The logic of parametric theories·

Terje Lohndal & Juan Uriagereka

University of Maryland

'Parameters in minimalist theory: The case of Scandinavian' (henceforth, PMT), by Anders Holmberg, is an interesting attempt to defend the view that 'deep parameters' should be part of the grammatical system. PMT argues that parameters should be seen as points of 'underspecification', and in part as 3rd factor effects in the sense of Chomsky (2005) – thus casting doubts on attempts to locate parametric variation in the lexicon and at the interfaces. In our view, however, PMT does not quite make the case for a 'deep parameter'. In examining this case, we also seek to clarify the relationship between parametric theories and acquisition, and more generally 'underspecification'.

1. 'Deep' parameters

Research in the 1980s took grammatical properties to swing in tandem. For instance, Rizzi (1982) attempted to build correlations between thematic null subjects, null expletives, free inversion and *that*-trace effects. A number of authors (e.g. Newmeyer (2005), Haspelmath (2008), Hornstein (2009) and Boeckx (in press)) have seen reason to retract from this position. PMT argues, however, that a pessimistic conclusion is premature, and that it is still possible to assume parameters in a deep sense within minimalism (cf. also Richards 2008 and Roberts and Holmberg 2010). PMT takes a 'deep' parameter to be one that has a 'range of surface effects' (p. 5). Strictly, this is not

· We are grateful to Marc Richards and Rozz Thornton for comments on this reply, and to Hans-Martin Gärtner for giving us the opportunity to react to Holmberg's topical piece.

quite the same sense in which parameters were taken to be 'deep' in the 1980's. With few exceptions (e.g. ideas that Richard Kayne attributed to Jean-Pierre Changeux, as discussed in Uriagereka 1994), classical parameters were seen as part of Universal Grammar – the origin of James Higginbotham famous 'switch-board' metaphor. In contrast, PMT takes the features that govern variation not to be 'dictated by UG' (p. 20). So although UG is taken to provide 'the mechanism […] for […] dependency relations', namely feature valuation (p. 20), relevant features to be valued vary from language to language.[1]

PMT shows how Insular Scandinavian languages (ISc) have unvalued number [uNr] and person [uPn], whereas the Mainland Scandinavian languages (MSc) do not. This difference is related to the agreement differences between ISc and MSc. Based on a number of assumptions, PMT accounts for the factual base that it sets to explain, although the piece does honestly admit in its conclusions that the relevant parameter 'in the big scheme of things […] is more of a microparameter than a macroparameter' (p. 40). In more conceptual terms, then, PMT attempts to build the case that 'deep parameters' can be contemplated within the Minimalist Program. From a minimalist perspective, however, it is actually not easy to decide whether, first of all factually, a single 'deep' property does have a range of surface effects, and if it does, second, whether this should be seen as part of narrow syntax (Chomsky 1995, 2001, Sigurðsson 2004). The matter does not seem to us to be resolved with PMT's case study.

All the properties that PMT discusses are salient, thus easy-to-observe by a child acquiring the language (e.g. rich subject-verb agreement, oblique subjects, stylistic

[1] This brings to mind the proposals in Fukui (2006), where differences between English and Japanese are attributed to the presence/absence of certain features, or Raposo and Uriagereka (2005), where differences in clitic placement are analyzed along similar lines.

fronting, null expletives, null generic subject pronouns, transitive expletives or heavy subject postposing). PMT actually manages to involve just two features to entail all these differences. But precisely the elegance of the analysis raises the question of whether this counts as resorting to a 'deep' parameter. The converse of easy-to-observe properties is hard-to-observe properties, meaning those that cannot be realistically abstracted away from the input. Consider for instance *that*-trace effects, illustrated with Standard English:

(1) a. Who do you think that John saw *t*?

b. Who do you think (*that) *t* left?

The problem is how a child comes to realize that (1b) is bad, under standard assumptions about negative data (see Lohndal (2009) for discussion). What made Rizzi's case interesting in 1982 was the fact that a child may have observed something about agreement (or whatever else allows her to easily detect subject-drop), but *that*-trace effects in embedded clauses are not part of a realistic data base for a learner to have. And yet, given some assumptions about UG, in the classical view of things the child is literally 'cornered' into the surprising co-occurrence of *that*-trace violations, perhaps after a period of hesitation, given the elaborate deduction.[2]

The variation in null generic subjects that PMT discusses, for instance, seems quite different in nature. The relevant examples are repeated here in (2) (PMT's (11)).

