English Language and Linguistics 3.1: 147–60. © Cambridge University Press 1999.

REVIEW ARTICLE

Back to the future

WOLFGANG STERNEFELD
University of Tübingen
(Received 25 June 1998)

Andrew Radford, *Syntactic theory and the structure of English: a minimalist approach*. Cambridge textbooks in linguistics. Cambridge: Cambridge University Press 1997. Pp. xii + 558. Hardback £50, US$69.95, ISBN 0 521 47125 7; paperback £16.95, US$24.95, ISBN 0 521 47707 7.

Until recently I was convinced that by far the best textbook ever written on Generative Grammar was Perlmutter and Soames' *Syntactic argumentation and the structure of English* (1979). Unfortunately, the theory advanced there died out. As concerns its successor, namely GB-theory, I still believe that Andrew Radford's pioneering *Transformational syntax* (1981) is one of the best introductions to Chomsky's Pisa lectures, despite the plethora of competing textbooks that have appeared over the years. Now that Radford has presented his most recent book, *Syntactic theory and the structure of English*, I am inclined to believe that it should be considered a competitor to Perlmutter and Soames' book. Not only the similarity to Perlmutter and Soames' title, but also the pedagogical and systematic orientation of the new book invites comparison. As both books meet the highest standards with regard to clarity of expression and exposition, I recommend Radford's book as the best textbook for up-to-date syntactic theory, and I am convinced that it will play the same influential role as an introduction to the Minimalist theory as did the 1981 book for GB theory.

What makes the new textbook unique among contemporary introductions is summarized by the author himself in the Preface. The six main features mentioned are the following: 1. the book does not presuppose any background knowledge of syntactic theory; 2. there is no discussion of the GB-framework; rather it deals directly with work from the 1990s within the *Minimalist Program*; 3. cross-linguistic variation is illustrated in terms of different varieties of English; 4. it contains a substantial *workbook section* at the end of each chapter, with fully worked-out *model answers*; 5. there is an extensive glossary at the end of the book, intended 'to alleviate the terminological trauma of doing syntax'; 6. the book is published alongside an abridged version called *Syntax: a minimalist introduction*.

Other features mentioned are its organization in terms of cumulative complexity; anyone who was impressed by Perlmutter and Soames' skilled arrangement of linguistic material and theoretical assumptions will also appreciate the same careful organization in Radford's book. The aim is to minimize theoretical complications

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

and to make the basic ideas and principles as easily accessible as possible. Yet each chapter becomes progressively more difficult, so that at the end the reader is in an excellent position to quickly understand Chomsky's writings and other work done within the Minimalist Program.

An equally well organized part of the book are the *workbook sections* at the end of each chapter. These comprise about 30 per cent of the entire text and provide another rich source of data. Almost all exercises deal with varieties of English, including child language and Creoles, and almost all have model answers. I have always found it difficult to make up appropriate exercises that go slightly beyond a purely mechanical application of theory to data, without at the same time leading to unsolvable research questions (such as those in Fanselow & Felix, 1987). Between mere repetition on the one hand and compilations of research programs on the other, Radford has found an almost ideal balance (although the exceeding explicitness of the model answers occasionally shifts the demands on the student's part more into the regurgital direction).

Consider, for example, exercise IX (p. 33 of ch. 1), which exemplifies the head parameter by presenting data from German like the following:

(1) (a) Hans muss auf seine Mutter stolz sein
Hans must of his mother proud be
(b) Hans muss stolz auf seine Mutter sein
Hans must proud of his mother be

The task is to identify which heads go before and which go after their complements. Exceptionally, no model answer is given, and the language in question is not English. An immediate (and correct) answer which results from a comparison of (a) and (b) is that (predicative) adjectives may vary with regard to their headedness. However, for the alert student there is much more to be detected in the further examples given, ultimately the verb-second property of German. I doubt that any student without knowledge of German will identify the correct solution, but even a partial awareness of the problems might motivate students to learn more about German and syntax in general.

The fact that such problems can already be discussed fruitfully at the end of chapter 1 shows how rapidly Radford develops the leading ideas.[1] The book has ten chapters, each of which divides into ten sections. Chapter 1 explains the goals of the generative enterprise, chapter 2 introduces lexical and functional categories; chapter 3 deals with higher constituents and phrase structure trees; chapter 4 begins with the discussion of empty categories (PRO, empty INFL, and null-determiners); chapter 5 on checking theory leads us right into Minimalist theorizing; chapter 6 discusses head movement; chapters 7 and 8 are about A- and A-bar movement respectively; and chapters 9 and 10 introduce the split-VP and split-INFL hypotheses.

