Foreword

Giuseppe Longobardi

Pages ix–xvii of
Principles of Syntactic Reconstruction
Edited by Gisella Ferraresi and Maria Goldbach
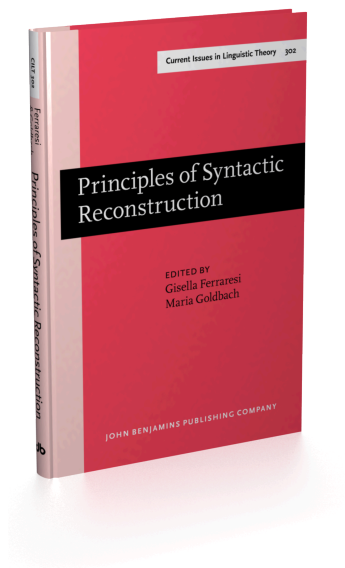[**Current Issues in Linguistic Theory**, 302]
2008. xvii, 219 pp.

Foreword

Giuseppe Longobardi

The publication of a volume of coordinated essays focusing on the problem of syntactic reconstruction is an extremely welcome initiative on the part of Gisella Ferraresi and Maria Goldbach, two scholars who have brilliantly participated for some years now in the effort of renovation of historical concerns within the framework of modern theories of grammar. The importance of this initiative in one of the most relevant, but least studied, subfields of historical syntax can hardly be overestimated: this is especially true in view of the contingent emergence of new insights, along with some misunderstandings and confusions, all beginning to surface neatly in the discussions of this volume, whose contributions range from the skeptical to the optimistic side of the debate over syntactic reconstruction. But I want to stress that, at this stage of the scientific research on language, there also exist deep epistemological reasons to seriously ask to what extent reconstruction is possible in syntax.

Reconstruction of past stages in order to explain the present is one of the most crucial tasks in any historical science: it is precisely by addressing such issues that historical syntax may cease to represent a relatively peripheral and somewhat outlandish subdomain of independently successful scientific paradigms, such as traditional historical linguistics or formal theories of grammar; rather, it may achieve a central and guiding role within a new paradigm of 'cognitive history', built as the intersection of historical (linguistic, anthropological, etc.) studies and the modern cognitive sciences. In fact, traditional historical linguistics, centered around lexical comparison and phonological change, has often regarded syntax as a field less able by itself to contribute significant generalizations to the most genuine of its concerns (comparison and reconstruction, indeed); while, within formal theories of grammar, historical syntax has been at best occasionally invoked to corroborate descriptive models of language variation and acquisition.

However, it is often the case that purely theoretical developments in one field enable that field to better serve the purposes of several related historical investigations, sometimes with exceptional advancements. For example, one of the most spectacular cases of such virtuous relation between progress in theoretical science and benefits for historical research has been the use of molecular discoveries in evolutionary biology and population genetics.

Now, two main developments have reshaped theoretical linguistics over the past decades, especially affecting syntax, and historical linguistics must come to terms with them and take advantage of their potential: one is the development of the abstract biolinguistic framework advocated by Chomsky and Lenneberg, among others, since the 1960s and 1970s; the other is the more specific hypothesis about language acquisition and the form of linguistic universals proposed by Chomsky around 1980 and called the Principles & Parameters model. As put by Chomsky (1995), the latter "[…] is in part a bold speculation rather than a specific hypothesis. Nevertheless, its basic assumptions seem reasonable […]". In particular, it seems reasonable that such assumptions may be tested with respect to their capacity to capture historical syntactic generalizations and thus, perhaps, independently corroborated as a historically realistic model of grammatical variation.

Both developments may contribute to raising new obvious questions and proposing ways of resolution in the domain of historical comparison and reconstruction. A proper start to introduce this important volume and help clarify the issues it aims to address seems precisely to be a brief overview of these conceptual contributions.

The most relevant consequence of adopting the abstract biolinguistic framework is the focusing on the biologically conditioned process of acquisition of individual linguistic competences, so called I-languages, rather than on the much vaguer and scientifically hardly definable idealization of E-languages, i.e., languages somehow shared by a community, in the more ordinary sense of everyday speech.