(2) a. Hér má ekki dansa. [Icelandic]

[2] Thornton (1990) shows how children struggle to acquire this property, and so in particular English children violate the *that*-trace filter at an early age.

here may not dance

'One must not dance here.'

b. Her må *(man) ikke danse. [Norwegian]

here may one not dance

Simply put, even if the presence of null generic subjects happened not to be part of the cluster of effects argued for in PMT, the variation in (2) would not raise any particular learnability problems. Silent subjects in SpecTP are never possible in MSc, so the learner will not develop a grammar that licenses null subjects of any kind. SpecTP in ISc seems to be different, at least on the assumption that stylistic fronting involves movement of an XP to SpecTP, in which case the elements that can occupy SpecTP in ISc are a superset of the elements that can occupy SpecTP in MSc. Presumably the primary linguistic data will also show Icelandic children that null generic pronouns are possible. Even if PMT's analysis happened not to be correct (which we are not claiming), the variation in (2) would be less serious a learning problem than acquiring *that*-trace effects.[3]

As it turns out, the correlation between *that*-trace effects and, e.g., null subjects or the difference between preverbal and postverbal extraction sites (Kenstowicz 1989) in the end fails, as Newmeyer (2005) and others have discussed. Unfortunately, one may say, because this was a case where we *could* have made good use of parameters in solving 'Plato's problem'. PMT's plausible correlation for the data set it deals with would require a further ancillary phenomenon as just discussed to constitute a defensible instance of a 'deep parameter'. This may well exist, but the present piece does not show it.

[3] Similar situations seem to us to arise in quantifier-scope readings of the sort described in Goro (2007), or variation in islands of the sort discussed in Yoshida (2006), among others.

Holmberg is of course aware that we are dealing with poverty-of-the-stimulus issues, and thus casts the discussion in terms of 'the logical problem of language acquisition' (p. 6). However, his piece also notes how 'purported absence of clusters of properties means that parameters, if they exist at all, are relatively superficial, with a limited number of effects' (p. 6). Now to the extent that this happens to be true (we do not know), parameters will simply be less interesting, and in the end less relevant to core linguistic theorizing. Note, in passing, that this is not quite the conclusion reached in the *locus classicus* of structural variation within the Principles and Parameters approach, namely Borer (1984). Although Borer conjectured that variation may be a property of the lexicon, given the sort of rich deductive system assumed at the time, minute differences in lexical specifications (such as the possibility of verbal inflection correlating with null subjects) could have quite drastic effects in the shape of particular grammars, as discussed above. The beauty in that sort of approach was not so much whether a language did or did not have, say, clitics. What was extraordinary was to be able to predict how a child whose grammar has clitics *knows where to place them*, in matrix or embedded clauses (see for instance Uriagereka 1988, 1995 for that sort of approach).

Again, perhaps deep parametric conditions simply do not exist, as a matter of fact, or at least we have not been able to detect them yet. But certainly reducing their range to properties that can be easily observed, in and of themselves, does not have the same theoretical clout.

2. Constraining variation

PMT actually claims that 'wherever the variation is located, it must be constrained so as

to allow only a limited number of options, or else we have not addressed the logical problem of language acquisition' (p. 11). However, we must again insist in differentiating easy-to-observe from hard-to-observe properties, from a learner's perspective. If we are only looking at easy-to-observe properties, the argument does not go through, at least not on strict learnability grounds. This is worth emphasizing from a different angle.

One can often see reference in the literature to the expression 'a theory of parameters', a locution PMT also employs (p. 3). From a learnability perspective that description actually makes relatively little sense. One can, of course, impose constraints on what a theory *of learning* should be (on negative data, direct or indirect, on degrees of learnability, on the instantaneous or developmental nature of the acquisition, etc.). But, from that point of view, imposing constraints on the parameters one should encounter would be like imposing constraints on a putative 'theory of principles' to reckon with. There are no 'theories of principles' because, in empirical science, we are in the business of *postulating* such principles, and as common practice tells us all too well, these are moving targets, as understanding grows. So all that one can require from this acquisition approach is that the observables to set putative parameters should be, well, observable – in terms that are feasible for a human learner. There is, then, no a priori sense in which variation 'must be constrained', for even if we have an *unlimited* 'number of options' these will pose no logical acquisition problem if they can be easily observed. The set of idiosyncracies in a lexicon, for instance, can be acquired that way, whether or not it has much internal structure, possibly full of unconstrained variation.

