Cambridge University Press
978-1-107-03483-9 - The Structural Design of Language
Thomas S. Stroik and Michael T. Putnam
Excerpt
More information

1 The biolinguistic turn

1.1 Introduction

Human language gives us the ability to express anything we can think, and to communicate these thoughts via a set of mutually comprehensible signals. Although all animals communicate, this boundless expressibility sets our species apart. (Fitch 2010: 5)

Human language is full of wonders, none more fascinating or arresting than its "boundless expressibility." The fact that any given human language can render a larger number of sentences than there are particles in the universe gives some indication of the magnitude of language. And the fact that humans can produce and use linguistic expressions that point, in some sense, to things that neither exist in the universe nor can be conceptualized by humans – such as "squared circles" or "colorless green ideas" – demonstrates the transcendence and the uncontainability of language. The fact of the matter is that human language is so expressive that it cannot be circumscribed by human experience or by human culture; it is, in some sense, larger than life. Given the sheer reach of human language, one cannot help but wonder how very young children can acquire the complex system of language (when they cannot master much simpler systems such as checkers or chess) and how our long-ago ancestors could have developed a biological system as complex as language in their evolutionary journey from early hominid to modern human.

In this book, we investigate the "boundless expressibility" of human language, attempting to explain how it evolved and how children acquire it. Needless to say, we are not the first scholars to pursue these investigations: numerous philosophers, scientists, and linguists have preceded us: Plato, Aristotle, Descartes, Humboldt, Darwin, and Chomsky, among others; and numerous philosophers, scientists, and linguists continue to study the evolution and development of human language (see, in particular, Hornstein 2009; Bickerton 2009; Fitch 2010; Lieberman 2010; Searle 2010; Chomsky 2010; Ramachandran 2011). However, we bring a unique perspective to this investigation: we focus not on the expressibility of language or on the linguistic structures that underwrite this expressibility, but on the design properties of language communication systems (how language systems align with conceptual and Sensorimotor performance systems) and we argue not only that human language possesses the same basic structural design properties of all other animal communication systems' but also that the "boundless expressibility" unique to human language surprisingly evolves out of these design properties.

Our analysis of human language builds on Chomsky's (1995, 2005, 2010) Minimalist assumption that the design of language is grounded in conceptual necessity. We depart from Chomsky, though, in the application of this assumption. For Chomsky, conceptual necessity applies primarily to the computational system (CS) of human language: that is, it applies to the operations humans use to build linguistic expressions larger than single lexical items and to the domains over which the operations apply. Chomsky invokes conceptual necessity to determine whether human language needs operations other than the Merge operation; and he invokes conceptual necessity to determine whether human language requires levels of syntactic representation beyond some level of logical form; and he invokes conceptual necessity to justify computational domains such as "phases." Although we agree with Chomsky that the computation system (the syntax) of human language should be conceptually justified, and not merely operationally justified – in fact our previous work takes a vigorous look at the sorts of mechanisms permissible in a conceptually necessary CS (see Stroik 1999a, 2009; Putnam 2007; and Putnam and Stroik 2009, 2010) – here we do not limit the assumption that human language is grounded in conceptual necessity to the CS; rather, we generalize this assumption, applying it to the overall design of language systems. In particular, we start with Chomsky's observation that a language system consists of Conceptual-Intentional and Sensorimotor performance systems and a variety of mechanisms that build and send structurally elaborate strings of words to these performance systems for interpretation, and we investigate what conceptual necessity can tell us about how the components of a language system must interrelate with the performance systems for the language system to work. If the biological systems of human language consist of two performance systems (one that oversees linguistic form and one that oversees linguistic meaning), together with a Lexicon and with a CS that puts words together into larger linguistic units, we need to determine how these components interrelate. Are these systems separate subsystems, external to one another (as Chomsky assumes)? Or are they intersective systems with overlapping properties? In this book, we view the structural design properties of human language through the lens of conceptual necessity, discovering that the language system is embedded within the performance systems and that this particular structural relationship stands at the heart of the evolvability of human language, as well as its "boundless expressibility."

