: As Frege put it, 'I call a thought something for which the question of truth arises' (Frege 1918–19/1956: 292). Intrinsically associated with any thought in the present sense there is a condition under which it is true. If I utter 'John left', then, if John left, the sentence is automatically true. We therefore cannot here claim to naturalize thought (or derive our cognitive phenotype) from its grammatical nature, if the connection between the grammaticalization of our brain and truth in the traditional philosophical sense has not been made intelligible. How *can* an account that only talks about grammar also capture what is most characteristic and most mysterious about thought, namely its being true or false?

However, as this challenge is posed, it *presupposes* that truth is a semantic notion, and I have argued that, like reference and for the same reason, it is a grammatical one instead. However, let it be the case that truth indeed is a grammatical concept: Truth-bearing entities uniquely arise in specific grammatical patterns, and there is no other known process that leads to them. It may still seem unclear, then, how 'truth' could still mean *truth*, i.e. the notion that it has meant in thousands of years of philosophical tradition in the West, where truth has been taken to be a 'semantic' or indeed metaphysical notion. Have we simply re-defined a concept?

The reason that we haven't relates to the persistent failure, in the same thousands of years, to explicate the notion of truth in any terms that do not presuppose it and to tell, in substantive terms, what truth actually is. Compare the case of our notion of 'water'. Here we can tell, in substantive and indeed scientific terms, what water is: H_2O, it so happens. So-called 'deflationists' in the philosophy of truth have made it one of the axioms of their position that *no* such substantive account is possible in the case of truth: It lacks such a kind of content. But this is exactly what we *expect* if truth is a grammatical notion. As stated initially, grammar in general exactly *lacks* the kind of content that the deflationist claims truth lacks.

In this sense, Un-Cartesian Linguistics explains and vindicates deflationism. Deflationists have nonetheless given an account of truth, avoiding any substantivist specifications of its content. This account is that all we need to know, in order to know what truth means, is how the notion is used (Horwich 2010); and in particular, that it will be the case for any sentence not itself containing the lexical truth-predicate, like 'John left', that the following holds:

(8) 'John left' is true if and only if John left.

Some prefer a slightly different version, which we will not need to distinguish here:

(9) That John left is true if and only if John left.

These equivalences are indubitably true: They capture a crucial fact about the English sentence 'John left'. Generalizing, we obtain the Equivalence Schema (ES); for any sentence S not containing itself the lexical truth predicate:

(10) ES: 'S' is true iff S
That S is true iff S.

But *why* does ES hold, and why are the above sentences so indubitably true? Again we can now answer: *Because* truth is a grammatical concept. For what ES shows is that in the two assertions flanking the equivalence sign 'iff', one contains the lexical concept of truth, while the other does not. Yet they are equivalent. Hence the lexical truth predicate is not required to assert a truth. This we know independently: 'John left' is, if asserted, asserted to be *true*—and nothing else: It is not asserted to be likely, to be desirable, or to be believed, which are all logically possible but not grammatically possible options. This makes sense if truth is a grammatical concept, but belief or likelihood are not. Neither for belief nor for likelihood do equivalences like ES exist, in which these lexical semantic predicates are present on one side and absent on the other.

In fact, using the lexical truth predicate is felicitous only if an act of assertion has been configured *before*, in a purely grammatical way: We can assert (re-confirm) *It's true that John left*, only if it was already asserted before that *John left*. Moreover, as already noted in passing, the addition of 'it's true' changes nothing in the content of the assertion of 'John left', or in that it was asserted: The grammatical process is already at its outer limits. In sum, while grammatical notions *can* be lexicalized—the words 'truth', 'reference', 'proposition', or 'identity' being examples—their lexical meaning is not where their secrets lie.

The ES thus again suggests what the Un-Cartesian program independently maintains: that truth is a grammatical notion. ES holds because the right hand side of the equivalence has the exact grammar to express truth configurationally or grammatically, while the left hand side does so additionally lexically. If truth as a concept cannot be grounded in any more fundamental notion, and ES is our key to what it means, and ES merely points us again to its grammatical nature, we cannot hope to illuminate truth in a way that goes *beyond* illuminating grammar. Grammar takes us to the limits of our conceptual scheme, and we cannot dig deeper. Beginning from sentence contents—propositions—we can move to evaluating these as true, and then we obtain what we call 'facts'. But nothing is more fundamental in our metaphysical scheme than facts: While a proposition evaluated as true is a fact, and facts in a sense are what makes propositions true, we cannot in turn ask what grounds a fact, or what makes a fact a fact. More fundamental than facts, things don't get.