It is mostly David Lightfoot who has developed all the consequences of this shift of focus for historical syntax in a series of books over the past 25 years, since Lightfoot (1979), along with a radical epistemological critique of the teleologism of certain long-term 'explanations' embedded in theories based on so scarcely manageable a notion as E-language. The new approach must instead exclude, for instance, the causal relevance of temporally non-local factors in the shaping of any particular I-language.

It remains to be discussed, then, whether the foundational concepts of historical linguistics, in particular those relevant for syntactic reconstruction (genetic kinship, derivation of a language from another, even at a great chronological distance), can indeed be satisfactorily defined in terms of the primitives of the biolinguistic framework, namely I-languages and the entities logically pertaining to the problem of their acquisition.

Thus, the biolinguistic framework is crucially inspired by the logical problem of language acquisition, a particular case of the philosophical paradox of human knowledge, pointed out several times in the history of Western thought from Plato down to Russell and Chomsky:

(1)   How can any child exposed to an *insufficient* and highly *individual* sample
      of utterances (a primary *corpus*) develop intricate linguistic knowledge (an
      I-language) both *complete* and relatively *uniform* with respect to that of other
      members of his/her community?

The described situation (poverty of stimulus with respect to the acquired knowl-
edge) suggests the existence of a rich, innate, and universal mental organ,
often called LAD (*Language Acquisition Device*), connecting the two types of
descriptive primitives of synchronic linguistic theory, namely primary *corpora*
and I-languages.

Crisma and Longobardi (to appear) argue that the same concepts are sufficient
as primitives to ground historical work in linguistics as well. However, historical
linguistics cannot obviously focus just on a single I-language, since it implies, vir-
tually by definition, the comparison between at least two individual languages,
sometimes remote in space and time, whose similarities are neither universal nor
due to chance: so they must be the residue of a historical relation, an event of con-
vergence between two I-languages or of partial divergence of one from another.
Thus, historical linguistics apparently needs to be founded on two basic notions:
(1) I-language; (2) historically significant relation between I-languages (hence-
forth H-relation). Now, can the latter notion be defined from acceptable primitives
without additional recourse to vague concepts such as E-language?

The answer proposed by Crisma and Longobardi is affirmative and relies on a
set of definitions like the following:

(2) L_2 derives diachronically from L_1 if and only if
      a. L_2 is acquired on a primary *corpus* generated by L_1
           or
      b. L_2 derives from L_3 and L_3 derives from L_1

(2b) is a recursive clause, correctly allowing an indefinite iteration of the relation
of diachronic derivation. (2a) alone, instead, defines the notion of *immediate* deri-
vation (diachronic contiguity).

These definitions were then generalized to finally capture the notion of
H-relation:

(3)   Two linguistic objects (I-languages or their fragments) X, Y are in a H-relation
      if and only if one derives from the other or there is a Z from which both derive.

As can be seen, the definition of H-relation suggested by Crisma and Longobardi
(hence, the conceptual foundation, among other things, of historical compari-
son and reconstruction in linguistics) is based on independent primitives of
synchronic linguistic theory and, most importantly, does not appear to require

resorting to E-language. Thus, the adoption of the basic tenets of the biolinguistic framework does not hinder a coherent formulation of the issues involved by the study of syntactic reconstruction.

The rise of the Principles & Parameters framework affects even more directly the theory and practice of syntactic reconstruction, because of the revolutionary model of grammatical variation it proposes. This model attributes to variation of grammar, as opposed to the lexicon, at least the following fundamental characteristics, the first two of which are virtually analytic in the definition of parameter:

(4)   a.   discreteness (i.e., the values of each parameter form no continuum and in the ideal case just amount to two)
       b.   finiteness
       c.   deductive depth (i.e., a limited number of quite abstract distinctions produce the many ones appearing on the surface: several superficial differences turn out to cluster together at the appropriate level of analysis).