Much of the attractiveness of Radford's exposition results from the sleight of

[1] This contrasts agreeably with his *Transformational grammar* (1988), which during the first half of the book somehow didn't get off the ground.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

hand with which he manages to introduce basic assumptions. His strategically clever arrangement of data and theory is surely to be admired (although it sometimes fools one into thinking that these assumptions do not require further justification). The development of ideas in each chapter proceeds carefully and piecemeal, often introduced by phrases like 'an obvious question to be asked is …', which help to complete particular analyses not fully developed in previous sections. Occasionally, however, some such obvious questions are not raised, leading to gaps in the argumentation. To illustrate, Radford faithfully recounts the standard literature on so-called floating quantifiers like *all* and *both*, arguing that these are stranded in potential subject positions as a result of moving the subject from its D-structure into its surface position. This has the (intended) effect of leading to a multiplication of empty subject positions for each potential position of the floating quantifier. However, there is no discussion of why objects do not seem to exhibit the same behavior as subjects. So why don't we find floating quantifiers with objects raising to AgrObj as in (2a), or with derived subjects as in (2b):

(2) (a) *She kissed $_j$ them$_i$ patiently [$_{VP}$ t_j [$_{DP}$ t_i all]]
(b) *They$_i$ were killed [$_{DP}$ t_i all]

It is simply presupposed that floating quantifiers must be base generated as parts of the DPs they anaphorically refer to. This, however, is an assumption that is not enforced by semantic considerations, and alternative semantic analyses – e.g., treating these expressions as modifiers of (extended) verbal projections – are readily available. Moreover, Torrego (1996) and Bobaljik (1995) have independently shown that the same syntactic restrictions that apply to floating quantifiers also apply to certain adverbials (e.g. *ever*), and crucially these do not involve movement.

It seems to me, then, that the evidence based on floating quantifiers is rather weak. The same appears to hold for other arguments in favor of VP-internal subjects in section 8.3, e.g. the argument based on the assumption that idiomatic expressions must be a constituent in D-structure. The question here is what one should conclude from the existence of idioms in German where the subject and the verb are idiomatic, while the object is not (cf. *ihn*$_{acc}$ *reitet der Teufel*$_{nom}$ 'him rides the devil' (he is possessed) and others cited in Reis (1982: 178)). As a matter of fact we observe that (3a) is much better than (3b):

(3) (a) weil die Frau der Teufel geritten hat
because the woman the devil ridden has
(b) ?*weil der Teufel die Frau geritten hat

But can we go so far as to conclude from this that the subject and the verb are base generated as a constituent that excludes the object, so that (idiomatic) German could be OSV underlyingly?

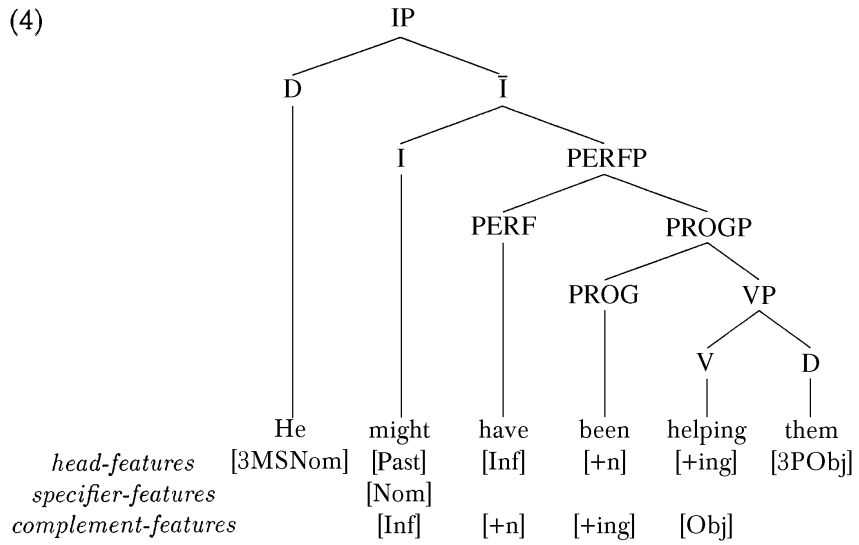
Apart from the difficulties with the *empirical* force of particular arguments (cf. also, as a paradigmatic case in point, Baker's 1991 critical discussion of negation and the unconvincing role of syntactic arguments in core grammar) I also had difficulties with the *theoretical* force of arguments, in particular those centring

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

around the notion of what Radford calls 'percolation'. The problem for me was that Radford's elegant exposition of feature checking occasionally leads to complications of the theory that might strike the astute reader as somewhat unnecessary. To illustrate what is at stake, let me briefly discuss the way Radford puts feature theory into work in chapter 5. After that I will illustrate the problem with two examples, one taken from the same chapter, and the second one (the analysis in (6) below) taken from the very end of the book, where case assignment is analyzed as specifier–head agreement.

