

A rejoinder to 'On the role of parameters in Universal Grammar: a reply to Newmeyer' by Ian Roberts and Anders Holmberg

Frederick J. Newmeyer
University of Washington

fjn@u.washington.edu

Revised version, February 2006

1. Introduction

My paper ‘Against a parameter-setting approach to typological variation’ (Newmeyer 2004; henceforth ‘*APS*’) contrasts parameter-based and rule-based approaches to cross-linguistic typological variation and opts for the latter. I note that as far as UG *per se* is concerned, a rule-based theory is in principle ‘less constrained’ (185) than a parametric theory. However, this circumstance is compensated for in three ways. First, in the rule-based theory argued for in *APS*, independently motivated performance principles do much of the work attributed to UG in parametric theory. Second, given parametric theory as it is actually put into practice, ‘a rule-based account is either more adequate than a parameter-based one or, when all the facts are taken into account, they are empirically indistinguishable’ (189). And third, every parametric theory of which I am aware posits rules (or their equivalent) anyway, side-by-side with parameters.

Roberts and Holmberg 2005 (henceforth ‘*R&H*’), in their reply to *APS*, argue that ‘Newmeyer’s arguments against [parametric theory] are based on misunderstandings either of theory or of data, are conceptually misconceived, illogical or simply false’ and ‘represent a retreat to observational adequacy’ (1).¹

In this rejoinder, I focus on the serious deficiencies in *R&H*, deficiencies which vitiate its critique of *APS*. I argue in turn that *R&H* sets the bar for motivating parametric theory so low that it is hard to imagine what sort of data might refute it (§2); that in more than one place *R&H* explicitly concedes the point argued for in *APS* (§3); that *R&H* implies that by claiming (and reclaiming) that parametric theory is superior *in principle* to rule-based theory, nothing more needs to be said (§4); that *R&H* adopts mutually incompatible assumptions (§5); and that *R&H* ignores a key ingredient of *APS*, namely the role of performance in accounting for cross-linguistic generalizations (§6). Section 7 is a brief conclusion.

2. A failed attempt to motivate parametric theory²

What does it take to motivate a parameter of UG? To read *R&H*, one would think that a sufficient criterion is the identification of a cluster of properties shared by a set of closely related languages, but not shared by a set of slightly less closely related ones. The one extended argument for parametric theory in *R&H* (pp. 9–14) involves a set of systematic differences between the Insular Scandinavian (ISC) and the Mainland Scandinavian languages (MSC). The following properties distinguish the two:

- (1) a. ISC but not MSC allow null nonreferential subjects.
- b. ISC but not MSC have non-nominative subjects.

¹ Page number references to *R&H* are to the manuscript version, not to the published version, which I have not had the opportunity to see.

² I would like to thank the following individuals for their helpful discussion of the typological issues taken on in this section: Michael Cysouw, Matthew Dryer, Martin Haspelmath, Anna Siewierska, and Halldor Sigurdsson.

- c. ISC but not MSC have stylistic fronting.
- d. ISC but not MSC are V2 in embedded clauses.
- e. ISC but not MSC have relatively rich subject verb agreement.

They claim that the differences fall out from the following ‘parameter’:

- (2) ISC has a nominal Agr in I, spelled out as subject-verb agreement, MSC doesn’t.

Let’s say for purposes of argument that (2) is sufficient to characterize the differences in (1a-e). That in and of itself is the weakest sort of evidence imaginable for the existence of a parameter and furthermore it has no bearing whatsoever on the adequacy of parametric theory vis-à-vis rule-based theory. A rule-based theory would simply posit a rule for ISC specifying that it has a nominal Agr in I, spelled out as subject-verb agreement, while MSC would lack such a rule.³ For the facts in (1a-e) to support parametric theory, it would be necessary to demonstrate that in *language after language* the constellation of properties particular to ISC and MSC reappear. But aside from some brief remarks about Middle English, Yiddish, and Old French, two of which, like Scandinavian, are Germanic, and the third of which was heavily influenced by Germanic, they ignore the typological dimension entirely.

All of the available evidence points to null nonreferential subjects, non-nominative subjects, stylistic fronting, V2 in embedded clauses, and relatively rich subject verb agreement as being *independent variables*, as far as their presence or absence in a language is concerned. Let us examine what research has revealed about each factor and its correlations.