This model provides a completely new and extremely manageable sort of theoretical entities for the procedures of syntactic comparison and reconstruction, namely parameter values. In fact, one of the causes of the mentioned minority status assigned to syntax by traditional historical linguistics had certainly to do with the difficulty of identifying precise *comparanda* and *reconstruenda* in syntax. A crucial condition for historical linguistics to exist is being able to establish, at least hypothetically, some precise correspondence (i.e., a postulated H-relation) between two entities (words, phonemes, rules, etc.) of two different I-languages. On this crucially rely the comparative method, reconstruction, scientific etymology, and, ultimately, the very idea of phylogenetic kinship. However, in traditional approaches it is often unclear what are the entities of a syntactic theory (rules, constructions, word sequences …?) and whether they really form a discrete and finite set, comparable across languages like, e.g., lexical units or the distribution of sounds within them. Thus, one may naturally wonder if Latin, Greek, Sanskrit, and English 'passive' or 'predication' do instantiate the 'same' entity in the sense and to the extent that we think Lat. *iugum*, ζυγόν, Skt. *yugám*, ModE *yoke* do.

It is obvious how parametric theories can solve the problem: parameters and parameter values should in principle not only form a finite set of discrete entities, but even a universal one, from which I-languages choose their configurations of values.

Thus, parameter values appear to provide perfect *comparanda* to establish syntactic correspondences across languages and, in the same sense, also make perfect *reconstruenda* for proto-syntaxes, thus solving in principle the long-standing problem of correspondence sets in historical syntax.

In principle, this opens the path to syntactic comparison and to comparative reconstruction of proto-syntaxes, as keenly pointed out by Roberts (1998). But Guardiano and Longobardi (2005) and Longobardi and Guardiano (to appear)

even take a further step, suggesting that this path may also lead to parametric linguistics as a tool for reconstructing phylogenetic relations among languages.

Whether these exciting developments may eventually be successful is an empirical question, and one of the most far-reaching challenges of historical linguistics for the next years. But another type of principled question needs to be addressed, namely if syntactic comparison and reconstruction can be founded on *comparanda* other than parameter values. Interestingly, Alice Harris in this volume and in other publications has suggested that a well definable notion, suitable for setting up syntactic correspondence sets, is that of *pattern*, a less abstract object than parameter values, more similar to the traditional notion of 'construction'.

As remarked by Roberts (1998), patterns are configurations which may be generated by different parameters in different languages, i.e., there is no necessary universal biunique relation between a certain parameter value and a certain pattern. Rather, it is increasingly clear that, owing to the enormously complex interaction possibilities among parameters, the set of surface constructions triggering a parameter value may be very rich and proteiform across languages. For this reason, it is highly unlikely that one may reconstruct an I-language by simply comparing scattered patterns. A serious empirical question is, however, whether by this procedure (which amounts to running the risk of establishing correspondence sets whose members are superficially alike but are determined by different mental objects in different languages) we may ever safely reconstruct fragments of E-language. In principle this possibility arises and has obvious successful parallels in phonology: by comparing a sufficient number of Italian and French words we would certainly be able to reconstruct the Latin accent of virtually every item (in most cognate words in the three languages the same syllable is accented); but the prosodic principles generally determining the placement of the accent in Latin, French and Italian are, synchronically, radically different rule systems, as is pretheoretically obvious to any Italian student learning Latin or French. In such cases the generalization historically connecting the three systems can be easily formulated at the level of primary *corpora*, i.e., of the generated E-language, not of speakers' mental structures (I-language).

However, it is conceivable that in many cases what is historically persistent and profitably comparable in a set of sister languages is rather a parametric value, in spite of the fact that slightly different 'patterns' may manifest it in each of them, owing to the interaction with their other syntactic or morphological properties. Reconstructing a parametric value this way will probably be less immediately useful to write, say, stories about horses and sheep in the proto-E-language (unless we are able to control for all the other morphosyntactic variables affecting its surface manifestations), but will directly target a deep feature of the proto-I-language.

The latter possibility, namely that parameter values exhibit a significant degree of persistence across time and through the storms of language transmission, is tentatively suggested by the recent attempts to use parametric linguistics for reconstructing phylogenetic kinship among languages, rather than the structure of a protolanguage, as mentioned above. The two tasks are different in goal, of course, but also in method. Recognizing this difference is important in order to avoid some frequent misunderstandings.