Consider first Radford's classification of features; these subclassify into head features, specifier features, and complement features, as in the following example (p. 180):

(4)



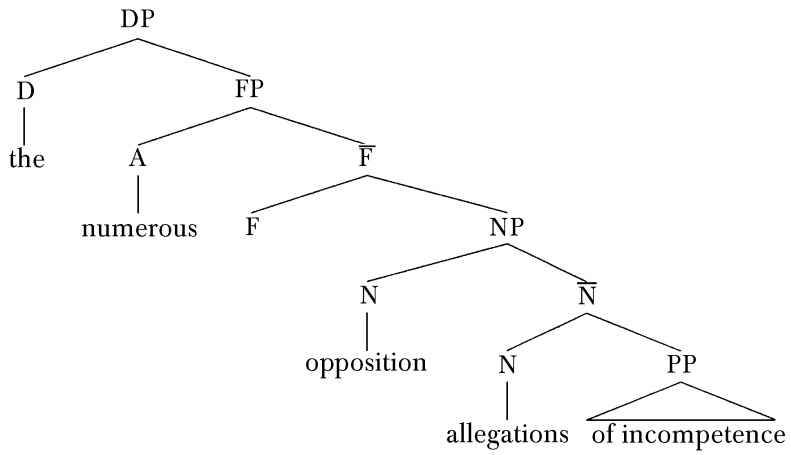
The complement features of a head *X* check the head features of *Y* iff *Y* is the head of the complement of *X*, and similarly for the specifier features. Head features need not be checked if they have a semantic interpretation, as is the case with the ϕ-features of pronouns and the tense features. Note that the tense feature of *might* is semantically misrepresented: considered as an interpreted feature, the tense feature of *might* should actually be [Present].[2] As demonstrated by Radford, a substantial part of grammatical structure can be described by an analysis that adheres to this pattern.

[2] It should also be mentioned, as Radford does on p. 199, that this kind of morphological feature mechanism makes the specification of syntactic categories in the tree almost redundant. Nonetheless one may wonder why it is that the *+ing*-VP *helping them* is not labeled as PROGP, so that *been helping them* would be a PERFP and *have been helping them* an INFINITIVE-P. The issue is simply not discussed, although an interesting point could be made: the labeling in (4) suggests that it is in fact the verb *be* that is the bearer of the progressive semantics, and similarly, it is the verb *have* that semantically expresses perfective aspect. This would indeed make some sense but would of course run counter to the morphological realization of the respective [+ing] and [+n] features. Thanks to Arnim von Stechow for having pointed this out to me.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

Consider next the analysis in (5) (p. 190):
Before going into the issue of unnecessary complications, let me briefly comment upon the perhaps unusual features of (5). Obviously, Abney's DP-analysis (Abney,

(5)



1987) is adopted right from the beginning, being developed already in chapter 2. Moreover, (5) exemplifies the fact that Radford makes extensive use of empty functional heads (like F); this will be commented upon below. Next, one might wonder why *numerous* is categorized as an A rather than an AP. The answer is that within Minimalism something like X-bar theory is almost nonexistent. The only relevant notion being the head of a construction, bar-levels have no theoretical status any more. This seems to drive Radford to the implicit notational convention that lexical items are never written as XPs, even when they are maximal projections (i.e. phrases). For the experienced syntactician this needs some adjusting to, but it shouldn't pose a problem for the unfamiliar beginner.

Turning now to some of the real problems posed by (5), note first that the analysis of *opposition allegations* – like that of *government plans* and *government attempts* in chapter 3 – is rather surprising, given the possibility that these are compounds. In fact, the usual syntactic tests – e.g. pronominalization with *one* –, the compound stress rule, and the lack of productivity typical for compounding (e.g. substitution of *university* for *opposition* seems a little bit obscure) would qualify *opposition allegation* as a single constituent, and one even finds potential awareness of the problem in exercise VI, p. 127: 'A distinction which we overlooked in the text is that between phrases and compound words.' It remains a mystery why Radford did not apply the distinction to the case at hand, and it is deplorable that the very concept of a specifier is introduced by such an unjustified analysis.[3]

[3] Fortunately, such questionable analyses are rare. It would thus be inappropriate to make it an additional exercise to find out how many errors one could discover in each chapter – as one of my colleagues used to do when teaching with Haegeman's 1991 textbook; cf. also the review in Müller & Sternefeld (1993).

