The historical reality of biolinguistic diversity

Giuseppe Longobardi˙, Cristina Guardiano^,
Luca Bortolussi°, Andrea Sgarro°, Giuseppina Silvestri°, Andrea Ceolin°*

˙Univ. York, ^Univ. Modena e Reggio Emilia, °Univ. Trieste, *Univ. Pisa

Goals. Arguing that the historical application of the biolinguistic model can complement molecular antrhopology to model out a 'grammatical anthropology' as a new discipline at the crossroads of cognitive, biological and historical sciences. **Background**. In many respects, P&P models of UG are a conceptually plausible answer to the problem of explanatory adequacy (Chomsky 1964). However, empirically, parametric theories are not yet sufficiently corroborated, since nobody has so far indisputably assessed their effectiveness to the acquisition of grammatical diversity by implementing a parameter setting system over a realistic collection of parameters (Fodor 2001, Yang 2003; cf. Chomsky 1995:7). It is therefore debatable that a P&P model has actually attained substantial explanatory adequacy, though progressing beyond language-specific descriptive adequacy. To address the need for more solid arguments in favor of P&P, Longobardi (2003) suggested the opportunity of: **i)** adopting a Modularized Global Parametrization strategy, aiming at studying together relatively many (closely interacting) parameters in relatively many languages within the circumscribed domain of small modules of grammar; **ii)** beginning to aim at further testing grounds and levels of success, i.e. at satisfactory accounts of the actual distribution of grammatical diversity in time and space (*historical* adequacy). **Methods**. Elaborating on previous work (Longobardi/Guardiano 2009), a sample of more than 50 carefully identified binary parameters in DP-syntax, set in over 30 languages, is focused on; it is complemented with a set of hypotheses about UG constraints, defining two levels of deductive structure: one determines the traditional covariation of properties following from the same parameter, the other encodes an extraordinarily rich implicational hierarchy among parameters themselves (more pervasive than hinted in Baker 2001), largely responsible for hierarchies of size (e.g. Biberauer/Roberts 2012). Phylogenetic programs of biostatistical derivation have been applied to this database to formally measure syntactic diversity and generate hypotheses of phylogenetic trees and networks. Specific mathematical procedures (a sampling algorithm capable of dealing with the universal constraints imposed on parameter setting) have been elaborated on purpose, to compute the width of potential diversity allowed by this fragment of UG and to evaluate the significance of the one observed in the actual language set. **First results**. The distribution of actual syntactic distances is statistically highly significant. The results have been measured against independently known historical data (from comparative linguistics, history, genetics), with largely correct correlations: given a non-trivial set of languages, the description of their variation provided by the systematic parametric analysis of a whole compact domain quite exactly recapitulates their known history and relationships. The reality of a P&P model of the language faculty, therefore, receives strong and original support from its historical adequacy. **Further testing**. Recently, the use of structural traits (superficial grammatical patterns) has been advocated and tested for conclusions on language phylogenies, the status of universals, and the modeling of grammatical evolution (Dunn *et al.* 2011). Now, since parameters try to represent 'abstract' differences, often exhibiting a high degree of deductive depth with respect to surface contrasts, counting similarities in patterns rather than in parameter values could turn out to provide different outcomes when quantitatively assessing areal or genealogical relatedness. In order to test this idea, the same experiments above have been repeated using a choice of the surface descriptive patterns derived from the parameters, rather than the parameter values themselves (i.e. comparing E-languages rather than I-languages). Again, the results were plotted against the same independently known historical variables. This experiment allows one to empirically test the parameter-pattern controversy and Dunn *et al.*'s alleged conclusion that implicational universals and

1

UG are not supported by the extant distribution of structural diversity in the world's languages. The first computations suggest that pattern-based phylogenies are by no means more significant or more revealing than those founded on abstract parameters and that the latter better represent actual historical linguistic relations. Further corroborating these results, we argue that biolinguistic models of diversity encode a higher level of reality than surface-oriented typologies.

References

Baker, M.(2001) *The Atoms of Language*. New York, Basic Books. **Biberauer** T., I. **Roberts** (2012) The significance of what hasn't happened. DiGS 14, Lisbon. **Chomsky**, N. (1964) Current Issues in Linguistic Theory. In: Fodor, J. A., J. J. Katz (eds.), *The Structure of Language: Readings in the Philosophy of Language*, Englewood Cliffs, Prentice Hall: 50-118. **Chomsky**, N. (1995) *The Minimalist Program*. Cambridge, MA, MIT Press. **Dunn**, M., S. J. **Greenhill**, S. C. **Levinson**, R. D. **Gray** (2011) Evolved structure of language shows lineage-specific trends in word-order universals, doi:10.1038/nature09923. **Fodor**, J. D. (2001) Setting syntactic parameters. In M. Baltin, C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*, Oxford, Blackwell. **Longobardi** G. (2003) Methods in Parametric Linguistics and Cognitive History. *LVY*, 3: 101-138. **Longobardi** G., C. **Guardiano** (2009) Evidence for Syntax as a Signal of Historical Relatedness. *Lingua*, 119, 11, 1679-1706. **Yang**, C. (2003) *Knowledge and learning in natural language*. Oxford, OUP.

