Introduction

The Minimalist Program and the concept of Universal Grammar

Jordi Fortuny Andreu

University of Groningen. Center for Language and Cognition Groningen
J.Fortuny.Andreu@rug.nl

Ángel J. Gallego

Centre de Lingüística Teòrica
Universitat Autònoma de Barcelona. Departament de Filologia Espanyola
Angel.Gallego@uab.cat

This volume offers a collection of brief and orginal articles that aim at developing the minimalist approach to language. In this introduction, we would like to discuss certain issues concerning the Minimalist Program and the concept of Universal Grammar as well as to present the contributions of the volume.

The term *minimalism* is typically associated to certain artistic tendencies whose work is said to be stripped down to its most fundamental features. In the linguistic agenda developed by Noam Chomsky and others since the early nineties, this term is endowed with the same aesthetic concern of searching for the most fundamental features of grammatical theory, but it crucially adds a substantive commitment to investigate to what extent language is—in a sense we clarify below—a perfect system.

In Martin & Uriagereka (2000) this bifurcated path to pursue the Minimalist Program (MP) is emphasized by establishing a distinction between *methodological* minimalism and *substantive* minimalism. The methodological facet is concerned with the "theory of language" under construction, whereas the substantive facet is concerned with how well designed "language itself" is. Methodological minimalism seeks those components of a linguistic theory that are redundant, stipulative, and idiosyncratic, in order to eliminate or reformulate them in the form of parsimonius, well-grounded, and general principles. Substantive minimalism explores whether an alleged property of language is a genuine property or an apparent property, and if genuine, whether it satisfies the *Strongest Minimalist Thesis* —SMT (see Chomsky 2000)—, i.e., whether it is an optimal solution to the requirements imposed by the external systems.

There are different case studies of methodological minimalism that have contributed to determine what the fundamental elements of the theory of Universal Grammar are. The most representative concerns are the reduction of the levels of

representation (D-Structure, S-Structure, and Logical Form; see Chomsky 1993; 2000), the elimination of Agr(eement) nodes (see Chomsky 1995), and the simplification of phrase structure and transformational theory.[1] Typically, almost any grammatical observation offers a case study for substantive miminalism. Just to mention a few questions for the SMT: What would the legibility conditions be for the existence of parasitic gaps in a language like English? Why is it that certain languages have grammatical functions? Would it not be enough to have thematic roles and illocutionary force and informational relations? In virtue of what legibility conditions different languages have different basic word orders or different Case patterns? Can cartographic effects be derived from semantic factors? Why are natural languages ambiguous? The list could go on and on.

We would like to emphasize that methodological and substantive minimalism are not to be confused. For instance, it is conceivable to construct a theory of how linguistic properties derive from external principles that postulates a multitude of idiosyncratic principles, instead of a relatively small set of general and fundamental principles. Although this would be an impressive achievement, it would not satisfy the usual scientific ideals of simplicity and solidness. However, the methodological and substantive facets may not be independent: quite plausibly, it is by struggling for an elegant and parsimonious theory of an object that one can trully understand to what extent such an object is an optimal solution to the external conditions it must face. We also stress the truism that these two facets are not exclusive of the MP or the generative tradition, but they are characteristic attitudes of any rational inquiry that is not merely descriptive or cumulative but searches for a 'principled account'—the substantive facet—that satisfies the usual 'simplicity desiderata'—the methodological facet—.

Clearly, the MP leads minimalist researchers to a view of Universal Grammar (UG) that is essentially different from the view in Principles and Parameters (P&P), from which the MP orginated. We can express the sharp contrast between the conceptions of UG in the MP and the P&P by reflecting about their opposite views on the logical problem of language growth and linguistic variation.

As to the logical problem of language growth (how can a child develop a natural language just by being exposed to some linguistic data for a relatively short period?), the P&P model would claim that UG is a very highly specified structure of exclusively linguistic principles with their respective parameters. Language growth would consist in setting a value for the parameters given by UG. The MP would investigate to what extent the principles belonging to UG, a virtually necessary cognitive component of humans, would belong to general principles, or in terms of Chomsky's (2005) factors, to what extent Factor I is a rewiring of Factor III; the process of language growth would consist in the interaction of linguistic principles (which should, ideally, be derived from general principles) and principles of

---

1. See Chomsky & Lasnik (1995) for a brief historical overeview of the simplification of phrase structure and Chomsky (1995) for the introduction of bare phrase structure. See Lasnik & Saito (1992) for an important simplification of the transformational component in terms of the single operation 'move α' which allowed much further work.

analysis (see Yang 2002; 2004, Gambell & Yang 2003) that apply to linguistic data (Factor II).