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

Returning next to the issue of empty functional heads, the analysis of adjectives in (5) is introduced on p. 157 as follows: 'If we suppose that agreement canonically involves a relation between a functional head and its specifier . . ., one way in which we might handle adjective–noun agreement . . . is to suppose (following Cinque, 1994) that the adjective [*numerous*] serves as the specifier of a higher functional head . . .' The above quote nicely illustrates the conditional and hypothetical style of the argumentation; Radford always explicitly makes clear the theoretical premises of an analysis and occasionally points out the hidden costs of an envisaged solution. At this point of the argumentation, however, it is not quite clear why one should buy into an empty projection (presumably semantically uninterpretable) and an additional technical device called 'percolation'. In order to better understand why adjective–noun agreement might turn out to be problematic, Radford should have reconsidered his classification of features. According to the traditional analysis of adjectives as adjuncts, the adjective is neither a complement nor a specifier, hence adjectival agreement seems to escape a treatment along the lines of the system in (4). However, postulating the structure in (5), the adjective would indeed become a specifier of an empty head, so that it becomes possible now to consider adjectival agreement as a subcase of a specifier–head relation.[4]

This alone, however, cannot be the end of the story, because the head F itself has to agree with N. Therefore Radford assumes the aforementioned 'percolation' mechanism, saying that the heads F and D 'acquire' the relevant agreement features from N via successive cyclic 'percolation'. This concept, however, does not fall into the head–complement–specifier tripartition and it remains unclear how it fits into the checking mechanism. The more relevant point here is, however, that in previous cases of determiner–noun agreement (p. 179), Radford proposed that it is the *complement features* of D which include the number features of its complement, so that the relation between D and F could just as well be captured by the complement features of D.[5] Likewise, the complement features of F could specify the features of N, so that complement features would suffice to do the job of 'percolation'. Following Minimalist reasoning, it thus transpires that 'percolation' is an additional entirely superfluous device. So why should it be adopted?

Percolation also plays a central role in the following chapter; there it is called

[4] Following Abney (1987: 293), *numerous* could also be analyzed as a specifier of N. This analysis would do without an empty head and would be in accord with the assumption that agreement involves a specifier-head relation. It is not in accord with the above quote, however, where agreement with a *functional* head is required.

[5] Something like this is also required in a language like German, as illustrated in (i):

(i) (a) der faule Beamte
the lazy officer
(b) ein fauler Beamter
a lazy officer

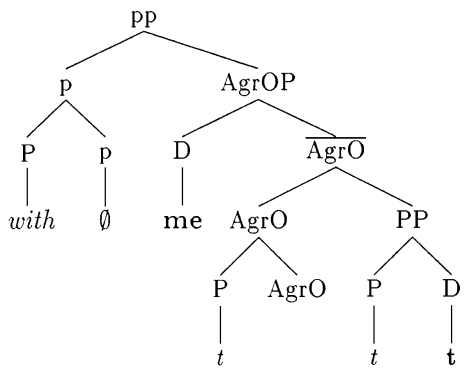
Here the determiner literally 'determines' the morphological form of adjectives and sometimes even that of some nouns. Since the inflectional forms of the adjectives are different from those of the determiners, the kind of percolation mechanism proposed by Radford would not work.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

attraction. Attraction is just another word for head-to-head movement of grammatical features, leaving behind the phonetic features needed for spellout. The paradigm case is movement of the agreement features of the finite verb into an empty INFL-head. It is argued that attraction is necessary in order to establish an agreement relation between the subject and the finite verb, again in terms of specifier–head agreement. However, given the formal mechanism introduced in (4), one may wonder again why such a mechanism is called for. After all, agreement could just as well be captured by an (empty) I-head whose specifier and complement features are identical, namely the agreement features of the subject as the specifier and the verb as the complement. As above, it seems that Radford's independently motivated feature mechanism renders percolation superfluous.

Turning now to the second example from the end of the book, consider the assumption that case assignment should also reduce to a specifier–functional-head relation (thereby denying a formally significant difference between agreement and government; indeed, the latter term no longer even exists in the Minimalist theory). For a simple PP like *with me* this assumption ultimately leads to the analysis shown in (6) (from p. 452):

(6)



Whereas the premises and the price to be paid (namely an abundance of empty heads) are fairly clear, it is less clear why the premises should be accepted.[6] Likewise, in the case of adjectival agreement, why shouldn't it be possible to establish a direct checking relation between the head N and an adjective adjoined to the projection of N? This way, empty (and often non-interpretable) projections could be got rid of in a straightforward way.