1.2 Theory-shaping

Modern biolinguistics, according to Boeckx (2006: 16–17), began with Chomsky's (1959) review of Skinner's *Verbal Behavior* and with his (1965) book, *Aspects of the Theory of Syntax*. It was in these works that the biolinguistic thesis started to take shape – the thesis that "the aim of a theory of grammar is to outline the biologically given cognitive structure that enables human children so reliably and effortlessly to project grammars from the [linguistic] data they receive" (Boeckx 2006: 19). Committed to the biolinguistic thesis, generative grammars from the Aspects model through the Government and Binding framework (see Chomsky 1981) and through the initial developments of the Minimalist Program (MP) (see Chomsky 1995, 2000) have continued to focus on ontogenetic language acquisition. All of these models of grammar have attempted to explain, first and foremost, how first-language grammars develop in children.

In the last decade, however, there has been a turn in the biolinguistic thesis. Even though early work by Chomsky did passingly raise questions about the evolution of language, recent work by Deacon (1997, 2012), Jenkins (2000), Hauser, Chomsky, and Fitch (2002), Chomsky (2005, 2009, 2010), Bickerton (2005, 2009, 2010), Boeckx (2006, 2009), Hinzen (2009a, 2009b), Reuland (2009), Hornstein (2009), and Fitch (2010), among others, has shifted the biolinguistic thesis to include serious consideration of phylogenetic language acquisition. Now the aim of a theory of language is to explain not only how children can biologically acquire a grammar, but also how our species could evolutionarily acquire human language.

As currently developed, the biolinguistic program seeks to answer five interrelated questions (discussed in Jenkins 2000 and in Boeckx 2008b, 2009):

1. What is the knowledge or Faculty of Language (FL)? (Humboldt's Problem)
2. How is this knowledge or Faculty of Language acquired? (Plato's Problem)
3. How is this knowledge put to use? (Descartes' Problem)
4. How is this knowledge implemented in the brain? (Broca's Problem)
5. How did this knowledge emerge in the species? (Darwin's Problem)

We cannot stress enough how much bringing Darwin's Problem (language evolution) into the biolinguistic program has changed the way we conceive of questions 1–4; it changes how we look at the structural design of FL and of the use and acquisition of FL. What Darwin's Problem adds to the biolinguistic program is the challenge that Saussure (1915, 1966) sought to avoid, the challenge of untangling the complex relationship between time and the development of language. Of the many tangles in the evolution of language, the two most pressing involve, as Bickerton (2005, 2009) and Gärdenfors and Osvath (2010) argue, first the emergence of symbolic communication and then the emergence of structured

Cambridge University Press
978-1-107-03483-9 - The Structural Design of Language
Thomas S. Stroik and Michael T. Putnam
Excerpt
More information

symbolic communication. The use of symbols, rather than signs, to communicate is, according to Bickerton (2009: 50), "the Rubicon that had to be crossed for our ancestors to start becoming human"; using symbols allowed our ancestors to extend their domain of communication beyond their immediate experiences, which in turn permitted them to engage in the long-range planning and the collaboration needed for survival. This Rubicon was crossed nearly 2 million years ago (see Bickerton 2009; Corballis 2010; and Bingham 2010), when a branch of hominids developed symbols, perhaps gestural symbols, to communicate. Although our ancestors appear to have mastered the use of symbols long ago, they do not seem to have developed the ability to combine these symbols in the structured ways we modern humans do until relatively recently. So while the development of a symbol-based protolanguage may indeed have been the essential first step towards the evolution of modern humans, it was not until our ancestors evolved the ability to combine their symbols in unbounded ways, thereby providing an unbounded stock of concepts about which and with which to communicate, that modern humans fully emerged.

