J. Linguistics **38** (2002), 347–374. © 2002 Cambridge University Press
DOI: 10.1017/S0022226701001408 Printed in the United Kingdom

REVIEW ARTICLE

Remarks on introductory syntax[1]

ROBERT FREIDIN

Princeton University

(Received 16 February 2000; revised 8 March 2002)

Peter Culicover, *Principles and parameters: an introduction to syntactic theory*. Oxford: Oxford University Press, 1997. Pp. xvi + 444.

James D. McCawley, *The syntactic phenomena of English* (2nd edn.). Chicago: University of Chicago Press, 1998. Pp. xxiv + 809.

Andrew Radford, *Syntactic theory and the structure of English: a Minimalist approach*. Cambridge: Cambridge University Press, 1997. Pp. xi + 558.

Ian Roberts, *Comparative syntax*. London: Arnold, 1997. Pp. ix + 307.

An introductory course on syntax is all about the nature of syntactic representations – providing a framework through which we interpret the facts at our disposal. However one approaches the teaching of introductory syntax, giving a clear presentation of the fundamental concepts of syntactic analysis is the first prerequisite. This requires a concern for a clear and rigorous formulation of the tools of linguistic analysis.[2] As Chomsky noted in the preface to *Syntactic structures* (1957: 5),

> [t]he search for rigorous formulation in linguistics has a much more serious motivation than mere concern for logical niceties or the desire to purify well-established methods of linguistic analysis. Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of the inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those

[1] Many thanks to Bob Borsley, Carlos Otero, Maggie Tallerman and two anonymous *JL* referees for helpful comments on the previous version.

[2] Note that the degree of formal rigor is only justified to the extent that it has empirical consequences; see Chomsky 1990.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

> for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects. I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory may have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it strictly to linguistic material with no attempt to avoid unacceptable conclusions by *ad hoc* adjustments or loose formulation.

It has always seemed to me that this point of view is just as important for teaching introductory syntax as it is for carrying out research at the frontiers of the field.[3] Thus both from the point of view of the beginning student as well as the teacher of introductory syntax, clarity in the presentation of the fundamental concepts of analysis requires formal rigor.

From the perspective of transformational generative grammar, the fundamental concepts can be divided into three distinct topics: phrase structure, transformations, and conditions on rules and representations.[4] These topics constitute the narrow focus of the discipline, where the wider focus treats these topics as part of the knowledge of language which underlies language use, hence a topic in cognitive psychology and human biology more generally. This review article will consider the four syntax texts listed above from the point of view discussed in the preceding paragraph. Although students can learn a great deal of syntax from each of these texts, they unfortunately contain certain weaknesses in interpretation and presentation that detract from their effectiveness as introductory textbooks.

Considered as introductions to transformational syntax, the four books

[3] Under the psycho–biological view of linguistic theory there is an even more compelling reason to adopt this formal perspective – namely, given that each person possesses a mental grammar of the language he or she speaks, there is a fact of the matter as to the exact character of that grammar. It is formulated in one precise way and not in a variety of conceivable other ways. Ultimately it is an empirical question what the correct formulation of each person's grammar may be, as well as the correct formulation of the initial state of the language faculty for the human species universally that allows each person to acquire a language.

[4] Notice that the distinction between phrase structure and transformations has been significantly blurred, if not entirely obliterated, under bare phrase structure theory (Chomsky 1995a), where phrase structure is constructed via the concatenation transformation Merge. When phrase structure was thought to be created via phrase structure rules, and the output of the phrase structure component was then subject to the transformational component, yielding what was called DERIVED CONSTITUENT STRUCTURE, there was a principled distinction between phrase structure and transformations. This is now virtually impossible to reconstruct, especially because derivations of sentences involve applications of Move (the apotheosis of the transformational component) interspersed with applications of Merge. The distinction between underlying and derived phrase structure is essentially meaningless from this perspective. Nonetheless, for the purposes of expository convenience I will continue to pretend that there is a distinction between phrase structure and transformations.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

under review show a remarkable diversity in both approach and material covered. The Culicover, Radford and Roberts texts are basically introductions to transformational grammar and contemporary syntactic theory, in contrast to the McCawley text, which claims to be 'primarily about the facts of syntax (mainly, indeed, those of English syntax), and only secondarily about linguistic theory' (1). The linguistic theory that McCawley develops is 'a highly revisionist version of transformational grammar' (which, he notes, 'probably no one other than myself accepts in all its details' (xvi)), in contrast to Culicover, Radford and Roberts, who all adopt a version of syntactic theory within the Principles & Parameters framework, though developed from distinctly different points of view. The most radical of the latter three is Radford, who proclaims his text to be 'an intensive introduction to recent work in syntactic theory (more particularly, to key concepts which are presupposed in works written within the broad framework of the minimalist program in the version outlined in Chomsky 1995b)' (xi). Roberts and Culicover, in contrast, are much more circumspect with respect to minimalist analyses. In Roberts the material relating to the minimalist program is limited to an eleven-page section on a checking theory for case plus a few additional references; Culicover's treatment is restricted to his tenth and final chapter, 'Head movement and Minimalism'. McCawley, perhaps unsurprisingly, ignores the minimalist program entirely – his text's references to Chomsky's published work ending with Chomsky (1986b) (not including Chomsky (1986a)).