Given that the matter is not an empirical one, the usual theoretical answer would

[6] This is all the more true for languages where the case of the DP depends on the lexical choice of the preposition. An analysis like (6) might suggest different AgrOs (e.g. one for genitive, accusative, and dative respectively) that select different classes of prepositions. Assuming this, however, it would no longer be the preposition that selects a certain case, but rather the case (AgrO) that selects a preposition. Such an analysis clearly puts the cart in front of the horse. Note that in Radford's analysis it is still the preposition itself that checks case when adjoined to AgrO. Therefore case checking cannot be a pure specifier–head relation, since the preposition is, after all, only adjoined to AgrO. The argument therefore rests on an additional stipulation, namely that the adjoined head somehow functions like the head itself.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

be this: the assumptions are formulated the way they are because the theory should be as *restrictive* as possible. For example, Chomsky, in one version of Minimalism (not explicitly addressed by Radford), dispenses with checking relations between heads and complements altogether. Complements are not in the 'checking domain' of a head, and it is for this reason that attraction takes over the role of government. (Strictly speaking, then, complement features would be disallowed in such a system.) Although this theory is briefly mentioned (p. 329), its consequences for Radford's previous analyses are not explained and the issue of restrictiveness is not explicitly addressed. As admittedly theory-internal considerations are often inaccessible even for the astute beginner, she might legitimately ask: why on earth should I accept a restriction like *No Left Turn*? Exemplifying that simplification in one part of grammar usually yields complications in another, such a restriction would necessitate three right turns instead of one left turn, and six right turns to perform a U-turn. Driving in a European city for the first time, you will, after only one or two right turns, inevitably get lost somewhere in a confusing system of small one-way streets.[7]

It should be stressed at this point that Radford often enough clarifies the provisional nature of the proposed analyses, admitting that 'important questions of detail and principle remain unresolved' (p. 461). In particular, 'work on checking theory is yet in its infancy, so that many of the ideas and descriptive details in this chapter are inevitably somewhat sketchy and speculative' (p. 170). On the other hand, the concept of checking, i.e. deletion or generation of features in certain environments, is the fundamental mechanism of feature grammars that were developed about twenty years ago. Government of a complement and agreement with a specifier encountered no *descriptive* difficulties at all in these theories. So the basic issues are well known; what is really speculated about nowadays is a *theoretical* issue, namely, whether certain (*ad hoc*?) restrictions should in fact lead to monsters like (6).

In contrast, a simple feature analysis like (4) would, not very long ago, presumably have been perceived as unrevealing and inferior to Chomsky's theory of Affix Hopping (as evidenced by the almost anachronistic high prestige of the Affix

[7] Another interesting issue is whether restrictiveness automatically guarantees learnability. Following the usual *credo* of Generativists in the Chomskian tradition, Radford states that Minimalism 'is motivated to a large extent by the desire to minimize the acquisition burden placed on the child, and thereby maximize the learnability of natural language grammars' (p. 6). I would like to express doubts on the view that restrictiveness can be directly related to learnability. The fact of the matter is that despite Chomsky's invention of the *Language Acquisition Device* none of his contributions to Generative Grammar have been directly motivated by learnability; rather, the driving force was almost always theory internal; witness his *Barriers* (Chomsky, 1986) as a typical example. The present Minimalist theory is no exception and was received with even more skepticism with regard to learnability than any of its predecessors. Given that it is usually agreed that the computational burden of transderivational economy constraints is very heavy – in fact so much that it threatens to undermine any hope for positive learnability results, the usual answer is that all this need not be learned, but is part of universal grammar. Unfortunately, however, almost nothing is known about the learnability of the remaining system, i.e. of what is to be considered as the Minimalist language-specific parameters.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

Hopping mechanism in chapter 4 of Freidin's 1992 textbook). In fact, (4) is based on precisely those generalizations that were formulated in a feature-driven framework almost twenty years ago, namely those concerning the behavior of head, complement, and subject (i.e. specifier) features, plus a number of further mechanisms assumed at that time, like feature spreading within NP (adjectival and determiner agreement) or anti-head percolation of *wh*-features (pied piping).[8] Then, as now, the aim was to find reasonable restrictions for a formalism that is basically as unconstrained as phrase structure grammars were before X-bar theory. As an illustration of what is at stake, let us discuss Radford's analysis of infinitival PRO-complements.

