A minimalist program for parametric linguistics?

Giuseppe Longobardi

1. Parametric linguistics[1]

Parametric linguistics may be viewed as a distinct subfield of generative grammar, or more generally of the abstract biolinguistic approach developed in the past decades of research, since at least Chomsky (1955, 1965) and Lenneberg (1967). It presupposes the crucial result acquired through this approach, namely that human language comprises a rich system of invariant innate knowledge, and concerns itself with the further problem of a formal and principled theory of grammatical diversity, a privileged testing ground for theories of the interaction between biologically shaped structures and culturally variable information.[2] Such a subfield is minimally defined by three fundamental questions:

(1) a. What are the actual parameters of UG?
    b. How do parameter values distribute in space and time?
    c. What is the form of possible parameters?

Question (1a) has been addressed in a rather inductive form by the hundreds of case studies proposing all sorts of morphosyntactic parameters over the past twenty years. Longobardi (2003) observed that hardly any significant module of grammar, small though we may take it, has so far attained a degree of parametrization with any pretension of typological exhaustiveness. To remedy this situation empirically it was suggested that linguists, rather than just existentially asserting the presence of certain parameters, should explicitly aim at a 'modularized global parametrization', i.e. at universal restrictive hypotheses on how widely languages may vary in very circumscribed grammatical modules.

[1] I am indebted to Michael Arbib, Hans G. Obenauer, and the participants in the Newcastle Conference on Parametric Variation (September 2005) organized by Anders Holmberg for inspiring discussions and comments on the topics of this squib.

[2] Cf. Longobardi (2003) for some more extensive remarks and methodological proposals.

Question (1b) is of a different nature and opens the way to the fascinating unification of two traditionally different domains, historical linguistics (with some connections reaching into paleoanthropology) and the cognitive approaches of formal grammar. This question has begun to be explored in Guardiano and Longobardi (2005), Gianollo, Guardiano, and Longobardi (2004), and will not be pursued in this paper.

Thus, the present note will deal with question (1c) and how its study may renovate our understanding of question (1a) and, indirectly, (1b) as well. The only well known restriction proposed on the format of parameters is so far the conjecture, stemming from Borer (1984), that parameters are always properties of functional heads of the lexicon. Fundamentally accepting this insight as a point of departure, in what follows I will suggest the possibility of a more articulated restrictive theory and point out some of its desirable consequences.

2. Parameters: problems and methods

In order to fruitfully address (1c) it may be useful to raise some more particular guiding issues, such as the following evolutionary ones, and try the adequacy of the proposals against them:

(2) a. Why is there so wide grammatical variation (so many parameters)?
b. Why is there grammatical variation (i.e. parameters) at all?[3]

Longobardi (2003) noticed that these problems have become more acute precisely with the development of parametric approaches.

Thus, the classical generative model (e.g. Chomsky 1965, call it the *Aspects* model for convenience) viewed the Language Acquisition Device (LAD) as consisting of a set of universal principles (Universal grammar, UG) and of an Evaluation Metrics for grammars, valuing grammars *freely* constructed by the language learner within the bounds posed by the Universal Principles. Within such a model, the existence of variation is potentially explained in terms of minimization of genetic endowment; for the 'freedom' of variation allowed can be construed as the complement of the principles made available by UG: the more such restrictive principles, the more the number of languages prohibited. Therefore, the width of variation could be due to the fact that the amount of universal restrictions

[3] The relevance of the latter question was first raised to my attention by Richard Kayne (p.c.) some years ago.

made available by human nature is limited by a sort of 'memory load' constraint on genetic transmission of cognitive information, a conceivable 'economy' condition on the architecture of the LAD, active through evolutionary history.

In the later Principles-and-parameters (P&P) model, LAD consists of a UG with both universal principles and parameters:

(3) UG = Principles + Parameters. Open parameters at S_0, closed parameters at S_S

In principle, then, grammatical variation is also innately given (*exhaustively* given, at the appropriate level of idealization), under the form of a presumably finite amount of discrete possibilities. Variability is already present at the initial state of the mind S_0 in the form of open parameters, actual varieties are represented by closed parameters at the steady state S_S. In this model the existence of variation is hardly explained, and certainly cannot be explained by the previous line of reasoning: for limiting the amount of transmittable genetic information, i.e. the size of the LAD, should presumably reduce the number of possible parameters as well, therefore increase, rather than decrease, the degree of invariance of the language faculty observable across individual languages.

