Universals in cognitive theories of language

commentary to Evans and Levinson to appear in BBS

Paul Smolensky[a] and Emmanuel Dupoux[b]

[a]*Department of Cognitive Science, Johns Hopkins University, Baltimore, MD 21218-2685.*

smolensky@jhu.edu http://www.cogsci.jhu.edu/faculty/smolensky.html

[b]*Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole des Hautes Etudes en Sciences Sociales, Département d'Etudes Cognitives, Ecole Normale Supérieure, Centre National de la Recherche Scientifique, 75005 Paris, France.* emmanuel.dupoux@gmail.com

http://www.lscp.net/persons/dupoux.html

Abstract: Generative linguistics' search for linguistic universals (1) is not comparable to the vague explanatory suggestions of the article, (2) clearly merits a more central place than linguistic typology in cognitive science, (3) is fundamentally untouched by the article's empirical arguments, (4) best explains the important facts of linguistic diversity, and (5) illuminates the dominant component of language's "biocultural" nature: biology.

1. *A science of cognition needs falsifiable theories.* Although the article's final seven theses include suggestions we find promising, they are presented as vague speculation, rather than as a formal theory that makes falsifiable predictions. It is thus nonsensical to construe them as superior to a falsifiable theory on the grounds that that theory has been falsified. Every theory is certain to make some predictions that are empirically inadequate, but the appropriate response within a science of cognition is to improve the theory and not to take refuge in the safety of unfalsifiable speculation. Insightful speculation *is* vital – not because speculation can *replace* formal theorizing but because speculation can be sharpened to *become* formal theory. Theory and speculation are simply not empirically comparable.

2. *In a theory of cognition, a universal principle is a property true of all human minds – a cog-universal – not a superficial descriptive property true of the expressions of all languages – a des-universal.* This is why generative grammar, with its explicit goal of seeking cog-universals, has always been more central to cognitive science than linguistic typology, which only speaks to des-universals. Unlike descriptive linguistic

typology, generative grammar merits a central place in cognitive science because its topic is cognition and its method is science – falsifiable theory formulation.

3a. Counterexamples to des-universals are not counterexamples to cog-universals. The des- universals of Box 1 must not be confused with the cog-universals sought by generative grammar. This general point applies to all cases addressed in the article, but we only illustrate with one example. That Chinese questions do not locate *wh*-expressions in a different superficial position than the corresponding declarative sentence (Box 1) is a counterexample to a *wh*-movement *des*-universal but, famously, generative syntax has revealed that *Chinese behaves like English* with respect to syntactically determined restrictions on possible interpretations of questions; this follows if questions in both languages involve the same dependency between the same two syntactic positions, one of them "fronted." In English, the fronted position is occupied by the *wh*-phrase and the other is empty, whereas in Chinese the reverse holds (Huang 1998; Legendre et al. 1998). It is the syntactic relation between these positions, not the superficial location of the *wh*-phrase, that restricts possible interpretations. Such a hypothesized *cog*-universal can only be falsified by engaging the full apparatus of the formal theory. It establishes nothing to point to the superficial fact that *wh*-expressions in Chinese are not fronted.

3b. ***There are two types of cog-universals: Architectural and specific universals.*** The former specify the computational architecture of language: levels of representation (phonological, syntactic, semantic, etc.) data structures (features, hierarchical trees, indexes, etc.), operations (rule application, constraint satisfaction, etc.). The authors correctly recognize these as "design features" of human languages, but they erroneously exclude them from the set of relevant universals. These architectural universals do not yield falsifiable predictions regarding typology, but they yield falsifiable predictions regarding language learnability. For instance, Peperkamp et al (2008) showed that without architectural universals regarding phonological rules, general-purpose unsupervised learning algorithms simply fail to acquire the phonemes of a language. The latter, specific universals, are tied to particular formal theories specifying in detail the architecture's levels, structures, and operations, thus yielding falsifiable predictions regarding language typology.