The limits of grammar are the limits of thought in this sense. Thought begins from where there are concepts distinct from percepts. Unlike percepts, concepts are not stimulus-controlled, and they can be activated in the absence of online sensory processing of a stimulus for them, unlike percepts, except in the case of hallucinations. Precisely for that reason, the question of reference arises for concepts, but not for percepts. As we perceive, we do not also refer, and need not do so. For the same reason, with concepts, intensionality effects systematiccally arise. The object of reference does not prescribe how we describe it. For any referent, there is an infinite number of concepts under which it falls. Therefore, two concepts can be chosen to be the descriptors of an object of reference by different speakers, and it may not be clear to them that they are referring, under these descriptions, to the same object: an intensionality effect. The same can arise in a single speaker, like Lois Lane, who does not recognize that Clark Kent is Superman.

Intensionality effects are thus evidence that thought is indeed involved: Reference takes place under descriptions whose lexical contents are provided by concepts. The effects in question are the footprints of the concepts involved. But they crucially also show that concepts, where involved in an act of reference, *are* indeed connected to external referents, of which they are true in an objective manner. It is *because* two concepts can objectively co-refer to the same thing that we may use one without realizing that the other identifies the same thing. Without there being something like 'applying to the same thing' (objectively), there is no intensionality. Two concepts are 'different' with *respect* to a referent, which they both identify.

Concepts, reference, grammar, and intensionality, are all *correlative* in this sense. Objectivity and truth are *factored into* our use of concepts. No person applying a concept can fail to know that concepts can *fail* to apply: That applying them to an object can be *false*, dependent on what the world is like, and not on what is believed or thought. To know that, a person needs to have a concept of world, and of truth, and of belief—distinguishing 'I think' from 'he thinks' and 'it is true'. Such knowledge, the Un-Cartesian hypothesis maintains, depends on grammar, and it does not affect the *concepts* to which it applies when it is generated in a grammatical process. These merely internally anchor the system of thought that we come to use as grammar comes to be at our disposal.

14. Conclusions

Grammar has been a domain of scientific inquiry for millennia, starting in Ancient India from the very onset of the scientific tradition in the Eastern world. In the most recent episode of universal grammar research, 'syntax' has been one particular choice of a theoretical abstraction with which to study this domain. The lenses with which we look at a domain, however, do not determine its ontological nature. In light of the explanatory task of telling where our system of thought, to the extent that it is *sapiens*-specific, comes from, the option should be explored that grammar is the path along which such a system can arise.

On this view, the grammaticalization of the *hominin* brain brought a new mode of thought into place. This is an empirical hypothesis refutable in at least two ways: (i) grammar might simply not be productively re-describable as a system governing how a thinking creature behaves; (ii) thought of the relevant kind and language might dissociate in our species or across species. With regards to (ii), this research program seeks to establish systematic links between specific linguistic profiles involving either lexical or grammatical changes, on the one hand, and particular changes in our cognitive phenotype, on the other (Hinzen & Sheehan 2013: Ch. 8, Hinzen & Schroeder 2014). As regards (i), the claim is that the organization of grammar systematically reflects the organization of *sapiens*-specific thought insofar as it is intentionally referential, with the formal ontology of thought arising as grammatical complexity is built.

It is not clear, in that case, what, when we are done with our description of the workings of grammar, is still left for a LOT to accomplish. Nor is there evidence that a system of thought of the same nature is available in a 'semantic component' that is architecturally separable from the grammatical process in which particular forms of grammatical meaning arise, and in particular the formal ontology of reference that is built up in it step by step. Formal-ontological distinctions do not appear to be independent features of the external world to which we use grammar to refer: They are distinctions arising at the level of how we decide to refer to the world, choosing both a concept and a way of carving out the formal ontology of reference. Formal-ontological distinctions like that between an object or an event, or an event and a proposition, do not track what aspects of our experience or the external world are independently given. Nothing 'semantic' determines this choice, either: It is a grammatical choice, existing where grammar does, availing us of a number of discrete ways of referring to the world. Qua persons, we are also endowed with a first-person perspective, but correlatively with that, a notion of the second person and the 'third' person as well, or the objective world as such. If form and content, or syntax and semantics, are separated architecturally, then the question of the existence of grammatical meaning that we have raised is begged: It won't exist.