The major survey of null nonreferential subjects is found in Gilligan 1987, where they are called ‘null non-thematic subjects’. Of the 102 languages in his sample, Gilligan found data on null non-thematic subjects for 41. As far as subject verb agreement is concerned, the best cross-linguistic data that we have is found in Haspelmath et al. 2005: 416-417 under the heading of ‘Verbal Person Marking’. Twenty-two languages are in both the Gilligan and the Haspelmath et al. sample, 12 of which have both null non-thematic subjects (NNTS) and subject person marking on the verb (SPMV), and 10 of which have NNTS, but not SPMV:

- (3) a. NNTS and SPMV: Basque, Chamorro, Finnish, Fula, Georgian, Hua, Icelandic, Italian, Mundari, Swahili, Turkana, Turkish
- b. NNTS, but not SPMV: Babungo, Burmese, Caviteño, Garo, Hausa, Mandarin, Tagalog, Thai, Vietnamese, Yoruba

In other words, the available data do not support a correlation between NNTS and SPMV.⁴ But things are much worse than that for R&H. Their parameter

³ APS states explicitly that ‘Essentially, [language-particular rules] are parameter-settings “detached” from the parameters themselves (which are hypothesized not to exist)’ (p. 183). So I am mystified why R&H would assume that the handling of the data in (1a-e) in a rule-based theory would require the complex machinery that they attribute to it (pp. 12-13).

⁴ Matthew Dryer has pointed out (p.c.) that some languages lack agreement affixes on the verb, but still have agreement in that they have pronouns that are separate words, but which are

unites NNTS with *relatively rich* subject verb agreement. Since languages with relatively rich agreement are (obviously) a proper subset of those with agreement, it seems reasonable to assume that the agreement manifested by some of the languages in (3a) is not ‘relatively rich’, resulting in an even weaker cross-linguistic correlation between (1a) and (1e).

As far as V2 is concerned, even in main clauses it is one of the rarest typological features in existence. V2 is attested only in Germanic, in languages in contact with Germanic (Sorbian, Old French and some other early Romance languages, Rhaeto-Romance, Breton, and Middle Welsh), and in Kashmiri (Raritätenkabinett 2003: 79). V2 in embedded clauses is rarer still. One would hardly wish the latter to fall out from the same parameter, if one exists, that yields (1a,b,c,e), all of which are rather common cross-linguistically.

There is no extensive typological survey, to my knowledge, of non-nominative subjects.⁵ However, in addition to ISC, they have been reported in Caucasian (Tsez, Lezgian), Arawak (Bare, Warekina, Baniwa), Japanese, Dravidian (Kannada, Tamil, Malayalam), Indo-Aryan (Konkani, Hindi-Urdu, Nepali, Kashmiri, Bangla), Tibeto-Burman (Newar), Australian (Iwaidjan), Papuan (Motuna), and Korean (see Aikhenvald et al. 2001; Bhaskararao and Subbarao 2004). Given the sparsity of cross-linguistic data, conclusions are necessarily tentative, but I see no reason to believe that non-nominative subjects are licensed by the same parameter, if one exists, that licenses (1a, c-e). For example, Indo-Aryan and Dravidian languages have rich subject-verb agreement, while Japanese, Korean, and Lezgian lack such agreement (Haspelmath et al. 2005: 416-417). For cross-linguistic correlations between non-nominative subjects and null nonreferential subjects, I have found no data at all.

Finally, no typological survey exists, as far as I know, of stylistic fronting. R&H attempts to deflect attention from the failed predictions of its parametric approach by claiming that:

[t]he theory does not predict that the ISC properties will always cluster. It allows for the possibility that a language has the positive setting for [(2)], yet does not, have, say, oblique subjects. This is so if allowing oblique subjects depends on other parameters in addition to [(2)]; to start with, the language should have non-structural case. Stylistic fronting, too, may depend on other parameters than [(2)]. Parametric theory predicts that a cluster dependent on a given parameter setting will reappear in language after language, all else being equal. But all else is seldom equal. The best chance of observing cross-linguistic correlations is in closely related languages or dialects, but even then other parameter settings may complicate the picture, masking the effects of a parameter setting. (12-13)

In the absence of even a modest attempt to characterize these ‘other parameters’ and in the absence of even the first steps to demonstrate a robust correlation

obligatory even when there is a separate nominal subject. One assumes that most generative linguists would consider them instances of Agr. However, Hausa is the only language under ‘NNTS, but not SPMV’ that has agreement via a separate word.