The classical historical-comparative method applied to the sound shape of lexical items is normally held very suitable for reconstruction both of protolanguages and of phylogenetic relations. Indeed, it is often claimed that some degree of reconstruction of the protolanguage is necessary to prove that two or more daughter languages are actually related. In other words, success at reconstruction is taken to be implied by assumed success at proving phylogenetic kinship. Other methods explicitly deny such implication: Greenberg's mass comparison method is claimed to be able to identify degrees of phylogenetic relationship without necessarily (and practically) reconstructing a single word of the postulated protolanguages. This is so because the supposed evidence of kinship is based on the stacking of several relatively similar lexical items in many languages, in order to reduce the probability of chance similarity, without any chance-proof demonstration that two items make a safe etymological pair, as can instead often be assessed by the classical method through precise sound correspondence sets. Whatever the (dubious and highly controversial) results of Greenberg's actual comparisons, the point is not methodologically incorrect: a statistically sufficient amount of similarity between two languages may in principle speak in favor of some genetic kinship, even if we are not always in the position to decide which exactly of the similarities can be attributed to the protolanguage. This is precisely one aspect of the reasoning behind parametric comparison, as advocated by Longobardi and Guardiano (to appear): no coincidence in any single parameter value alone may speak for kinship between two languages, but coincidence in a (probabilistically definable) huge number of them may. It follows that we can be relatively confident in assessing a degree of relation between languages on the grounds of the comparison between sets of parameter values, without necessarily being able to hypothesize for each of these parameters how it was set in the protolanguage. Therefore, reconstruction of the latter is not always a necessary precondition to deciding issues of kinship. This is another reason why the two goals, reconstruction of protolanguages and of phylogenetic relations, must not be confused in historical syntax.

Syntactic reconstruction of phylogenetic kinship offers us the opportunity to briefly return to the parameters vs. patterns controversy, to simply notice that, if the same parameter may be responsible for the co-variation of several distinct surface pattern, then measuring the syntactic distance (hence the probability of

relatedness) in terms of patterns rather than parameter values might be arith-
metically somewhat misleading from the viewpoint of a phylogenetic enterprise
(let alone for a sound typological theory). In this sense, at least, it seems that an
abstract parametric analysis is not dispensable in a serious historical approach
to syntax.

An introduction to the issues of reconstruction in historical syntax and its
conceptual backgrounds could hardly be complete without mentioning a third
recent idea, which appears quite promising for a sound epistemological founda-
tion of the study of syntactic change: the notion of Inertia, stemming from Edward
Keenan's brilliant intuition at the beginning of the 1990s.

The relevance of Inertia for syntactic reconstruction derives from the fact
that proper reconstruction depends on a restrictive theory of possible changes.
An example from phonology, again, is helpful: just comparing Germanic f-, Latin
p-, Armenian h-/Ø-, Celtic Ø-, or any similar array of data, one would hardly
reconstruct anything other than PIE p-, even in spite of the majority rule. This
is so because a sufficiently restrictive theory of phonological change considers
highly improbable a change f > p or h > p, let alone Ø > p, as opposed to the
reverse changes.

In principle, the same should be true for syntax. Now, Keenan's notion of Iner-
tia provides precisely the potentially most restrictive framework for a theory of
grammatical change. In Keenan (1994, 2000), the principle of Inertia is phrased in
the following way:

(5)   INERTIA: Things stay as they are, unless acted upon by an outside force or
       DECAY

Let us call this formulation *absolute* Inertia: if 'decay' in this formulation can be
understood as a restricted class of sound changes (e.g., p > Ø, but not viceversa),
the other major class of causes of change ('outside force') is essentially represented
by interference factors, a major part of which is probably to be construed *lato
sensu* as the inevitable heterogeneity of primary *corpora*. That interference, though
important, cannot be regarded as the only primitive factor in language change, had
already been assumed by the Neogrammarians, especially in the light of works like
Winteler (1876), describing a putative isolated dialectal community, still affected,
though, by some (perfectly regular) sound change. Therefore, it is improbable
that we may ever achieve the most ideal theory of linguistic change, one where all
primitives for change are just external factors, but – here lies in my view the revo-
lutionary core of Keenan's insight – one should try to limit internal unmotivated
causes for change to the bare minimum.

