A Prospect for Evolutionary Adequacy: Merge and the Evolution and Development of Human Language

Koji Fujita

Biolinguistic minimalism seeks a deeper explanation of the design, development and evolution of human language by reducing its core domain to the bare minimum including the set-formation operation Merge. In an attempt to open an avenue of research that may lead to an evolutionarily adequate theory of language, this article makes the following proposals: (i) Merge is the elementary combinatorial device that requires no more decomposition; (ii) the precursor to Merge may be found in the uniquely human capacity for hierarchical object manipulation; (iii) the uniqueness of the human lexicon may also be captured in terms of Merge. Empirical validations of these proposals should constitute one major topic for the biolinguistic program.

Keywords: action grammar; anti-lexicalism; evolutionary adequacy; FLN/FLB; unbounded Merge

1. The Logical Problem of Language Evolution

Language is a biological/mental organ of extreme perfection and complication. Following the standard practice in biology, we can set up three distinct but interconnected levels of investigation for this uniquely human organ:

(1) a. Design
b. Development
c. Evolution

This article is based on my presentation at the *BALE 2008* conference as well as on many other occasions including the following: the *8th Annual Meeting of the Society of Evolutionary Studies*, Japan (August 2006), the *8th Annual International Conference of the Japanese Society for Language Sciences* (June 2006), and the *23rd National Conference of the English Linguistic Society of Japan* (November 2005). The research has been partially supported by the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (Challenging Exploratory Research), grant number 21652037. In preparing the manuscript, I have benefited a lot from personal communications with Naoki Fukui, Kazuko Harada, Masayuki Ike-uchi, Tom Roeper, and Juan Uriagereka. I am also deeply indebted to two anonymous reviewers for their helpful comments and in-depth criticisms on an earlier version of this article. Thanks go to Terje Lohndal, too. I claim full responsibility for all the inadequacies that remain.

biolinguistics
© The Author (2009)
ISSN 1450–3417
Biolinguistics 3.2–3: 128–153, 2009
http://www.biolinguistics.eu

In classical generative grammar, the theoretical goals of descriptive adequacy and explanatory adequacy were neatly distinguished for the studies of language design and language development, respectively, while the topic of language evolution remained somewhat afield. Current biolinguistic program goes beyond this tradition and elevates language evolution as a central issue, for which we may define a new, higher theoretical goal of 'evolutionary adequacy'.[1]

Let us say that a theory of UG is evolutionarily adequate if it explains how it was possible for the human faculty of language (HFL) to emerge during our evolutionary history. With the advent of the minimalist program (MP), it is now well understood that UG is not so much the *explanans* for the logical problem of language acquisition (LPLA), as it is itself the *explanandum* in the context of evolutionary linguistics. The proclaimed species-specificity of UG, in tandem with the observed dysfunctional and maladaptive nature of its highly modularized principles as envisioned back in the GB era of the early 1980s, kept the topic of its origin and evolution as a kind of Pandora's box, a very delicate and almost intangible issue within the framework of the Neo-Darwinian orthodoxy and the Modern Synthesis. Without any trace of its possible precursors found in the whole biological world, and without strong evidence for its reproductive fitness, the emergence of HFL seemed to be the most unlikely biological event on earth.

How, then, was HFL able to come into being at all? This is the core essence of the new problem we have to face squarely in biolinguistic investigations, which may be dubbed the 'logical problem of language evolution (LPLE)' (Fujita 2007, 2009; see also Christiansen & Chater 2008, where the same term is used for a different purpose) or 'Darwin's problem' (Fujita 2002, Boeckx 2009, Hornstein 2009).[2] The LPLE stands as the clear indication that adaptation by natural (or sexual) selection cannot be the whole explanation of the evolution of HFL, and/or that UG need not be such a complex cognitive system comprising many domain-specific and non-adaptive grammatical principles interacting intricately. In short, the LPLE tempts us to consider a drastic reorganization of UG and biological evolution in general.

In such a context, the MP offers a very promising research strategy for the biological study of language evolution (the biolinguistic minimalism). In an important sense, the MP is an attempt to reduce the internal mechanism of UG/HFL to the bare minimum by shifting the focus of inquiry to domain-general physico-mathematical principles and constraints working on the evolution and

---

1 Besides Fujita (2007, 2009), the term 'evolutionary adequacy' has already appeared in Longobardi (2004: 103), where it is proposed that explanatory adequacy and evolutionary adequacy are adequacy levels, respectively corresponding to the questions "What are the biologically possible human languages?" and "Why do we have precisely these biologically possible languages?". The usage of the term there is largely the same as in this article, though Longobardi seems to be more concerned with an explanation of historical/cultural variation of language by a parameter theory. I thank a reviewer for calling my attention to Longobardi's work.

2 The term 'Darwin's problem' may be something of a misnomer, however. This is because Darwin's major concerns were the apparently limitless biological variations that can be found in nature and their explanation by means of natural selection, whereas we are more interested in the uniformity of the biological organ in question, that is, HFL/UG, and how it may be accounted for in terms of natural/physical laws. Thus the problem may be more appropriately dubbed '(D'Arcy) Thompson's problem', for example.

development of any complex system in the natural world. To the extent that apparently language-specific properties can be derived from those 'third factors' (Chomsky 2005a), the genetically determined component of UG becomes smaller, and this has the effect of rendering the topic of language evolution more accessible. In its departure from classic genetic determinism and its emphasis on epigenetic processes under structural constraints through which immense phenotypic diversity will arise, the general idea behind the MP is in perfect harmony with the *evo–devo* paradigm in biology (see, among many others, Arthur 1997 and Hall & Olson 2003), which fact forces us to seriously reconsider topics such as modularity, autonomy, domain-specificity, evolvability, and the relation between evolution and development within the generative framework.

The idea that developmental processes, rather than genetic information *per se*, is responsible for the observed language variations was already obvious in the formative years of the Principles-and-Parameters approach, the basic tenet of which it is that a slight change in a parametric value will bring about tremendous cross-linguistic differences that are apparently limitless but only within a tightly restricted range. Adopting the *evo-devo* perspective for language evolution and language development suggests that we can proceed much further and take the universality of human language also as a phenotype, not directly encoded in the human genome as such. If this should prove to be the case, then ultimately there will be nothing left to be ascribed to UG — the final stage of the minimalist inquiries but for the time being too remote a goal even to speak of.

The term LPLE is in part intended to point to the parallel nature of the problems surrounding language evolution and language acquisition. The latter issue has been traditionally associated with the 'poverty of the stimulus' (PoS) argument, the observation that what is circumstantially available to the learner (primary linguistic data, PLD) alone is not sufficient to make language acquisition possible at all. Likewise, given the qualitative difference (insurmountable gap) between HFL and non-human primate and non-primate communication systems, language evolution seems to present a kind of 'poverty of the precursors' (PoP) argument, that is, what our common ancestors had already had (pre-existing capacities) before the formation of HFL was not sufficient to allow its emergence only in the human lineage. For the LPLA, UG has been assumed to bridge the gap between PLD and the attained steady state of FL (I-language). For the LPLE, UG is obviously of no help since it is the end product, not a pre-existing condition, of language evolution.[3]

---

[3] A reviewer suggests that something like 'poverty of selective pressures' (PoSP) would be a better evolutionary counterpart to PoS, stressing that it is selective pressures as a 'species-external' factor that are on a par with PLD as an 'individual-external' factor. While I fully appreciate the merit of the reviewer's alternative, which makes every sense especially in light of Yang's (2002) variational model of language acquisition that sees linguistic input from the environment as a selectional pressure on competing grammars in the learner's brain, my contention here is that PoP presents a qualitatively different, more general problem than PoSP: While a pluralist solution (which claims that natural selection is just one among many driving forces of evolution) is readily available for PoSP even within Darwinian theorizing, such is not the case with PoP, which calls for a serious reconsideration of an adaptationist program.

The same reviewer also questions the compatibility of the PoP argument with the exaptive scenario of language evolution adopted in this article. Let us just note that while

Human language is essentially a system for connecting meaning and sound (including the alternative externalization by means of signing and, sporadically, writing) via syntactic structure. The phonetic and semantic interpretation of a linguistic expression (sentence) is to some extent determined by its syntactic computation and the resulting hierarchical phrase structure. Structure dependency in this general sense sharply distinguishes human language from animal communication. Structure building in the syntax module is at the same time sending instructions to the external systems of sensory-motor (SM) and conceptual–intentional (C–I) capacities. These three systems constitute equally autonomous components of HFL.