Since government no longer plays a role in Minimalist theorizing, PRO is characterized as having a 'null case' checked by *to*. Control verbs are then characterized as selectors of an IP complement headed by null case *to* (p. 193), whereas exceptional case-marking verbs select a variant of *to* that cannot assign case. In terms of the feature analysis pursued by Radford, this implies that control verbs (as opposed to raising or exceptional case-marking verbs) are characterized as having complement features that select the variant of *to* that checks null case. This means that the control verb must be capable of not only checking the *head* feature of *to* but must also somehow have access to the *specifier* feature of *to*, i.e. the one that checks null-case. This clearly goes beyond Radford's initial architecture of the feature mechanism. But should such an extension be formally permitted? And if so, what formal means could prevent us from extending and iterating the procedure, so that heads can select for complement features of heads, which in turn may specify complement features of complements, or of specifiers, or of specifiers of complements of complements, etc.? Introducing such a recursion would clearly undermine the strict locality of the checking procedure.

Such questions were exactly the ones that puzzled me almost twenty years ago. At that time, GB theory (with its much celebrated PRO-theorem and the mechanism of S-bar deletion as a tool unavailable to the surface-oriented feature grammarian) presented itself as an attractive alternative which somehow managed to get around these murky questions and at the same time appeared to be much more explanatory. So what makes feature theory so attractive today? Is it the combination with 'restrictive' analyses like (6) that go well beyond anything imaginable twenty years ago?

Returning to the book under review, I have heard professional linguists express their disappointment because Radford only too superficially explains what Chomsky's theory is really all about. Admittedly, Radford's general characterization of Minimalism as 'a theory of grammar whose core assumption is that grammars

[8] Cf. von Stechow & Sternefeld (1981). Other cronies of feature theory had emerged with Generalized Phrase Structure Grammar (GPSG). For example, should we allow for gap-features that themselves contain such features? Are we allowed to subcategorize for categories with gaps? What combinations of features are admissible, and why and how are others to be ruled out?

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

should be described in terms of the minimal set of theoretical and descriptive apparatus necessary' (p. 515) just sounds like good methodology – any good theory will be minimalistic in that sense. As concerns specifically Minimalist issues like derivational and transderivational economy, these are touched on more or less *en passant*, and Radford is careful enough not to go into these matters too deeply. For example, there is no technical discussion or even mention of Reference Sets, Numerations, Procrastinate, etc., let alone the more problematic aspects of cyclic movement in connection with Greed. But should the beginner really be confronted with such conceptually difficult ramifications of the theory?

My answer is: no. On the other hand I found it somewhat disturbing that the appeal to economy frequently remains on a general and diffuse level, sometimes pretending that there is one single unifying 'economy principle'. Rather, economy in Radford's book amounts to basically three applications: the classical (pre-Minimalist) appeal to *do*-insertion as a last-resort device; a very brief description of Superiority and cyclical movement as instances of some kind of Shortest Move principle; and, most importantly, a characterization of weak features in terms of economy of movement, to which I will turn below. Of course, the sketchiness of the discussion is the intentional result of simplification, but still the hardcore Minimalist might complain that leading ideas, which constitute much of the intellectual appeal of Minimalism, are not really conveyed in Radford's 'light' version of Minimalism. Let me discuss some further examples.

When introducing the derivational T-model in its present Minimalist form – with its phrase formation operations, spellout, and LF- and PF-operations – Radford announces that he has little to say about PF and focusses instead on how LF-representations are computed. In fact, the only relevant LF-operations are movements of grammatical (i.e. syntactic) features. For these, however, it seems immaterial whether or not they are moved on the way to spellout or afterwards. As a matter of fact, then, spellout and LF collapse, so that, contrary to appearance, neither LF nor PF plays any substantial role in the model.

Personally, I consider this a welcome simplification. On the other hand, much of previous Minimalist reasoning (e.g. case checking *as an LF operation*, the contrast between Pesetsky's *Earliness Principle* and Chomsky's *Procrastinate*, the idea that parametric variation is the result of *at which stage* in the derivation *of LF* the operation of spellout applies, etc.) either does not fit into the conception of the book or has to be reformulated in an alternative way. Parametric variation, for example, is recast as the question of whether a certain movement operation, namely the one triggered by 'strong' features, involves pied-piping of PF-features (which results in *overt* movement) or whether grammatical features, namely the 'weak' features, move alone, leaving behind PF-features for the purpose of spellout (which is correctly called *covert* movement, rather than LF-movement). In consequence, the difference between spellout and LF becomes quite meaningless, as does the idea that parametric variation has something to do with the language-specific stage of a derivation where spellout applies. Again, I would like to stress that I do not want to

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

object to this result, though I could imagine that it might deprive Minimalism of some of its intellectual appeal.[9]

