

NOTES AND DISCUSSION

**Response to David Adger's 'Remarks on
Minimalist feature theory and Move'¹**

ASH ASUDEH & IDA TOIVONEN

Carleton University

(Received 21 January 2006; revised 27 February 2006)

David Adger raises some interesting issues and makes several valuable points in his 'Remarks on Minimalist feature theory and Move' (henceforth MFTM), a response to our review article 'Symptomatic imperfections' (henceforth SI) in this journal (Asudeh & Toivonen 2006), which was in part a review of his *Core syntax* (Adger 2003). In this response, we address some of the points in MFTM. We would also like to set the record straight about some points in SI which we feel have been misrepresented. In several instances, MFTM argues against claims that were not made in SI. Whatever the independent merit of these arguments, we do not wish to defend viewpoints we did not propose in the first place.

We argue in SI that the low standard of formal rigour and explicitness in the Minimalist Program (MP) is problematic.² Adger agrees with our assertion that various formalisations of Minimalism, including Stabler (1998), have been ignored in the mainstream MP literature, including *Core syntax* (CS). However, he claims that this is 'beside the point'. Adger writes:

[T]he informal presentation I gave in CS ... maps rather directly onto Stabler's formalization; that is, for the most part there is a one-to-one correspondence between the concepts I informally present and Stabler's formalization ... This is unsurprising, as both Stabler and I essentially use the notions developed by Chomsky. (MFTM: 663–4)

There are several problematic aspects to this passage. First, it does not make any sense to state that a one-to-one correspondence holds only 'for the most part'. In that case, it is not a one-to-one correspondence and the relevant issue is what the differences are and what their properties are. Second, there is no way for a reader of CS to know that its formalisation 'maps rather

[1] We would like to thank Chris Potts and Geoff Pullum for brief comments on an earlier draft. Thanks also to Bob Borsley and Ewa Jaworska at *JL* and to our anonymous *JL* referee. Last, but not least, many thanks to David Adger for a stimulating debate and for his comments on this response. We accept all responsibility for any errors.

[2] We are by no means the first to point this out; see SI for some relevant references.

directly onto Stabler's formalization', since Adger never shows this to be the case in *CS* (e.g. Stabler's work is not cited in the book). Third, Stabler (1998) is itself a formalisation of only some of the ideas in *MP* and is therefore by no means a complete formalisation. Fourth, the fact that *CS* and Stabler (1998) share a common genesis in Chomsky's work is no guarantee that they have interpreted the latter's ideas in the same way. Chomsky sketches a research program which can be and has been interpreted in various ways. For example, the other volume we review in *SI* – Andrew Radford's *Minimalist syntax* (Radford 2004) – also uses the notions developed by Chomsky, and yet it develops a framework that is strikingly different from that of *CS*.

A more important general point is that explicitness and formalisation are separable issues. It is possible to be fully explicit about a theory or sub-theory (e.g. the theory of grammatical features) without providing a full formalisation. Conversely, a formalisation is not itself sufficient to render an analysis explicit: inexplicit analyses may still be constructed if they do not clearly refer to the formalisation. We think that the possibility of separating explicitness from formalisation is what lies behind the following passage in section 2 of *MFTM*:

I think that it is perfectly legitimate to do (theoretical) linguistics in the absence of a formalization and that one can gain a great deal of insight in this way. (*MFTM*: 664)

We do not in principle disagree with this. However, our contention in *SI*, which we reiterate here, is that Minimalist analyses are informal AND inexplicit. It is highly problematic to claim that an informal and imprecise theory or analysis is simple. We cannot know that the initial simplicity or elegance will remain once the details are worked out. To quote ourselves: 'explicit analyses tend not to seem as minimal or elegant as analyses that leave out the details' (*SI*: 409).

The Minimalist Program rests on purely intuitive understandings of inexplicit notions such as external/internal Merge, feature interpretability, feature strength, feature checking, feature copying, feature valuation, etc. However, Minimalism still lacks clear definitions of these notions that closely correspond to their use in actual Minimalist accounts. Individual Minimalist analyses interpret these notions completely differently, which an explicit theory would preclude. *CS* takes important first steps towards rendering its theory explicit, as noted and commended in *SI*, but it is still inexplicit about many details, no doubt partly due to its nature as a textbook. Furthermore, Adger points out that his theory differs in important ways from *MP* as it is practised elsewhere (*CS*: xi–xii). The general reference in *MFTM* to an existing formalisation of *MP* (Stabler 1998) is not useful in the absence of an account of how that formalisation would model Adger's own theory.