⁵ According to Martin Haspelmath (p.c.), ‘non-nominative subject’ is a cover term for a number of quite disparate phenomena. R&H would need to demonstrate that such subjects of the type manifested by ISC correlate cross-linguistically with (1a, c-e).

among (1a-e) based on a reasonably-sized sample of the world's languages, their caveat amounts to no more than hand-waving. In fact, 'observing cross-linguistic correlations in closely related languages or dialects' is the absolutely *worst* methodology for pinning down parameters and their settings, since closely related languages and dialects typically remain in contact with each other over time, thereby introducing a complicating variable not found between two unrelated speech varieties spoken at a great distance from each other. In particular, speakers of ISC have not only kept up contact with speakers of MSC over the centuries, but a significant percentage of speakers of the former have been (and are) bilingual, speaking some variety of the latter. No advocate of parametric theory has, to my knowledge, attempted to control for the effects of contact and bilingualism. A focus on a large number of unrelated languages as opposed to a small number of related ones is also necessary to properly sort out parametric change from less systematic nonparametric change (for discussion of the two types, see Lightfoot 1991, Lightfoot 1999). For example, there are literally dozens of syntactic differences between American and British English. How might one know which of these differences are parametric and which idiosyncratic without extensive cross-linguistic investigation?

R&H close their section on the differences between ISC and MSC by expressing the 'doubt that any such [performance] principles can explain the correlations in [(1a-e)]' (14). But since they have not demonstrated the robustness of these correlations cross-linguistically (nor do they even seem aware that such a demonstration is necessary), there is no reason to think that there anything for a performance-based theory, or any other type of theory, to explain. The treatment of ISC and MSC in R&H is a textbook example of the nineteenth century Eurocentrism that still pervades linguistics. One observes a constellation of features pertaining to some subfamily of Indo-European (typically Germanic or Romance) and, in the absence of evidence, one draws conclusions about human language as a whole. Let us hope that as the twenty-first century progresses, practice of that sort will become increasingly rare.

3. R&H's concessions to *APS*

The first critique of parametric theory in *APS* — and, by far the most important — is that 'there is [no] gain in descriptive simplicity with parameters' vis-à-vis a rule-based theory (189) and that in crucial cases "'parameter" has simply become a synonym for "rule"' (191). Astonishingly, R&H agrees, noting correctly that 'Newmeyer's theory includes rules which are equivalent to parameter settings' (2), that 'a parameter can always be reformulated as a set of competing rules' (3), and hence that 'the model sketched by Newmeyer is not formally simpler than standard principles-and-parameters theory' (3). Precisely my point! Any parameter-setting can be trivially reformulated as a rule, but given that parametric theory needs to characterize properties of rules anyway (a point not contested by R&H), 'on the basis of Occam's razor, one would be forced to renounce the idea of any a priori desirability of a parametric theory' (191).

APS goes on to note that 'parameters do not clearly have binary settings, and in that respect parameters do not differ from rules' (193). R&H devotes a few paragraphs to defending binary parameter settings, but then implies that thinking of settings as binary may be 'useful', but is not really very important,

since 'the question of binarity is arguably more a matter of formulation than anything else. ... The only really substantive claim behind a binary formulation of parameters is that the values are discrete: there are no clines, squishes or continua' (4). But *APS* never challenges the discreteness of rules or of any other construct of grammatical theory (for a defense of discreteness see Newmeyer 2003; Newmeyer 2005), so once again R&H, in effect, concedes the essence of the section to *APS*.

To make a methodological point, any n -ary contrast ($n > 2$) can trivially be decomposed into a larger set of binary contrasts. To the only question is, for any particular case, how motivated such a decomposition might be. In a parametric theory, some oppositions are clearly binary and some are not; in a rule-based theory, some difference between pairs of rules are clearly stateable in binary terms and some are not. In this respect, the two approaches seem quite similar.

4. Principled versus actual benefits of parametric theory

The original vision of parametric theory was a marvelous one — from the interaction of a small number of simply-formulated parameters, the observed complexity of human language grammar would fall out as a matter of course. Unfortunately, things have not worked out that way. The past quarter century has seen the number of proposed parameters increase exponentially, a problem that has only been aggravated by the move to 'microparameters'. As I reported in *APS*, Lightfoot 1999: 259 has written that 'a single issue of *Linguistic Inquiry* may contain 30-40 proposed parameters' (193). If fewer than 1000 parameters can be found in the literature, I would be very surprised. The fact, noted by R&H, that 'only twenty independent binary parameters are need to produce [millions of possible grammatical systems]' (4) is utterly irrelevant to the issues that divide *APS* and R&H. If Roberts and Holmberg really believe that 50-100 is 'a plausible, if conservative, conjecture' as to the number of parameters, they should have no trouble spending a few hours drafting a provisional list of them and demonstrating that the parameters on the list derive the totality of attested grammatical systems. The fact that neither they nor any other advocate of parametric theory has ever done that — or, I predict ever will do that — is indicative of the lack of seriousness with which the question of parameters has been approached in the literature. After 25 years we need more than promissory notes and bombast. We need the results that have never been forthcoming. It is simply not enough to write 'Language A differs from Language B with respect to Parameter P' and think that an interesting empirical claim has been made. Of course it is true that no theory in the process of development needs to be held accountable for all of the data arguably within its domain. But it is an ABC of scientific investigation that if a theory is on the right track, then its overall complexity decreases with time as more and more problematic data fall within its scope. Just the opposite has happened with parametric theory. Year after year more new parameters are proposed, with no compensatory decrease in the number of previously proposed ones. At the same time, there is more confusion