A side effect of taking this line of thought is that considerations of economy come into play in a very specific way, namely in order to prevent PF-features from travelling along with weak features. Although there is of course nothing incorrect with this application of economy, I have doubts whether it may serve to illustrate the more general conception in an optimal way. The problem is that the beginner might conceptualize the distinction between 'strong' and 'weak' as two different operations, one that pied-pipes PF-features, and another one that – *by definition* (sic!) – does not. So once having understood these two different options, why do we need any recourse to economy? Once the concept of pied piping is introduced as a mechanism to implement overt movement, why not introduce something like its counterpart, say 'monadic' movement, which is simply defined in such a way as to leave other features behind? As mentioned above, the distinction between 'weak' and 'strong' represents the most frequent illustration of economy; it is therefore deplorable that the role of economy in this context is in danger of being undermined by a likely misunderstanding of the difference between strong and weak.[10]

Another feature of Radford's analyses – to be deplored by the radical Minimalist – concerns the fact that Chomsky's cyclical theory of *Bare Phrase Structure* (cf. Chomsky, 1995a) is not implemented, so that most derivations are presented in a quasi 'anti-cyclic' way, as if Deep Structure were still existent. This seems again to be the result of intended simplifications, and indeed Radford might have had good reasons for not adopting Chomsky's system. For example, it could not be maintained that V raises to I (in Old English) for reasons of greed, i.e. in order to check features *with the subject*, because at that stage of the derivation, i.e. after V-to-I movement, the subject would not yet be generated in Chomsky's system (in fact, V checks its features against I via adjunction to I, which is much more complicated than Radford's simplifying exposition would like to have it).

On the other hand, bearing in mind Bare Phrase Structure, it will become

[9] One might object that the very idea that makes such a result possible, namely that certain features can, so to speak, travel on their own, is a major achievement of Minimalism. Strictly speaking this is not quite correct, as the idea has been around for quite a while. It was formulated in a seemingly paradoxical way in Baker's theory of *Incorporation without Incorporation*, cf. Baker (1985). At that time, a number of linguists (including Baker and myself) were already discussing the possibility of doing LF movement 'at S-structure', but the idea was probably considered too far-fetched (and inferior to the well-established mechanism of LF-movement) to be taken seriously and put in print. What is in fact new and specific to Minimalism is the extent to which the difference between abstract and overt incorporation-like processes is exploited in the explanation of language variation.

[10] It is sometimes hard to make the terminological difference precise. This is because the term is applied ambiguously to features and to categories. Categories are said to be strong if they cannot remain empty. I also found it strange that the difference between strong and weak is duplicated by an abstract distinction between bound and free forms. That is, movement of α to a functional projection β can be enforced by either saying that α has strong features, or by saying that it is a bound form that needs β as an affix. The latter distinction can potentially replace the former, in particular, since it seems to me that free forms cannot attach to a functional projection precisely because they have to remain free. But if this is so, why make such a fuss about features and economy, one might ask.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

apparent that Radford's 'theoretical justifications' of traces no longer holds once the system is adopted. The concept of a trace is introduced by reference to the movement of the modal *will* into the C-position, as in ***Will*** *we* ***t*** *finish the rehearsal?*. Radford's first argument for assuming a trace is that the trace 'must have the same complement features as *will*, since the head V constituent of the VP has to be in the infinitive form . . .' (p. 220). In other words, complement features on the trace are needed for the purpose of subcategorization. But this is not really true in *Bare Phrase Structure*, where these features are checked and therefore disappear as soon as the complement merges with the head, that is, *before movement*. So there is no motivation for there being a grammatical feature on a trace. Turning to his second theoretical motivation, the existence of traces is justified by the c-command requirement on movement, which is assumed in order to guarantee raising and to block lowering. This requirement is treated as an independent axiom, although it would follow from the theory of Generalized Transformations as a theorem. In that theory, movement is always upwards and would still be so even if movement were allowed not to leave a trace. Finally, his only 'empirical' argument in favor of traces is rather weak; it relates to the contrast in (7) from p. 221:

(7) (a) (i) We have finished the rehearsal
(ii) We've finished the rehearsal
(b) (i) *Will* we *t* have finished the rehearsal?
(ii) **Will* we've finished the rehearsal?

