

Basquing in Minimalism

Piattelli-Palmarini, Massimo, Juan Uriagereka & Pello Salaburu (eds.).
2010. *Of Minds & Language: A Dialogue with Noam Chomsky in the Basque Country*. Oxford: Oxford University Press, 472pp.

by Alex Drummond & Norbert Hornstein

In 1975, in an old abbey in Royaumont, France, Massimo Piattelli-Palmarini gathered linguists, philosophers, psychologists and computationalists for the first biolinguistics get-together to discuss what the linguistics of the period told us about the structure of the mind/brain. The participants defended two contrasting viewpoints. One group, centered on Piaget's work, claimed that language mastery supervened on very general laws of cognitive development, language acquisition and use reflecting mental powers continuous with those witnessed in other areas of cognition. A second cohort, with Chomsky's work as cynosure, countered that the specificity of linguistic competence argued against any general psychological processes that encompassed both linguistic knowledge and other forms of cognition. It is fair to say that Chomsky's side got the better of the argument. The main problem for the Piagetians was explaining in non-metaphoric detail how their proposed general cognitive and developmental mechanisms could result in the particular kinds of phenomena linguists had discovered (e.g., how does the structure dependence of grammatical operations follow from Piaget's constructivism?). Chomsky's challenge was simple: Deduce the "laws" of grammar from the more general laws of psychology / development, or concede that the mind/brain contains very specific linguistically dedicated mental structures that guide the emergence of linguistic competence and performance.

The Royaumont volume (Piattelli-Palmarini 1980) serves as an interesting backdrop to the one under review here. Once again, **Massimo Piattelli-Palmarini** serves as impresario (this time in concert with **Juan Uriagereka** and **Pello Salaburu**) and convenes a group of distinguished linguists, psychologists, biologists, and neuroscientists in a very scenic spot (San Sebastian in Spain's Basque Country) to discuss what current linguistic theory suggests about the structure of the mind/brain. This time round, however, the relevant linguistic perspective is the one offered by the Minimalist Program (Chomsky 1995 *et seq.*). Consequently, much of the discussion aims to minimize the degree to which language mastery requires specifically linguistic mental powers and emphasizes the continuation between linguistic competence and other cognitive capacities.

It is tempting to conclude from this that, although the Piagetians lost the earlier battle, they decisively won the peace by converting their antagonists.



Tempting but inaccurate. At Royaumont, Chomsky & Co. had no principled objections to Piaget's conclusions. Rather they argued that the Piagetian claim was too thin to be of interest for it failed to show *how* the postulated general cognitive principles could explain the specific properties of the faculty of language.¹ It is this rather large hole that the Minimalist Program aspires to plug, and unless it does so, minimalist aspirations will prove to be no harder than the Piagetian ones that Chomsky & Co. so successfully routed at Royaumont. In what follows, we discuss both how to specify the gap that needs filling and how minimalism is going about filling it.

Like the earlier Royaumont volume, the papers here are uniformly thoughtful and provide an excellent guide to some of the best thinking on biolinguistic themes. The discussion sections at the end of each paper are particularly intriguing (and invariably amusing). They give the reader the feel of being part of the festivities. It is clear that the participants had a really good time and this makes for a very good read. We recommend that your reading be accompanied by snacks — preferably tapas and a good wine (a nice Rioja would suit) — to help recreate the atmosphere. In what follows we will touch on some of the themes discussed. However, a magisterial review of all the papers herein is beyond our abilities. As excuse, we borrow Chomsky's (p. 379):

I've tried to think a little about how to organize some comments. An awful lot of fascinating material has been presented here, some of which I understood, some of which I didn't. What I'll try to do is pick out some points that come to mind [...] expressing an apology in advance to everyone whose work I misrepresent.²

1. The Minimalist Program

The volume begins and ends with some useful history (Noam Chomsky's Opening Remarks in chapter 2 and Conclusion in chapter 23), part of which rationally reconstructs the generative enterprise. The generative enterprise has its origins in a 17th century question: In what way is language a window on the mind? By giving the question some contour, we can generate a modern research program as follows: (i) Specify the properties of human natural languages, (ii) figure out what kind of mind/brain could assimilate and manipulate objects with these properties, and (iii) determine which mind/brain properties are proprietary to dealing with objects like these and which are features of a more generally competent cognitive device. (i)–(iii) correspond to the three principle epochs in generative research: the early years, which investigated the kinds of rules and constructions characteristic of natural languages (recursive, structure-dependent, with both local and non-local dependencies); the middle years, when these rule systems were simplified by factoring out their general properties (A/A'-move-

¹ The same holds for claims reducing linguistic knowledge to a species of "general intelligence". It is not so much that these positions are wrong as they are of dubious utility if one's aim is to understand the attested structure of cognitive capacities. For some relevant discussion of these themes, see Rochel Gelman's discussion of core domains versus HoWs (hell-on-wheels) in chapter 15.

² Our addition: And those that we ignore or only briefly mention.

ment dependencies, local and non-local construal systems, domains of opacity for movement and construal like islands and binding domains, core notions like c-command, abstract licensing conditions such as case and ECP); and finally the current minimalist project, which tries to segregate the general properties (e.g., identified by Principles-and-Parameters theories like Government-and-Binding Theory) into those requiring specifically linguistic principles and operations and those reflecting general aspects of cognition/computation.

Chomsky's remarks sketch how the current concerns of the minimalist program "arose in a natural way from the successes of the P&P approach". They did so in two ways. First, the results of P&P theories allowed "the basic contours of an answer to the problem of explanatory adequacy" to come, at least faintly, into view and thus allowed research to "turn more seriously to the 'why' questions that transcend explanatory adequacy". Second, Chomsky remarks that P&P theories removed "the major conceptual barrier to the study of the evolution of language [...]. [By] divorc[ing] the principles of language from acquisition, it no longer follows that the format that "limits admissible hypotheses" must be rich and highly structured to satisfy the empirical conditions of language acquisition, in which case inquiry into evolution would be virtually hopeless" (p. 25).