One of the truly amazing facts about the development of modern language from protolanguage is how quickly this development took place. As Boeckx (2009: 46) observes, "everyone who has thought about the evolution of language seems to agree that it . . . happened within a very short space of time." The consequence of this compressed evolutionary time line, for Hornstein (2009), is that it places significant constraints on language development:

> A common assumption is that language arose in humans in roughly the last 50,000–100,000 years. This is very rapid in evolutionary terms. I suggest the following picture: FL is the product of (at most) one (or two) evolutionary innovations which, when combined with cognitive resources available before the changes that led to language, delivers FL (2009: 4).

Hornstein's speculations severely limit what evolutionarily plausible theories of grammar should look like. Given his speculations, theories of grammar that assume complex sets of language-specific rules (such as Chomsky's 1965 Standard Theory or Gazdar et al.'s 1985 Generalized Phrase Structure Grammar and its aftermath) or complex sets of filters (such as Chomsky and Lasnik's 1977 work) or the optimality theoretic syntax of Pesetsky (1998) or complex sets of language-specific modules (such as Chomsky 1981 and Chomsky and Lasnik 1993) will be evolutionarily implausible because the development of these complex structures would be too time-intensive. Instead, a theory of grammar should be a theory of FL that, as Hauser, Chomsky, and Fitch (2002) maintain, consists of pre-linguistic cognitive processes (called Faculty of Language broad, FLB) able to subserve language processes and an extremely small set of language-specific processes (called Faculty of Language narrow, FLN).

But how exactly do we go about developing a theory of FL? For Chomsky (2005: 11), this question is "The general question . . . How far can we progress in showing that all language-specific technology is reducible to principled explanation, thus isolating the core properties that are essential to the language faculty, a basic problem of linguistics?" Chomsky argues that there are three factors that reduce language-specific technology in FL: (i) genetic endowment, (ii) experience, and (iii) principles not specific to the FL. Of these three factors, it is, according to Chomsky, third-factor principles such as "language-independent principles of data processing, structural architecture, and computational efficiency . . . [that provide] some answers to the fundamental questions of biology of language, its nature, use, and perhaps even its evolution" (2005:9).

Investigating the third-factor principles underlying FL, however, requires, from Chomsky's perspective, that we approach language as a natural, biological object. The consequence of treating language as part of the natural world is that the design structures of language must obey physical laws (which we call Turing's Thesis following Jenkins 2000 and Ott 2009) and that the computational processes involved in language must, as Galileo's Thesis requires, follow mathematical laws (see Boeckx 2006 for a discussion of Galileo's Thesis). This means that the structural mechanisms and the computational operations posited for FL must meet thresholds of simplicity, generality, naturalness, and beauty, as well as of biological and conceptual necessity. Or, as Hornstein (2009: 3) states, "FL will be natural if it is based on principles and operations that promote computational tractability, that are built from parts that are cognitively general and atomic, and that are basic enough to be (plausibly) embodied in neural circuitry." The goal of this book is to investigate how Hornstein's challenge can be met.

Since our response to Hornstein's challenge is largely shaped with Turing's Thesis in mind (and the constraints this thesis places on the structural design of human language), we would like to take a close look at Turing's Thesis and how it applies to our analysis of language. In his study of morphogenesis, Turing (1952: 37) makes the observation that "The purpose of this paper is to discuss a possible mechanism by which the genes of a zygote may determine the anatomical structure of the resulting organism . . . [the] theory does not make any new hypothesis; it merely suggests that certain well-known physical laws are sufficient to account for many of the facts." As Turing admits, his thesis (which grounds biology in physical, and arguably structural, laws) is not a new thesis. Versions of this thesis have been advanced by others, such as Thompson, who notes that

> Cell and tissue, shell and bone, leaf and flower, are so many portions of matter, and it is in obedience to the laws of physics that their particles have been moved, moulded, and conformed . . . Their problems of form are in the first instance mathematical problems,

their problems of growth are essentially physical problems, and the morphologist is, *ipso facto*, a student of physical science (1917/1992:10).