Un-Cartesian linguistics makes the claim that nothing else than grammatical organization is similarly well placed to explain the remarkable transition that has transformed the hominin mind when, after six million years of evolutionary experimentation with the genus *Homo*, our species suddenly appeared, changing the surface of the earth globally and in a remarkably short time. The same perspective may shed new linguistic light on a number of cognitive disorders in this species; as of now, the Un-Cartesian claim is that the organization of grammar systematically *reflects* the organization of a *sapiens*-specific mode of cognition. If so, this is to return a role to language that it has lost, but that could inspire philosophy as much as it could inspire linguistics, neurology, and psychiatry.

References

Arnauld, A. & C. Lancelot. 1966. Grammaire generale et raisonnee ou La Grammaire de Port Royal (1660/1676). Bad Cannstatt: Frommann-Holzboog.

Baddeley, A.D. 1990. *Human Memory: Theory and Practice*. Hillsdale, NJ: Erlbaum & Associates.

Berwick, R.C. & N. Chomsky. 2011. The biolinguistic program: The current state of its evolution and development. In. A-M. di Sciullo & C. Boeckx (eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, 19–41. Oxford: Oxford University Press.

Bianchi, V. 2006. On the syntax of personal arguments. *Lingua* 116, 2023–2067.

Binder, J. & R. Desai. 2011. The neurobiology of semantic memory. *Trends in Cognitive Sciences* 15(11), 527–536.

Carey, S. 2009. *The Origin of Concepts*. Oxford: Oxford University Press.

Cheng, L.L.-S. & R. Sybesma. 1999. Bare and not so bare nouns and the structure of NP. *Linguistic Inquiry* 30, 509–542.

Chomsky, N. 1957/2002. *Syntactic Structures*, 2nd edn. The Hague: Mouton de Gruyter.

Chomsky, N. 1966. *Cartesian Linguistics*. New York: Harper & Row.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

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

Chomsky, N. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, N. 2007. Approaching UG from below. In U. Sauerland & H.-M. Gärtner (eds.), *Interfaces + Recursion = Language?*, 1–29. Berlin: Mouton de Gruyter.

Collins, A.M. & M.R. Quillian. 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8(2), 240–247.