For those who found these last remarks as cryptic as we did, here is what Chomsky meant.³ Pre-P&P syntactic theory consists of a "format for rule systems and an evaluation metric (EM) to rank them". This EM is "language internal", not a domain independent "general simplicity measure".⁴ EM seeks to specify "what [...] a legitimate generalization in language" is. This account permits a possible theory of language acquisition — "run through all possible grammars from the shortest on and stop when you have one conforming to PLD [primary linguistic data — NH & AD]. But that's unfeasible [...]. So it's not really a theory". The P&P theory is feasible as it requires "answering a finite questionnaire". Pre-P&P theories, in contrast, require the language acquisition device to make global decisions about the structure and format of entire grammars, and this, though logically coherent, is computationally unfeasible as it involves powers beyond what it is reasonable to suppose that children (or adults) have. Thus, the intractability of EM and its linguistic specificity made evolutionary considerations quite hopeless. Put positively, P&P's distillation of a set of general and useable generalizations ("laws of grammar") makes it possible to address the question of the etiology of these generalizations.

A quick digression: The idea that P&P theories should be computationally tractable because acquisition involves "answering a finite questionnaire" (i.e. setting a finite number of parameters) has proven to be very difficult to demonstrate. **Janet Fodor** (chapter 17) reviews the issues and explains why the problem has proven so refractory. The main difficulty is that the parameter values are not independent, their correct values often depending on the values of other parameters. This makes it very hard to set them correctly one at a time and this makes

³ Thanks to Noam Chomsky (p.c.) for clarifying discussion. Quotes are from correspondence with Chomsky. See also Fodor's paper in chapter 17, where she outlines and evaluates the *Aspects* theory (Chomsky 1965).

⁴ It is unlikely that non-trivial versions of this kind of simplicity measure exist. At least no successful one has been proposed to date.

incremental learning very difficult as parameters once set need not stay set. In effect, having a finite set of questions to answer does not simplify the acquisition problem if the answers require a global evaluation of the data. Fodor suggests that an adequate answer will require reconsideration of evaluation metrics with all of the (apparent⁵) language particularities of the one in *Aspects* (Chomsky 1965). The model she describes, however, has the virtue of being feasible in the sense that “neither online computation, nor memory is overtaxed” (p. 269).⁶

Back to the main thread: Chomsky’s historical reconstruction highlights two important points: First, the generative enterprise has been cumulative with results from prior periods of research conditioning later investigations. Second, with each epoch, the generative enterprise added a new question to the research agenda. This, of course, is exactly what one would expect from a successful program; break general questions down into a series of manageable bite size very specific research accessible topics, feed ensuing results into another set of related more ambitious research questions, then reiterate.

Though this is what one should expect, it is useful to rehearse this simple point at this time. Why? Because outsiders (and even some insiders) apparently look at the generative enterprise and see a blooming buzzing confusion motivated by seemingly irrational changes in fashion, MP simply being the latest rage. The logic of the enterprise has clearly passed them by. A good remedy is a concise Whig history, and Chomsky provides one. Second, because understanding how linguistic theory has developed affects how we should now understand the latest minimalist phase. For example, it modulates a widespread (if often inchoate) view that takes MP and GB to be competitors.⁷ Given the discussion above, they cannot be, at least in one sense (but see below), for the results specified by P&P (GB being P&P’s most auspicious incarnation) serve as the data to be analyzed by MP. As noted above, minimalist inquiry starts from the assumption that the “laws of grammar” as specified by GB are roughly empirically correct. They correctly describe key features of linguistic competence, at least to a good first approximation. MP embraces this conclusion and adds the following item to the research agenda: Why *these* laws? Why do we find the locality principles we find (e.g., A-over-A, islands, binding domains, minimality, cyclicity)? Why are dependencies coded in chains with hierarchical requirements (e.g., the head is a case position, the tail a theta position, and the links locally c-commanding one another)? Why are operations structure-dependent? Why are phrase markers hierarchically organized endocentrically? These and many other ‘facts’ are the explananda. MP’s aim is to furnish an explanans. If so, MP will not

⁵ But see her discussion on pp. 269–270.

⁶ Dresher & Kaye (1987) is the first work to observe this problem with parameter-setting models. They propose a solution that requires a linguistically specific learning algorithm (one with specified paths through the parameter space; cf. Dresher 1999), in contrast to the general learning algorithms generally deployed. Their discussion involves stress learning and so it is not clear that their solution generalizes to the syntactic case. At any rate, it appears that there is now a general consensus that the P&P picture, though alluring, does not easily translate into a workable model of acquisition.

⁷ Following the approach of Hornstein *et al.* (2005), we will take GB to be the best worked out P&P theory, and we will tentatively use it as proxy for the accepted wisdom. However, for what follows, the results of other frameworks would do just as well.

(and cannot) in general challenge these GB ‘facts’ (though they may modify them at the margins — recall that the assumption is that these are *roughly* empirically correct) but must assume them to be accurate. As corollary, MP cannot render obsolete GB-style research for there will always be value in discovering additional laws of grammar, new kinds of operations and generalizations. Rather, MP adds new questions to the research agenda; what principles (if any) do these generalizations follow from (e.g. economy, inclusiveness) and are these principles specific to linguistic computations (e.g. Merge) or are they reflections of more general principles of cognitive computation (e.g. No Tampering, Minimal Search)? In this sense, then, MP does not and cannot replace GB, any more than GB replaced the Standard Theory’s project of specifying the formal properties of linguistic constructions. In both cases, the results of the prior investigations were to a large extent conserved and served to set the scene for the next set of questions. Indeed, how could it have been otherwise! It is hard to factor out the common features of grammar rules without specifying any grammar rules and it is hard to ask what is specifically linguistic in a law of grammar without having any laws to consider.⁸

This said, there *is* a sense in which MP aims to replace GB. The MP problematic has two parts. The first is that the generalizations that generative grammar has discovered over the last 50 years are roughly empirically accurate. The second is that the generalizations unearthed are *not* ontologically fundamental. The first assumption sets a target for the second: how to deduce the specific properties of the faculty of language described by the laws of GB from more fundamental principles of computation and cognition. If so, there is a sense in which MP and GB *do* conflict. What GB takes as constitutive of UG, MP takes as accurate description but ultimately derivative. Physicists have good terms for this. They distinguish ‘effective’ theories (which are empirically roughly right) from ‘fundamental’ theories, which limn the metaphysical basics. Given this distinction, we can say that though GB is an ‘effective’ theory, it is not fundamental.

