Biological Foundations and Origin of Syntax

Edited by

Derek Bickerton and Eörs Szathmáry

Program Advisory Committee:
Derek Bickerton, Balázs Gulyás, Simon Kirby,
Luc Steels, and Eörs Szathmáry

The MIT Press

Cambridge, Massachusetts
London, England

Series Editor: J. Lupp
Assistant Editor: M. Turner
Photographs: U. Dettmar
Typeset by BerlinScienceWorks

The book was set in TimesNewRoman and Arial.
Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Ernst Strüngmann Forum (2008 : Frankfurt, Germany)
Biological foundations and origin of syntax / edited by Derek Bickerton and Eörs Szathmáry.
p. cm. — (Strüngmann Forum reports)
Forum held July 13–18, 2008, in Frankfurt, Germany.
Includes bibliographical references and index.
ISBN 978-0-262-01356-7 (hardcover : alk. paper)
1. Grammar, Comparative and general—Syntax—Congresses. 2. Biolinguistics—Congresses. I. Bickerton, Derek. II. Szathmáry, Eörs. III. Title.
P291.E75 2009
612.8'2336—dc22
2009023478

1

Syntax for Non-syntacticians

A Brief Primer

Derek Bickerton

Abstract

Some of the most basic concepts and processes in syntax—Merge (and the hierarchical structures it creates), binding, control, movement, and empty categories (elements that are "understood" but not phonetically expressed)—are briefly and simply described and illustrated. The chapter concludes with some suggestions regarding possible avenues of approach towards a fuller understanding of how syntax is instantiated in the human mind/brain.

Introduction

For the average lay person, word order is the most significant thing about syntax. For some, it's all of syntax.

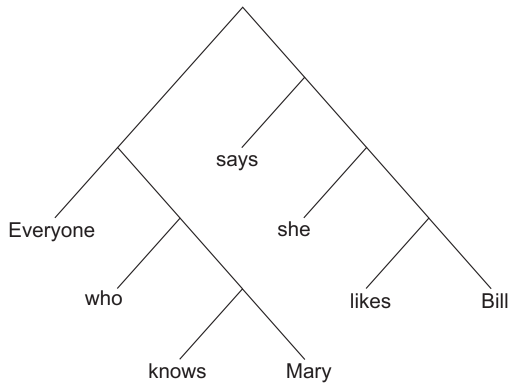
Nothing could be further from the truth. In fact you could argue that word order is an epiphenomenon, necessitated by the fact that we have only a single channel of speech, forcing words to be produced in a linear order, and that word order falls out merely from reading the terminal nodes of a hierarchical tree structure (such as that below) from left to right.

The most significant thing about syntax is its hierarchical structure. Syntactic trees are not just heuristic devices; they reflect how sentences must be constructed. Take the following example:

(1) Everyone who knows Mary says she likes Bill.

In linear order, *Mary* and *says* are adjacent and seem to show evidence of subject–verb agreement; indeed, *Mary says she likes Bill* is in itself a complete and fully grammatical sentence. In hierarchical structure, however, *Mary* and *says* are far apart from one another and not in a subject–verb relationship.

The proper relationship between *Mary* and *says* is shown when we create a simplified hierarchical tree structure for the sentence (many details omitted):



It will be noted that, as stated above, a reading of the node terminals from left to right yields the sentence described by, and analyzed in, the syntactic tree.

Merge as the Central Process in Syntax

As the above suggests, what one might intuitively describe as the "closeness" of words is illustrated by their positions in the tree. The closest relationship is that known as "sisterhood"—one between two words exclusively dominated by a single node, such as *knows* and *Mary*, or *likes* and *Bill*. Although *Mary* and *says* have a linear adjacency similar to that between *knows* and *Mary*, the relationship between them is remote, since the only node that dominates both is the node that dominates the entire sentence. The degree of closeness is further shown by the fact that words cannot normally be inserted between sisters: we can say, *She obviously likes Bill* or *She likes Bill, obviously*, but not **She likes obviously Bill*. (An asterisk in front of a sentence indicates that the sentence is ungrammatical.)[1]