Collins, A.M. & E.F. Loftus. 1975. A spreading activation of semantic processing. *Psychological Review* 82, 407–428.
Cottingham, J. 1998. 'The only sure sign…': Thought and language in Descartes. In J. Preston (ed.), *Thought and Language*, 29–50. Cambridge: Cambridge University Press.
Crow, T.J. (ed.). 2002. *The Speciation of Modern Homo Sapiens*. Oxford: Oxford University Press.
Crow, T.J. 2008. The 'big bang' theory of the origin of psychosis and the faculty of language. *Schizophrenia Research* 102(1–3),31–52.
Crow, T.J. 2010. The nuclear symptoms of schizophrenia reveal the four quadrant structure of language and its deictic frame. *Neurolinguistics* 23, 1–9.
Davidson, D. 1967. Truth and meaning. *Synthese* 17, 304–323.
Davidson, D. 2004. *Problems of Rationality*. Oxford: Oxford University Press.
Davidson, D. 2005. *Truth and Predication*. Cambridge, MA: Belknap Press.
DeVilliers, J. 2007. The interface of language and theory of mind. *Lingua* 117(11), 1858–1878.
DeVilliers, J. In press. Which concepts need language? *Language Sciences*.
Devitt, M. & K. Sterelny 1987. *Language and Reality*. Cambridge, MA: MIT Press.
Evans, N. & S. Levinson. 2009. The myth of language universals. *Behavioral & Brain Sciences* 32, 429–492.
Fitch, T. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.
Fodor, J.A. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.
Fodor, J.A. 2001. Language, thought, and compositionality. *Mind and Language* 16(1), 1–15.
Fodor, J.A. 2008. *LOT 2*. Oxford: Oxford University Press.
Fodor, J.A. & Z. Pylyshyn. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 3–71.
Frege, 1918–19/1956. Der Gedanke: Eine logische Untersuchung. *Beiträge zur Philosophie des Deutschen Idealismus I*, 58–77. The translation 'Thought: A logical Inquiry' by P. Geach appeared in *Mind* 65(259), 289–311.
Gallistel, R. 1998. Symbolic processes in the brain: The case of insect navigation. In D. Scarborough & S. Sternberg (eds.), *Invitation to Cognitive Science, vol. 4: Methods, Models, and Conceptual Issues*, 1–51. Cambridge, MA: MIT Press.
Hassabis, D. & E.A. Maguire 2007. Deconstructing episodic memory with construction. *Trends Cognitive Sciences* 11, 299–306.
Hinzen, W. 2006. *Mind Design and Minimal Syntax*. Oxford: Oxford University Press.
Hinzen, W. 2007. *An Essay on Names and Truth.* Oxford: Oxford University Press.
Hinzen, W. 2014. The rationality of case. *Language Sciences* 46B, 133–151.
Hinzen, W. & K. Schroeder. 2014. Is 'first person' a linguistic concept? To appear in M. Guta (ed.), *The First Person*, volume in preparation.
Hinzen, W. & M. Sheehan. 2013. *The Philosophy of Universal Grammar*. Oxford: Oxford University Press.
Hobson, R.P., R.M. García-Perez & A. Lee. 2010. Person-centered (deictic) expressions in autism. *Journal or Autism Developmental Disorders* 40(4), 403–414.
Hodges, J. & K. Patterson 2007. Semantic dementia: A unique clinicopathological

syndrome. *Lancet Neurology* 6, 1004–1014.
Holmberg, A. 2010. How to refer to yourself and when talking to yourself. *Newcastle Working Papers in Linguistics* 16, 57–65.
Holmberg A. 2013. The syntax of answers to polar questions in English and Swedish. *Lingua* 128, 31–50.
Horwich, P. 2010. *Truth, Meaning, Reality*. Oxford: Oxford University Press.
Kircher, T. & A. David. 2003. *The Self in Neuroscience and Psychiatry*. Cambridge: Cambridge University Press.
Longobardi, G. 1994. Reference and proper names. *Linguistic Inquiry* 25, 609–665.
Longobardi, G. 2001. How comparative is semantics? *Natural Language Semantics* 9, 335–369.
Longobardi, G. 2005. Toward a unified grammar of reference. *Zeitschrift für Sprachwissenschaft* 24, 5–44.
Longobardi, G. 2008. Reference to individuals, person, and the variety of mapping parameters. In H.H. Mueller & A. Klinge (eds.), *Essays on Nominal Determination*, 189–211. Amsterdam: John Benjamins.
Lowe, E.J. 2008. *Personal Agency*. Oxford: Oxford University Press.
Lycan, W.G. 2008. *Philosophy of Language*. Oxford: Routledge.
Martín, T. & W. Hinzen. 2014. The grammar of the essential indexical. *Lingua* 148: 95–117.
McKenna, P. & T. Oh. 2005. *Schizophrenic Speech*. Cambridge: Cambridge University Press.
Mausfeld, R. 2011. The perception of material qualities and the internal semantics of the perceptual system. In L. Albertazzi, G. van Tonder & D. Vishwanath (eds.), *Perception beyond Inference. The Information Content of Visual Processes*, 159–200. Cambridge, MA: MIT Press.
Merchant, J. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27, 661–738.
Newton, A. & J. DeVilliers 2007. Thinking while talking. *Psychological Science* 18 (7), 574–579.
Newmeyer, F.J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford University Press, Oxford.
Olson, E.T. 2007. *What Are We? A Study in Personal Ontology*. Oxford: Oxford University Press.
Penn, D.C., K.J. Holyoak & D.J. Povinelli 2008. Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral & Brain Sciences* 31, 109–178.
Perry, J. 1993. *The Problem of the Essential Indexical*. Oxford: Oxford University Press.
Pettito, L. 1987. On the autonomy of language and gesture. *Cognition* 27, 21–52.
Pettito, L. 2005. How the brain begets language: On the neural tissue underlying human language acquisition. In J. McGilvray (ed.), *The Cambridge Companion to Chomsky*, 84–101. Cambridge: Cambridge University Press.
Pietroski, P. To appear. *Conjoining Meanings: Semantics without Truth Values*. Oxford: Oxford University Press.
Renoult, L., P. Davidson, D.J. Palombo, M. Moscovitch & B. Levine. 2012. Personal semantics: At the crossroads of semantic and episodic memory. *Trends in Cognitive Sciences* 16(11), 550–558.