Chomsky’s remarks in both the introduction and the conclusion make a second important point: Not only is the time ‘ripe’ for minimalist inquiries, given the discovery of plausible laws of grammar whose properties we can investigate, but logic dictates that the linguistic specificity and complexity of the GB conception of UG *must* be illusory. Here’s the argument: Assume that complexity is the product of natural selection (NS), and assume that long time spans are required for natural selection to work its magic (say, on the order of many hundreds of thousands or millions of years). If so, UG’s properties cannot be the work of NS for there is reason to think that the human faculty of language is about 50–100,000 years old at most. Conclusion: The basic architecture of the Faculty of Language (FL) is not the result of an arduous and long process of selection but reflects an at most small addition to an already existing system of computations. This argues against the internal modularity of FL and against

⁸ A corollary: It cannot be a criticism of a proposed analysis that it is “not minimalist” because it invokes GB-style technology. In fact, if MP succeeds in reducing GB to a more basic theory (see below), then any decent GB analysis of a given phenomenon will have a smooth translation into a more fundamental minimalist account.

wildly idiosyncratic operations or principles of organization.⁹ The minimalist project aims to factor out those operations and principles that are part of cognitive computation more generally and see what small bit has to be added to get what we in fact see. In effect, on this view, FL/UG is an interaction effect, the output of the interaction of one (or two) linguistically bespoke operation(s) in interaction with more general, cognitively extant principles of computation.

This animating general picture suggests at least two specific kinds of research projects. The first is reductive. If complexity bespeaks long periods of evolutionary time, then short time scales suggest simple organization. If so, the complex modular structure that P&P accounts like GB postulate for FL/UG must be illusory. In other words, an FL/UG with seven different modules, with many different grammatical operations, defined over different primitives, and invoking different domains of application must be illusory. The illusion is made manifest by reducing the different modules to a common set of operations and principles.

The earliest versions of MP (e.g. Chomsky 1993) initiated this line of inquiry by reducing case checking to movement theory, case checking being a special instance of A-movement (subject to minimality restrictions) to a dedicated case position. More (indeed, far more) controversial instances of the same logic have aimed to reduce control and anaphora to movement as well.¹⁰ Assuming Chomsky's recent proposal that phrase structure-building and movement are special instances of the same basic operation — Merge (External and Internal Merge, respectively) — there is no fundamental distinction between movement and phrase structure construction either.

The logic mimics Chomsky's reduction of Ross's (1967) islands to a common set of operations (Move α), subject to a common set of conditions (Subjacency). If successful, this envisioned reduction significantly reduces the internal modularity of FL/UG. Case, control, binding, agreement, and movement are at bottom all the same kind of thing (*viz.* instances of Internal Merge), subject to the same restrictions and conditions in much the same way that topicalization, relativization, question formation, and focus movement are all instances of A'-movement, subject to Subjacency in Chomsky (1977). The best possible outcome is the elimination of all FL/UG-internal modules, and this possibility is clearly in sight theoretically.

The second project aims to build up the basic operations from computationally simple, well-behaved natural operations. Merge, which Chomsky has proposed is responsible both for phrase formation *and* displacement (aka movement), is the minimalist operation of choice. If we assume that Merge is compu-

⁹ This point has been put well in Fodor (1998: 12):

If the mind is mostly a collection of innate modules, then [...] it must have evolved gradually, under selection pressure. That's because [...] modules contain lots of specialised information about the problem-domains that they compute in.

Thus, a highly modular FL goes in tandem with an extended period within which NS operates. Conversely, a short period for NS to operate argues against a highly structured and modular FL.

¹⁰ See, among others, Lidz & Idsardi (1998), Hornstein (2001), Kayne (2002), Zwart (2002), Grohmann (2003), and Boeckx *et al.* (2010).

tationally well behaved (methodologically, the best assumption) we expect it to respect certain plausible conditions such as No Tampering and Inclusiveness. If we are clever (and lucky), we will discover that Merge, subject to its nice computational restrictions, will suffice to describe all the different kinds of dependencies that are treated as separate in GB; in other words, we will discover that, once an operation like Merge arises, which is able to combine atomic elements together in a computationally simple and natural way, language-like systems will emerge, able to represent (all and only) dependencies of the kind described by GB-like versions of FL/UG. In other words, once phrase structure emerges, displacement subject to the restrictions we find also arises. We are not at that point yet, but as Chomsky notes we are at least able to contemplate what such accounts might look like, and that is very exciting.

2. Some Disputations

This is the bare bones MP problematic. For the past two decades, Chomsky has offered a variety of different proposals to add meat to these bones. Some of these have proven to be controversial even among those moved by the animating picture. And this is how it should be! There are many ways to pursue the goals of the Minimalist *Program*. Inquiry will advance through the exploration of different specific minimalist *theories* — and over time, it is hoped, normal scientific practice will produce a fuller and fuller outline of the truth. Some provocative additions to the ongoing debate can be found in this volume. Let's consider a couple.

There is decent evidence that phrases in natural language are endocentric. Minimalists have coded this fact by supposing that Merge is a complex operation that involves two sub-operations, one that unites two elements merged and one that labels the subsequently created object in terms of one of the two mergees. The element that provides the label is the head. This results in a phrase that is both complex and categorized as having the same type as its head. The properties of endocentricity, it is hoped, reflect (indeed, follow from) this labeling process. The natural question to ask is why natural languages have labeling? This volume provides two different answers.

Chomsky denies that labels are fundamental objects or that labeling is a basic grammatical operation. Rather, he suggests that labels/labeling reflect minimal search, a design feature of a well-behaved computational system. The idea is that a computational system that embodies minimal search will isolate one element of a combined pair (i.e. roughly a binary-branching phrase) as more accessible than the other. This element is the head. Or, as Chomsky says (p. 393):

It's close to true that Merge is always a head and another object (a head is just a lexical item, one of the atoms, so Merge is a lexical item and some other object). To the extent that this is true — and it is overwhelmingly true — you eliminate the last residue of phrase structure grammar (projections or labels) because the head is just the thing that you find by minimal search.