Adger conjectures in MFTM that for him theory comes first and for us formalisation comes first. He characterises his view as follows:

Once some theoretical understanding of the phenomenon is developed, one can ask what a good formalization of the theory is and, when the theory changes, one can see whether the underlying formalization needs to be revised (extended or reduced). (MFTM: 664)

We agree with Adger that the theory should constrain the formalism and not the other way around. This does not mean, though, that it makes sense to theorise inexplicitly and in the absence of a formalism that actually models the theory as it is practised.

In addition, his characterisation of our position is incorrect. We do not claim that formalisation comes first: formalisation makes no sense in the absence of theory. Our actual stance, which is clear from SI, is that it is not possible to have a useful theoretical understanding of a phenomenon, nor to individuate changes in the theory properly, without an explicit and formal theory. This is a standard conception of the relationship between formalism and theory in linguistics, forcefully articulated in Chomsky (1957: 5). It is telling that Adger himself uses formalisation in MFTM when making his theoretical points about compositionality, complexity, and constrainedness.

In our review article, we criticised Adger's feature theory on the following three grounds: it allows noncompositionality, it is complex, and it is unconstrained. In MFTM, Adger responds that the feature theory of 'Core syntax' can be changed and extended to render it compositional and more constrained. We never claimed that it could not. In fact, we indicated in SI one particular way in which compositionality could be maintained (SI: 415). With respect to constrainedness, MFTM does not actually give details of how the feature theory of CS can be further constrained (see below), but we nevertheless accept that it probably can. However, although it is true that Adger could have written a different book (which seems to be what is being stated in section 2 of MFTM), this is not really a valid response to our criticism of the actual book.

Adger begins his section on compositionality (section 2.1) by questioning our assertion that his system permits noncompositionality. He argues that a close reading of his book makes it clear that his system is purely compositional. We see no reason to revise our conclusion from SI, which was based on a very close reading of the book. The reader is invited to inspect the relevant passages in CS (28–32, 39) and the relevant passages in our review (SI: 413–416). Adger accepts partial responsibility for the 'misunderstanding' about compositionality, but defends himself by pointing to a passage in CS where he indicates that the rule for interpreting the feature plural 'needs to be much more carefully worked out (so that the semantics of dual forms is correct, for example)' (CS: 31). Two points need to be made here. First, the

statement that the rule for plural is not correct as stated is not any indication of what the intended correct version is and how one might go about correcting it. Second, it is noteworthy that, having been pressed on the point, the formalisation that Adger provides in his section 2.1, based on Harbour (2003), does not in fact correct the rule for plural on page 31 of *CS*. The formal statement in *MFTM* entails that plural means ‘more than 1’, which is perfectly commensurate with the rule in the book (‘Interpret a noun specified with [plural] as referring to a group of entities’). In fact, it is the semantics of the feature [singular] that is revealed to be at issue in the formalisation of *MFTM*. Thus, the kind of formal consideration that we encouraged and that Adger carries out in *MFTM* has revealed an error in the informal presentation in the book, where Adger incorrectly points to a complication in the interpretation of [plural] being necessary. This is itself an example of the value of formalisation in clarifying theoretical claims and how unformalised claims can be misleading or incorrect.

As mentioned above, we provide one compositional analysis of the Hopi dual data that lies behind this discussion.³ The analysis rests on exceptionally interpreting singular on Hopi verbs as fixing the subject’s number as ‘1 or 2’ (SI: 415 ex. (11)), rather than as ‘exactly one’. In other words, our analysis suggests that what seems to be singular marking on the verbs actually turns out to be something other than what is normally meant by singular. The analysis is purely compositional because all the relevant parts are assigned meanings and the meanings of compound expressions made up of these parts are determined solely by the meanings of the parts and their combination (conjunction).