Rey, A., P. Perruchet & J. Fagot. 2012. Centre-embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons (*Papio papio*). *Cognition* 123, 180–184.
Rochester, S. & J. Martin. 1979. *Crazy Talk*. New York: Plenum Press.
Sheehan, M. & W. Hinzen. 2011. Moving towards the edge. *Linguistic Analysis* 37, 405–458.
Stringer, C. 2011. *The Origin of Our Species*. London: Penguin.
Tattersall, I. 2008. *The World from Beginnings to 4000 BCE*. Oxford: Oxford University Press.
Terrace, H. 2005. Metacognition and the evolution of language. In H. Terrace & P. Metcalfe (eds.), *The Missing Link in Cognition*, 84–115. Oxford: Oxford University Press.
Tomasello, M. 2008. *Origins of Human Communication*. Cambridge, MA: MIT Press.
Tulving E. 1972. Episodic and semantic memory. In E. Tulving & W. Donaldson (eds.), *Organization of Memory*, 381–403. New York: Academic Press.
Vigliocco G., D.P. Vinsona, J. Druksa, H. Barberb & S.F. Cappa. 2011. Nouns and verbs in the brain. *Neuroscience and Biobehavioral Reviews* 35, 407–426.
Wilson, B., M.G. Collison, H. Slater, K. Smith, W. Marslen-Wilson & C. Petkov. 2011. Behaviour and functional imaging of 'artificial grammar' sequence learning in rhesus macaques. *Society of Neuroscience-Abstract* 173.18, 53.
Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*. London: Kegan Paul.
Wu, Y. & A. Bodomo 2009. Classifiers ≠ Determiners. *Linguistic Inquiry* 40, 487–503.
Zahavi, D. 2006. *Subjectivity and Selfhood*. Cambridge, MA: MIT Press.

Wolfram Hinzen
ICREA & Universitat de Barcelona
Department of General Linguistics
Gran Via de les Corts Catalanes, 585
08007 Barcelona
Spain
hinzen@ub.edu

★ FORUM ★

Notice

I would like to use this opportunity to briefly thank all those involved in creating the eighth volume of *Biolinguistics*, a turning point for the journal with Cedric Boeckx' departure from the editorial office. I am thankful for everything he has done over the past 8 years—conceiving of the idea, pulling it off, and seeing it through for all this time. Thank you very much, Cedric!

There will be some changes to the journal, to be settled in early 2015 with volume 9. For now, my special gratitude goes to the reviewers that have served us throughout 2014, who are listed below (colleagues who reviewed more than one submission are suffixed by an asterisk). For everything else, I thank all supporters as well as the members of the *Biolinguistics* Advisory Board, the Editorial Board, and the Task Team that are not specifically mentioned by name for active participation and feedback.

Reviewers

Koji Arikawa
Robert C. Berwick
Derek Bickerton
Norbert Hornstein
Alistair Knott
Víctor M. Longa
James McGilvray
Lisa S. Pearl
Karl-Magnus Petersson
Paul Pietroski
Volker Struckmeier
Robert Truswell
Juan Uriagereka
Jeffrey Watumull*

ISSN 1450–3417
Biolinguistics 8: 258, 2014
http://www.biolinguistics.eu

Biolinguistics

Volume 8

2014

TABLE OF CONTENTS

ISSN 1450–3417

Biolinguistics 8: 001–258, 2014
http://www.biolinguistics.eu