than ever about which phenomena should be excluded from the domain of parameters altogether by being consigned to the P-syntax.⁶

By the way, it is simply false that ‘parameters allow us to collapse cross-linguistic differences into single abstract properties of grammars, while language-specific rules are just that: language-specific’ (4). It has been known since the earliest days of transformational grammar that rules are both abstract and often shared by more than one language (just as parameter (2) is probably best interpreted as a rule shared by the ISC languages).

Throughout their piece, Roberts and Holmberg write as if a parametric approach is somehow the ‘default’, with rules needing special motivation. Just the opposite is the case. Rules are historically antecedent to parameters and continue to be proposed side-by-side with parameters (see the discussion of Hixkaryana in *APS*, p. 190). Since parameters add an extra mechanism to the theory, without dispensing of mechanisms already in currency, it is they that need special motivation. In particular, a convincing argument is needed that parameter-settings both differ empirically from rules and lead to an overall more elegant theory. That argument is not to be found in R&H.

5. The mutually inconsistent assumptions in R&H

R&H defends Mark Baker’s Parameter Hierarchy (PH) (Baker 2001) from the critique in *APS*, which ‘points out one or two empirical difficulties with it’ (5). By my count, *APS* pointed out ten empirical difficulties, but no matter. The PH is formulated in purely GB terms, that is, in terms of a theory that was abandoned by virtually all workers in the principles-and-parameters orbit over a decade ago and, crucially, by Roberts and Holmberg themselves, who have adopted the MP. This is not just a ‘nice’ point. The PH is literally unformulable in the MP, for the simple reason that parameters in that framework are no longer considered (as they were in GB) to hold for entire languages. So in the MP, a language is not necessarily null subject nor non-null subject. Rather, the inventory of functional categories and the properties of the features associated with those categories can be chosen to allow a language to be null subject in some respects, but not in others. The same point can be made with respect to other GB-style parameters. Such a move (which, as noted in *APS*, is a move toward a rule-based theory) results in a more flexible — and hence more empirically adequate — theory, but at the same time renders obsolete the PH. Reading R&H, one is left wondering precisely what theory it assumes. Throughout their piece, Roberts and Holmberg skip inconsistently between GB and the MP. For example, they endorse ‘Baker’s PH as a precursor to a periodic table for languages’ which they feel ‘is an optimistic view but one worth striving to maintain’ (14). Yet at the same time, they adopt the utterly incompatible view that ‘parameters are associated with lexically encoded properties of functional heads’ (9).

⁶ It is worth pointing out that few principal players in minimalist syntax seem to share Roberts’ and Holmberg’s passion for parametric theory. For example, discussion of parameters is all but absent in Chomsky 2002 and Lasnik 1999 and the (arguably) major textbook on minimalism, Adger 2003, devotes less than a page to them. The indexes of two anthologies of minimalist work that I pulled off my shelf at random, Martin et al. 2000 and Hendrick 2003, point to one page each on parameters.

6. The role of performance in accounting for cross-linguistic generalizations

A central aspect of *APS* is the shift in accounting for many, if not most, typological generalizations from competence to performance. Ten pages of the article are devoted to arguing that the parsing-based theory of Hawkins 2004 accounts neatly for many of the generalizations that, in the past, had been attributed to differences in parameter settings. Yet this aspect of *APS* is all but forgotten by R&H, which alludes to performance explanations only briefly in its final page. By downplaying such explanations, it ends up magnifying the ‘language-particularity’ of the theory argued for in *APS*. For example, R&H writes:

So Newmeyer’s example neatly illustrates how a reliance on language specific rules will lead one to miss significant cross-linguistic generalizations. A parameter-based approach, on the other hand, will naturally lead one to look for similarities and differences in other systems.
(6)