That human language has this kind of modular architecture may be a good indication that language first came into being through a process known as exaptation. It is natural to assume that the evolution of these three systems predated the emergence of language, that they had, if any, separate original functions whose connection with language was thin, and that language suddenly appeared as a result of their integration within the human brain, dated about 50,000-200,000 years ago, when the pre-existing SM and C–I capacities were mediated by the so far isolated computational system of syntax.

The view that the emergence of human language was a sudden and unexpected event is often misunderstood and criticized by opponents who believe that evolution is always gradual. As a matter of fact, such an instantaneous model of language evolution, quite like the also controversial generativist instantaneous model of language acquisition, is a form of abstraction and idealization, one that purports to make otherwise too complex an issue less so for the purpose of investigation. Chomsky (2004: 395) states: "Plainly, the faculty of language was not instantaneously inserted into a mind/brain with the rest of its architecture fully intact. But we are now asking how well it is designed on that counterfactual assumption. How much does the abstraction distort a vastly more complex reality?" Not very much, perhaps.

The now famous distinction between HFL in the narrow sense (FLN) and in the broad sense (FLB), advocated by Hauser *et al.* (2002) and defended by Fitch *et al.* (2005), is of immense import when we seek to attain evolutionary adequacy by first identifying what more is necessary, other than the pre-existing capacities, for human language to come into existence. By definition, FLN is that part of HFL which is unique to the humans and human language, and FLB includes all other components which are more or less shared among other species or among other human cognitive faculties. We can assume that the core of the LPLE lies in the origins and evolution of FLN and its integration into the rest of FLB.

What constitutes FLN, then? Hauser *et al.* suggest: (i) the recursive computational operation of human syntax which gives rise to the property of discrete infinity, and (ii) the two interface systems connecting syntax to the C–I- and SM-faculties, respectively. In minimalist theorizing, the former has been recognized as the combinatorial operation of unbound Merge, the sole structure-building device of HFL. Following the criticism by Pinker & Jackendoff (2005) and

---

language as a system of distinct subcomponents does not seem to have any precursors, those subcomponents, taken in isolation, may well be exaptations of pre-existing capacities.

Jackendoff & Pinker (2005), we may tentatively add the lexical system as yet another component of FLN, for the obvious reason that in its productivity and profligacy the human lexicon is unique to the species.

(2) *Ingredients of FLN (tentative)*

a. Recursion (unbounded Merge)
b. Interfaces (C–I and SM)
c. Lexicon

"To create is to recombine," said the French biologist François Jacob, whose insight applies to the process of the evolution of language, too. HFL is an evolutionary novelty resulting from a recombination of FLN and FLB, the origins and evolution of the former being unclear for the time being.

Importantly, if the three components of FLN listed in (2) are phylogenetically unrelated capacities and their evolutionary origins have to be sought independently of each other, the LPLE has to remain as hard as before. If, on the other hand, it is shown that they can be traced back to some common precursor, or alternatively that one of them serves as the precursor to the other two, then the LPLE becomes easier to approach. Biolinguistic minimalism suggests that this kind of reductive thinking is a possibility we should not discard immediately.

As a first step toward this reduction, I will argue in the remainder of this article that (2a) may be (part of) what made (2b) and (2c) possible at all, and that (2a) may find its precursor in the uniquely human capacity for hierarchical object manipulation, the Subassembly strategy of Greenfield's (1991, 1998) grammar of action (Action Grammar). For that purpose, I will first clarify the true nature of the elementary syntactic operation Merge (section 2), next point out the striking formal resemblance between Merge and Action Grammar (section 3), and then suggest that Merge is crucially responsible for the formation of the C–I-interface and the lexicon (section 4).

As is often the case with theoretical studies of language evolution, much of the following arguments have to remain speculative. It is worth noting, however, that descriptive linguistic research sometimes sheds light on the core property of HFL and consequently on its origins and evolution. Evolutionary adequacy on one hand and explanatory and descriptive adequacy on the other are not two separate goals but are tightly interconnected and should be pursued in parallel.

2. The True Nature of Merge

In its simplest form, Merge is a set-formation operation that takes two objects and combines them into an (unordered) set.

(3) Merge $(\alpha, \beta) = \{\alpha, \beta\}$

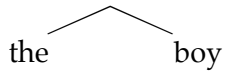
Merge is then a binary and symmetric operation. For heuristic purposes alone, it may be safely assumed that the recursive, unbounded application of Merge is the only generative device of HFL. To the extent that this assumption is tenable, we

can begin our discussion of language evolution by focusing on the possible evolutionary scenario(s) of Merge. In this context, it seems fair to ask whether Merge *per se* is the most elementary operation or whether it is a complex operation that can be further decomposed into more fundamental operations. This is because, if the latter supposition turns out true, the target of the evolutionary explanation is not Merge but rather those fundamentals, Merge being a later innovation (biological or cultural) resulting from their (re)combination.

Decomposition of Merge into smaller units is a rather natural minimalist move, and such attempts are not new in the literature, including works by Fukui (2006, to appear), Boeckx (2009), and Hornstein (2009). These authors share the view that labeling, which gives rise to endocentricity, one prominent property of human language, and which therefore seems to be an indispensable component of Merge, should be conceived of as a distinct syntactic operation detachable from the core part of Merge (Merge in the narrow sense, say core Merge). Fukui calls the labeling operation 'Embed', while Boeckx gives it the different name of 'Copy'. Hornstein also proposes to derive the effect of Merge from the combination of concatenation (our core Merge) and labeling. Rather surprisingly, however, their proposals in fact serve to confirm that no additional operation other than (core) Merge is necessary to label a set formed by a prior application of Merge, as I will now show. For expository purposes here I use Fukui's Merge+ Embed system for discussion.[4]

As a simple exemplar, suppose we build *the boy* by Merging *the* and *boy*. In accordance with Fukui's framework, we first combine the two lexical items into an unordered set by means of core Merge.

(4) a. Merge (the, boy) = {the, boy}

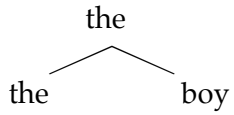
b. 

At this point, there is as yet no label for this new object, and its structure lacks endocentricity. Fukui argues that the Merge operation in (4a) has the effect of defining the Base Set (BS) in (5), to which further syntactic operations may apply.

(5) BS = {the, boy}

It is to this BS that Embed now applies, to form a labeled structure, as in (6).

(6) a. Embed (the, {the, boy}) = the ∪ {the, boy} = {the, {the, boy}}

b. 

[4] To be more precise, Fukui's approach differs from both Boeckx's and Hornstein's, in that instead of decomposing Merge into smaller operations, it adds Embed as a distinct new operation available for syntactic computation while Merge is kept intact. I thank Naoki Fukui (p.c.) and a reviewer for reminding me of this distinction.

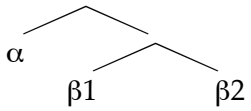
In general terms, Embed takes BS and one of its members, combines them, and forms a set union of the two. In other words, it is an operation that embeds a given object in a larger set that contains it as a member. In (6), BS = {the, boy} is embedded in the larger set {the, {the, boy}}, in which *the* acts as the label of the resulting phrase structure, now yielding endocentricity (and symmetry is broken, so to speak). Recursion is equally relevant to Merge and Embed. Both operations may or may not apply recursively, and the type of recursive structure we usually associate with natural language results from a successive application of Merge followed by Embed.

For Fukui, the distinction between Merge and Embed is crucially relevant for evolutionary studies, too, because he assumes that while Merge is probably not specific to human language, it is Embed that is truly unique to it. If so, the key to solving the LPLE may be further narrowed down to the origin of Embed and its combination with the independently existing Merge. While I fully appreciate the merit of such reasoning, my suspicion is that there is a way of making a better sense of Fukui's proposal. That is, contrary to his conclusion, I take it reasonable to think of Embed as nothing more than another sub-class of Merge (on a par with Move as Internal Merge).