This puzzle will constitute a good start for an investigation of restrictions on parameters. Another prerequisite to attempt a restrictive theory of parameters is a sufficient collection of structured data about variation and a way of representing such information in a perspicuous form. To approximate this objective realistically, it has been suggested in Longobardi (2003) to adopt the strategy of Modularized Global Parametrization, alluded to above. Trivializing matters a little, this method can be summarized in the following formula: studying relatively *many* parameters across relatively *many* languages in a *single* module of grammar. Considering a certain number of parameters together is obviously necessary to attempt any generalization; observing more than just a pair of contrasting languages for each parameter is required of a theory with some ambition of typological completeness; and concentrating on a single module makes the enterprise more realistically feasible but also allows one to explore a major formal feature of parameter sets, as already emerging from the works of Fodor (2000), Baker (2001), and Guardiano and Longobardi (2005), namely their widespread interdependence (cf. immediately below). The MGP method seems thus to be an appropriate compromise between depth and coverage. Following this method, Gianollo,

Guardiano, and Longobardi (2004) have set up a grid of 49 parameters affecting the internal structure of Determiner Phrases and tentatively stated their values in 23 languages and their partial dependencies, i.e. the absolutely pervasive situations where choosing one of the two values of a parameter neutralizes the relevance of the valuing of another parameter. All such parameters could be formulated as binary and their values have been marked in the adopted formalism as + and -. When the value of a parameter depends entirely on the value of (one or more) other parameters it is marked with a 0 (cf. Gianollo, Guardiano, and Longobardi 2004 for further details). This approach and formalism produce parametric grids summarizing large amounts of empirical information and theoretical hypotheses, highly valuable for further speculation on the theory of parameters itself.

3. Parametric minimalism

On the grounds of the empirical material so collected it becomes possible to address question (1c) and, more generally, raise an issue like (4):

(4) Can we subject parameters and their formats to minimalist critique?

Consider now that, as a first approximation, it turned out that at least 41 out of the 49 parameters of DP-internal structure investigated in Gianollo, Guardiano, and Longobardi (2004) can be reduced to a form which falls into one of only 4 abstract **parameter schemata**, listed below, where F and X,Y are variables over features and categories:

(5) a. Is F, F a functional feature, grammaticalized?
b. Is F, F a grammaticalized feature, checked by X, X a lexical category?
c. Is F, F a grammaticalized feature, spread on Y, Y a lexical category?
d. Is F, F a grammaticalized feature checked by X, strong (i.e. overtly attracts X)?

By ‘grammaticalized’ in (5a) it is meant that the feature must obligatorily occur and be valued in a certain structure, e.g. definiteness is obligatorily marked in argument DPs in certain languages (say English), not in others (say Russian). This does not mean that even the latter languages cannot

have lexical items occasionally used to convey the semantic meaning of definiteness (presumably demonstratives can convey such a meaning in every language), but in this case the feature 'definiteness' would be regarded as a lexical, not a grammatical one.

(5b) asks whether a certain feature requires establishing a relation with a specific (optionally or obligatorily present) category in the structure, creating a dependency (acts as *probe* searching a certain syntactic space for a *goal*, in Chomsky's 2001 terminology).

(5c) asks if a feature which is interpreted in a certain structural position has uninterpretable occurrences, depending in value on it, on other categories.

Finally, (5d) corresponds to the traditional schema inaugurated by Huang (1982), asking whether the dependency mentioned in (5b) involves overt displacement of X, i.e. remerging of X next to F, or not.

I will leave the question open, for the time being, whether, considering other domains, a fifth recurrent parameter schema needs to be added to such formats, namely one asking whether a certain category or feature may be phonetically null in a certain situation (e.g. the well known 'null pronoun' schema, as inaugurated by Taraldsen 1978 and especially refined in Rizzi 1986); it is possible that, in some cases at least, this sort of parametrization is reducible to environmental factors and combinations of other parametric schemata above, therefore does not instantiate a separate schema.

Let us then suppose, very speculatively, that these are the only possible 'core' parameter schemata: from this approach it already follows that certain conceivable types of variation are excluded; there follows, e.g., a conclusion with far-reaching consequences, such as (6):

(6) The locus of interpretation of each grammatical feature is universal, not parametrized

To exemplify the functioning of the parameter schemata system, consider the answers (7a-d) provided, on the grounds of the evidence tentatively built into the specific parametric grid of Gianollo, Guardiano, and Longobardi (2004), to the four respective questions asked in (5) once the values for the variables over features and categories have been fixed:

(7) For F = definiteness, X = demonstratives, Y = relative clauses
   a. Yes: Italian, No: Russian
   b. Yes: Italian, No: Greek
   c. Yes: Wolof, No: Italian
   d. Yes: Bulgarian, No: Norwegian

This way, parameter schemata of the type of (5) derive actual parameters, which can be literally constructed out of functional features, lexical categories, and parameter schemata, and set under usual assumptions. If this approach is correct in its essentials, it becomes unnecessary to suppose that the initial state of the mind consists of highly specific parameters, but just of an incomparably more restricted amount of parameter schemata, which combine with the appropriate elements of the lexicon (features and categories) under the relevant triggers in the primary data to both yield the necessary parameters and set their values for each language:

(8) Principles&Schemata model: UG = principles and parameter schemata. Parameter schemata at S_0, closed parameters at S_S

It is then conceivable that parameters which seem to end up being set on 0 (according to the formalism of Gianollo, Guardiano and Longobardi alluded to above) in a particular I-language have actually never been present at any state in the mind attaining that I-language.