On this view of things, (i) endocentricity is a product of minimal search and (ii) minimal search suffices to capture “everything about headedness”. Given (i) and (ii), we can reasonably conclude that (iii) endocentricity is a defining property of

well-designed computational systems and (iv) we should expect to find that most natural (and artificial?) computational systems exploit endocentric structures.

Cedric Boeckx presents a different view in chapter 3.¹¹ He proposes that the process of labeling a complex via one of its constituents is an operation unique to language and one of its basic defining properties.¹² As he says (pp. 47–48):

[T]his [labeling — AD & NH], as far as I can tell, is very, very specific to language as a kind of hierarchical structure. If you look elsewhere in other systems of cognition [...] you find a lot of evidence for hierarchical structuring of systems, possibly recursive ones, but as far as I can tell, these hierarchical structures are not headed or endocentric [...]. That [...] is very specific to language, so while you find hierarchies everywhere, headed or endocentric hierarchies seem very central to language.

On this view, whereas a “grouping procedure that puts X and Y together” is likely to be very common across cognitive modules, the operation of “selecting one of these two members and basically using that member as the next unit for recombination” (p. 48) is unique to language. Thus, we do not expect it to be part of computational systems in general but to arise predominantly (perhaps, exclusively) in linguistic structures.

We are partial to the second view of labels for several reasons: First, Boeckx’s observation that we do not find endocentric structures in other areas of cognition argues against analyzing it as a product of minimal search for were this its source such structures should be ubiquitous features of any moderately well designed computational system, apparently contrary to fact.¹³ Second, it is not clear to us how minimal search suffices to derive the properties associated with endocentricity. Take, for example, head-to-head selection or subcategorization, e.g. the relation between *interest* and *in* in phrases like *interest in Bill*. This relation is severely restricted. Nouns might select/subcategorize for the prepositional head in a PP complement, e.g. *interest* requires *in*, but there is never selection for the complement of the prepositional head. Thus, there are no nouns like *interest* that require that the nominal complement be *fish* (or a DP headed by *the*) but allow for any preposition whatsoever, e.g. **interest (over/under/to/with) fish*. Why not? Because only the head within a phrase can be seen by an outside head. And this is not because the preposition is less complex than its complement for make the complement of *in* as simple as possible, e.g. *it*, and the complement will still

¹¹ The reader should be warned that one author is a partisan here, as this view is also proposed and defended in Hornstein (2009).

¹² Hornstein (2009) proposes that it is the adventitious addition which, when added to the rest of cognition, results in a grammatical distillate.

¹³ It is not easy to find examples of artificial computational systems that code endocentricity. Basic arithmetic seems to do well without it, for example. Programming languages also tend to avoid endocentric means of combination (though here there is more room for interpretation). In both cases, the most common (and in arithmetic, only) form of combination is homogenous: Multiple expressions of the same kind combine to form another expression of that kind. For example, the integer expressions 2, -2, and 3 combine to form the expression $[2 * [-2 + 3]]$, or a sequence of statements combines to form a single complex statement. In the case of programming languages, one also finds clear examples of non-homogenous but nonetheless non-endocentric combination. For example, a list of variable and method declarations combines to form a class (which itself is neither a variable nor method declaration).

never be visible to an outside selecting head. This suggests that minimal search is besides the point for, hierarchically speaking, the preposition and its simplex complement are equidistant from any higher element.¹⁴ As such, searching for either should be equally ‘minimal’. Third, labels make computational sense when one considers the fine structure of Merge. Labeling closes the ‘combine’ operation in the domain of the lexical atoms. As Chomsky (2008) has noted, mergeable items must have edge features. All lexical atoms have these inherently. Furthermore, when two atoms combine to form a unit, that unit must also have an edge feature for it is further combinable. Whence *this* edge feature? It cannot be a lexical property as complex units are not primitive atoms. Thus, it must inherit this property, presumably from the elements combined. In other words, even if we define Merge as the operation that combines lexical atoms, we need to explain how this combinatorial power extends to lexical complexes. Given standard assumptions, it is natural to think that labels endow complexes with the powers of atoms: A lexical atom inherently has an edge feature and a phrase inherits an edge feature insofar as it shares the properties of its head.¹⁵ In other words, edge feature inheritance supervenes on labels as labeling is what allows complex units to combine in just the way that simplex lexical atoms do. In effect, labels create an equivalence class of expressions centered on primitive lexical atoms. This is the secret to the kind of hierarchical combination natural languages exploit.

There is, no doubt, more to say on either side of this ‘debate’. However, what is important here is the recognition that neither proposal is *a priori* more or less ‘minimalist’; both fit the spirit of the program. Which is correct is (ultimately) an empirical matter, not an ideological one. The Minimalist Program is compatible with various minimalist theories, and which versions should prevail is an empirical (albeit, no doubt, complex) matter.

Luigi Rizzi’s discussion of two notions of locality (chapter 11) offers a second illustration of a nice, ripe minimalist research question. Movement is currently subject to what appears to be two different locality restrictions, Relativized Minimality (RM) and Phase Impenetrability (PI). The former forbids dependencies between two elements X and Y over an intervening Z of the same type, with type being defined featurally. The latter prevents dependencies that span the complement domains of specific elements, *v* and C being (at least) two. Both notions are part of the current minimalist technical apparatus and both have empirical virtues. Rizzi asks the obvious minimalist question: Do we really need both?¹⁶ More specifically, given that locality is a way of bounding domains of computation and given that bounded computations are invariably efficient and

¹⁴ Note that endocentric labels will help define minimal dependencies and heads will be more prominent than anything else in a phrase. However, this reverses the explanation: Minimal search supervenes on headedness, not the other way around.

¹⁵ This follows from what a label is. For example, it is the standard assumption that the label is the head (cf. Chomsky 1995). Moreover, in virtue of this, X'-theory assumes that “a head and its projections share *all* [our emphasis — AD & NH] properties apart from bar-level [...]” (Chomsky 1986: 18). Thus, if a head has an edge feature, the phrase that it labels does too. This suffices to project edge features in the required way: A phrase inherits an edge feature from one of its daughters and the daughter it inherits it from is the head.