As to linguistic variation, the minimalist reasoning would invite us to investigate to what extent linguistic variation can be derived from the Articulatory-Perceptual (A-P) system (which includes phonetics, morphology and phonology), in opposition to the P&P view, according to which variation between two languages would reflect different parameter settings in the process of language growth. According to the minimalist working hypothesis, linguistic variation would be no more than the effect of the relative freedom allowed by the A-P system, and the important question would be to determine the principles governing variation and their relationship to the principles of the A-P system.

This new perspective is better captured in the following quote, which summarizes the reductionist—or, to use Chomsky's words, 'from below'—twist of minimalism:

> Throughout the modern history of generative grammar, the problem of determining the character of FL [Faculty of Language] has been approached "from top down": How much must be attributed to UG to account for language acquisition? The MP seeks to approach the problem "from bottom up": How little can be attributed to UG while still accounting for the variety of I-languages attained, relying on third factor principles? The two approaches should, of course, converge, and should interact in the course of pursuing a common goal. [from Chomsky 2007:4]

What Chomsky suggests here is that most devices and properties that were attributed to UG in the GB era—Factor I—should be recast as economy and efficient computation principles—Factor III—. In a nutshell: "UG is what remains when the gap has been reduced to the minimum, when all third factor effects have been identified. UG [only] consists of the mechanisms specific to [the faculty of language]" (Chomsky 2007:5). The goal of the present volume is precisely to help reduce the gap by considering what properties can truly be adscribed to UG, and what properties cannot. If the SMT is seriously entertained, then the faculty of language boils down to a computational system and the interface systems, and the crucial question is to determine what features of the computational system and the interfaces are primitive elements of language belonging to Factor I or derivative from general principles belonging to Factor III.

Before introducing the contents of the papers of this volume, we shall sketch Chomsky's (2007; 2008) most recent essentialist view of UG. Chomsky's very restrictive conception of the computational system focuses on *Merge*, "the most elementary property of language [...] [giving rise to] a system of discrete infinity consisting of hierarchically organized objects" (Chomsky 2008:137). Much literature has been dedicated to investigate the properties of this operation, and discuss its triggers, its consequences, and its subcomponents (see Boeckx 2008, Hornstein 2009, and references therein). Although there have been proposals to motivate Merge, attributing some Agree-like nature to it, the defining property of this operation concerns the hypothesis that it is 'unbounded' and feature-free, an idea that has been expressed by means of a special property of lexical items: their *edge fea-*

ture (EF).[2] As Chomsky puts it: "The fact that Merge operates without limit is a property of LIs [...] EF articulates the fact that Merge is unbounded, that language is a recursive infinite system." Chomsky (2008:139). Merge therefore operates by combining two syntactic objects, α and β—either simple or complex—. If α and β are independent—not having participated in previous applications of Merge, or being directly taken from the lexicon—, Merge is external; if they are not, Merge is internal. Or in different terms,

(1)   a. External Merge: Merge (α, β) when α is an outcome of a previous application of Merge (or selected from the lexicon) and β is selected from the lexicon

    b. Internal Merge: Merge (α, β) when α is an outcome of a previous application of Merge and β is selected from the domain of α[3]

A further characteristic attributed to Merge in some recent literature (see Boeckx 2008; 2009, Chomsky 2005; 2007; 2008) is its symmetric nature. Typically, this is represented by using sets, which do not—by definition—impose any asymmetry to their members. The only exception appears to come from adjuncts, which have been approached by invoking a pair-creating version of Merge—pair Merge—, which establishes an asymmetry between Merge-mates. The reason for this double treatment appears to have, in Chomsky's (2004) account, an empirical motivation: adjuncts do not have an impact on the syntactic computation—they do not participate in theta-role configurations, do not receive Case, are potentially unbounded, etc.—, but they make an interpretive contribution. It is, according to Chomsky (2004), as if adjuncts were not there, apart from semantic interpretation—where the 'were not there' part has been taken to suggest that adjuncts occupy a parallel plane or are late-merged into the derivation (see Chametzky 2000, Lebeaux 1991). In Chomsky's (2004) treatment, the parallel plane metaphor has been implemented by invoking (ordered) pairs for adjuncts, because only pairs establish asymmetries between Merge-mates. This can be seen in (2).