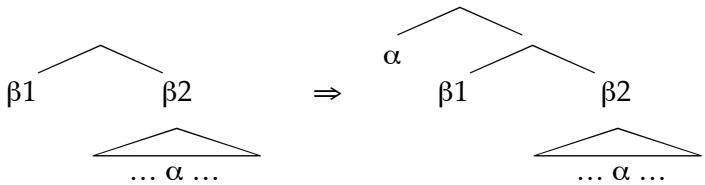
Let us begin by noting the striking formal parallelism between Merge, Move and Embed:

(7) Given $\beta = \{\beta 1, \beta 2\}$,

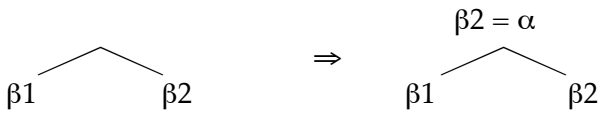
a. Merge: $(\alpha, \beta) \Rightarrow \{\alpha, \beta\}$, where α is external to β.



b. Move: $(\alpha, \beta) \Rightarrow \{\alpha, \beta\}$, where α is internal to a member of β, here $\beta 2$.



c. Embed: $(\alpha, \beta) \Rightarrow \{\alpha, \beta\}$, where α is a member of β, here $\beta 2$.



These three computational devices share the property of being a binary set-forming operation, the sole difference lying in where α is chosen from. In particular, for Move, it is to be found inside $\beta 1$ or $\beta 2$, while for Embed, it is $\beta 1$ or $\beta 2$. Thus Embed can be understood as a strictly localized version of Move, in the sense that its search domain is minimized. This way of reformulating Embed as a kind of Move (and therefore of Merge) is in harmony with the independently

suggested similarity between chains (product of Move) and projections (product of Embed) (Uriagereka 1998 and Boeckx 2008).[5]

We are therefore inclined to suspect that just as Move naturally follows from Merge, so does Embed, too. In short, Embed as well as Move comes for free once we have Merge, and consequently there is no need to speculate on their distinct evolutionary origins in addition to that of Merge.

There remains one obvious discrepancy between Merge and Embed, of course, and that is whether the operations are symmetric or not. By definition, Merge has been understood to be symmetric, given {A, B} = {B, A}, whereas labeling by Embed is obviously asymmetric, given {A, {A, B}} ≠ {B, {A, B}}. This discrepancy becomes only apparent when we take into consideration the trigger of Merge in each application and not just the resultant unordered set. In (6), for example, it is the selectional feature (or the edge feature) of *the* that gets *boy* Merged to it: In familiar terms, *boy* is the complement of the head *the* and not vice versa. We can safely assimilate this asymmetric relation to the attractor–attractee relation (Chomsky 1995) or the more recent probe-goal relation (Chomsky 2001) involved in the application of Move. And it is precisely on the basis of the choice of the attractor/probe in the preceding Merge operation that the subsequent Embed operation correctly picks out the label, the one object in which to embed the other.

To exemplify, assume we have arrived at the derivational point where the structure (8a) is formed by Merge, where *v*P contains DP (as a subject or an object):

---

[5] Incidentally, one might wonder whether there can be something like 'non-local Embed', which will have the effect of choosing the label of a phrase from inside its immediate constituents. Such an operation is formally indistinguishable from Move in the present proposal, and the resulting structure will be exocentric in nature. One relevant case that comes to my mind immediately is the oft-discussed internally headed relative clauses (IHRCs). Here is an example from Japanese:

(i) [Taro-ga ronbun-wo toukou-sita] no-ga kyakka-s-are-ta.
Taro-NOM paper-ACC submission-did NMNL.NOM rejection-do-PASS-PAST
'The paper which Taro submitted was rejected.'

In (i), the intended head noun *ronbun* 'paper' occupies the canonical object position inside the relative clause while the clause itself functions as the matrix subject DP, as if headed by *ronbun*. This phenomenon will receive a simple explanation if non-local Embed is at work here, which picks up *ronbun* as the label of the relative clause. Provided that non-local Embed is non-distinct from Move, this proposal may be taken as a simple reformulation of the movement-based analysis of IHRCs without actual constituent movement. Because of the complex nature of the potential problems such a new analysis will necessarily face, I refrain from pursuing it any further here.

Let us also note that endocentricity does not seem to be an essential property of linguistic structure in general. Morphological root compounding has been known for its exocentricity. In Japanese, for example, both *takai* 'high' and *hikui* 'low' are adjectives, but when combined together in the form of *takai-hikui*, this compound behaves as a noun.

(ii) Tatemono-no takai-hikui-ga juuyoo-da.
building-GEN high-low-NOM important-is
'The height of the buildings matters.'

(8) a. {T, *v*P} ⇒
b. {T, {T, *v*P}} ⇒
c. {DP, {T, {T, *v*P}}} ⇒
d. {T, {DP, {T, {T, *v*P}}}}

To the extent that the Merge operation in (8a) is triggered by the selectional feature on the part of T, the next step is necessarily to Embed (8a) in T, yielding (8b). Subsequent application of Move to the *v*P-internal DP, again triggered by the EPP feature of T, will form (8c), and because of this, the next step is again to Embed (8c) in T, as in (8d). Note that the order of (8b) and (8c) cannot be reversed to incorrectly build #{T, {DP, {T, *v*P}}}, because minimal search always prefers Embed to Move where applicable. The Merge–to–Embed–to–Move–to–Embed sequence depicted in (8a-d) is itself a nice illustration of Merge applying recursively, given that both Embed and Move are subtypes of Merge. In short, labeling by Embed, and the resulting endocentricity of linguistic structure, are natural consequences of recursive Merge. It can now safely be concluded that Merge needs no further decomposing, and it is the most elementary computational operation of HFL. As a consequence, the origin and evolution of Merge remains at the core of the LFLE.

3. From Action Grammar to Merge

For evolutionary studies of human language, reducing the human syntax to the most elementary operation Merge has the great advantage of making it possible to compare Merge with other human and non-human capacities that bear formal resemblance to it in search of the precursor(s) to Merge and the human syntax. This kind of comparative studies were certainly out of the question in earlier days of generative grammar when people spoke of phrase structure rules and X-bar theory to explain the hierarchical phrase structure of human language together with its endocentric nature. No one would ever dream of discovering a homologue or an analogue of X-bar schemata in the non-linguistic behaviors of non-human primates, for example. Biolinguistic minimalism has brought the viability of such a comparative method to the attention of interested researchers for the first time in the long history of generative grammar, reminding them that they can (and must) approach the issue of language evolution in the same way as evolutionary biologists explore biological evolution in general.

So what other capacities may be comparable to Merge? There have already been a lot of proposals in the literature, ranging from navigation and foraging to music and songs, gestures, manual dexterity, and social intelligence (including theory of mind (ToM), reciprocal altruism, Machiavellian intelligence etc.).[6] Unfortunately, our current (lack of) knowledge does not allow us to tell which of these proposals has more or less plausibility than others, and it appears that each of them has its own problems. Take the idea of ToM as a precursor to recursion,

[6] See Corballis (2002) for some discussion on the possible link between ToM and linguistic recursiveness.

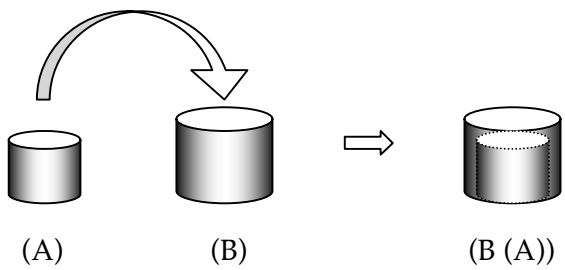
for example. Although the view that some kind of mind reading is involved in syntactic recursion may sound convincing in light of the iterated complementation as in *I know that you know that I know that…*, the correlation must remain rather illusionary, largely (i) because there are instances of syntactic recursion that have nothing to do with mind reading (*Theory A proves that Theory B proves that… Theory Z is wrong*), and (ii) because mind reading can be expressed linguistically without clausal complementation, as in *He likes her idea.*

Such being the case, instead of examining these alternatives any further I will here add just one more conceivable (in my view highly plausible) candidate, by referring to and bringing to the fore the now classic developmental studies by cognitive psychologist Patricia Greenfield.

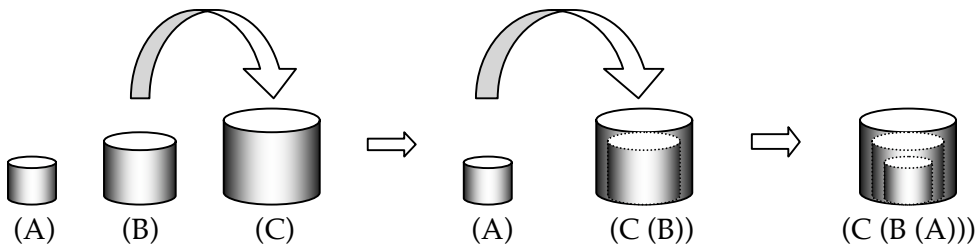
Greenfield (1991, 1998) builds on her earlier work (Greenfield *et al.* 1972) and argues that young children's developing skills in hierarchically organized object manipulation, as typically exhibited in cup nesting skills and tool use (such as using a spoon), precede their language development and serve as a pre-adaptation for it, and furthermore that a similar situation may hold true of the evolution of language in the species. In connecting the ontogeny and phylogeny of human language via their common precursor, her pioneering studies were precedent to current *evo–devo* approach to language evolution.