Adger criticises this solution as follows:

[The compositional analysis in SI] is actually impossible in the *CS* system as it makes use of meanings of uninterpretable features – in *CS* uninterpretable features do not have meanings. So the *CS* approach is in fact more restrictive, as well as being compositional. (*MFTM*: 666)

This argument does not follow. First, in stating our alternative we do not have to accept the assumptions of *CS* regarding the ‘interpretability’ of

[3] Despite the implication in *MFTM* that Adger is concerned with empirical facts and that we are concerned primarily with formalism, Adger sets aside the ‘putative problems’ for the analysis of Hopi that we raise in *SI*, where we investigate the empirical facts of Hopi agreement in more detail than *CS* or *MFTM*. We point out in *SI* that Adger’s treatment of dual, given the theory subsequently developed in *CS*, incorrectly predicts that a singular subject and plural verb should equally give rise to dual interpretation. Such examples are in fact ungrammatical (SI: 414 ex. (10’)). Furthermore, Kalectaca’s (1978) grammar of Hopi reveals the original data to be more restricted than anticipated. Adger does not deal with these empirical issues in *MFTM* due to a lack of space. However, we hope that the passage on Hopi will be clarified in future editions of *CS*. We think that an expanded discussion is warranted: the Hopi discussion is a good example of Adger’s nice use of cross-linguistic data to make theoretical points.

features. Second, the claim of greater restrictiveness is somewhat extravagant: we could dispense with the entire distinction between uninterpretable and interpretable features and thus claim to be even more restrictive, because our system cannot even appeal to the distinction between the two feature types. Of course, such a claim would be empty in the absence of an explicit theoretical understanding of its claim to restrictiveness, as is the claim in the paragraph just quoted.

In MFTM, Adger provides an alternative formalisation that he claims to be compositional; but we question whether this is really true, under normal conceptions of compositionality.⁴ Janssen (1997: 419) observes that the usual presentation of the principle of compositionality is so vague that ‘anyone can put his own interpretation on the principle’. In other words, it is quite difficult to tell if different linguists are really operating under the same conception of compositionality. We will not here get into a lengthy presentation of what we mean by (non)compositionality; nor will we attempt to analyse what Adger understands compositionality to be. It suffices to demonstrate that the Adger/Harbour treatment of the semantics of singular is formally identical to a rather uncontroversial instance of noncompositionality (or at the very least, extremely weak compositionality).

MFTM (p. 665) provides the following interpretation for singular:

- (1) [singular, A]_N if and only if $\neg \exists N' \subset N, [A]_{N'}$

Note that the feature bundle is made up of the feature [singular] and some other, arbitrary feature [A], which would be [plural] in the case at hand. There is no independent meaning given for the feature [singular]; it only has a meaning when some other feature, A, is present. This is noncompositional, we maintain, because the feature singular is contributing no meaning on its own, but only in construction with something else. Consider a parallel case. Suppose we said that the adjective *red* has no meaning on its own, but that *red dress* has a meaning, *red book* has a meaning, and so on. We might have given a meaning for *red* in some sense, but it would not be thought of as a compositional meaning, except under an extremely weak interpretation of what compositionality means.

Let us now turn to the issue of complexity. In section 2.2 of MFTM, Adger argues that his feature theory is simpler than the feature systems of LFG and HPSG. He is again contesting a point we never make: we do not argue that LFG or HPSG has a simpler feature system than CS. Our actual claim is that a feature system that adopts only attribute-value features is arguably

[4] MFTM is unfortunately slightly misleading about SI's discussion of compositionality. The implication is that SI claims that Adger plainly STATES in CS that noncompositionality is allowed by his feature theory (hence his defense '[b]ut nowhere do I say this'; MFTM: 666). We do not make such a claim. Our claim is that CS's system has the actual consequence of allowing noncompositionality, regardless of what its author states or implies about the system.

simpler than a system that adopts a mixed privative/attribute-value system. Adger contends that a system that is ‘only privative ... is more elegant’ (MFTM: 666). However, his system is not ‘only privative’, it is mixed. Furthermore, he has given no argument for his notion of ‘elegance’ and, as such, it is only a purely aesthetic judgement.