Pursuing this line, Longobardi (2001) has proposed a slightly less commit-
ting formulation, which could be considered a version of Inertia relativized to the

module of syntax, as viewed in particular in the minimalist program. It claims that, given two I-languages connected by a relation of diachronic derivation, as defined above, at least narrow *syntax* is actually completely immune to primitive changes. It should only change by interference or as a predictable reaction to changes taking place or represented at its interfaces: e.g., changes in phonological rules and in phonological or semantic representations of lexical items. Inertial hypotheses imply, of course, that the LAD is, at least in the relevant modules, an essentially deterministic machine: which, in isolation from disturbing factors, like other primitive changes and multiplicity of I-languages generating the relevant primary *corpus*, it should exactly replicate, in a new I-language, the syntax of the I-language from which it immediately derives.

Theoretically, inertial theories, no matter how extensively applied, are similar to the Neogrammarian *Ausnahmslosigkeit* hypothesis: they are superficially challenged by a large number of observations, which one should try to reconcile with the hypothesis by reducing them to some of the affecting factors above (outside force, decay, or primitive changes in other modules …).

Whatever the eventual feasibility of such a research program, it is clear that it strongly aims toward much higher restrictiveness for theories of diachronic change. From these general concerns, therefore, we may hope to anyway derive the discovery of many more local constraints on possible syntactic changes, analogous to the phonological cases just exemplified. If so, the whole task of syntactic reconstruction, both comparative and internal, could substantially benefit from a better apprehension of which previous unobserved conditions may or may not have given rise to a certain state of affairs in the syntax of attested languages.

All these insights and questions lie behind the problem of syntactic reconstruction and are touched on, sometimes explicitly, in other cases more implicitly, as the careful reader will certainly notice, in the papers contained in this volume, which mark some definite progress in the depth of our understanding with respect to the state of the debate of only two or three decades ago, as laid on especially by Watkins (1976) and some of the essays in Ramat et al. (1980).

References

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

Crisma, Paola & Giuseppe Longobardi. to appear. "Change, relatedness and inertia in historical syntax". In press in Paola Crisma & Giuseppe Longobardi eds. *Historical Syntax and Linguistic Theory*. Oxford: Oxford University Press.

Guardiano, Cristina & Giuseppe Longobardi. 2005. "Parametric Comparison and Language Taxonomy". *Grammaticalization and Parametric Variation* ed. by M. Batllori, M.L. Hernanz, C. Picallo & F. Roca, 149–174. Oxford: Oxford University Press.

Keenan, Edward. 1994. "Creating Anaphors. An historical study of the English reflexive pronouns". Unpublished manuscript, UCLA.

Keenan, Edward. 2000. "An Historical Explanation of Some Binding-Theoretic facts in English". Unpublished manuscript, UCLA.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Longobardi, Giuseppe. 2001. "Formal Syntax, Diachronic Minimalism and Etymology: The history of French *chez*". *Linguistic Inquiry* 32.2: 275–302.

Longobardi, Giuseppe & Cristina Guardiano. to appear. "Evidence for Syntax as a Signal of Historical Relatedness". In press in *Lingua,* spec. issue *The Forests behind the Trees*, ed. by John Neesanne & Tranz Maumi.

Ramat, Paolo et al., eds. 1980. *Linguistic Reconstruction and Indoeuropean Syntax*. (= *Current Issues in Linguistic Theory*, 19.) Amsterdam & Philadelphia: John Benjamins.

Roberts, Ian. 1998. Review of Alice Harris & Lyle Campbell, *Historical Syntax in Cross-Linguistic Perspective* (Cambridge: Cambridge University Press, 1995). *Romance Philology* 51.363–370.

Watkins, Calvert. 1976. "Toward Proto-Indoeuropean Syntax: Problems and pseudoproblems". *Diachronic Syntax* ed. by Sanford B. Steever, Carol A. Walker & Salikoko S. Mufwene, 305–326. Chicago: Chicago Linguistic Society.

Winteler, Jost. 1876. *Die Kerenzer Mundart des Kantons Glarus in ihren Grundzügen dargestellt.* Leipzig & Heidelberg: C.F. Winter.