Greenfield observes that there are three distinct developmental stages in children's 'Action Grammar', from the simplest Pairing strategy via the Pot strategy to the most complex Subassembly strategy. It is extremely interesting to note that these strategies neatly correspond to the development of linguistic structure, in particular to the different modes of the application of Merge. In terms of nesting cups, these three combinatorial methods can be represented as follows (Greenfield *et al.* 1972 and Greenfield 1991; see also Maynard Smith & Szathmáry 1995):

Strategy 1: Pairing Method



Strategy 2: Pot Method



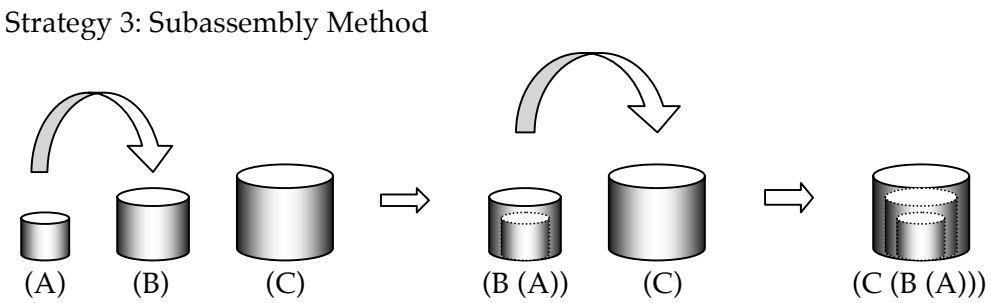


Figure 1: Three Stages of Action Grammar

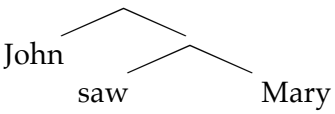
In the Pairing method, we just combine the two cups into one object by putting the smaller cup (A) into the larger one (B). In the Pot method, this same procedure applies twice (or more), combining three (or more) cups into one object, first by putting the middle-sized cup (B) into the largest (C), then by putting the smallest (A) into (C), which now contains (B), too. The third strategy is the crucial one. This Subassembly method may at first appear not very different from the second strategy, but in fact their gap is immense. In this case, we first put (A) into (B), and then we take the complex object consisting of (A) and (B) as a subunit for further operation, putting this subassembly into (C). Anticipating the comparison of Action Grammar to Merge below, it may be noted here that the Pot method requires just one constant attractor, whereas the Subassembly method has to switch attractors at each step.

Studies in comparative cognitive ethology inform us that the Subassembly strategy is almost uniquely human (the only exception seems to be chimpanzees trained linguistically in captivity), while the Pairing and Pot strategies are shared among other primates and non-primates equally. For example, Tokimoto & Okanoya (2004) demonstrate that even Degus (*Octodon degu*) have the capacity for hierarchically organizing objects by using the Pot strategy. Thus it seems natural to suspect that the uniquely human Subassembly strategy in Action Grammar plays some important role in the formation of human language, in particular of syntax.

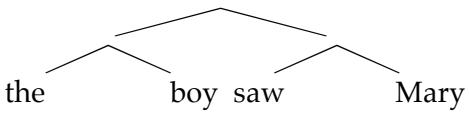
Greenfield's (1991) contention was that Action Grammar corresponds to phoneme combination in word formation, but in response to criticism by other researchers (including Tomasello 1991 and Swan 1998), Greenfield (1998) agrees to return to the initial insight of Greenfield *et al.* (1972) and admits that the proper object of combination is not the phoneme but the word. This makes it possible to project the three strategies of Action Grammar directly onto different modes of application for Merge.

For example, consider building the VP structures (9a) and (9b) (here the *v*P projection is omitted for simplicity).

(9) a. John saw Mary.



b. The boy saw Mary.



(10) a. Merge (saw, Mary) = {saw, Mary}
Merge (John, {saw, Mary}) = {John, {saw, Mary}}

b. Merge (saw, Mary) = {saw, Mary}
Merge (the, boy) = {the, boy}
Merge ({the, boy}, {saw, Mary}) = {{the, boy}, {saw, Mary}}

In both derivations, Merging (saw, Mary) into {saw, Mary} is analogous to the simple object combination by the Pairing method. Here *Mary* is attracted by *saw*, as if the smaller cup goes into the larger cup. Note that Action Grammar and Merge share the property of being symmetric in principle but asymmetric in practice. That *saw* contains the selectional feature that attracts *Mary*, and not vice versa, determines that *saw* but not *Mary* counts as 'the larger cup'.[7]

The next step crucially differentiates the two derivations. In (9a), with *saw* attracting *John*, Merge applies to (John, {saw, Mary}) to yield {John, {saw, Mary}}, as in (10a). Here *saw* remains the constant attractor, and the operation now counts as an instance of the Pot strategy. In (9b), however, something different must take place, as depicted in (10b). In order for the subject DP *the boy* to be properly attracted by *saw*, this DP must first be constructed by an independent application of Merge: *the* functions as the attractor, triggering Merge to form {the, boy}. In other words, this DP acts as a subassembly in the whole derivation. Obviously, the derivation of (9b) corresponds to the Subassembly strategy.[8]

Throughout the history of generative grammar, certain nodes or phrases have been known to block extraction from within, and have been subject to different forms of formulation under the rubrics of islands, barriers, phases, and so on. The notion of a phase is particularly interesting in this connection, as it functions as a subassembly unit in the derivational process. The Phase Impenetrability Condition (PIC, Chomsky 2001), whatever its precise definition may be, is very presumably a reflection of the fact that a derivational subunit, once completed, cannot be probed into by later operations. On the face of it, PIC is a highly language-specific principle that appears to defy a deeper explanation,

---

7 Di Sciullo & Isac (2008) argue that the asymmetry of Merge can be captured in terms of the proper inclusion relation that holds between the relevant feature bundles of the two objects undergoing Merge. Without going into the details of their analysis, we can say that their intuition is fully compatible with the present observation. Here *saw* is 'larger' or 'heavier' than *Mary* because of the selectional feature carried only by the former.

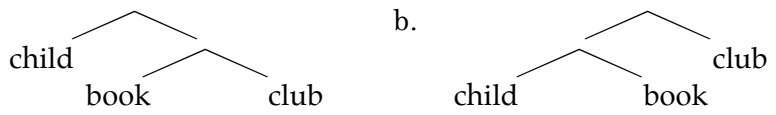
8 A reviewer objects that the analogy drawn here between Subassembly-type Action Grammar and Merge is quite arbitrary and that it would be equally valid to assume that the Pot strategy more closely reflects the full human syntax, by reversing the attractor–attractee (or Probe–Goal, in his or her words) relation. More specifically, this reviewer asks why the largest cup has to be chosen as the Probe and not as the Goal. A simple answer would be to point to the fact that in Move (Internal Merge), the Probe is (part of) the category being moved into (like the larger cup) and the Goal (part of) the category being moved (the smaller cup), rather than vice versa. To claim that the moved cup acts as the Probe in Action Grammar would require further justification beyond this simple formal correspondence.

This same reviewer also questions the validity of nesting cups as an analogue of syntax on grounds that the former always gives rise to a total inclusion/dominance relation, whereas syntactic structure exhibits such a relation only sporadically. This type of objection is based on the failure to notice that Action Grammar is observed in a large variety of object combining behaviors, not restricted to cup nesting actions. For example, consider the nut-hammer relation in nut cracking, which can hardly be assimilated to an inclusion relation.

but when seen this way, it may turn out that PIC can be given a natural place in the evolution of HFL.

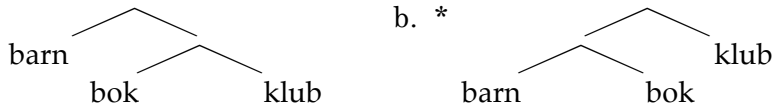
That the distinction made here between Subassembly-type Merge (henceforth, Sub-Merge) and Pot-type Merge (Pot-Merge) reflects some important aspect of syntax is supported by an important observation made by Roeper & Snyder (2005) with respect to the cross-linguistic variation in root compounding patterns. In English, a compound like *child book club* is structurally and semantically ambiguous: Both the right-branching structure (11a) and the left-branching structure (11b) are permitted.