Now granted: one doesn't have unlimited time and funds to proceed with science, so one makes wagers on what *could* be the shape of conditions to encounter, whether they

be principles or parameters. The Minimalist Program is one such wager, for better and for worse. One may, then, conjecture that, in fact, variation happens to be both systematic and constrained, so as to allow only a limited number of options. But unless among the options there are hard-to-observe conditions, this will have nothing to do with language acquisition. PMT correctly argues that it does not matter *where* variation is located within the system, if it exists (fn. 7); however, it does matter *what kind* of variation we are dealing with. In an interesting architectural way, it is not the same to say that variation arises because of learning considerations than to say that it arises because of design specifications in the design of grammar – or for that matter for superficial reasons that do not affect core grammar in the least. These are all very different ontological claims.

3. Underspecification

The most interesting idea in PMT is not developed beyond the programmatic statement that 'the right way to think of parameters is in terms of underdetermination by UG' (p. 3), and that 'UG provides for the possibility, but does not dictate, that a given grammatical feature has an unvalued counterpart, and also does not dictate the distribution of the unvalued feature […]' (p. 36). Underspecification in this sense has not been a prominent alternative in linguistics. After Richard Kayne played with it two decades ago, the idea was revived in Biberauer and Richards (2006), Richards (2008), Roberts and Holmberg (2010) and Berwick and Chomsky (in press). For Richards (2008: 153) "an underspecified UG relying on third-factor principles yields points of indeterminacy where the system 'no longer cares' owing to a loss of information or lack of specification."

Back in 1981, Jean-Pierre Changeux was saying that:

> As a compromise between these [the innate and the empiricist] attitudes, we have postulated ("selective stabilization hypothesis") that the genetic program directs the proper interaction between main categories of neurons […]. However, during development within a given category, several contacts with the same specificity may form; in other words, a significant but limited "redundancy" of the connectivity exists. The early activity of the circuits, spontaneous (in the embryo) or evoked (after birth), would increase the specificity or the organization of the system by reducing this transient redundancy (Changeux 1981: 193).

We can easily change the terms in this quotation and talk about principles and parameters (see also Vercelli and Piattelli-Palmarini 2009). The core idea here is that even if we do not encode the range of variation as part of UG, UG may still determine variation through its interaction with the environment, viz. the input. This is underspecification, which stands in radical opposition to overdetermination, as e.g., in Yang (2002). It also stands in opposition to the idea that all languages have the same universal set of features as part of UG (Sigurðsson 2004, Richards 2008, Boeckx in press).

It is important to emphasize the significance of this idea from a minimalist perspective: Even if we want narrow syntax to be entirely uniform, this does not preclude that the genetic component, aka. UG, can still yield variation in the sense described by Changeux (1981). We are not attempting to defend this particular view here, as opposed to the overspecification approach (which we also believe has serious merits and a definite

place within current biology)[4]. All we would like to see is a thorough discussion of these different perspectives, to determine whether one trumps the other or they are both part of the system, actually a possibility as well.

4. Conclusion

PMT is an important contribution to the present debate on the nature of parameters, in that it shows that minimalism is also able to capture correlations by looking closely at how various features come to play. However, we have argued that the approach does not make some necessary distinctions in order to clarify what the acquisition problem is, and whether further dimensions exist to the problem of variation. We have also attempted to contextualize PMT's appeal to underspecification and argued that it is an important research program, which connects current linguistic theorizing to contemporary biology.

References

Berwick, Robert C. and Noam Chomsky. In press. The Biolinguistic Program: The Current State of its Evolution and Development. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, Anna-Maria Di Sciullo and Cedric Boeckx (eds.). Oxford: Oxford University Press.