Furthermore, as discussed above, arguments of simplicity or elegance are impossible to evaluate in the absence of an explicit theory and a formalisation. This brings us to the formalism that Adger uses to make his claims about simplicity. He proposes that a purely privative feature system can be characterised with a rule that expands a lexical item as a set of features. First, it should be noted that an attribute-value feature system can be characterised just as easily as a set of ordered pairs, i.e. a relation. Perhaps a set is in some mathematical sense ‘simpler’, but we are not aware of any such argument and a relation is, in any case, certainly a very simple mathematical object, too.

The following claim by Adger is initially puzzling, given that he realises that a binary system can be characterised like his privative system:

A non-recursive attribute-value system is of broadly the same level of complexity [as a binary system – AA&IT], but we need to index values to the attributes they are values of. We can do this using context-sensitive rules. (MFTM: 667)

Taken at face value, this is simply false: even a regular language would suffice to characterise an attribute-value feature system, let alone a context-sensitive one. Furthermore, none of the attribute-value systems that Adger specifically discusses are context-sensitive in this manner, which suggests that the claim that context-sensitivity is necessary is inaccurate.

In personal communication, Adger has clarified to us that what is intended by the passage just quoted is the claim that an attribute-value feature system will overgenerate in the absence of constraints on possible attribute-value pairs. A similar line of reasoning applies to MFTM’s subsequent consideration of recursive attribute-value structures:

[T]here is simply no straightforward way of ensuring that values can be restricted to particular attributes ... We can’t write a context-sensitive rule to do this, since non-adjacent information is required ... there is no possible context-sensitive rule which allows us to say that an attribute that is a value of a higher attribute, can be keyed to the category of that higher attribute, as the higher attribute is never its sister. (MFTM: 667–8)

It should be pointed out that in theories which use attribute-value features to model recursive syntax, such as HPSG or LFG, there is no reason to bar the kind of recursive structures under discussion. However, Adger (p.c.) again intends this claim to be about overgeneration: it is not the case that the relevant recursive structures should in general be impossible, but rather that not all such structures should be possible.

However, the argument from overgeneration holds only for a system, such as *CS*'s, in which feature valuation is unrestricted and applies freely (discussed as *FEATURE-VALUE UNRESTRICTIVENESS* and *FREE VALUATION* in *SI*: 417–420). In addition, *MFTM*'s claim that the necessity of 'non-adjacent information' renders even a context-sensitive rule insufficient is again simply false on the face of it. It was demonstrated some time ago that even a context-free grammar can capture non-local dependencies (Gazdar 1981, Gazdar et al. 1985).⁵

With these points in mind, let us consider the conclusion of Adger's argument, which no longer follows:

The solution is, then, either to allow unrestricted rewrite rules, or to type the features outside the formal feature system itself, and go with a more powerful feature theory [examples given are *HPSG* and *LFG – AA&IT*] ... [*HPSG* and *LFG*'s solutions] are perfectly reasonable solutions, but they require a theory beyond what is expressible via these simple feature grammars ... A system which does not allow feature recursion is, then, at least from this perspective on complexity, far simpler. The theory of lexical items and the theory of features proposed in *CS* is actually therefore markedly less complex than that proposed in the unification-based lexicalist theories the authors of *SI* advocate, in contrast to what is claimed in *SI*. (*MFTM*: 668)

There are three problems with this conclusion. First, the premises leading to it are questionable – at least unestablished, if not outright false. Therefore, the conclusion has not been established (the argument is not sound). The choice is not between arbitrarily powerful rules and a more powerful feature theory. It is not true that a typed feature theory (e.g. *HPSG*'s) is more powerful than an untyped one, because the typed feature theory places additional constraints on what is expressible. It is similarly not true that *LFG*'s feature theory is more powerful due to 'explicitly stating which functions can be arguments of which others' (*MFTM*: 668). As we noted in *SI* (419), *LFG* has free valuation, like *CS*'s feature theory. However, unlike the *CS* feature theory, *LFG* does not generate all attribute-value pairs freely, only those that are specified according to actual rules and lexical entries. In particular, *LFG* parses always yield the *minimal* well-formed feature structure (Kalpan & Bresnan 1982). Adger has therefore not established either of the disjuncts in his conclusion, since it not necessary to have arbitrarily powerful rules and it is not true that the alternative kind of feature