(11) a. b.



In Swedish, however, the corresponding compound *barn bok klub* is not ambiguous and only the right-branching structure (12a) is possible.

(12) a. b. *



Roeper & Snyder (2005) offer their own elaborate account of this discrepancy between the two languages, but here let us just note that only the left-branching structure requires Sub-Merge to apply. (11a) and (12a) need only one attractor, but (11b) and (12b) need two.[9]

The fact that there is at least one language that utilizes Pot-Merge but not Sub-Merge for compounding, can be understood as an indication that the latter type is computationally more complex, probably echoing the species-specificity of the Subassembly strategy in Action Grammar: After all, it is the last strategy to emerge in child development. Importantly, that Swedish bans Sub-Merge in compounding does not mean at all that the language lacks it altogether: Otherwise, even a simple sentence like (9b) would be excluded. All we can infer from above is (i) that Sub-Merge is a universal option of the human syntax, and (ii) that each language may have a different range of its actual application. While I admit that when stated this way, the universality of Sub-Merge becomes virtually indisputable, this may be the right way to look at things first before we jump at the opposite conclusion.

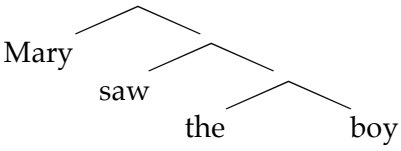
I emphasize this point because the same consideration is highly relevant in assessing the controversial Pirahã data. Everett (2005, 2007) has famously demonstrated that this Amazonian language lacks clausal complementation, relativization and other hallmarks of embedded structure, and suggested that the language is without recursion in general. The phenomena he brings to our attention are each very interesting in their own right, and I totally agree with him that cultural factors have considerable influence on the grammar of a language.

[9] If modifiers are also attracted by a head, then selectional features cannot be the only relevant attracting force.

But that may be where we should stop for the moment. I see no deep conflict between his Pirahã data and the generativist claim that recursion or embedding is an innate and universal property of human language. In any case, for something to be part of UG does not require that it be observed in every particular language, extant or extinct.

The above illustrations may have given the reader the (wrong) impression that Sub-Merge is something very special, a trick to be resorted to under very limited conditions. As a matter of fact, instances of Sub-Merge can be found in every bit of linguistic expressions and it is indeed what makes HFL worthy of the name. Consider, for example, the derivation of *Mary saw the boy*. In this case after the object DP is first built in (13b), Sub-Merge applies and attracts this whole subunit to *saw* as in (13c). (13d) is a case of Pot-Merge.

(13) a. Mary saw the boy.



b. Merge (the, boy) = {the, boy}
c. Merge (saw, {the, boy}) = {saw, {the, boy}}
d. Merge (Mary, {saw, {the, boy}}) = {Mary, {saw, {the, boy}}}

In general, every head–complement merger must take place in the form of Sub-Merge, unless of course the complement is also a zero-level lexical item. Sub-Merge thus seems to be at the core of phrasal syntax, as a default option. This last point may be relevant in searching for the reason why its application is sometimes restricted in root compounding, as in the Swedish data quoted above. For example, it can be assumed that in root compounding, to the extent that it belongs to the domain of non-phrasal syntax, Sub-Merge counts as extraneous, and that some grammars prefer to deter its application if it is only for the purpose of compound formation.

The discussion so far has amply demonstrated the formal parallelism between the elementary syntactic operation and manual object manipulation. This alone, of course, is not evidence for the evolutionary and developmental link between grammar and action, nor does it show that Action Grammar is the precursor to Merge, a possibility that needs to be explored in a multidisciplinary endeavor by researchers from every relevant field of cognitive sciences. The point I would like to make is that, by reducing the surface complexity of the human syntax to its bare minimum in the form of Merge and its recursive application, generative grammar now takes a prominent role in such an enterprise, a situation that was not easy to envision before the advent of biolinguistic minimalism.[10]

[10] To claim that Merge is evolutionarily linked to Action Grammar or any other cognitive capacity that is not species- or domain-specific does not entail that Merge is not part of FLN, contrary to what a reviewer seems to believe. Merge as the definitive syntactic operation of human language is both species- and domain-specific, but its origin may still be found in some pre-existing domain-general faculty. The alternative possibility that there is nothing in FLN remains, of course, which is well worth pursuing as an ultimate minimalist hypothesis.

Before we proceed, let us note that the crucial property of Merge, its unboundedness, has been left untouched so far. Apparently, Action Grammar applies in a bounded manner, and even if it turns out to be the precursor to Merge, how the shift from boundedness to unboundedness took place calls for an independent explanation. In this respect, Chomsky's (2007b: 23) following comment must be carefully considered. He states: "[F]or both evolution and development, there seems to be little reason to suppose that there were precursors to unbounded Merge." In his view, Merge is unbounded from the beginning. It is to everyone's knowledge that young children go through the developing stages of one word and two word utterances before they exhibit the full expressive power of unbounded Merge. This is primarily due to their limitation in language-independent cognitive and physical capacities, and does not argue against the innateness of unbounded Merge.

In evolution, however, it seems more natural to suppose a transitional process from bounded to unbounded Merge, a transition made possible by various factors including the enhancement of working memory in the enlarged brain. If Action Grammar is linked to the evolution of Merge, then it is likely to be the precursor to bounded Merge, so it seems. Note, however, that the distinction between bounded Merge and unbounded Merge is largely for theoretical purposes. In actual practice, Merge is of course bounded for familiar reasons: Life is short, and no one would ever produce a sentence that could only be generated by applying Merge 2^{100} times! But if Merge is unbounded only in theory, the same can be said of Action Grammar, too. With an infinite number of cups and infinite time (and strength) to do the nesting, the Pot and Subassembly strategies could be repeated endlessly in theory. The distance from Action Grammar to unbounded Merge may not be too remote.

4. Anti-Lexicalism and Evolutionary Adequacy

With so much discussion on Merge in mind, let us turn now to the other two components of FLN listed in (2): The two interfaces and the lexical system. Given that Merge is an indispensable ingredient of FLN, one possible research direction along the minimalist guideline is to ask to what extent these additional systems can be related to, or even reduced to, this elementary combinatorial operation. With ample abstraction, it can be taken for granted that any system that creates something must be equipped with a Merge-like device: Recall Jacob's remark that creation is recombination. In this section, I will suggest that at least some fragments of the C–I-interface can be trivialized if we take seriously the prospect of *anti-lexicalism*, which may at the same time allow us to largely dismiss the problem concerning the evolution of the human lexicon.

By anti-lexicalism, I mean the following general picture:

In fact, Fitch *et al.* (2005: 203) state: "If future empirical progress demonstrates that FLN represents an empty set, so be it."

(14) *Basic claims of anti-lexicalism*

a. Words are generated by syntax.
b. The lexicon can be decomposed into FLN (Merge) and FLB (sound and meaning).
c. Consequently, there is virtually no lexicon.[11]

There can be many different implementations of anti-lexicalist theorizing, the Distributed Morphology (DM) framework (Halle & Maranz 1993 and Marantz 1997, among others) and the studies on 'l(exical)-syntax' (Hale & Keyser 1993, 1998 and many other works) being two representatives. Other important works include Baker (2003), van Hout & Roeper (1998), and Roeper & van Hout (1999). Here I will not commit myself to any particular theory but only keep to the general idea that words and sentences are equally outputs of syntactic computation; in other words, syntax (recursive Merge) is the sole generative engine of HFL and the uniquely human generative lexicon is part of syntax.

In my view, anti-lexicalism can offer a profound account of language evolution, because obviously one cannot assume the existence of a rich lexicon from the start. In evolutionary contexts, words were not what had already been given to humans, but they had to be created through a process of synthesis and analysis, which is essentially what syntax does. This point was clear to Bronowski (1977: 120), for example, when he discussed human and animal languages and stated: "It cannot be true literally that 'In the beginning was the word': On the contrary, in the beginning was the sentence."