¹⁶ Our impression from the dearth of discussion of minimality in Chomsky’s latest papers is that he too is attracted to the idea that there is only one operative locality condition, *viz.* PI.

well behaved, we expect on conceptual grounds that linguistic computations should be bounded and local. However, bounding the computations twice over via two different locality conditions does not obviously make computations *more* well-behaved or *more* efficient. So, why two?

The question becomes more insistent when one notes that current versions of RM and PI are somewhat redundant. So, for example, each only kicks in when two or more expressions of the same type are lurking and neither applies to put an absolute bound on a grammatical domain. Here's what we mean. Consider a case of multiple movement as in (1):

- (1) John_i seems to be believed to be likely to be kissed *t*_i.

There is a long-distance relation (potentially unbounded) between *John* and its base position indicated by *t*_i. RM places no bound on this dependency as there is no relevant intervening element, i.e. no other DP in this domain. PI, at least Chomsky's version, treats (1) as a single phase.¹⁷ So, both RM and PI countenance arbitrarily long distance dependencies in similar circumstances. Moreover, both block them in similar cases. So for example, in (2) both RM and PI block movement of the object to the higher subject position:

- (2) *John_i seems that it was told *t*_i that Bill left.

For RM, *it* intervenes between *John* and its base position and with PI the complement of *that* is a phase and hence its complement is impermeable. This suggests that in different ways, the two conditions apply to the same configurations and hence are partly redundant. This only strengthens the above-mooted why-question: why *these two* locality restrictions?¹⁸

One of the more interesting arguments for PI comes from successive cyclic movement. Indeed, the derivation of successive cyclicity from Subjacency was one of its more endearing properties. There is evidence that in examples like (3) we have moved C-to-C:

- (3) How_i did you think [_{CP} *t'*_i that we will solve Plato's Problem *t*_i]?

Note that if C is a strong phase, then PI requires movement of *how* via the phase edge. This results in successive cyclic movement. It is less clear how to accommodate these data given RM. Rizzi proposes a solution. In effect, he postulates features in finite C (Q-features) that via RM will prevent *wh*-operators like *how* from traversing them. He further provides empirical evidence from Chamorro for the selective nature of this restriction, some *wh*'s able to cross the finite C — others not.

¹⁷ This is not quite right, there is no *strong* phase between *John* and *t*. There are several weak phases, but as these fail to block grammatical intercourse, they do not bound computations.

¹⁸ That RM and PI should roughly overlap is suggested by the fact that, on both, transitive clauses are expected to be local domains for A-movement. This is true for RM because at least two DPs live in a transitive clause and hence one should block movement of the other. In PI this is because transitive *v* is strong. We will see momentarily that there are ways of endowing finite C, which is a strong phase head within PI, with features that block non-local dependencies across them in an RM-approach.

The discussion is subtle and interesting, albeit more suggestive than conclusive.¹⁹ The paper is too cursory to settle the question of whether all unbounded dependencies can be regulated using just RM.²⁰ For example, the discussion is limited to weak island effects and the classical cases of locality are those involving strong islands.²¹ However, the question *is* a very good one and the kinds of considerations that Rizzi brings to bear in the discussion well illustrate the empirical utility of raising such minimalist concerns.

Tom Bever (chapter 18) offers a third example in his discussion of the EPP, the requirement that sentences have subjects. This principle has resisted attempts at rationalization or deeper explanation, and it stands to date as a theoretical diacritic enforcing the requirement that sentences have subjects. Bever's paper proposes to remove the EPP from FL/UG and instead locate it in learning theory. It is an expression of the Canonical Form Constraint (CFC), which requires that "sentences [...] must sound like they are sentences of the language to afford the individual child a statistical entrée into acquiring it" (p. 279). If correct, this serves the minimalist ambition of tidying up FL/UG by repositioning a requirement that is an *ad hoc* idiosyncrasy when understood grammatically to a principled effect when understood as part of a more general theory of learning.

The ambition and form of the argument Bever presents is impeccable. Less clear is whether it covers all the relevant cases. The chapter concentrates on cases of the CFC and the EPP in main clauses. The discussion ranges over expletive constructions and passives and one can see how the CFC could be deployed to account for the presence of expletive subjects and movement of objects under passivization. As Bever notes, "the vast majority of sentences and clauses have a canonical form with a subject preceding a correspondingly inflected verb" (p. 289). Expletives serve to gather sentences that have no apparent logical subjects under the same statistical rubric, as would moving objects to subject position in cases like passive, where subjects are left unexpressed, or unaccusatives, where there are no subject-like arguments to be had. This all fits nicely together, especially in finite matrix domains. Note, however, that in *these* domains, there are also plausible and natural grammatical candidates to explain the EPP, for example, the Inverse Case Filter; the requirement that there be lexical expressions able to support the (nominative) case that finite T⁰ assigns to their specifiers. Here, no "conspiracy" (p. 290) of interacting constraints is required, only the natural assumption that morphological case needs to be assigned, and thus something needs to be there to receive it.²² The real testing ground for the EPP is non-finite

¹⁹ In addition, there exists some evidence that arguments like *who/what* trigger inversion in languages like French and Spanish but adjuncts like *why/how* do not. This might at first blush be taken to indicate that only *who/what* undergo successive-cyclic movement. This raises non-trivial variation issues if correct, since in the work Rizzi cites, Chung (1994: 29) has argued that both adjuncts and arguments trigger *wh*-agreement in Chamorro.

²⁰ This is *not* a criticism of the paper, given the venue in which it was presented. We are sure that Rizzi would agree that the argument he provides was intended to be illustrative (and in this it succeeds admirably) rather than definitive.

²¹ It is not clear that the approach of adding features to heads can be applied to yield strong island effects. Then again, extending PI to accommodate strong islands seems to require adding D to the inventory of phases, and this sits poorly with Chomsky's interest in tying the inventory of phases to propositional heads.

²² Note that we are not saying that this is the correct explanation. Rather, we are observing

embedded clauses; do these also require subjects? A case-based account would answer in the negative, as would, we surmise, an account based on the CFC.

At present, there is not overwhelming evidence that non-finite clauses require subjects.²³ However, this is where the consequences of the EPP are the most interesting and least obviously tied to what one finds in matrix clauses. This said, Bever's proposal is interesting for it provides a picture of how one might try to give a non-grammatical account of what would appear to be a paradigmatic example of a grammatical restriction. This is the sort of thing that minimalists should welcome for simplifying UG it makes it easier to explain its remaining properties.