(2)   a. Set Merge: $\{\alpha, \beta\} = \{\beta, \alpha\}$

    b. Pair Merge: $\langle\alpha, \beta\rangle \neq \langle\beta, \alpha\rangle$

Along with edge features that trigger Merge, it is a fact—against the SMT, at first glance—that UG has uninterpretable features, the so-called φ-features. If UG has φ-features, then it will optimally activate some valuation procedure—which

---

2.   Chomsky's (2007; 2008) EF must be seen as a Merge-inducing property, not a *bona fide* feature. There are at least three traits that distinguish EFs from other features: they do not involve feature-Match, they have no values, and they do not delete. For these reasons, EFs should not be treated like other features.

3.   See Chametzky (2000), Boeckx (2008), Fortuny & Corominas-Murtra (this volume), and references therein for an accurate definition of the distinction between internal and external Merge.

is referred to as Agree—, and a deletion mechanism. Chomsky (2007; 2008) argues that deletion is part of a transfer operation, which is responsible for reducing computational load by 'periodically forgetting' chunks of the derivation. On this view, φ-features are not an imperfection, but a way to signal what the transfer points are. Chomsky calls the relevant transfer units phases, and claims that φ-features are located in the functional heads C and v* (see Gallego 2007, Fortuny 2008, and Richards 2004 for discussion).

(3)   Transfer points are signaled by φ-features.

Once transfer has applied, the syntactic objects generated by Merge are cashed out to the external systems—through the SEM and PHON interfaces—, which are supposed to incorporate Factor III conditions that UG must meet. The following conditions have been explored in the minimalist literature:

(4)   *Factor III Conditions*
     a. Inclusiveness Condition
     b. Extension Condition / No Tampering Condition
     c. Full Interpretation
     d. Phase Impenetrability Condition
     e. Minimal Link Condition

First of all, it is worth noting that it is at the interfaces that asymmetries emerge. So, whereas Merge treats α and β as equal, in a symmetric fashion, the interfaces seem to reinterpret the syntactic outputs, providing asymmetry. This is so with different interface phenomena: if α binds β, then β does not bind α; if α is an argument of β, then β is not an argument of α; if α precedes β, then β cannot precede α; etc. A further relevant issue about the interfaces and their conditions can be formulated as follows: are these principles 'substantive' or 'methodological'? In other words: are the conditions in (4) genuine conditions of UG that can be derived from Factor III or mere aesthetic devices? Putting aside Full Interpretation, which genuinely appears to fall within Factor III, it is rather plausible that (4a) and (4b) are just methodological, as they restrict the number of formatives and operations that one can invoke in order to approach the faculty of language. (4d) and (4e) are, on the other hand, not obviously just methodological: they are not especially concerned with the minimal elements or conditions that must be assumed in the theory, but they are principles that attempt to account for the locality of movement, a property that has always been related to economy conditions (see Chomsky 1973; 1986; 1995 for different illustrations). A different and important issue is to what extent these two principles attributed to Factor I can be derived from Factor III, thereby showing that they (i.e., the principles that constraint/regulate locality effects) are an optimal solution to some external requirement, in which case (4d) and (4e) offer a case study for substantive minimalism.

On the basis of the above given essentialist considerations about Factor I, one may be tempted to go a step further and claim that the content of UG is solely

Merge and those features that trigger its applications (namely, EFs and φ-features)—all other linguistic properties being a byproduct of the external systems. There may be, though, several qualms for such a claim. Note, for instance, that it would confuse the syntactic component and UG (Factor I), i.e., the set of innate elements of grammar which are necessary to interpret some given external data as linguistic experience and constrain the set of possible growing paths. For these reasons, one could envisage the possibility that UG be composed of several components other than Merge, such as the knowledge of phonological structure (which seems necessary for statistical learning to be successful in a realistic setting; see Yang 2002, 2004, Gambell & Yang 2003) or, quite likely, the coarticulation mechanisms that transform a sequence of discrete phonological units into a continuous acoustic sign (see Harcastle & Hewlett 2000 for a broad discussion and illustration), alongside those mechanisms that reconstruct a sequence of discrete phonological units from a received continuous acoustic sign, and the interpretative mechanisms that allow humans to refer to displaced events (see Hockett 1960), i.e., to events taking place at a time different from the present one and in a world different from the actual one. All this, and quite likely much more, may be necessary if one wants to characterize what allows humans to develop the knowledge of grammar and how.[4]