The enormous number of possible core parameters depends, in principle, on the more limited numbers of functional features F and of lexical categories X, Y, combined with the tiny class of parameter schemata. Notice, however that it is not necessary that all parameter schemata are realized for every possible functional feature and all potentially relevant categories: specific principles of UG might forbid variation of an *a priori* admitted format for particular combinations of features and categories. Descriptively claiming, for example, as is often done, that the so called EPP feature is 'universally strong' amounts to preventing a widespread schema of variation among languages from determining differences as to whether the Spec of T is overtly filled or not.

4. The issues of variation

Accepting the Principles&Schemata model immediately determines the possibility of huge arithmetic simplification in the primitive axioms of the theory of grammatical variation: exactly like parameters were adopted

(also) as cross-constructional generalizations, significantly reducing the amount of apparently atomic points of variation, parameter schemata, in the intended sense, are more abstract, cross-parametric entities, allowing further simplification of the set of primitives. This begins to provide a sensible answer to problem (2a), because the amount of variation itself to be explained is drastically reduced: it will be sufficient to justify the existence of a certain parameter schema through justification (e.g. reduction to 'virtual conceptual necessity', in Chomsky's 1995 sense) of a single parameter of that schema, in order to explain the possibility (ultimately, the evolutionary rise) of the whole family of parameters of the same format.

But such an approach already relieves the burden of the explanation for the very existence of language diversity (issue (2b)) as well: for, within the proposed model, variation could largely be explained as in the first, pre-P&P, model. As we have just noted, once the introduction of a parameter schema into the language faculty is justified (e.g. evolutionarily explained, perhaps reducing it to conditions of efficiency on language transmission and use) for one case, it will be admitted and cause proliferating potential variation for all possible combinations of relevant entities of the lexicon (features and categories). This, unless a further particular principle of UG prohibits certain types of variation: in other words, once a schema has entered UG, then reducing variation essentially requires adding to the size of LAD, exactly as in the *Aspects* model. The kind of explanation in terms of 'economy of UG size' usable in that model can therefore be reproposed in the Principles&Schemata approach.

Of course, in order for a full minimalist program to be pursued within this approach it is necessary to show each of the parameter schemata to be indispensable, i.e. reducible to virtual conceptual necessity, or at least to be significantly related to architectural/computational properties present in other biological systems.

This whole, crucial, part of the program cannot be seriously addressed now, especially within the narrow limits of the present work. Only some exemplification of the required direction of research can be provided.

For example, (5a) could be motivated again by 'economy' constraints on performance (no language could grammaticalize the full set of conceivable functional features) to be spelt out by specific research; (5c) could perhaps be ultimately related to an acquisitional strategy of formal preservation of morphological content under the pervasive diachronic phenomenon of categorial shift or reanalysis (say, of a lexical item from a class where the occurrence of certain features is interpretable to another one where it is not).

Although this sketch is very preliminary, it should provide a reasonable idea of the heuristic power of the program advocated here and of the perspective of applying Chomsky's (1995, 2005) groundbreaking minimalist suggestions to the domain of grammatical variation.

References

Baker, Mark. 2001. *The Atoms of Language*. New York: Basic Books.

Borer, Hagit. 1984. *Parametric Syntax*. Dordrecht: Foris.

Chomsky, Noam. 1955. *The Logical Structure of Linguistic Theory*. Ms. (published in 1975, New York: Plenum).

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

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

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass: MIT Press.

Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36.1: 1-22.

Fodor, Janet D. 2000. Setting syntactic parameters. In *The Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin and Chris Collins. Cambridge, MA and Oxford, UK: Blackwell, 730-767.

Gianollo, Chiara, Cristina Guardiano, and Giuseppe Longobardi. 2004. Historical Implications of a Formal Theory of Syntactic Variation. Paper presented at DIGS VIII, Yale. In press in *Proceedings of DIGS VIII*, ed. by Stephen Anderson and Dianne Jonas. Oxford: Oxford University Press.

Guardiano, Cristina and Giuseppe Longobardi. 2005. Parametric Comparison and Language Taxonomy. In *Grammaticalization and Parametric Variation*, ed. by Montserrat Batllori, Maria-Lluïsa Hernanz, Carme Picallo, and Francesc Roca, 149-174. Oxford: Oxford University Press.

Huang, C.-T. James. 1982. Logical Relations in Chinese and the Theory of Grammar. Doctoral dissertation, MIT, Cambridge Mass.

Lenneberg, Eric H. 1967. *Biological Foundations of Language*. New York: John Wiley & Sons.

Longobardi, Giuseppe. 2003. Methods in Parametric Linguistics and Cognitive History. *Linguistic Variation Yearbook* 3, 2003, 101-138.

Rizzi, Luigi. 1986. Null Objects in Italian and the Theory of *pro*. *Linguistic Inquiry* 17.3:501-557.

Taraldsen, Knut Tarald. 1978. On the NIC, vacuous application, and the *that-t* filter. Ms. MIT, distributed by Indiana University Linguistics Club, Bloomington IN.