The importance of Turing's Thesis to our work is that it brings physical constraints and structural design into biology. It calls for, as Fodor and Piattelli-Palmarini (2010: 74) note, a constraint-based biology that questions "blind trial and error followed by natural selection" as a basis for biological growth and assumes instead that "It's vastly more plausible to suppose that the causes of these [biological] forms are to be found in the elaborate self-organizing interactions between several components that are indeed coded for genes (protein complexes, morphogenetic gradients, hormones, cell–cell interactions, and so on) and the strictures dictated by chemical and physical forces." Given Turing's Thesis, we should expect biological form and growth to be shaped by physical laws and structural design, an expectation supported by Deacon (2012: 68–69), who asserts that

Although the features that distinguish these different, specialized cell types depend on gene expression, the geometry of the developing embryo plays a critical role in determining which genes in which cells in which ways . . . Many popular accounts of DNA . . . ignore (that) the information ultimately embodied in the elaborate patterns of interaction among cells and how these affect which genes get expressed.

According to this line of argumentation, the evolution of organisms is not primarily governed by extrinsic forces, but rather is driven by the design properties internal to the organism in question. Or as Fodor and Piattelli-Palmarini (2010: 27) put it, "the whole process of development, from the fertilized egg to the adult, modulates the phenotypic effects of genotypic changes, and thus 'filters' the phenotypic options that ecological variables ever have a chance to select from." We accept the assumptions of constraint-based theories of biology and are committed to looking at the development of language in light of Turing's Thesis, investigating language as a biological system that emerges out of the interplay of internal, phenotypical constraints placed on the structural design of human language.

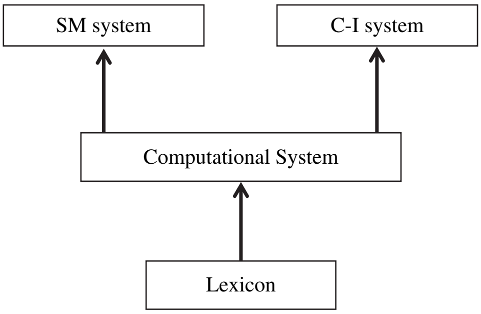
1.3 Minimalism and the design of the Faculty of Language

Minimalist inquiries into FL seek to formalize unified answers to questions 1–5 that also satisfy the constraints imposed by both evolutionary biology and conceptual necessity. These answers embrace "[t]he central idea . . . that our human language capacity . . . shows signs of optimal design and exquisite organization that indicate the inner workings of very simple and general computational (physical and mathematical) laws in a mental organ" (Boeckx 2006: 3). The challenge that Minimalism sets for itself, then, is to design an architecture for FL built on the principles of simplicity and economy.