Biberauer, Theresa and Marc Richards. 2006. True Optionality: When the grammar

[4] Changeux's sort of perspective has crystallized into the Evo-Devo paradigm in biology (see for instance Carroll 2005). In turn, a remarkable version of the overspecification approach is offered in Sherman (2007), in terms of his concept of a 'universal genome' for all metazoan life.

doesn't mind. *Minimalist Essays*, Cedric Boeckx (ed.), 35-67. Amsterdam: John Benjamins.

Boeckx, Cedric. In press. Approaching Parameters from Below. *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty*, Anna-Maria Di Sciullo and Cedric Boeckx (eds.). Oxford: Oxford University Press.

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

Carroll, Sean B. 2005. *Forms Most Beautiful: The New Science of Evo-Devo*. New York: Norton.

Changeux, Jean-Pierre. 1981. Genetic Determinisim and Epigenesis of the Neuronal Network: Is There a Biological Compromise between Chomsky and Piaget? *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, Massimo Pattelli-Palmarini (ed.), 185-197. Cambridge, MA: Harvard University Press.

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

Chomsky, Noam. 2001. Derivation by Phase. *Ken Hale: A Life in Language*, Michael Kenstowicz (ed.), 1-52. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36: 1-22.

Fukui, Naoki. 2006. *Theoretical Comparative Syntax: Studies in Macroparameters*. London: Routledge.

Goro, Takuya. 2007. Language-Specific Constraints on Scope Interpretation in First Language Acquisition. Doctoral dissertation, University of Maryland.

Haspelmath, Martin. 2008. Parametric versus functional explanations of syntactic

universals. *The Limits of Syntactic Variation*, ed. by Theresa Biberauer, 75-107. Amsterdam: John Benjamins.

Hornstein, Norbert. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.

Kenstowicz, Michael. 1989. The null subject parameter in Modern Arabic dialects. *The Null Subject Parameter*, ed. by Osvaldo Jaeggli and Kenneth Safir, 263-275. Dordrecht: Kluwer.

Lohndal, Terje. 2009. Comp-t Effects: Variation in the Position and Features of C. *Studia Linguistica* 63: 204-232.

Newmeyer, Frederick J. 2005. *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Raposo, Eduardo P. and Juan Uriagereka. 2005. Clitic Placement in Western Iberian: A Minimalist View. *The Oxford Handbook of Comparative Syntax*, Guglielmo Cinque and Richard S. Kayne (eds.), 639-697. Oxford: Oxford University Press.

Richards, Marc. 2008. Two Kinds of Variation in a Minimalist System. *Varieties of Competition*, Fabrian Heck, Gereon Müller and Jochen Trommer (eds.), 133-162. Linguistische Arbeits Berichte 87, Universität Leipzig.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Roberts, Ian and Anders Holmberg. 2010. Introduction: parameters in minimalist theory. *Null Subjects: the structure of parametric variation*, Theresa Biberauer, Anders Holmberg, Ian Roberts and Michelle Sheehan (eds.), 1-57. Cambridge: Cambridge University Press.

Sherman, Michael. 2007. Universal Genome in the Origin of Metazoa: Thoughts About Evolution. *Cell Cycle* 6: 1873-1877.

Sigurðsson, Hálldor Á. 2004. Meaningful silence, meaningless sounds. *Linguistic Variation Yearbook* 4, 235-259.

Thornton, Rosalind. 1990. Adventures in Long-distance Moving: The Acquisition of Complex *Wh*-questions. Doctoral dissertation, University of Connecticut.

Uriagereka, Juan. 1988. On Government. Doctoral dissertation, University of Connecticut

Uriagereka, Juan. 1994. A Conjecture on the Form of Parameters. *UConn Papers in Linguistics* 4: 247-270.

Uriagereka, Juan. 1995. Aspects of Clitic Placement in Western Romance. *Linguistic Inquiry* 26: 79-123.

Yang, Charles. 2002. *Knowledge and Learning in Natural Languages*. New York: Oxford University Press.

Yoshida, Masaya. 2006. Mechanisms in Long-Distance Dependency Formation. Doctoral dissertation, University of Maryland.

Vercelli, Donata and Massimo Piattelli-Palmarini. 2009. Language in an Epigenetic Framework. *Of Minds and Language: A dialogue with Noam Chomsky in the Basque Country*, Massimo Piattelli-Palmarini, Juan Uriagereka and Pello Salaburu (eds.), 97-107. Oxford: Oxford University Press.