As with many other so-called empirical arguments we need a further (hitherto not introduced) theoretical assumption, namely that traces block cliticization. Accordingly, the trace in (7bi) blocks (7bii). My point is not that this argument already presupposes the existence of traces; it is rather that an alternative description would do the same job. Since clitics are drawn from the lexicon as bound affix-like forms – as Radford also suggests in the case of the clitic *not* = *n't* – one might argue that, for some reason or other, verbal clitics always bear agreement (and possibly tense) features (i.e. features that help to identify them). This assumption immediately explains why (8) is ungrammatical:

(8) *We may've finished the rehearsal (with /v/ rather than /əv/)

More importantly, it also accounts for (7b), simply by the fact that an infinitive form of '-*ve* with vowel-less /v/ does not exist in the lexicon.[11] But if so, none of the arguments in section 6.3 ('Traces') survives. And in fact, apart from *wanna*-contraction (cf. Pullum (1997) for the most recent counterargument against the role of traces in preventing contraction) no further arguments are presented. In particular, the classical arguments from binding theory cannot be applied, because

[11] As far as I can see (as a non-native speaker) the '-*ve*-form (with vowel-less /v/) only cliticizes to pronouns, thereby exhibiting selectional properties of affixes (with the additional complexity that phonology also plays a crucial role). If so, arguments in favor of traces in other positions do not hold either, e.g. the one on p. 270:

(i) *Which students would you *say've* got most out of the course?

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

binding theory is not mentioned in the book. Of course, I wouldn't doubt the existence of traces in general, e.g. for semantic reasons, but these naturally lie outside the scope of the book.

To summarize, I have tried to make clear that many of the objections that could be raised against particular features of the analysis are the result of skillful, intentional simplifications. For the beginner, there is thus every reason to greet this book with enthusiasm. On the other hand, Radford's excellent presentation ironically begs the question of how much of Minimalism should be simplified away in order to stay consistent with common sense. It also raises the question of how much theoretical progress has been made, in particular, when considering (a) the elimination of government, (b) the renaissance of subject-to-object movement in the guise of movement of the subject into some AgrO-projection, and (c) the lack of motivation for traces. Seen from this perspective, the theoretical distance between Minimalism and the cyclic theory of Perlmutter and Soames is much shorter than one might expect. On the other hand, comparing Radford (1981) with Radford (1997) there is not much theoretical overlap, so that it is not easy to say which theoretical assumptions are the more explanatory. Both books seem to exhibit roughly the same explanatory depth, but when measured in terms of the amount of data covered by them, the new book might win out. Whether any substantial theoretical progress has been made we might be in a position to evaluate perhaps in another twenty years.

Author's address:
Universität Tübingen
Seminar für Sprachwissenschaft
Wilhelmstr. 113
D-72074 Tübingen
Germany
Sternefeld@uni-tuebingen.de

References

Abney, S. (1987). The English noun phrase in its sentential aspect. PhD thesis, MIT, Cambridge, MA.
Baker, C. L. (1991). The syntax of English *not*: the limits of core grammar. *Linguistic Inquiry* **22**: 387–430.
Baker, M. (1985). Incorporation: a theory of grammatical function changing. PhD thesis, MIT, Cambridge, MA.
Bobaljik, J. D. (1995). Morphosyntax: the syntax of verbal inflection. PhD thesis, MIT, Cambridge, MA.
Chomsky, N. (1986). *Barriers*. Cambridge, MA: MIT Press.
Chomsky, N. (1995a). Bare phrase structure. In G. Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell. 383–439.
Chomsky, N. (1995b). *The Minimalist Program*. Cambridge, MA: MIT Press.
Fanselow, G. & S. Felix (1987). *Sprachtheorie. Band 2: Die Rektions- und Bindungstheorie*. Tübingen: UTB Francke.
Freidin, R. (1992). *Foundations of generative grammar*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press

Haegeman, L. (1991). *Introduction to Government and Binding Theory*. Oxford: Basil Blackwell.
Müller, G. & W. Sternefeld (1993). Review article: Liliane Haegeman: *Introduction to Government and Binding Theory*. *Linguistics* **31**: 1111–39.
Perlmutter, D. & S. Soames (1979). *Syntactic argumentation and the structure of English*. Berkeley: University of California Press.
Pullum, G. K. (1997). The morphological nature of English *to*-contraction. *Language* **73**: 79–102.
Radford, A. (1981). *Transformational syntax: a student's guide to Chomsky's Extended Standard Theory*. Cambridge: Cambridge University Press.
Radford, A. (1988). *Transformational grammar*. Cambridge: Cambridge University Press.
Reis, M. (1982). Zum Subjektbegriff im Deutschen. In Abraham, W. (ed.), *Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*. Tübingen: Gunter Narr Verlag. 171–212.
Torrego, E. (1996). On quantifier float in control clauses. *Linguistic Inquiry* **27**: 111–26.
von Stechow, A. & W. Sternefeld (1981). A modular approach to German syntax. Unpublished MS, Konstanz University.

https://doi.org/10.1017/S1360674399000167 Published online by Cambridge University Press