After this brief sketch of the current minimalist approach to language and some critical remarks about the concept of UG in the MP and the P&P, we shall finally present the contributions of this volume to the development of the MP. In the first paper, **Abels and Bentzen** discuss the nature of successive cyclic movement. These authors defend the hypothesis that movement has a punctuated (i.e., non-uniform) nature, only targeting the so-called phase edges: SPEC-v* and SPEC-C. Apart from data from wh-movement effects, Abels and Bentzen investigate the interaction of quantifier phrases and adverbs in Norwegian and ellipsis licensing, providing new evidence in favor of their claim.

Boeckx argues for a system that sticks to an architecture in the spirit of what we have discussed so far. According to this author, the asymmetries that prominently appear in the language faculty can be seen as a consequence of the phase-based transfer that Chomsky (2000 and subsequent work) advocates for. Assuming that Merge is inherently symmetric, the distinctions that were established by invoking the notion head/label are recast by a label-free algorithm that takes the second application of Merge to be the key factor to establish Probe-Goal dependencies, cyclic transfer, the internal-external argument cut, and the nature of phenomena such as ellipsis, pied-piping, or islands.

Richards' paper addresses the possibility that, just like there is external and internal Set Merge, there is an analogous duality affecting Pair Merge. Richards provides interesting pieces of evidence that this missing mode of movement, which

---

4.   Plausibly, this would correspond to a broad interpretation of UG (in the sense of Hauser et al. 2002).

he calls internal Pair Merge, may well be behind a number of operations (e.g., focalization, scrambling, successive cyclicity) that are usually addressed by postulating dedicated features that are not of the agreement type—criterial or EPP features—. In brief, internal Pair Merge is the variety of movement that is featureless—or pure-EPP driven—according to Richards.

Vicente's contribution focuses on A-bar chains—paying special attention to its quantificational nature—, discussing how the Copy Theory of Movement and Multidominance approaches deal with them. As Vicente observes, the distinctions are not particularly notorious empirically, as there is no evidence favoring one of the approaches over the other. What this investigation does show is that the device one adopts does have an effect on the mapping between the computational component and the external systems, with specific consequences for domains like linearization and the composition of logical forms.

Fortuny and Corominas-Murtra propose to construct the core of the transformational generative syntactic theory of language on the basis of the set-theoretical concept of 'nest', on which Kuratowski's (1921) general theory of order is based. Accordingly, one does not need to appeal to idiosyncratic linguistic constructs (such as X'-theory) to represent hierarchical properties of linguistic expressions, but elementary set-theoretical notions are sufficient.

Hornstein and Pietroski articulate a minimalist conception of syntactic and semantic composition by identifying fundamental operations of language. They take as a starting point the operation COMBINE(A, B), which is claimed to be responsible for combining expressions in an I-language. COMBINE is decomposed into two simpler operations, CONCATENATE and LABEL. Each application of COMBINE(A, B) is argued to be a semantic instruction in such a way that the meaning of a complex structure is the meaning of a labelled concatenation.

Irurtzun is concerned with the general design of grammar, and more precisely, with the comparison of two different views of how syntax is related to the external systems: the classic inverted-Y model, according to which syntax is responsible for creating structures interpreted by the Articulatory-Perceptual (A-P) system and the Conceptual-Intentional (C-I) system with no further connection between A-P and C-I, and the more articulated alternative proposed by Jackendoff (1997), which postulates a non-syntactic stretch A-P → C-I. On the basis of the particular case study of the grammar of focus, Irurtzun argues for the superiority of the simpler inverted-Y model.

Zwart explores a top-down model of syntactic computations that yields, at each step of the computation, an ordered pair of elements relevant to the expression of order, information and grammatical features. A crucial feature of Zwart's proposals is that the output of a previous derivation appears as an atom in the resource for the next derivation. Zwart suggests that the existence of layered derivations, and not conditions on movement, would be the source of opacity effects.