Take a simple case of forming a nominal from a verb, say *destruction* from *destroy*. Following the practice of DM for concreteness, once *destroy* is decomposed into the verbalizing element (v) and the category-neutral root ($\sqrt{}$), replacing v with the nominalizing element (n) in the combination $v+\sqrt{}$ will give rise to the nominal form $n+\sqrt{}$, to be realized as *destruction*. This sort of extraction and recombination must have underlain the formation of a rudimentary vocabulary from segmentation of distinct alarm calls (to be assimilated to unstructured sentences) used by some species of birds and monkeys for different types of predators. It seems reasonable not to posit the whole lexicon as a totally distinct component belonging to FLN, but rather to decompose it into the generative component and the list of surface morpho-phonological forms associated with conceptual–semantic properties, with the former falling under the proper domain of recursive Merge.

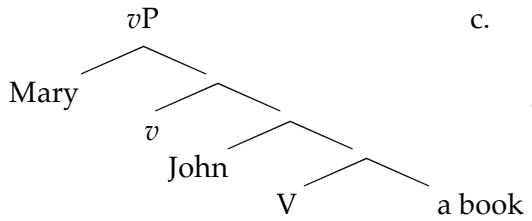
This kind of unitary approach to words and sentences by means of basic syntactic computation may find its roots in the traditional idea of lexical decomposition dating back to Generative Semantics in the early 1960s, now partially

[11] That is, as an independent module of grammar responsible for word formation. Needless to say, there has to be a universal pool of features in the human brain, different combinations of which will ultimately yield a different set of lexical items or words (sound-meaning pairings) available in particular I-languages. These are a residue of the lexicon that may safely be assumed to be part of FLB. Hauser *et al.* (2002: 1576) discuss some "key aspects of words" that may be "distinctively human" — including the astonishing "scale and mode of acquisition" by children, and the absence of "straightforward word-thing relationship." Whether these uniquely human properties must be stated as such as part of FLN or they may be better explained by other equally unique capacities, is another matter, of course.

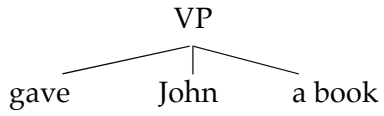
reincarnated by the split VP structure generally assumed in minimalist syntax. As a representative case, compare the two versions of VP structure proposed for double object verbs like *give* and *show*:

(15) a. Mary gave John a book.

b.



c.



(15b) is the familiar split VP format, while the flat structure in (15c) is adopted, in particular, in the Simpler Syntax framework (Culicover & Jackendoff 2005). The structure (15c) is said to be simpler than (15b) because, as Culicover & Jackendoff (2006) explicitly define it, the structural complexity is determined on the basis of sub-constituents and invisible structure. The problem is that they assume a non-derivational, representational model of phrase structure, which radically differs from our strictly derivational model, and therefore that any straightforward evaluation in terms of simplicity between (15b) and (15c) is in fact impossible. The claim here is that, seen from the derivational viewpoint, it is certainly (15b) that is simpler, since its derivation involves only binary Merge, whereas (15c) would require a more complex operation of tertiary Merge (or quaternary Merge, to derive *gave John a book in the park*, for example). The purported simplicity of (15c) is valid only on highly theory-internal grounds.

Our major concern here is, which of them is more helpful in mapping syntax to the C–I-system, that is, which makes the topic of C–I-interface 'simpler'. Recall that the split VP structure like (15b) is an embodiment of the general conception of lexical decomposition, one important interpretation of which being that the fundamental part of conceptual/semantic structure is directly encoded in syntax. Compare the putative conceptual structure (16) with (15b).

(16) [Mary CAUSE [John HAVE a book]]

The abstract causative function CAUSE corresponds to the small verb *v* in (15b), and HAVE to the large verb V. The actual word *gave* will be the morpho-phonological realization of the *v*–V amalgam (plus the past tense value) formed by syntactic Merge (head movement) and/or morphological merger. This virtually isomorphic relation between syntactic structure and (core) conceptual structure I take to be the essential foundation on which the C–I-interface is established.[12]

[12] We may want to structurally encode the distinction between Agent subject and non-agentive Causer subject, the necessity of which is shown by oft-cited examples like *The exam gave John a headache*. In terms of conceptual structure, the distinction can be made by using another function DO as an abstract agentive verb, while reserving CAUSE for the non-agentive interpretation, as in:

(i) [x DO [x CAUSE [y HAVE z]]]

To a certain degree, it can be said that syntactic structure building by recursive Merge is at the same time a parallel hierarchical conceptual structure formation by Merging semantic atoms successively (say, conceptual Merge). This proposal, by no means, is intended to suggest that syntactic structure and semantic structure are the same, as was once claimed falsely by Generative Semantics. On the contrary, full semantic interpretation requires much more information than syntactic structure provides (in particular where the compositionality principle fails to capture the vastly multifaceted and flexible syntax-semantics relations), and syntax and semantics remain two autonomous modules as before.

Seeing the C–I-interface in the suggested way allows us, however, to reconsider the origin and evolution of the connection between these two modules in a more gradual manner than is often hinted at within the generative camp. Instead of taking them as derivatives of unrelated origins, we may speculate that they come from a single root and their mutual autonomy is a matter of later cladistics, or even that either one of them is an exaptation of the other, co-opted for new functions. In short, it is advisable to take into consideration Darwin's concept of 'descent with modification' in discussing the evolutionary scenario of the C–I-interface, and in making this point clear to us, biolinguistic minimalism is successful in bringing theoretical linguistic research into the broader context of evolutionary biology.

Turning back to the comparison of (15b) and (15c), it may be noted that while the layered structure (15b) is optimized for the C–I-interface, the flat structure in (15c) is optimized for the SM-interface instead. That is, (15c) looks closer to the surface linear sequence and is able to make the linearization task trivial. In an important sense, then, this structure can be assimilated to the Surface Structure representation of classic generative grammar, and (15b) to the Deep Structure representation of core thematic relations. It seems to me that this contrast is tightly connected to the supposed primary (evolutionarily older) function of language. The flat structure is more adapted for external functions including communication, while the layered structure is fitter for internal functions (thought, planning, etc.).

Both inside and outside linguistics, researchers fall into two groups: Those who take the original function of language to be that of communication and those who take it to be that of thought. By arguing for the layered structure, I am adopting the view here that it was thought and not communication that was initially facilitated by the emergence of language: That was enough to make those who happened to obtain this capacity reproductively more successful, and communicative functions along with many others were later accommodated

---

In a series of earlier works (Fujita 1996 and references cited therein), I have proposed a three-layered VP structure roughly of the form (ii), to provide a structure-based account of various syntactic and semantic peculiarities of double object and dative object verbs, middle and ergative verbs, and psychological verbs.

(ii) [Agent V1 [Causer V2 [V3 Theme …]]]

Here the mapping between (i) and (ii) becomes more straightforward, rendering the C–I-interface even more accessible accordingly.

through exaptation. A simple thought experiment clearly shows that such must have been the actual situation: Supposing that you were the first individual in the population to obtain language, perhaps by some point mutation, how could you put this new faculty to communicative use when there was no one else who shared it with you?

In short, if something like anti-lexicalism is on the right track and actual words are generated by syntax, we can minimize the proper domain of FLN by saying (i) that both the C–I-interface and the lexicon are subserved by Merge, and therefore (ii) that only recursive, unbounded Merge constitutes the genuine part of FLN. The list in (2) may safely be updated as in (17).[13]

(17) *Ingredients of FLN (less tentative)*

a. recursion (unbounded Merge)
b. nothing else

I have said nothing so far about the status of the SM-interface from the perspective of anti-lexicalism. To the extent that linearization is a matter of deriving order from unordered hierarchical structure (for example, by mapping the derivationally determined c-command relations onto linear ordering; cf. Epstein *et al.*'s 1998 reformulation of Kayne's 1994 Linear Correspondence Axiom), the SM-interface is also crucially dependent on the recursive application of Merge. Also, derivational and inflectional morphology reflects the hierarchical relations of relevant heads to a considerable extent. Whether the same can be said about other aspects of morpho-phonological interpretation remains largely unclear, but pursuing such a possibility will be one major issue in the generative studies of language evolution. My conviction is that anti-lexicalism offers a rich avenue of research towards an evolutionarily adequate theory of HFL.

Note incidentally that the success or failure of anti-lexicalism is not directly related to the controversial issue of whether proto-language was analytic (holophrastic) or synthetic. According to the former (Wray 2000, Arbib 2003), proto-language started with sentence-like holistic units which were later analyzed and segmented into what looked more like our modern words. The latter view (e.g., Bickerton 2003, 2007, Tallerman 2007) holds that proto-language initially had only individual words which were only to be combined in a meaningful way thanks to the later emergence of syntax.