[5] Adger (p.c.) notes that his sort of feature grammar does not allow the kinds of complex category that this sort of solution rests on; but there does not seem to be any a priori reason for this. Furthermore, if there is no method for constructing such categories in the course of processing (i.e. they can only be lexically specified), then the complex categories can just be considered to be atomic like other categories and are in fact not more 'complex' at all.

theory is more powerful in the relevant sense. Second, we pointed out above that an attribute-value system can be characterised just like MFTM's privative feature system, a fact that we noted Adger tacitly acknowledges. Therefore, it is not even clearly true by his own criterion of complexity that his privative features are simpler. Third, as we pointed out above, the original point of contention was whether a mixed privative/attribute-value system of the kind Adger adopts or a purely attribute-value system is more complex. Even if the points about complexity of attribute-value systems that Adger raises were true, which they are not, then his system would, by his own admission, suffer whatever problems attribute-value systems putatively have AND it would also have the additional mechanism of privative features.

In the passage just quoted, Adger claims that we 'advocate' HPSG and LFG.⁶ We certainly did not intend our review article as an advertisement for LFG or HPSG. The following passage, we thought, makes it clear that we are not advocating any particular theory in SI (420, footnote 18):

Notice that we do not mean 'a unification-based, lexicalist framework' to have a specific denotation: we are not advocating that Minimalists shift to a particular, existing alternative. Rather, the resulting sort of theory could be purely Minimalist in spirit and keep many of its insights and, crucially, its programmatic assumptions, which are not shared by other theories of the kind advocated.

This passage indicates that we are advocating a certain KIND of theory, not particular instances, and that we judge Minimalism to be largely of this kind. We use HPSG and LFG as useful points of comparison in certain cases and, in doing so, attempt to build bridges between those theories and Minimalism. We point to some problems within MP and we offer some (to our mind, helpful) suggestions for how to address those problems. In doing so, there is not necessarily a need to adopt the rest of the theoretical machinery of HPSG or LFG, and there is also no need to abandon all the assumptions connected with MP. Adger's statement that he has demonstrated something in contrast to what we advocate in SI is thus perplexing, but is perhaps due to a lack of clarity on our part in SI.

Turning to the final point, concerning constrainedness, we note in our review article that the feature system in CS is unrestricted with respect to what values attribute-value features may receive (SI: 417–420), as already mentioned above. A concrete example from Adger's book is that *past* and *singular* can value the same feature. We also note that valuation applies

[6] MFTM, and possibly SI, might give readers not familiar with LFG or HPSG the incorrect impression that the two theoretical frameworks are very similar. LFG and HPSG both have a detailed and mathematically well-understood theory of the lexicon, but they are nevertheless in many respects very different from each other.

freely, subject to Locality of Matching. Valuation is so free that privative features can value an attribute-value feature, and a feature can value a feature with a different attribute (e.g. [*uPhi*:singular] can value [*uInfl*:]). (This is all discussed in section 4.3.2.3 of *SI*; see *CS* for more discussion of these examples and further examples.) Adger responds that ‘the system as presented is pretty constrained’ (MFTM: 670). We are guessing from reading section 2.3 of MFTM that Adger means that his theory is constrained because learners can only acquire features that they find evidence for. (‘I take (9) to be facts which English acquirers learn. I don’t think that this is unreasonable and it seems pretty restricted to me.’; MFTM: 669) Granted; but we think one can set the bar higher than this. Adger objects to the possibility of constraining features based on their content.⁷ However, there are other possibilities. It would be reasonable to allow only two features with the same attribute to value each other, for example.

Lastly, it is noteworthy that in making MFTM’s claims about constrainedness, Adger adopts fully context-sensitive rules (MFTM: 669). General context-sensitive rules are quite unconstrained and, more importantly, known to be beyond the power required by natural language, which is adequately captured by the weaker class of ‘mildly context sensitive’ formalisms (Joshi et al. 1991). However, as Adger (p.c.) points out to us, the context-sensitive rules in MFTM are not fully general, since they are non-recursive. This sort of limited context-sensitive rule (if written as the equivalent, e.g. *past* → *uInfl*:) is in fact known to denote a regular language (Kaplan & Kay 1994), and is therefore quite weak. However, to use Adger’s own reasoning from MFTM, the question then becomes how to formally constrain the system to allow only limited context-sensitivity. By the reasoning in MFTM, such a theory would not be as ‘simple’ because it would need extra constraints.