The basic syntax-creating process proposed by Chomsky's Minimalist Program (Chomsky 1995) is Merge: a process that takes any two units (words, phrases, clauses…) and forms them into a single unit, subject to "feature-matching." For instance, *likes* has the feature [+transitive] that must be satisfied by a Theme argument in object position, while a noun like *Bill* must seek a verb that still requires an argument. The unit [*likes Bill*] requires a subject that is third-person singular; *she* supplies this deficiency and is therefore merged with [*likes Bill*] to yield [*she* [*likes Bill*]]. (Note that bracketing is simply a

[1] It should be noted that while in English a verb and its direct object are generally sisters, in many European languages an adverb or other verbal modifier, if present, may be a sister of the verb (at least in surface structure; see Rizzi, this volume).

notational alternative exactly mirroring the structure of a syntactic tree.) The process is repeated until the sentence is complete, and naturally it cannot be completed until all the requirements of the lexical items it contains have been satisfied. Here are a few simple examples of sentence failures:

(2) (a) *broke the bottle (lacks an agent argument in subject position)
(b) *Mary see John daily (third-person agreement required)
(c) *Reciprocity invited Bill (verb needs [+human] agent)
(d) *There seems someone to be asleep (*someone* inappropriately merged; dummy subject *there* appropriate only if *someone* is first merged with *asleep*)

The order in which the various operations of Merge are performed will, of course, determine the linear order of constituents. In English, the general rule is that verbs will first be merged with their direct objects (if present); then the resultant verb–object combination will be merged with indirect objects (if there are any) and adverbial phrases, subjects being merged last into the structure. Other languages may adopt different orders. The overriding consideration is that the argument structure of the sentence be recoverable—in other words, that it is possible to determine who did what to whom, when, and where (if the latter are specified). Languages may adopt strategies other than sequential ordering of Merge (e.g, case marking) in order to achieve a greater degree of freedom in ordering constituents. (In general, earlier in the sentence equals old information, later means newer, so some degree of flexibility in ordering may be communicatively advantageous.)

Although this was not the purpose for which it was originally designed, sentence construction via Merge forms a plausible model of how the brain may actually operate in creating sentences. Much evidence suggests that, rather than sending individual words directly to the motor organs of speech, the brain combines neural signals representing several words (at least up to the level of phrases or short clauses, perhaps demarcated by intonational phenomena) before it dispatches them for utterance. Such segments may well correspond to Chomsky's proposed "phases" (Chomsky 2001) which are dispatched individually to Phonetic Form for Spell-out, after which such phases are inaccessible for further syntactic computation. It should be apparent that Merge (as distinct from the purely linear bead-stringing process underlying protolanguage, where it is assumed, on the basis of phonological behavior by speakers of pidgin [Bickerton 2008, chap. 8; Bickerton and Odo 1976], that single words are dispatched separately to the speech organs) automatically creates hierarchical structures.

A crucial question for anyone examining the evolutionary origins of syntax is whether Merge involves recursion. It has been claimed in one widely cited paper (Hauser et al. 2002) that it does, and that recursion is unique to humans, requiring us to assume either a special mutation or the exaptation of some task-specific mechanism that predated—hence originally had nothing to do with—language. In linguistics, recursion is generally defined as the ability to insert one structure inside another of the same kind. However, this concept seems to have originally arisen from early (and long-abandoned) generative analyses (Bickerton 2009, chap. 12). According to a reviewer of an earlier version of this paper, it is only if one adopts an unlabeled version of Bare Phrase Structure (as in, e.g., Collins 2002) that the usual definition of recursion fails to apply. This view, however, seems to confuse notational with operational criteria. The issue is not whether, in some descriptive notation, the label for one structural type does or does not fall within the boundaries indicated by the label of a similar structural type (CP within CP, NP within NP…). It is whether (assuming Merge to be an actual operation performed by the brain) the brain, in the course of constructing sentences, actually inserts a member of one structural type within another member of the same type. Were this the case, the claim by Hauser et al. (2002) that some uniquely human adaptation is required for recursion might stand a chance of being true. However, if the brain obeys Merge, it does not insert anything within anything, but merely merges ever-larger segments of lexical material with one another until a complete sentence is achieved.