All three discussions illustrate another important point. The take-home message from San Sebastian is that it takes a lot of work to show how linguistic competence relates to cognition in general. Serious illumination requires identifying specific linguistic features (e.g., endocentricity, selective islands, EPP), isolating some plausibly general cognitive operations or computational restrictions and showing how the former follow from the latter. The examples above attempt this. If they succeed then we have managed to isolate the cognitively / computationally general from the linguistically specific. Unless we do this, what we have is the sound of one hand waving.

3. Words and Concepts

Chomsky's introduction touches on a second very important feature of linguistic systems. In addition to the unique kind of recursion we find in language, human linguistic systems contain atoms, lexical items (words) that exhibit a distinctive (perhaps unique) set of properties. They are remarkably flexible things, at least when compared with what we find in other biological communication systems. As Chomsky (p. 27) notes:

The basic problem is that even the simplest words and concepts of human language and thought lack the relation to mind-independent entities that has been reported for animal communication: representational systems based on a one-one relation between mind/brain processes and "an aspect of the environment to which these processes adapt the animal's behavior," to quote Randy Gallistel [Gallistel 1990]. The symbols of human language are sharply different. [...] Communication relies on shared cognoscitive

that this is a natural grammar-internal explanation, that it is not theoretically idiosyncratic, and that it serves to remove the diacritical nature of the EPP. This kind of proposal exists in the literature, cf. Castillo *et al.* (1999) and Epstein & Seely (2006).

²³ In our opinion the best argument concerns cases like (i):

- (i) John_i appears to Mary_k to seem to *herself_k/himself_i to be intelligent.

The unacceptability of *herself* and the acceptability of *himself* follows if there is a trace of *John* in the subject of the first embedded clause, which in turn follows if the EPP holds in all clauses, including non-finite ones. The main problem with such examples concerns the relative poor status of (ii). It should be perfect given the unacceptability of (i) with *Mary* antecedent *herself*, given that pronouns are licensed by the binding theory precisely where reflexives are barred. To our ears, (ii) is not particularly good. This renders the datum in (i) hard to interpret.

- (ii) John_i appears to Mary_k to seem to her_{77k} to be intelligent.

powers, and succeeds insofar as shared mental constructs, background, concerns, presuppositions, etc. allow for common perspectives to be (more or less) attained. These semantic properties of lexical items seem to be unique to human language and thought, and have to be accounted for somehow in the study of their evolution.

It is not at all clear what the cognitive basis of these distinctive features of words is. It is not even clear whether this is a specifically linguistic property of words or a more general cognitive feature of human concepts. It appears, however, that this is a unique characteristic of human words/concepts and is key to any future understanding of the creative aspects of language use. To date, we lack a systematic description of the basic semantic properties of words/concepts or a systematic comparison between human words/concepts and those found in other animals.

An interesting line of current inquiry prompted by recent minimalist concerns compares animal and human cognition to isolate those features that are specifically 'human'.²⁴ **Marc Hauser's** contributions (chapters 5 and 19) directly advance this project, as does **Randy Gallistel's**, albeit more indirectly. Animal cognition is very elaborate. Gallistel's paper (chapter 4) describes the elaborate "foundational abstractions of time, space, number and intentionality [that] inform the behavior of birds and bees (p. 61)". It appears that birds can reason about time: They "compute elapsed intervals and compare them to other intervals in memory (p. 63)". They can reason about number: They are able to "first subtract the current number from a target number in memory and then compare the result to another target number in memory (p. 63)". And they can evaluate "the likely intentions of others and reason from their own actions to the likely future actions of others (p. 64)". All of this requires rather formidable computational powers as well as primitive foundational (innate) abstractions concerning time, number and other minds of a robust sort. Interestingly, these elaborate capacities appear to exist independently of a facility for language and so we can conclude that linguistic capacity is not necessary for a rich mental life (see the discussion of Gelman's chapter below).

In chapter 5, Hauser pursues this comparative methodology to consider concepts that are characteristic features of linguistic systems. He reports work indicating that rhesus monkeys distinguish singular from plural concepts as well as mass from count. If this is correct, then despite the grammatical exploitation of these notions, their etiology is independent of FL. This does not imply that the concepts operative grammatically are in every way identical to those that Hauser's monkeys have. He points out that infants do less well than rhesus monkeys when forced to enumerate objects versus masses (p. 81). Thus, even if the count/mass distinction does not require linguistic underpinnings, it may well be that absorption into FL endows the distinction with distinctive properties. Nonetheless, these results support an important minimalist intuition; that language is a very complex system and that linguistic competence is an interaction effect with various non-linguistic parts of our cognitive apparatus making

²⁴ See the distinction between the broad and narrow faculty of language in Hauser et al. (2002), where the minimalist sources of this kind of research is clearly outlined.

important contributions.

If concepts do indeed predate language, we might ask what else was present in the pre-linguistic conceptual-intentional (C-I) systems. This question is raised by **Wolfram Hinzen** in a provocative chapter 9, which argues for the abolition of the C-I interface. Chomsky believes that the structures derived in the syntactic component are fed into a set of highly intricate interpretative processes; let us term this the ‘interpretavist’ position. These interpretative processes are standardly taken to fix certain aspects of, for example, scopal and anaphoric interpretation. As such, they must have logical as well as conceptual resources at their disposal. In particular, they must be able to ‘read’ the recursive hierarchical structures offered up by the syntax. This, Hinzen argues, is a fatal flaw in Chomsky’s conception of the relation between the syntax and the C-I interface. Chomsky has famously claimed that recursion is a, or perhaps even *the*, property specific to language.²⁵ But, of course, this is not a coherent position if his conception of the C-I interface presupposes recursion in other cognitive faculties.

There are various technical moves that could be made at this point in defense of the interpretavist position. For example, it may be that Hinzen is inferring too much from the fact that the C-I interfaces are able to read the recursive structures generated by the syntax. This does not necessarily imply that “recursive thought” was possible prior to the development of language. Rather, it may be that the capacity to interpret recursive structures lay dormant until the corresponding ability to generate them developed. In any case, we will not quibble further on these points, since we think that Hinzen has hit on a genuine puzzle. Our main concern is with Hinzen’s sketch of a solution to this puzzle, which we find problematic.