Before moving forward we want to consider the plausibility of addressing the challenge above. Is this strong challenge one that can actually be met? Is it really possible to design an architecture for the Faculty of language that is grounded, in accordance with Minimalist assumptions, in both evolutionary biology and conceptual necessity? Kinsella (2009) thinks not. She contends that conceptual necessity in Minimalism is built on a notion of "perfection," in which "[t]he perfection of language is . . . understood to result from the simple uncluttered way that the system operates" (48), but such perfection simply "does not arise in nature as the result of gradual adaptive evolutionary processes" (60). Operational perfection of the sort advanced by Minimalism, as Kinsella rightly observes, runs afoul of evolutionary processes; that is, "perfection directly contradicts the bias for redundant, uneconomical systems, the trend for evolution to meliorize" (160). Kinsella makes a very important point: evolutionary processes are driven by contingencies not by conceptual necessity, and the structures that arise from adaptive processes are not conceptually necessary structures. So, if "language is a system that is grounded in biology" (91), it will not be a perfect system or even, as is often claimed about Minimalism, "a perfect solution to meeting the interface conditions" (57). Does this mean that the Minimalist challenge given above, as well as the Minimalist Program itself, should be abandoned as untenable? Kinsella thinks so; we think not. We think that the challenge only needs to be reformulated. While adaptive evolutionary processes might not yield optimal or perfect systems, they can yield optimal or perfect output. We certainly see optimal output in the petal-patterns of flowers and the structural symmetries of cells in honeycombs.[1] Given Kinsella's arguments, neither bees nor flowers have perfect designs. However, these imperfectly designed organisms can still produce perfect output. The same can be true of FL. That is, if the "perfection" of FL is in its output – in the fact that the linguistic structures produced by the CS of FL are necessarily usable by the performance systems – then FL, though an "imperfect" and adaptive system, can be grounded both in evolutionary biology and in conceptual necessity. But is the computational output of FL "perfect" in the sense of being necessarily usable by the performance systems?[2] It has been a long-held assumption in generative grammar that the computational output of the grammar is perfect – that, as Chomsky (1965) maintains, a grammar derives structural descriptions for all and only the sentences of a given language. Minimalism incorporates this assumption into its design of FL by having the CS meet (not feed) the interface conditions imposed by the performance systems. We, too, accept what we call the Strong Computational Thesis (SCT) (that the computational output of language is necessarily usable by its performance systems) as a working hypothesis, because denying the SCT opens the possibility of having an unknowable CS – one that in principle could generate unusable output that could not be tracked, identified, or studied. Given the SCT, language can be a biological system that is a perfect system grounded in conceptual necessity. This, however, would be true not only of the human language system, but of all animal communication systems. In other words, the perfection of the language system is not chauvinistically restricted to the human FL; rather, as Bickerton (2009) argues, all animal communication systems are designed to perfectly meet the communicative needs of the users. Of note, this notion of "perfection" applies to human protolanguage, which means that human language is not more perfect than human protolanguage – the former is merely an adaptive variant of the latter.

With the above discussion in mind, let's look at the Minimalist design for FL. Since its inception (see Chomsky 1995), the Minimalist Program has hypothesized that FL requires four independent (but connectable) subsystems: a Sensorimotor (SM) system, a Conceptual-Intentional (C-I) system, a Lexicon (LEX), and a computational system (CS). Broadly described, the SM system emits and receives linguistic signals; the C-I system interprets linguistic signals; LEX stores information about words; and CS generates complex linguistic structures from the words stored in LEX. Under standard Minimalist assumptions, the four subsystems of FL produce linguistic structures in the following way: lexical items (LIs) are placed in short-term memory (in a Numeration (NUM)); the LIs are brought into the CS and are structurally put together (via various syntactic operations to be discussed later); and these structures are sent to the external performance systems – SM and C-I – for phonetic and semantic-pragmatic interpretation. The architecture for FL, under these assumptions, looks like (6).

(6)



Of special note here, the performance systems in (6) are pre-linguistic systems that can "use" the linguistic structures produced by CS. These performance systems evolutionarily precede the linguistic systems (LEX and CS) and they are external to the linguistic systems. Because the performance systems are external to the linguistic system, linguistic structures produced by CS must be "transferred" to, or "shipped" off to, the performance systems for interpretation, as Chomsky (2005: 16) and Hornstein, Nunes, and Grohmann (2005: 46), among others, propose.

As we state above, the performance systems are pre-linguistic systems. They are ancestral systems that are millions of years old. The SM system, which is involved in the production and reception of gestures and/or acoustic signals, certainly pre-dates the evolution of homo sapiens; and the C-I system, which is responsible for the inferencing processes about the world, including our ability to project a Theory of Mind for others in which we can "read" their intentions from their actions, is so ancestral that it is something we share with other species of great apes.[3] Under Minimalist assumptions, the linguistic systems developed by humans use these older performance systems; however, they do so in compliance with established performance system conditions. That is, the linguistic system's "internal mechanisms and operations are not arbitrary, but a least-effort solution to minimal design specifications" (Ott 2009: 255). In slightly less technical language, this means that the linguistic system must be linked optimally and naturally with the performance systems it will utilize. Importantly, Chomsky (2000) argues that investigations into FL should be guided by the assumption that there is an optimal linkage between the linguistic system and the external performance systems – he calls this assumption the Strong Minimalist Thesis (SMT).[4]