It is, of course, possible to adopt a looser definition of recursion, one that makes Merge itself a recursive process, by defining the latter as any process that uses the output of one stage as the input to the next. This solution is adopted by Rizzi (this volume). However, recursion defined in these terms could apply to almost any process, including processes routinely executed by other species—a bird building its nest, for example:

Step 1: Weave two twigs together.
Step 2: Interweave a third twig with the interwoven pair.
Step 3: Interweave a fourth twig with the interwoven three, etc.

With this second definition, any justification for claiming recursion as a uniquely human component of the narrow language faculty simply disappears. In short, Merge itself can be treated as an iterative rather than a recursive operation, requiring no specialized development in the human brain.[2]

[2] It is useful to relate these two approaches to linguistic recursion to concepts of recursion in mathematics and computer science, especially since the modeling approaches are gaining momentum. With some simplification, Rizzi's second definition is like any primitive recursive function that can repeatedly be applied to its own output; most well-known functions are like this (such as addition, multiplication, exponentiation, etc.). In computer science, solving a problem using recursion means the solution depends on solutions to smaller instances of the same problem. A loop in a computer program, when executed, calls itself again and again until a certain condition is met, when the program jumps out of the loop. The very definition of a factorial is recursive in this richer sense, since fact(n) = 1 if n = 0, and fact(n) = n fact(n – 1) if n > 0. Merge could be regarded as an operation that calls itself in the course of sentence construction: the accomplishment of Merge at the lowest levels of the tree requires some linguistic features that trigger a further application of Merge, until the sentence is completely assembled. However, it should be noted that the recursivity of the process depends crucially on the materials merged, i.e., words. It is the fact that words (and combinations of words) have

However, while Merge is central to syntax, it far from exhausts the phenomena that an evolutionary account of syntax must explain. Among the most salient of these are binding, control, movement, and the reference of empty categories. All of these processes (with minor variations specific to individual languages) are found universally in language, and have in common the fact that they can only operate over limited domains. In other words, all seem to involve some principle or principles of "locality" that demarcate the sections of trees over which processes can operate.

Binding

Binding is a relationship that exists between anaphors (e.g., reflexives, reciprocals, pronouns) and their antecedents. Typically, pronouns are free in reference. Take the sentence we already examined:

(3) Everyone who knows Mary thinks she likes Bill.

On pragmatic grounds, we prefer an interpretation that identifies *she* as *Mary*. However, this is not necessarily the case; *she* could, in principle, refer to any individual female. However, in the sentence *Mary likes her*, *Mary* and *her* cannot refer to the same person, whereas in *Mary likes herself*, *herself* can only refer to Mary.

In general, anaphors cannot occur outside the clause that contains their antecedent. For instance, in the sentence *Mary asked Susan to wash herself*, the meaning cannot be that *Mary asked Susan to wash Mary*. However, there are exceptions in both directions. The following sentence, for example, is grammatical, with *himself* co-referring outside its clause with *John*.

(4) John believes that stories about himself are exaggerated.

But then so is:

(5) Mary saw a snake near her.

Here, a pronoun within the same clause may refer to *Mary*. Definition of a binding domain has therefore not proven easy, even for English (without taking the complications introduced by other languages into account). Consequently, the fact that native speakers of any language can acquire, without explicit training, distinctions that have eluded trained linguists suggests that there must be a limited set of possible algorithms for binding, innately established, and that

dependencies that must be filled which drives repeated applications of Merge until the problem is solved—that is, until the complete grammatical sentence is generated. In other words, it is not simply Merge, but rather Merge + lexical material that constitute the recursive process, as well as force it to result in a hierarchical structure. This suggests that Pinker and Jackendoff (2005) were correct when, contra Hauser et al. (2002), they listed words as a uniquely human part of the language faculty.

children can select among these on the basis of very limited data (what kind of data, how much, and much more remain topics for future inquiry).

Control

The subjects of certain nonfinite verbs in subordinate clauses are "controlled" (given obligatory reference) by constituents of higher clauses. Typical are sentences of the following kind:

(6) Mary expects to arrive shortly.