Hinzen’s approach is in the tradition of ‘abstract syntax’. On the interpretavist model, as developed within the Minimalist Program, syntactic processes are for the most part semantically blind. Some syntactic units may receive no interpretation at all (e.g. dummy prepositions such as *of*), and the same syntactic configuration may admit of multiple interpretations. In contrast, Hinzen argues that syntactic structures are inherently meaningful. They do not *receive* interpretations; they *bear* meanings.

We see two primary difficulties with Hinzen’s approach. The first problem is one that his approach has in common with its historical antecedents. There is a multitude of gnarly technical difficulties involved in attempting to capture semantic generalizations in syntactic terms. Consider, for example, judgments of entailment and synonymy. Some of these might plausibly be read off syntactic structure directly (e.g., that if John met Bill on Tuesday, he met Bill). Others can be farmed out the conceptual component (e.g., that if John ate breakfast, he ate some food). There remain, however, awkward intermediate cases. For example, though competent English speakers know that (4a) is near-synonymous with (4b) the contrast between (5a) and (5b) suggests (ingenious arguments to the contrary notwithstanding²⁶) that this relationship cannot be captured in syntactic terms:

²⁵ Or to be more precise, a/the property specific to the faculty of language in the narrow sense (Hauser *et al.* 2002).

²⁶ See Jackendoff (1969: chap. 5) and references cited therein for relevant discussion.

- (4) a. John rarely makes mistakes.
b. It is rare for John to make mistakes.
- (5) a. John deliberately makes mistakes.
b. *It is deliberate for John to make mistakes.

Attempts to give syntactic analyses for phenomena of this sort have inevitably led to the postulation of highly complex and abstract syntactic structures. Complex and idiosyncratic rules are then required to ensure that, for example, (4b) can surface as (4a), but that (5b) cannot surface as (5a).

This brings us to the second difficulty, which unlike the first is proprietary to minimalism. Suppose that the abstract syntax project is one day executed successfully. (The technical difficulties may, after all, be overcome.) Could the resulting theory be in any sense minimalist? A virtue of the interpretativist approach is that it permits us to separate out a small domain — a relatively modest syntactic component — in which a reductionist approach is feasible. On almost anybody's account, the interpretative systems are far less tractable. A cautiously optimistic person might hope to see the basic properties of A-movement derived from "minimal search" in his or her lifetime. But it seems fantastic to suppose that the whole semantic circus will submit to this sort of treatment. In our view, then, minimalism is quite fatal to any attempt to break down the barrier between syntax and the C-I system — or to put it in less guarded terms, between language and thought. Whatever may be the correct division of labor between mental faculties, we suspect that the bit that does the merging and moving will turn out to be a lot smaller than the bit that does the thinking.

Chomsky's remarks (especially in the conclusion) touch on his own views regarding the relation of language, meaning and thought in general. He briefly mentions classical conundra, such as the ship of Theseus, and Kripke's (1979) puzzle regarding the meaning of belief-ascribing sentences. What he does not belabor here, but has explicitly noted elsewhere, is that more often than not, natural language runs roughshod over classical semantic distinctions; temperatures can be 90 and rising, books that weigh five pounds can be engrossing, the average man can have 5.4 children, hexagonal European states can be republics, London, a city of 5 million, can burn down and be rebuilt three miles down the Thames, there can be flaws in arguments, etc. These constructions seem to conflate values and functions, refer to impossible entities, run-together geographical and constitutional conceptions, treat abstract and concrete dimensions on a par etc. All of this should discourage the jump from linguistic or conceptual analysis to metaphysics, since, "if you give a metaphysical interpretation to these things, you run right off into impossible conundrums" (p. 382). If we are interested in what is really distinctive about human language, it behooves us to find out just how labile our words/concepts are.²⁷

Several other papers in the volume directly speak to this enterprise. For example, **Rochel Gelman** (chapter 15) distinguishes between *core* and *non-core* domains of knowledge. The former consist of "skeletal" innate mental structures

²⁷ Paul Pietroski has recently emphasized this (cf. Pietroski 2010).

that “direct attention and permit the uptake of relevant data in the environment” (p. 226). Such a skeleton seeds the acquisition process by providing “a set of coherent principles that form a structure and contains unique entities that are domain specific” (p. 226). Gelman provides a concise description of the properties of core domains (pp. 228–229), which should look familiar to readers of this journal as it well describes how linguists have been thinking of principles of UG for the last 50 years, as Chomsky notes (p. 384). Gelman also provides an illustration for the domain of natural number and isolates three principles that seed the counting process (p. 231). It would be interesting to consider how these specific principles may or may not be related to those that undergird linguistic knowledge, especially given speculations that “the numbering system is just a trivial case of language” (p. 33) and the apparent fact noted in Gallistel’s chapter that birds are pretty good arithmeticians.²⁸

Lila Gleitman (chapter 16) offers a second illustration, closer to home. She has long observed how hard word learning actually is, in part because of the “richness of the stimulus” (p. 210). The hard problem, she persuasively demonstrates, is to understand how the infant is able to “represent [...] context ‘in the right way’” (p. 207). How do children know to ignore some features of a scene as irrelevant to lexical meaning in some contexts while attending to them correctly in others? She illustrates the problem by reviewing some experiments by Gordon (2003) on 10-month-olds (who as yet utter no words) in which a stuffed bear passed between two participants is gauged relevant in scenes depicting givings but not huggings. This suggests that the capacity to index participants to event types is ‘epistemologically prior’ to word learning and serves as a boost to that process. Similarly, later on the richness of the stimulus is similarly tamed by considering how discourse and thematic roles are mapped into syntactic frames and how these can be used to constrain mappings to meaning. For example, it appears that both adults and children use an interlocutor’s eye gaze to pick out the subject of a sentence and this in turn helps to disentangle fleings from chasings (p. 245). Mapping words to concepts is a very complex affair in which “a mosaic of conspiring cues [...] are exploited [...] to converge almost errorlessly on the lexicon of the native tongue” (p. 211). Not surprisingly, these cues appear culled from various parts of our cognitive apparatus and interact with the specifically linguistic properties of FL/UG in very complex ways.