Satisfying the SMT has spawned a rash of putatively necessary mechanisms for the linguistic system. The most widely embraced of these mechanisms is the Merge operation – an operation that "takes two elements, α and β, and combines them into a set [α, β]" (Boeckx 2008a: 28); this operation is seen as conceptually necessary because it captures the compositional nature of human language. Although all versions of Minimalism accept some variation of the Merge operation, there is considerable disagreement about the properties of Merge. For one, Chomsky (2009: 29) and Ott (2009) propose that Merge is unbounded, allowing any two elements to merge together, as in (7a); on the other hand, some theorists such as Adger (2003, 2008) and Stroik (2009a) propose that Merge requires the two merged elements to have matching features, as in (7b).

(7) a. Merge $\{\alpha, \beta\} \rightarrow [\alpha\ \beta]$
b. Merge $\{\alpha[+F], \beta[+F]\} \rightarrow [\alpha\ \beta]$

For another, Chomsky (2005: 15) contends that Merge is the basic operation of FL, while Hornstein (2009) and Boeckx (2009) suggest that Merge is the product of two even more basic operations – Concatenate and Copy – that are not specific to FL.[5] Third, Chomsky (2005) argues that there are two flavors of Merge: External Merge, which introduces an LI into a syntactic derivation (SD), and Internal Merge,[6] which re-combines an element already in a syntactic

Cambridge University Press
978-1-107-03483-9 - The Structural Design of Language
Thomas S. Stroik and Michael T. Putnam
Excerpt
More information

derivation with the derivation; Hinzen (2009b), Stroik (2009a), and Putnam and Stroik (2009) argue, contra Chomsky, that Internal Merge is not a conceptually necessary operation.[7] And finally, there is a great deal of disagreement about the output of the Merge operations. Some theorists, such as Chomsky (2005) and Epstein and Seely (2006), permit the Merge operations to produce output that is not usable by the performance systems (output that "crashes" at the performance systems); others (for example Stroik (1999, 2009a), Frampton and Gutmann (2002), and Hinzen (2006)) contend that the CS must be crash-proof, sending only usable output to the performance systems.[8]

In addition to the Merge operation, some versions of Minimalism (see Chomsky 2000 and Boeckx 2008a, 2008b) have an Agree operation that allows one element (a Probe) to check the matching features of another element (a Goal). The Agree operation, which, as Boeckx (2008a: 77) states, "can take place at a distance," is used to explain the agreement relations between the main verb "appear" and the embedded subject "six very hungry children" in (8).

(8) There appear to be six very hungry children sitting at the kitchen table.

Despite rather wide acceptance of the Agree operation, there is dispute about the properties of this operation. Some theorists, such as Lopez (2007), Hornstein (2009), Preminger (2009) and Sigurðsson (2006), accept the Agree operation, but they contend that this operation cannot apply at a distance; rather, it is restricted to local domains (Hornstein (2009: 126–130) offers a conceptual argument against long-distance Agree, claiming that this sort of Agree is unnecessary because it duplicates the Internal Merge operation). Stroik (2009a) takes an even more unsympathetic position on the Agree operation; for him, the only operation necessary in the CS is External Merge. Hence, he rejects the Agree operation.

One final CS operation proposed by various theorists is the Transfer operation, an operation that ships derivations, or portions of derivations, to the performance systems. Given that the CS is outside of the performance systems under standard Minimalist assumptions, it seems conceptually necessary that the output of the CS be sent along to the SM and C-I systems; hence, a Transfer operation would seem to be a conceptually necessary part of FL.[9] However, when does this operation apply? After each Merge operation applies, as Epstein and Seely (2002, 2006) and Boeckx (2008a) propose? After a phase head (C or *v**) is merged, as Chomsky (2001, 2005) and Ott (2009) propose? Or after the SD is complete, as Stroik (2009a) proposes?

As we can see, although there is some agreement about the CS operations in FL, there is also substantial disagreement (a matter that requires a conceptual resolution, under Galileo's Thesis). A similar situation arises with the Lexicon. Most theorists would agree with Tallerman's (2009: 181) claim that "The