Although there is no overt subject for *arrive*, that verb is understood to be coreferential with the subject of the higher verb, *Mary*. However, if the higher verb also has an object, this usually supplies the (understood) subject of the lower verb:

(7) Mary told Bill to leave immediately.

Here, *Bill* is taken as the subject of *leave*. There are some exceptions to this rule, for example:

(8) Bill promised Mary to leave immediately.

Here, *Bill* is understood as the subject of *leave*.

Control, as compared to binding, seems to involve a relatively straightforward algorithm involving locality: select the "closest" (in terms of tree structure) possible antecedent for the verb in question, subject to lexical exceptions (like *promise*) that are, in turn, determined by purely semantic considerations.

Movement

Do constituents of sentences "really" move around, or can the empirical data that suggest they do be explained some other way? This question, in one form or another, has concerned generativists since the dawn of generative grammar, even though Chomsky himself once suggested that movement and alternative explanations might merely be "notational variants" of one another. (The question's most recent incarnation involves attempts to derive Move from—or incorporate it somehow into—Merge.)

Phenomena that suggest the existence of movement involve cases where some constituent seems to have been "displaced" from its "normal" position:

(9) Who did Mary see___?

Since *who* clearly functions as the object of *see*, movement theorists hypothesize that *who*, originally in the object position, has been moved to the left periphery of the sentence. Seemingly confirming this is the parallel evidence of "express surprise/request confirmation" pseudo-questions such as:

(10) Mary saw who?

What is referred to as the copy theory of movement removes some, but not all, of the problematic features of movement by assuming an intermediate stage in the derivation consisting of *Who did Mary see who*? followed by deletion of the lower occurrence. (Other possible solutions—the leaving of “traces” at movement sites, or the assumption that those sites contain “empty categories” subject to general rules of interpretation for such categories—need not concern us here, although we return to the second issue below.) What concerns us here is that the relationships between “extraction sites” and overt positions—often referred to as long-distance dependencies (LDDs)—fall under locality restrictions somewhat different from those that apply in cases of control and binding.

Some pre-minimalist versions of generative grammar sought to capture those restrictions with a notion known as subjacency (originally “lying nearby but lower”), which defined the domain within which LDDs could hold as (roughly) the clause that had the moved constituent at its left periphery plus any chain of subcategorized clauses below that clause in the tree. Thus a *wh*-word could be moved to the left periphery of any sentence along the lines of:

(11) Who do you think Bill believes that Mary saw ___?

A subcategorized clause is one that is required by a higher verb; *think* and *believe* both take sentential complements. However, if a subordinate clause is not subcategorized, movement results in ungrammaticality:

(12) (a) *Who did Bill have a bad cold ever since he met ___?
(b) *What did John sing ___ and Mary accompanied him?
(c) *Who did the fact that Mary likes ___ irritate Bill?
(d) *What did Mary see the man who played ___?

Note that

(13) (a) Bill has had a bad cold ever since he met John.
(b) John sang German folksongs and Mary accompanied him.
(c) The fact that Mary likes John irritates Bill.
(d) Mary saw the man who played the tuba.

are all perfectly grammatical sentences that would be appropriate answers to the questions in (12), if those questions could be asked. Items in the underlined positions, however, cannot be questioned, at least not in the ways shown.

The ungrammaticality of sentences like in (12) cannot be dismissed as a result of semantic or pragmatic factors. Sentences are assembled automatically, at very high rates of speed—too fast for any conscious monitoring (or even conscious awareness) of the process. There must surely be some kind of algorithm the brain executes to ensure that (except through very occasional error) sentences like the starred examples do not occur.

Reference of Empty Categories

As we saw above, hypothesized movement creates empty categories—sites where you would expect to find an overt constituent but do not, where you have to search among overtly expressed items to determine what these empty categories refer to. But empty categories also occur in sentences where no movement—or at least, no overt movement of the kind discussed in the previous section—has occurred. In many of these latter cases, in contrast to overt movement cases, empty categories may alternate with pronouns, yielding pairs of sentences both of which are fully grammatical but which differ sharply in meaning. Consider examples like the following:

(14) (a) Mary is too angry to talk to.
(b) Mary is too angry to talk to her.