Our own perspective on constrainedness reflects a perspective that Adger takes on complexity in section 2.2:

I’d rather see how simple a system we can get away with in the structure of lexical items ... I see the right level of complexity of the feature theory as an empirical issue, and I’d rather explore the strongest hypotheses first. (MFTM: 668–9)

We agree in principle with what Adger says here, although we disagree with his specific claims about simplicity, as discussed above. We think

[7] For example, Adger sees no problem in treating [singular] and [past] as possible values of the same feature, as long as the feature in question is uninterpretable. However, he points out that interpretable features are in this respect completely different from uninterpretable features: values of interpretable features ARE restricted, by semantics (MFTM: 669). Exactly how interpretable features are restricted by the semantics still remains to be worked out, though.

that his philosophy can be extended to the issue of constrainedness. To paraphrase Adger: we think it would make sense to see how *CONSTRAINED* a system we can get away with. We would rather explore the strongest hypothesis first.

The last section of *MFTM* (section 3) addresses our suggestion that movement is no longer necessary in the Minimalist Program. Adger first addresses the issue of a derivational versus representational conception of Minimalism. It is true that certain purely representational approaches to Minimalism (e.g. Brody 1995, 2002; Koster 2003) do not have a transformational component and therefore lack anything that is clearly the operation *Move*. However, the derivational approach to Minimalism is by far the dominant one and it is as part of this approach that Chomsky (2005: 12) has declared *Move* to be ‘a virtual conceptual necessity’. Adger writes that ‘[i]n such a situation, however, a debate is surely a better way to proceed than simply removing one of the options by metatheoretical fiat’. This is a peculiar comment, given that we have indeed been engaging in a debate, stating our position clearly in *SI* and again here. Furthermore, it is Chomsky, if anyone, who has attempted to remove an option by meta-theoretical fiat with the declaration of internal *Merge* (i.e. *Move*) as a ‘virtual conceptual necessity’.

Adger goes on to write that he is ‘at a loss to understand the point *SI* makes’ (*MFTM*: 671) on the question of movement. We will reiterate our point here, since we take it to be a fairly simple one. The mechanism of *Move* adds further formal power and complexity to a grammar. Minimalists claim to be concerned with complexity. Minimalism has developed in such a way that *Move* no longer seems necessary. Therefore, *Move/internal Merge* could, and perhaps should, be eliminated in the Minimalist Program, rather than being retained as a ‘virtual conceptual necessity’. An anonymous *JL* referee suggests that another way of framing the debate is to observe that although there are feature-based or category-based theories with no transformations, Minimalism has thus far not offered a transformational theory with no features. This makes the necessity of *Move* highly questionable, since it seems that there is no basis for the claim, but substantial basis for its negation. In fact, one could ask whether all of the restrictions on features in Minimalism that are argued to be due to considerations of simplicity are truly about constraining the feature component, or whether they are rather a way to force a role for *Move* in the system. We agree with Adger that ‘*Move* [is] a perfectly reasonable hypothesis’ (*MFTM*: 671), but the thing about hypotheses is that they are supposed to be revised and, if necessary, discarded. We of course strongly disagree with *MFTM*’s contention that ‘*SI*’s statement that Minimalism should just get rid of *Move* really can be reduced to not much more than a statement that people doing Minimalism should just do *HPSG* or *LFG* instead’ (*MFTM*: 672). This is just not true, according to what we actually

wrote (see the footnote from SI quoted above), and statements like this do not help to move the debate forward.

Adger points out in MFTM that something or other is needed to capture unbounded dependencies and that any mechanism that does so is equivalent to Move. This is not true, for two important reasons. First, the mechanisms that Combinatory Categorical Grammar (CCG), HPSG and LFG use (continuing with Adger's examples) are not special mechanisms that deal only with unbounded dependencies and other transformational phenomena. The mechanism of Move is of course also used, in modern transformational theories, for phenomena where there is no displacement (in other words, 'string vacuous movement' is no longer frowned upon). So Move is in some sense not a special mechanism any more than unification is in unification-based frameworks. However, Minimalism adopts a PAIR of mechanisms that work together: Move (or internal Merge) and some type of feature matching (which some authors seem to conceptualise as, essentially, unification). We repeat our original point: a system that uses only feature matching is to be favoured (by Ockham's razor) over a system that uses feature matching AND movement, since the latter has been demonstrated to be dispensable.