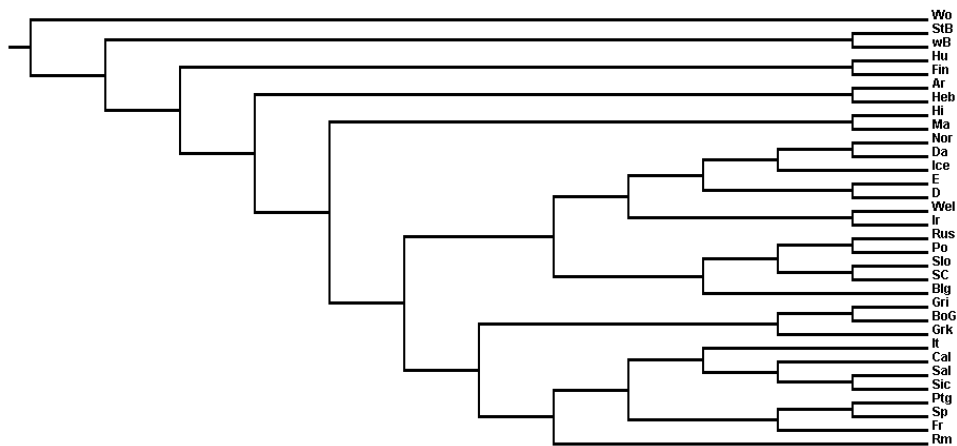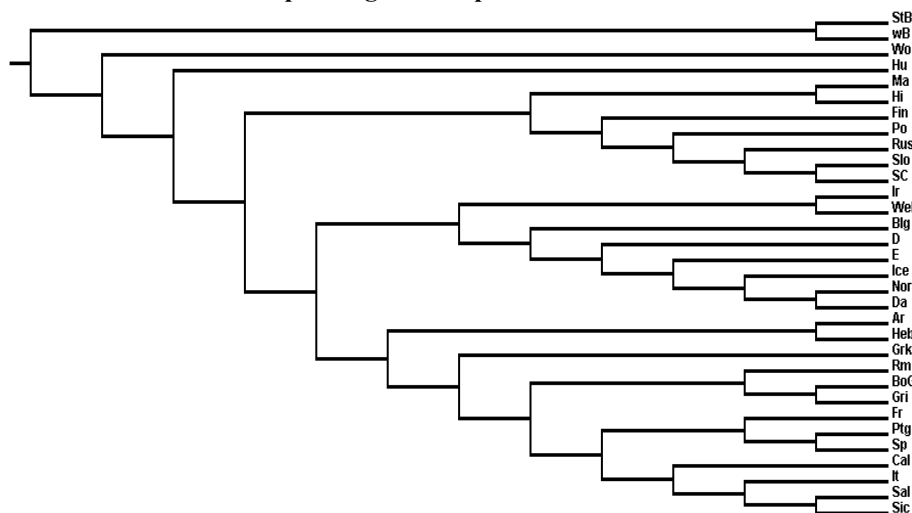**Fig. 1: UPGMA Tree from 56 syntactic parameters[1]**

Fig. 2: UPGMA Tree from 113 corresponding surface patterns

---

[1] **Legenda**. **Wo** = Wolof; **StB** = Standard Basque; **WB** = Western Basque; **Hu** = Hungarian; **Fin** = Finnish; **Ar** = Arabic; **Heb** = Hebrew; **Hi** = Hindi; **Ma** = Marathi; **Nor** = Norwegian; **Da** = Danish; **Ice** = Icelandic; **E** = English; **D** = German; **Wel** = Welsh; **Ir** = Irish; **Rus** = Russian; **Po** = Polish; **Slo** = Slovenian; **SC** = Serbo-Croat; **Blg** = Bulgarian; **Gri** = Grico (Salentino Greek); **BoG** = Bovese Greek (Calabria, Grecanico); **Grk** = Greek; **It** = Italian; **Cal** = Calabrese; **Sal** = Salentino; **Sic** = Sicilian; **Ptg** = Portuguese; **Sp** = Spanish; **Fr** = French; **Rm** = Rumanian.