References

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

Boeckx, Cedric (2009). "Aspects of a Theory of Phases". Ms., ICREA / CLT-UAB.

Chametzky, Robert (2000). *Phrase structure. From GB to minimalism*. Oxford: Blackwell.

Chomsky, Noam (1973). "Conditions on Transformations". In *A festschrift for Morris Halle*, S. Anderson and P. Kiparsky (eds.), 232-286. New York: Holt, Renehart and Winston.

Chomsky, Noam (1986). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1993). "A minimalist program for linguistic theory". In *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, K. Hale and S. J. Keyser (eds), 1-52. Cambridge, Mass.: The MIT Press. [Reprinted in Chomsky, N. 1995. *The Minimalist Program*, 167-217. Cambridge, MA: The MIT Press].

Chomsky, Noam (1995). "Categories and Transformations". In *The Minimalist Program*, 219-394. Cambridge, MA: MIT Press.

Chomsky, Noam (2000). "Minimalist Inquiries: The Framework". In *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, R. Martin et al. (eds.), 89-155. Cambridge, MA: MIT Press.

Chomsky, Noam (2004). "Beyond Explanatory Adequacy". In *Structures and Beyond. The Cartography of Syntactic Structures (vol. 3)*, A. Belletti (ed.), 104-131. Oxford, NY: Oxford University Press.

Chomsky, Noam (2005). "Three factors in language design". *Linguistic Inquiry* 36: 1-22.

Chomsky, Noam (2007). "Approaching UG from below". In *Interfaces + Recursion = Language? Chomsky's minimalism and the view from syntax-semantics*, U. Sauerland and H-M. Gärtner (eds.), 1-30. Berlin: Mouton de Gruyter.

Chomsky, Noam (2008). "On Phases". In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, C. Otero et al. (eds.), 134-166. Cambridge, MA: MIT Press.

Chomsky, Noam; Lasnik, Howard (1995). "The Theory of Principles and Parameters". In *The Minimalist Program*, 13-127. Cambridge, MA: MIT Press.

Epstein, Samuel; Seely, Daniel (2002). *Derivation and Explanation in the Minimalist Program*. Malden, MA: Blackwell.

Fortuny, Jordi (2008). *The Emergence of Order in Syntax*. Amsterdam: John Benjamins.

Gallego, Ángel J. (2007). Phase *Theory and Parametric Variation*. PhD dissertation, UAB.

Gambell, Timothy; Yang, Charles D. (2003). "Scope and limits of statistical learning in word segmentation". Ms., Yale University.

Hardcastle, William J.; Hewlett, Nigel (2000). *Coarticulation: Theory, data and techniques*. Edited by. Cambridge: Cambridge University Press.

Hauser, Marc D.; Chomsky, Noam; Fitch, Tecumseh (2002). "The faculty of language: What is it, who has it, and how did it evolve?". *Science* 298: 1569-1579.

Hockett, Charles F. (1960). "The orgin of speech". *Scientific American* 203: 88-101.

Hornstein, Norbert (2009). *A Theory of Syntax: Basic Operations and the Minimalist Program*. Cambridge: Cambridge University Press.

Jackendoff, Ray (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kuratowski, Kazimierz (1921). "Sur la notion de l'ordre dans la théorie des ensembles". *Fundamenta Mathematicae* 2: 161-171.

Lasnik, Howard; Saito, Mamoru (1992). *Move Alpha: Conditions on Its Application and Output*. Cambridge, MA: MIT Press.

Lebeaux, David (1991). "Relative clauses, licensing, and the nature of the derivation". In *Perspectives on phrase structure: heads and licensing*, S. Rothstein (ed.), 209-239. San Diego: Academic Press.

Martin, Roger; Uriagereka, Juan (2000). "Some Possible Foundations of the Minimalist Program". In *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*, R. Martin et al. (eds.), 1-29. Cambridge, MA: MIT Press.

Richards, Marc D. (2004). *Object Shift and Scrambling in North and West Germanic: A Case Study in Symmetrical Syntax*. PhD dissertation, University of Cambridge.

Yang, Charles D. (2002). *Knowledge and learning in natural language*. New York: Oxford University Press.

Yang, Charles D. (2004). "Universal Grammar, Statistics or both?". *Trends in Cognitive Sciences* 8: 451-456.