Although anti-lexicalism may at first seem to be in harmony with the holophrastic view (since it asserts that words cannot exist in the absence of syntax), it is intended to capture the richness and productivity of the lexicon of full human language which should be qualitatively different from that of proto-language. The word-like elements of proto-language (proto-words) may have existed as a primitive conceptual unit before the advent of Merge, but it certainly

[13] A tacit claim made here is that Merge *per se* is not part of FLN but only its recursive nature is. Likewise, Agree can be seen as belonging to FLB, since it is a form of general pattern recognition as can easily be found in animal cognition or even in molecular biological phenomena like immune reaction. A reviewer suggests that Agree may also be an instance of internal Merge (applying to values of features), theoretical consequences of which deserve a careful examination.

took Merge to elevate them to the full-fledged modern words with their internal composition. What anti-lexicalism suggests is that the great leap from the proto-lexicon to the full human lexicon, so to speak, could not have occurred in the absence of the capacities for synthesis and analysis afforded by syntax.

The above scenario of lexical evolution may receive support from the observation on the parallel development of lexico-syntactic knowledge in young children. Tomasello (1992) famously proposed the Verb Island Hypothesis (VIH), according to which during the first two years of life children use verbs in an item-based manner, without any general knowledge of argument structure or categorization. Where in adult grammar two verbs belong to the same subclass and behave syntactically in the same way, for example as transitive verbs that take Agent and Theme, children treat them as two distinct, unrelated entities, each with its own organization and appearing in different frames. Abstract generalizations concerning categories, schemas and thematic roles emerge in children at later stages of development only gradually, according to Tomasello.

The VIH has subsequently been critically reexamined by other researchers and effectively rebutted, for example, by Ninio (2006), who concludes that "no verb is an island" (p. 59). Ninio instead proposes a lexicalist analysis of children's early syntactic knowledge. According to this latter view, at the earliest stage item-specific lexical rules regulating the syntactic combination patterns of each verb are sufficient to allow for the later development of full syntactic knowledge.

Both Tomasello's and Ninio's empiricist positions seem to be incompatible with the observation that young children begin to learn verbs by fully utilizing lexical decomposition very early. Viau (2006), for example, points out on the basis of data from CHIDLES a tight correlation between their acquisition of double object and dative object verbs on one hand, and that of the atomic elements like HAVE, GO, and CAUSE that these verbs can be decomposed into.

(18) a. DP1 gives DP2 DP3
b. [DP1 CAUSE [DP2 HAVE DP3]]
c. DP1 gives DP3 to DP2
d. [DP1 CAUSE [DP3 GO-TO DP2]]

Omitting the details, the mean ages of their acquisition can be shown sequentially as in (19), adapted from Viau (2006):

(19) CAUSE (2;0.4) ≥ HAVE (2;0.7) ≥ Double Obj Verbs (2;1.6) >
GO (2;4.0) ≥ Dative Obj verbs (2;4.9)

It is obvious from this result that children acquire double object verbs immediately after they acquire CAUSE and HAVE, while the acquisition of dative object verbs occurs significantly later, only after the acquisition of GO. My interpretation of this interesting fact is as follows: (i) By the time children reach the stage of two word utterances, they have fully activated recursive Merge, and (ii) this same capacity enables them to construct new verbs (including the well-documented overgeneration such as *Daddy giggle me*) by Merging basic conceptual units they already have. Tomasello's VIH, if it is correct at all, is appli-

cable only to the earliest stage of development before Merge comes into force, and the insular behaviors of the limited verbs at that stage, without systematic organization, is analogous to the supposed property of proto-verbs (primitive verb-like elements of proto-language).

Although we know that Ernst Haeckel's theory literally does not hold biologically and that ontogeny does NOT recapitulate phylogeny, I think one practical application of the *evo–devo* approach to language evolution is to adopt the working hypothesis that language development in the individual proceeds more or less analogously to language evolution in the species. Anti-lexicalism proposes to treat the full human lexicon, but not the primitive proto-lexicon, in terms of syntactic computation by recursive Merge. Where those island-like atomic units, to which Merge applies to form words, first came from is another issue, and they very likely belong to the domain of FLB.

On this latter issue, Emonds (2004) has already pointed out the apparent paradox that syntax is uniquely human but some of those features that drive syntactic derivation, notably ϕ-features, are not uniquely human. The paradox is lost once we notice that language evolution, just like all other instances of biological evolution, is a process of recruiting some old traits for a new combination. The lexicon, just like HFL itself, is a product of such reorganization. Although not necessarily the only possibility, anti-lexicalism seems to be one promising framework in which we can seek a theory of HFL that may attain a certain level of evolutionary adequacy.

Notice finally that to the extent that simple words are syntactically complex objects, it follows that Sub-Merge (Subassembly-type Merge) is always involved even in the derivation of two word utterances. This is so since to Merge *milk* and *cup* to form *milk cup*, for example, each of the two nouns must first be formed by Merge. In addition to corroborating the central role of Sub-Merge in the evolution and development of human language, this fact forces us to take a new look at the Swedish compound data (12), because it must be the case that (12a) is also formed by Sub-Merge. One can imagine quick solutions, such as allowing for (literally) root compounding of categorially unspecified roots, as in {MILK, CUP}, before this new object is specified as N and realized as *milk cup*. In the absence of any direct empirical evidence for (or against) such a move (but see fn. 5), I will not discuss this possibility any further. Compounding continues to be a highly important phenomenon in understanding the nature, development and evolution of Merge (see Roeper 2007 for a close examination of Merge and compounding as they develop in children's minds).

5. Final Remarks

As theoretical linguists, we certainly understand more about the internal machinery of human language than researchers in other fields, and it is in this respect that we make our own contribution to the multidisciplinary study of language evolution. In fact, it makes almost no sense to try to approach the topic without first establishing a working model of HFL, and in this sense generative biolinguistics constitutes the most productive, if not ultimately correct, frame-

work for research.

Against this theoretical linguistic approach, it is sometimes objected that a full understanding of language evolution must take into consideration various 'external' or environmental factors affecting the way language evolved in one way or another, most notably the human 'life history'. For instance, humans have prolonged infancy, and one may want to suggest that language evolved in order to firmly maintain the mother-child bonding as a means of communication.

This explanation has every kind of weakness a typical adaptationist scenario does. First and foremost, communicative utility, whether it is for social bonds, courtship or competition, is not enhanced by the formal device of unbounded Merge, and therefore its evolutionary explanation must be sought elsewhere. Second, in order to be usable at all as a communicative tool between mother and child, HFL must have already existed in their brains. Natural selection only chooses among existing variations, some of which happen to be fitter than others, and whatever is to be chosen must be present before selectional pressures work. It is the 'arrival' of the fittest, not its survival, that truly matters in biological evolution. Specifying the benefit and advantage conferred on us by the existence of language is no explanation of its initial arrival.

Generative biolinguistics, by contrast, succeeds in asking the right kind of questions, if not in answering them, by decomposing language into distinct components or modules. It explores the nature, origins and evolution of each of these components before they were interconnected to form what later became HFL, without confusing their respective original adaptive functions (not linguistic at all) with their current utilities in the organization of language.

This point should be kept in mind when we discuss the 'communicative function' of Internal Merge (Move), too. It is sometimes suggested that the dual function of language as a tool for thought and communication is served by External Merge and Internal Merge, respectively (Chomsky 2005a). External Merge establishes θ-relations and argument structure, whereas Internal Merge defines discourse-related information structure. While these descriptions may correctly characterize the functional motivations for applying these operations, they are not to be understood as explaining their origins and evolution, for the obvious reason that the capacity for them had to exist in the human brain before they were co-opted in the suggested dual manner.

Last but not least, thinking about the origins and evolution of language, especially from the viewpoint of exaptation and recombination of preexisting faculties, leads to a serious reconsideration of the proclaimed species-specificity and domain-specificity of language. Taken in isolation, each of the capacities that jointly constitute our language faculty, including Merge, is not strictly specific either way at the levels of genes and neural substrates, for evolution and development alike. The pursuit of evolutionary adequacy invites us to seek an integration of generative grammar into the multidisciplinarity of, say, evolutionary developmental linguistics (*evo–devo* linguistics; see Locke 2009 for one such proposal). The future of generative biolinguistics largely depends on the success of such a unification.

References

Arbib, Michael A. 2003. The evolving mirror system: A neural basis for language readiness. In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution*, 182–200. Oxford: Oxford University Press.