In the first example, there are two empty categories, respectively subject and object of *talk to*:

(15) Mary is too angry ___ to talk to ___.

At first sight, the first empty category might seem to fall under "control," as described above; however, the algorithm that determines empty-category reference has to start with the most deeply embedded empty category (the one nearest the bottom of the syntactic tree). This must be identified with an overt referential constituent in the sentence, and there is only one, *Mary*, which must accordingly be identified as the object. However, since no two items in the same clause may co-refer (unless that fact is marked by a reflexive pronoun or some other explicit marker of co-reference), *Mary* cannot be interpreted as the subject in this case. Thus the first empty category must be assigned arbitrary reference, i.e., interpreted as *people* (unspecified) or *anyone*, hence the meaning of the sentence is *Mary is too angry for anyone to talk to Mary*.

The presence of a pronoun in the second member of the pair changes the dynamics of the sentence. Since the first empty category (subject of *talk to*) is now the only one in the sentence, it becomes free to co-refer with *Mary*, and since clause-mates cannot co-refer, *her* can then no longer be interpreted as *Mary* but must be read as *some female person other than Mary*.

Note that, in a superficially similar pair of sentences, the references of pronoun and empty category are reversed; the pronoun can be (and will be, with a very high degree of probability) interpreted as *Mary*, while the empty category cannot be thus interpreted.

(16) (a) Mary needs somebody to talk to.
(b) Mary needs somebody to talk to her.

As before, the first sentence contains two empty categories, in the same positions and with the same functions as before:

(17) Mary needs somebody ___ to talk to ___.

In this case, in contrast to *Mary is too angry to talk to*, there are two referential constituents in the sentence: *Mary* and *somebody*. As before, the deepest-embedded empty category is identified first, and since its nearest possible antecedent is *somebody*, it is thus interpreted, leaving the first empty category to select the next-closest referential unit, *Mary*, as its antecedent. In the second member of the pair, however, the presence of the pronoun leaves the subject of *talk to* as the only empty category requiring a co-referent. Since *somebody* is the closest referential item, it is interpreted as the subject, leaving *Mary* as a possible (and for pragmatic reasons, by far the likeliest) antecedent.

It would be satisfying if we could take all cases of empty-category reference assignment, including those cases involving movement and control, and subsume them under a single algorithm. However, this does not seem to be possible, since the item co-referent with an extraction site may be found indefinitely far from that site, while the site itself must obligatorily remain empty (i.e., cannot alternate with a pronoun, unlike the cases just discussed). For our present purposes, it is unnecessary to attempt to find a common algorithm; we need merely to understand that in order to handle syntactic processes, the brain must be sensitive both to the nature and extent of specific domains (e.g., subcategorized and non-subcategorized clauses and phrases) and to relative distances (in terms of position in a tree, not serial order) between referential constituents.

For those who would claim that the aspects of syntax discussed above could constitute learned behaviors, rather than the results of automatic, autonomic neural processes, the reference of empty categories represents the most difficult counterevidence to settle. How do you learn a nothing? Here the burden of proof clearly lies with those who would argue for learning as an adequate explanation of syntactic processes.

Beyond Analysis

To summarize, syntax consists of a process for progressively merging words into larger units, upon which are superimposed algorithms that determine the reference of items (in various types of structural configuration) that might otherwise be ambiguous or misleading. Many other factors—far more than can be mentioned here—are implicated in syntax, but the processes described above are central to it, and if these can be accounted for in evolutionary and neurobiological terms, we will have taken a massive step forward.

Although strictly speaking this lies outside the scope of the present paper, it may be worthwhile to briefly glance at some further considerations that affect the problems facing us, and suggest some lines of inquiry that might prove fruitful, or at worst, provide a null hypothesis that may serve as a basis for further investigation.