There is also a second reason why we disagree with the statement that any mechanism that deals with unbounded dependencies is equivalent to Move. Move is not equivalent to equality, which is the underlying formalisation of unbounded dependencies in various other frameworks. Equality is a stronger relation than a movement transformation and is therefore more constraining. The sort of structure-sharing that HPSG and LFG use to model local and non-local dependencies is a form of token equality: the very same structure fills two positions. The function composition of CCG for unbounded dependencies similarly results in token equality. In contrast, Move is at best a weaker notion of TYPE equality, whereby different tokens of the same kind of structure fill multiple positions (the copy theory of movement). Adger (p.c.) argues that construing Move as internal Merge can be interpreted as having token equality between the input and output. This is a step towards a version of MP that we argued for in SI and we look forward to seeing an explicit and formal presentation of internal Merge with token equality.

It is true that we are of the opinion that the Minimalist Program suffers from certain weaknesses. However, we want to end this commentary by reiterating a point from SI: a genuine strength of Adger's *Core syntax* is the very fact that it is at all possible to take it to task the way we do in SI. Adger does an admirable job of turning suggestions and intuitions into explicit analyses. He creates a theory of features that is at least explicit enough to be criticised. This is a necessary step towards creating a solid theory of features, something that Minimalism very much still needs.

- Adger, D. (2003). *Core syntax: a Minimalist approach*. Oxford: Oxford University Press.
- Asudeh, A. & Toivonen, I. (2006). Symptomatic imperfections. *Journal of Linguistics* 42. 395–422.
- Brody, M. (1995). *Lexico-logical form: a radically Minimalist theory*. Cambridge, MA: MIT Press.
- Brody, M. (2002). On the status of derivations and representations. In Epstein, S. D. & Seeley, T. D. (eds.), *Derivation and explanation in the Minimalist Program*. Oxford: Blackwell. 19–41.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry* 36. 1–22.
- Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.
- Gazdar, G., Klein, E., Pullum, G. K. & Sag, I. A. (1985). *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.
- Harbour, D. (2003). *Elements of number theory*. Ph.D. dissertation, MIT.
- Janssen, T. M. V. (1997). Compositionality. In van Benthem, J. & ter Meulen, A. (eds.), *Handbook of logic and language*. Cambridge, MA: MIT Press. 417–473. [Co-published with Elsevier Science B.V., Amsterdam.]
- Joshi, A., Vijay-Shanker, K. & Weir, D. (1991). The convergence of mildly context-sensitive formalisms. In Sells, P., Shieber, S. M. & Wasow, T. (eds.), *Foundational issues in natural language processing*. Cambridge, MA: MIT Press. 31–81.
- Kalectaca, M. (1978). *Lessons in Hopi*. Tucson, AZ: The University of Arizona Press. Edited by R. W. Langacker.
- Kaplan, R. M. & Bresnan, J. (1982). Lexical-Functional Grammar: a formal system for grammatical representation. In Bresnan, J. (ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press. 173–281.
- Kalpan, R. M. & Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics* 20. 331–378.
- Koster, J. (2003). Move, merge and percolate are one! On the elimination of variables in grammar. In Delsing, L.-O., Falk, C., Josefsson, G. & Sigursson, H. A. (eds.), *Grammar in focus: a Festschrift for Christer Platzack* [*Grammatik i fokus: festskrift till Christer Platzack*] (vol. 1). Lund: Department of Scandinavian Languages, University of Lund.
- Radford, A. (2004). *Minimalist syntax: exploring the structure of English*. Cambridge: Cambridge University Press.
- Stabler, E. (1998). Acquiring languages with movement. *Syntax* 1. 72–97.

*Authors' address: Institute of Cognitive Science & School of Linguistics and Applied Language Studies, 2201 Dunton Tower, Carleton University, Ottawa, Ontario K1S 5B6, Canada.
E-mail: ash_asudeh@carleton.ca
ida_toivonen@carleton.ca*