Arthur, Wallace. 1997. *The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology*. Cambridge: Cambridge University Press.

Baker, Mark C. 2003. *Lexical Categories: Verbs, Nouns, and Adjectives.* Cambridge: Cambridge University Press.

Bickerton, Derek. 2003. Symbol and structure: A comprehensive framework for language evolution. In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution*, 77–93. Oxford: Oxford University Press.

Bickerton, Derek. 2007. Language evolution: A brief guide for linguists. *Lingua* 117, 510–526.

Boeckx, Cedric. 2008. *Bare Syntax*. Oxford: Oxford University Press.

Boeckx, Cedric. 2009. The nature of merge: Consequences for language, mind, and biology. In Massimo Piattelli–Palmarini, Juan Uriagereka & Pello Salaburu (eds.), *Of Minds & Language: A Dialogue with Noam Chomsky in the Basque Country*, 44–57. Oxford: Oxford University Press.

Bronowski, Jacob. 1977. *A Sense of the Future: Essays in Natural Philosophy*. Cambridge, MA: MIT Press.

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

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language,* 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Language and mind: Current thoughts on ancient problems. In Lyle Jenkins (ed.), *Variation and Universals in Biolinguistics*, 379–405. Amsterdam: Elsevier.

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

Chomsky, Noam. 2005b. Some simple evo–devo theses: How true might they be for language? Paper presented at the *Alice V. & David H. Morris Symposium on Language and Communication: The Evolution of Language*, Stony Brook, NY. [Stony Brook University, 14 October 2005.]

Chomsky, Noam. 2007a. Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies* 15, 1–21.

Chomsky, Noam. 2007b. Of minds and language. *Biolinguistics* 1, 9–27.

Chomsky, Noam. 2007c. Approaching UG from below. In Uli Sauerland & Hans–Martin Gärtner (eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax–Semantics*, 1–29. Berlin: Mouton de Gruyter.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean–Roger Vergnaud,* 133–166. Cambridge, MA: MIT Press.

Christiansen, Morten H. & Nick Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31, 489–558.

Corballis, Michael C. 2002. *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford Uni-

versity Press.
Culicover, Peter W. & Ray Jackendoff. 2006. The simpler syntax hypothesis. *Trends in Cognitive Sciences* 10, 413–418.
Di Sciullo, Anna Maria & Daniela Isac. 2008. The asymmetry of Merge. *Biolinguistics* 2, 260–290.
Emonds, Joseph. 2004. What humans have that animals don't have. In Takao Gunji & Taisuke Nishigauchi (eds.), *Kotoba-no Kagaku Handobukku* [A Handbook of Linguistic Science], 239–250. Tokyo: Kenkyusha.
Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations.* New York: Oxford University Press.
Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46, 621–646.
Everett, Dan. 2007. Cultural constraints on grammar in Pirahã: A reply to Nevins, Pesetsky, and Rodrigues. [http://ling.auf.net/lingBuzz/000427.]
Fitch, W. Tecumseh. 2005. The evolution of language: A comparative review. *Biology and Philosophy* 20, 193–230.
Fitch, W. Tecumseh, Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: Clarifications and implications. *Cognition* 97, 179–210.
Fujita, Koji. 1996. Double objects, causatives, and derivational economy. *Linguistic Inquiry* 27, 146–173.
Fujita, Koji. 2002. Review of *Biolinguistics: Exploring the Biology of Language* by Lyle Jenkins. *Gengo Kenkyu* 121, 165–178.
Fujita, Koji. 2007. Facing the logical problem of language evolution. *English Linguistics* 24, 78–108.
Fujita, Koji. 2009. Dimensions of language evolution: A view from generative biolinguistics. *Studies in Language Sciences* 8, 45–58.
Fukui, Naoki. 2006. Merge and/or Embed. Ms., Sophia University/Harvard University.
Fukui, Naoki. To appear. Merge and bare phrase structure. In Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press.
Greenfield, Patricia M., Karen Nelson & Elliot Saltzman. 1972. The development of rulebound strategies for manipulating seriated cups: A parallel between action and grammar. *Cognitive Psychology* 3, 291–310.
Greenfield, Patricia M. 1991. Language, tools, and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14, 531–595.
Greenfield, Patricia M. 1998. Language, tools, and brain revisited. *Behavioral and Brain Sciences* 21, 159–163.
Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expressions of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 53–109. Cambridge, MA: MIT Press.
Hale, Ken & Jay Keyser. 1998. The basic elements of argument structure. *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect — MIT Working Papers in Linguistics* 32, 73–118.

Hall, Brian K. & Wendy M. Olson (eds.). 2003. *Keywords and Concepts in Evolutionary Developmental Biology*. Cambridge, MA: Harvard University Press.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT 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. 2009. *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge: Cambridge University Press.

van Hout, Angeliek & Thomas Roeper. 1998. Events and aspectual structure in derivational morphology. *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect — MIT Working Papers in Linguistics* 32, 175–200.

Jackendoff, Ray & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language. *Cognition* 97, 211–225.

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

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

Locke, John L. 2009. Evolutionary developmental linguistics: Naturalization of the faculty of language. *Language Sciences* 31, 33–59.

Longobardi, Giuseppe. 2004. Methods in parametric linguistics and cognitive history. In Pierre Pica with Johan Rooryck (eds.) *Linguistic Variation Yearbook* 3 (2003), 101–138. Amsterdam: John Benjamins.

Lorenzo, Guillermo & Víctor M. Longa. 2003. Minimizing the genes for grammar: The Minimalist Program as a biological framework for the study of language. *Lingua* 113, 643–657.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *UPenn Working Papers in Linguistics* 4(2), 201–225.

Maynard Smith, John & Eörs Szathmáry. 1995. *The Major Transitions in Evolution*. Oxford: Oxford University Press.

Nevins, Andrew, David Pesetsky & Cilene Rodrigues. 2007. Pirahã exceptionality: A reassessment. [http://ling.auf.net/lingBuzz/000411.]

Ninio, Anat. 2006. *Language and the Learning Curve: A New Theory of Syntactic Development*. Oxford: Oxford University Press.

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

Roeper, Tom. 2007. *The Prism of Grammar: How Child Language Illuminates Humanism*. Cambridge, MA: MIT Press.

Roeper, Thomas & Angeliek van Hout. 1999. The impact of nominalization on passive, *–able* and middle: Burzio's generalization and feature-movement in the lexicon. *Papers from the UPenn/MIT Roundtable on the Lexicon — MIT Working Papers in Linguistics* 35, 185–211.

Roeper, Thomas & William Snyder. 2005. Language learnability and the forms of recursion. In Anna Maria Di Sciullo (ed.), *UG and External Systems: Language, Brain and Computation*, 155–169. Amsterdam: John Benjamins.

Swann, Philip. 1998. Greenfield on language, tools, and brain. *Behavioral and Brain Sciences* 21, 155–159.

Tallerman, Maggie. 2007. Did our ancestors speak a holistic protolanguage? *Lingua* 117, 579–604.

Tokimoto, Naoko & Kazuo Okanoya. 2004. Spontaneous construction of 'Chinese boxes' by Degus (*Octodon degu*): A rudiment of recursive intelligence? *Japanese Psychological Research* 46, 255–261.

Tomasello, Michael. 1991. Objects are analogous to words, not phonemes or grammatical categories. *Behavioral and Brain Sciences* 14, 575–576.

Tomasello, Michael. 1992. *First Verbs: A Case Study of Early Grammatical Development.* Cambridge: Cambridge University Press.

Uriagereka, Juan. 1998. *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, MA: MIT Press.

Uriagereka, Juan. 2008. *Syntactic Anchors: On Semantic Structuring.* Cambridge: Cambridge University Press.

Viau, Joshua. 2006. *Give* = CAUSE + HAVE/GO: Evidence for early semantic decomposition of dative verbs in English child corpora. *Boston University Conference on Language Development* (*BUCLD*) 30(2), 665–676.

Wray, Alison. 2000. Holistic utterances in protolanguage: The link from primates to humans. In Chris Knight, Michael Studdert–Kennedy & James R. Hurford (eds.), The *Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, 285–302. Cambridge: Cambridge University Press.

Yang, Charles D. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Koji Fujita
Kyoto University
Graduate School of Human and Environmental Studies
Sakyo-ku, Yoshida
Kyoto 606-8501
Japan
K.Fujita@fx4.ecs.kyoto-u.ac.jp