If we are looking at biological foundations and origins, we would do well to forget about the distinction between competence and performance. This distinction proved useful, in the early days of generative grammar, to determine what was valid empirical evidence and what was mere accident and happenstance. The distinction between ungrammatical and grammatical sentences was crucial to the enterprise—no one previously had thought about what *couldn't* be said, merely about what could—and the distinction seemed best made on the basis of a presumed body of invariant knowledge, either Universal Grammar or the grammar of a specific language. So was born the distinction between competence (knowledge of language) and performance (how that knowledge was expressed in the creation of actual sentences, subject to all the potential confounds of memory limitations, slips of the tongue, and so forth), with competence as the primary focus of research and performance generally downgraded, ignored, or postponed for future inquiry.

There is good reason, however, to believe that the distinction between competence and performance has outlived its usefulness. It is surely not by accident that a standard reference work on the cognitive sciences (Wilson and Keil 1999) contains no entry for "competence–performance distinction" (even though most other issues in generative grammar are extensively documented) but merely refers the reader to three other entries, none of which mentions the competence–performance distinction. Any evolutionary account surely demands that we treat language as an acquired behavior rather than a static body of knowledge, a behavior with deep roots in our biology but expressed through real-time actions of neurons, axons, synapses, and speech organs. Our focus should be firmly fixed on what humans, and their brains, had to be able to do in order to rapidly, automatically, and unconsciously produce sentences that would fall within the quite narrow bounds that delimit human language. What was required for these tasks might in principle represent (or be derived from) either capacities shared with other animals but selected for novel purposes, or novel, purpose-built capacities unique to the human species. But the null hypothesis clearly is that given symbolic units to combine, the processes used to combine them were ones of a fairly general nature, already present in the genome and shared with other mammalian (and perhaps even avian) species. While some rewiring of the brain may well have been required, it was most likely a rewiring of existing areas (areas that had previously had a variety of functions), rather than one that required the superimposition of some task-specific mechanism unique to (and uniquely used for) language.

Aside from what was required to produce Merge (in all probability a domain-general, iterative mechanism already used to integrate sensory inputs within single modalities and perhaps also cross-modally), it is worth considering the likeliest mechanism for use in dealing with the other syntactic processes outlined above. These, as we have seen, require sensitivity to distance between individual constituents and sensitivity to boundaries between syntactic units. The most parsimonious solution would, as with Merge, employ one (or more than one) already-existing domain-general mechanism, and a plausible candidate would be the brain's ability to establish sequence by means of the fading of neural signals.

All neural signals degrade over time, so that *ceteris paribus* the equations stronger = later and weaker = earlier should hold over all brain operations. Since Merge combines constituents in a temporal series of sequential operations, it should be possible for the brain to keep track of sequence by exploiting this aspect of neural signaling.

Note that focusing on processes and how the brain might execute them, rather than on hypothesized linguistic knowledge, might help to restrain what has been one of the most recurrent (as well as the most frustrating) developments in generative grammar. That theory has passed, since its inception, through at least five avatars: Chomsky's original (1957) formulation, the Standard Theory, the Extended Standard Theory, the Principles and Parameters framework, and the Minimalist Program. Each new avatar has undertaken a radical revision of its predecessor, and each revision has been prompted by the fact that the preceding version had proliferated and complicated itself to an inordinate extent. However, each of these versions eventually succumbed to the very failings it was created to correct. The latest, the Minimalist Program, is no exception. In its origins the leanest and most elegant expression of generative grammar (see Rizzi's admirably clear and concise description in this volume), it has become a jungle of arcane technicalities, daunting for any neurobiologist who might hope to reconcile it with the workings of the human brain.

The need is obvious: to nail down the biological foundations of syntax demands scholars who combine a thorough training in both linguistics and neurobiology, enough so that they give neither precedence over the other in their thinking. Given the traditional Balkanization of science, this may be a lot to ask. But until such scholars are available, researchers in both communities should remain as open to one another as possible. There is a great temptation to say things like, "That linguistic model could never be realized by the brain," or "That neurological model is far too simplistic to account for language." Such remarks, however plausible they may seem, are no more than assumptions. The issues themselves are empirical. The brain creates sentences somehow, and the way it actually does this must constitute the real and only true grammar of language.