But that would be the case only if the sole constructs provided by linguistic theory, broadly defined, were rules. However, performance principles lead one as well ‘to look for similarities and differences in other systems’. The point can be appreciated by a look at the facts that led to the above quote. Following Cinque 1999, R&H notes that there is a robust, but not exceptionless, cross-linguistic generalization regarding the order of tense, mood, and aspect (TMA) markers. R&H takes this fact as support for parametric theory, but that is not necessarily the case. Typologists have been arguing for over twenty years that the ordering of TMA morphemes has an external basis (see Foley and Van Valin 1984: 208f., Bybee 1985a, Bybee 1985b, and Hengeveld 1989). In brief, functional pressure to maintain an ‘iconic’ relationship between form and meaning predicts that the more ‘relevant’ an inflection is to a head, the closer it is likely to be to that head and the more likely it is to be bound to that head (as opposed to occurring as an independent word) and the more distant the element is from the head, the wider its scope.

So *now* contrast the handling of TMA markers in a parametric theory and in a rule-based theory. In the former, unmarked parameter settings yield the most common ordering of TMA markers, while marked parameter settings, ancillary parameters, or (possibly) language-particular rules derive the ordering for those languages deviating from the norm. In the latter, UG provides rules whose interaction allows for the universally possible orderings of TMA markers, and performance principles explain why some of these rules are called upon more than others. Does the parametric theory ‘win’ this contest? I see no reason to think so.

7. Conclusion

It is more than a little ironic that Roberts and Holmberg would begin and end their reply to *APS* by dismissing a rule-based theory as being merely ‘observationally adequate’. That is precisely the charge that has been leveled in

Bouchard 2003 (which is cited in *APS*, but ignored by R&H) against parametric theory. As Bouchard notes, if every difference, or even if most differences, distinguishing one language from another are attributed to a difference in 'parameter setting', then saying that two languages have different parameter settings simply becomes an obfuscatory way of saying that they have different rules. After a quarter-century of its well-documented failures and retreats, one is forced to conclude that the parametric program, and the defense of it in R&H, is little more than an exercise in wishful thinking.

REFERENCES

- Adger, David. 2003. *Core Syntax: A minimalist approach*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra, R. M. W. Dixon, and Masayuki Onishi eds. 2001. *Non-canonical marking of subjects and objects* Amsterdam: John Benjamins.
- Baker, Mark C. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.
- Bhaskararao, Peri and K. V. Subbarao eds. 2004. *Non-nominative subjects*. *Typological Studies in Language* vol. 60 and 61. Amsterdam: John Benjamins.
- Bouchard, Denis. 2003. The origins of language variation. *Linguistic Variation Yearbook* 3:1-41.
- Bybee, Joan L. 1985a. Diagrammatic iconicity in stem-inflection relations. In *Iconicity in syntax*, ed. by John Haiman, 11-48. Amsterdam: John Benjamins.
- Bybee, Joan L. 1985b. *Morphology: A study of the relation between meaning and form*: *Typological Studies in Language*, vol. 9. Amsterdam: John Benjamins.
- Chomsky, Noam. 2002. *On nature and language*. Cambridge: Cambridge University Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Foley, William A. and Robert D. Van Valin. 1984. *Functional syntax and universal grammar*: *Cambridge Studies in Linguistics*, vol. 38. Cambridge: Cambridge University Press.
- Gilligan, Gary M. 1987. A cross-linguistic approach to the pro-drop parameter. Unpublished Ph. D. dissertation, University of Southern California.
- Haspelmath, Martin, Matthew S. Dryer, David Gil, and Bernard Comrie eds. 2005. *World atlas of linguistic structures* Oxford: Oxford University Press.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hendrick, Randall ed. 2003. *Minimalist syntax* Oxford: Blackwell.
- Hengeveld, Kees. 1989. Layers and operators in functional grammar. *Journal of Linguistics* 25:127-158.
- Lasnik, Howard. 1999. *Minimalist analysis*. Oxford: Blackwell.
- Lightfoot, David W. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
- Lightfoot, David W. 1999. *The development of language: Acquisition, change, and evolution*: *Blackwell/Maryland Lectures in Language and Cognition*, vol. 1. Oxford: Blackwell.

- Martin, Roger, David Michaels, and Juan Uriagereka. 2000. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press.
- Newmeyer, Frederick J. 2004. Against a parameter-setting approach to language variation. *Linguistic Variation Yearbook* 4:181-234.
- Newmeyer, Frederick J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79:682-707.
- Raritätenkabinett. 2003. <http://ling.uni-konstanz.de:591/universals/searchrara.html>.
- Roberts, Ian and Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In *Organizing grammar: Linguistic Studies in honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, and Jan Koster. Berlin: Mouton de Gruyter.