4. There Is Much More

There are many more interesting papers in this volume that we do not have the competence or space to discuss in detail. All of these are summarized in the useful and detailed introduction by **Massimo Piattelli-Palmarini, Pello Salaburu and Juan Uriagereka**. In chapter 6, **Gabriel Dover** argues strongly against the idea the “laws of form” significantly constrain biological variation, and makes some speculations regarding free will.²⁹ **Donata Vercelli and Massimo Piattelli-**

²⁸ See the very brief discussion between Gelman and Higginbotham on p. 236 as well.

²⁹ With regard to free will, we are not persuaded that Dover really addresses the problem at all. Dover’s focus is on the first-person subjective problem of free will. He suggests that free will is “a situation of rapidly and subtly changing outcomes as degenerate neuronal net-

Palmarini (chapter 7) give a useful introduction to epigenetics (the study of heritable changes which are not caused by changes in DNA), and discuss with Dover its implications for the biology of language. **Christopher Cherniak** outlines the case for “non-genomic nativism” with regard to the optimization of neural connections (chapter 8). **James Higginbotham** makes a case for a robustly truth-conditional approach to semantics, and presents a brief analysis of typological variation in the availability of resultative constructions (chapter 10). **Juan Uriagereka** discusses the problems posed by uninterpretable Case features, and a possible solution in terms of the “viral” theory of Case (chapter 12). **Angela Friederici** presents fMRI data suggesting that different regions of the brain are responsible for processing local dependencies vs. hierarchical structures (chapters 13 and 22). Chapter 14 contains a fascinating round table discussion with **Cedric Boeckx**, **Janet Dean Fodor**, **Lila Gleitman**, and **Luigi Rizzi** on “Language Universals: Yesterday, Today and Tomorrow”. **Itziar Laka** clarifies certain questions regarding the domain-specificity of UG principles (chapter 20). **Núria Sebastián-Gallés** discusses neuro-imaging studies which shed light on the question of why there is so much variation in the extent to which second language acquisition is successful (chapter 21).

It is fashionable nowadays to lament the current state of generative grammar in general, and minimalist syntax in particular. This volume serves as an excellent antidote. In fact, it is a pretty good advertisement, in our view, for the intellectual vitality of the current enterprise. Minimalism has added a new set of questions to the research agenda, both within syntax proper and in the neighboring domains of psychology, neuroscience, and biology. Some of these links are more tenuous than others, as we should expect. Nonetheless, it is clear, at least to us, that research on the structure of FL is thriving, and that this is partly due to the fecundity of the core ideas of the Minimalist Program. From where we sit, syntax has rarely looked healthier.

References

- Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.
- Castillo, Juan Carlos, John E. Drury & Kleanthes K. Grohmann. 1999. Merge over move and the Extended Projection Principle. *University of Maryland Working Papers in Linguistics* 8, 63–103.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

works switch from one quasi-stable state of topology to others [...]. At the level of biology [...] there is an unbroken route of cause and effect [...] but at the level of our sense of what happened, we feel that at the threshold of the final step [...] is one for us alone to decide”. Leaving aside the decorative references to degenerate neural networks, this is just a statement of the standard determinist position, i.e. that for any given pair of a pre-decision state S1 and post-decision state S2, S1 lawfully and deterministically brings about S2, even though we (erroneously) believe that S2 is contingent both on S1 and some additional act of volition. It remains unclear on Dover’s account why we feel as if we’re making free decisions. (Why don’t we just feel as if we’re switching between quasi-stable states?)

- Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal Syntax*, 71–132. New York: Academic Press.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale & Samuel J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
- Chung, Sandy. 1994. Wh-agreement and ‘referentiality’ in Chamorro. *Linguistic Inquiry* 25, 1–44.
- Dresher, B. Elan. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30, 27–67.
- Dresher, B. Elan & Jonathan Kaye. 1987. A computer-based learning theory for metrical phonology. Ms., University of Toronto.
- Epstein, Samuel David & T. Daniel Seely. 2006. *Derivations in Minimalism*. Cambridge: Cambridge University Press.
- Fodor, Jerry A. 1998. The trouble with psychological Darwinism. *London Review of Books* 20, 11–13.
- Gallistel, C. Randy. 1990. Introduction. *Cognition* 37, 1–2.
- Gordon, Peter. 2003. The origin of argument structure in infant event representations. *Boston University Conference on Child Language (BUCLD)* 28, 189–198.
- Grohmann, Kleanthes K. 2003. *Prolific Domains: On the Anti-Locality of Movement Dependencies*. Amsterdam: John Benjamins.
- Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
- Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Malden, MA: Blackwell.
- Hornstein, Norbert. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.
- Jackendoff, Ray. 1969. Some rules of semantic interpretation for English. Cambridge, MA: MIT dissertation.
- Kayne, Richard. 2002. Pronouns and their antecedents. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 133–166. Malden, MA: Blackwell.
- Kripke, Saul. 1979. A puzzle about belief. In Avishai Margalit (ed.), *Meaning and Use*, 239–283. Dordrecht: Reidel.
- Lidz, Jeff & William J. Idsardi. 1998. Chains and phono-logical form. *PLC* 22—*University of Pennsylvania Working Papers in Linguistics* 5.1, 109–125.
- Piattelli-Palmarini, Massimo. 1980. *Language and Learning: The Debate Between Noam Chomsky and Jean Piaget*. Cambridge, MA: Harvard University Press.
- Pietroski, Paul. 2010. Concepts, meanings, and truth: First nature, second nature, and hard work. *Mind and Language* 25, 247–278.

- Ross, John R. 1967. Constraints on variables in syntax. Cambridge, MA: MIT dissertation.
- Zwart, Jan-Wouter. 2002. Issues relating to a derivational theory of binding. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 269–304. Malden, MA: Blackwell.

Alex Drummond
 University of Durham
 Department of Philosophy
 50 Old Elvet
 Durham DH1 3HN
 United Kingdom
alex.drummond@durham.ac.uk

Norbert Hornstein
 University of Maryland
 Department of Linguistics
 1401 Marie Mount Hall
 College Park, MD 20742
 USA
nhornste@umd.edu