4a. ***Optimality Theory (OT), mentioned in the article as a promising direction, contains the strongest architectural and specific universals currently available within generative grammar.*** According to OT's architectural universals (Prince & Smolensky 1993/2004; 1997), grammatical computation is optimization over a set of ranked constraints. This strong hypothesis (more than the hypothesis of "parameters"), has contributed insight into all levels of grammatical structure from phonology to pragmatics and has addressed acquisition, processing, and probabilistic variation (http://roa.rutgers.edu hosts more than 1,000 OT papers). In a particular OT *theory*, specific universals take the form of a set of constraints (e.g., C1 = "a sentence requires a subject"; C2 = "each word must have an interpretation," and so on. A *grammar* for a particular language is then a priority ranking of these constraints. For instance, C1 is ranked higher than C2 in the English grammar, so we say "it is raining," although expletive "it" contributes nothing to the meaning; in Italian, the reverse priority relation holds, making the subjectless sentence "piove" *optimal* – grammatical (Grimshaw & Samek-Lodovici 1998).

4b. ***OT's cog-universals yield theories of cross-linguistic typology that generally predict the absence of des-universals.*** Each ranking of a constraint set mechanically predicts the possible existence of a human language. OT therefore provides theories of linguistic typology that aim, as rightly urged by the article, to grapple with the full spectrum of cross-linguistic variation. OT makes use of a large set of specific universals (i.e., constraints), but because of the resolution of constraint conflict through optimization, they do not translate into des-universals. In the preceding example, C1 is violated in Italian and C2 in English. Some des-universals can emerge, however, as general properties of the entire typology, and they can be falsified by the data (as, perhaps, the existence of onsetless languages). This does not entail abandoning the Generative Linguistics program nor the OT framework, but revising the theory with an improved set of specific universals.

5. ***Language is more a biological trait than a cultural construct.*** The authors do not provide criteria to determine where language is located on the continuum of biocultural hybrids. Lenneberg, quoted in the target article, presented four criteria for distinguishing biological traits

from cultural phenomena (universality across the species, across time, absence of learning of the trait, rigid developmental schedule) and concluded that oral (but not written) language is a biological trait (Lenneberg 1964). The validity of this argument is ignored by the authors. Ironically, OT is more readily connected to biology than to culture: the f-universals of OT are emergent symbolic-level effects of subsymbolic optimization over "soft" constraints in neural networks (Smolensky & Legendre 2006), and Soderstrom et al. (2006) derive an explicit abstract genome that encodes the growth of neural networks containing connections implementing universal constraints.

References

Grimshaw, J. & Samek-Lodovici, V. (1998) Optimal subjects and subject universals. In: *Is the best good enough? Optimality and competition in syntax*, ed. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, pp. 193–219. MIT Press.

Huang, C.-T. (1998) *Logical relations in Chinese and the theory of grammar*. Garland.

Legendre, G., Smolensky, P. & Wilson, C. (1998) When is less more? Faithfulness and minimal links in *wh*-chains. In: *Is the best good enough? Optimality and competition in syntax,* ed. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky, pp. 249–89. MIT Press.

Lenneberg, E. (1964) The capacity for language acquisition. In: *The structure of language,* ed. J. A. Fodor & J. J. Katz, pp. 579–603. Prentice Hall.

Peperkamp, S., Le Calvez, R., Nadal, J. P. & Dupoux, E. (2006) The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101:B31–B41.

Prince, A. & Smolensky, P. (1993/2004) *Optimality theory: Constraint interaction in generative grammar*. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Revised version published by Blackwell, 2004. Rutgers Optimality Archive 537.

Prince, A. & Smolensky, P. (1997). Optimality: From neural networks to universal grammar, *Science* 275(5306):1604–10.

Smolensky, P. & Legendre, G. (2006) *The harmonic mind*. 2 vols. MIT Press.

Soderstrom, M., Mathis D. & Smolensky, P. (2006) Abstract genomic encoding of universal grammar in optimality theory. In: *The harmonic mind,* ed. P. Smolensky & G. Legendre, pp. 403–71. MIT Press.