The coverage in these texts reflects various perspectives and purposes. Thus several chapter headings in the McCawley text appear to focus on syntactic phenomena in English (in particular, chapters 8 (Auxiliary verbs), 9 (Coordination), 11 (Anaphora), 12 (The structure of noun phrases), 13 (Relative clauses), 14 (Interrogative clauses), 16 (Other types of nondeclarative sentences), 17 (Negation), 18 (Scope of quantifiers and negation), 19 (Adverbs), 20 (Comparative constructions), 21 (Discourse syntax), and 22 (Patches and syntactic mimicry)). The other chapters concern McCawley's theoretical framework for analysis (in particular, chapters 2 (Overview of the scheme of syntactic analysis adopted below), 3 (Tests for deep and surface constituent structure), 4 (Some subject changing transformations), 5 (Complements), 6 (Rule interaction), 7 (Syntactic categories), and 15 (Principles restricting and extending the application of transformations)). Thus, in spite of McCawley's claim to the contrary, his book is equally about English syntax and the version of linguistic theory he uses to analyze the phenomena covered. McCawley's text covers more syntactic phenomena than the other texts; but this does not make it any less theoretical.[5]

[5] See David Lightfoot's excellent review of the 1988 first edition in this journal (1991: 267–274) for further discussion of this issue.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

Moreover, the emphasis on describing the phenomena IN THEORETICAL TERMS without careful explication of the theoretical apparatus in which the descriptions are formulated works to the detriment of the reader.[6]

In stark contrast, the Culicover and Radford texts make no reference at all to the descriptive categories that serve as chapter headings in McCawley. Instead, the chapter headings refer primarily to theoretical constructs that serve as descriptive categories at a higher level of abstraction, as can be seen from table 1.

Chapter	Culicover	Radford
1	Foundations & methods	Principles & parameters
2	Arguments, government, & case	Categories & features
3	Binding theory	Syntactic structure
4	A-Movement	Empty categories
5	X′-Theory	Checking
6	A′-Movement	Head movement
7	Barriers	Operator movement
8	LF representations	A movement
9	Binding and logical form	VP shells
10	Head movement & minimalism	Agreement projections

Table 1

The Roberts text is more narrowly focused, having only five chapters: 'Categories and constituents', 'Case and agreement', 'Binding', 'Locality' and 'Principles, parameters, and language acquisition'. Exactly how the chapter headings of the three relate to those of McCawley's text is a nontrivial puzzle. Moreover, the Culicover, Roberts and Radford texts are also disjoint in their coverage. Perhaps the most obvious difference is that Culicover devotes two chapters (3 and 9) to the topic of binding and Roberts, one, whereas in Radford it is barely mentioned – as a three-page illustration of the c-command relation in chapter 3. On the other hand, Radford devotes an entire chapter to VP shells, whereas Culicover spends less than ten pages (mostly tree diagrams) discussing them and Roberts does not mention them at all. The point of all this comparison is simply to illustrate that there is a significant disjointness of coverage among the four texts.

[6] This applies to both the beginner with little if any knowledge of syntactic analysis and readers who have taken a course on transformational grammar. The reason this may hold for the latter group is that McCawley's syntactic descriptions are out of tune with most current work in generative grammar, for reasons to be discussed below.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

Given this general, albeit superficial, characterization of the texts, we turn now to the ways they present the fundamental concepts. Before considering the three technical topics, it is worth looking at how the four texts situate the study of syntax in its broader context, thereby identifying the object of inquiry.

1. The object of inquiry

The thrust of most work in transformational generative grammar for at least the past quarter century has been to show that all languages share a common structure and that variation across languages can be explained on the basis of a limited range of parameters. Furthermore, the work on comparative grammar within the principles and parameters framework has led to the bold conjecture that apart from Saussurean arbitrariness and lexical idiosyncrasies, 'there is only one computational system and one lexicon' (Chomsky 1993: 3) – essentially that there is only one human language.

It is therefore surprising to find the introduction to *Comparative syntax* announcing that the 'core of the matter for us' is the question 'what is different about the minds of speakers of different languages?' Roberts employs a brilliant pedagogical strategy to demonstrate the differences between languages (in the teleological, normative sense). First he cites a normal English sentence; next, the words of that sentence scrambled into word salad; and then three more scrambled versions of the English sentence that are clearly not English, but are more comprehensible than the word salad example. He then shows that if we translate the English words into German, Welsh, and Japanese respectively, that's exactly how the first English sentence would be expressed in those languages. The emphasis is on parametric variation instead of properties shared in common in spite of apparent variation. Universal Grammar is introduced a few paragraphs later to account for the universality of the mental capacity that allows native speakers to recognize sentences of their language, which is not exactly the notion that human languages are essentially the same when viewed from the appropriate level of abstraction.

The introduction also covers the important I-language/E-language distinction, first explicated in Chomsky (1986a). Unfortunately, here too what we read seems a fair distance from what we find in Chomsky. Consider for example, 'I-language, then, is the mental faculty that underlies the knowledge and use of language' (5). First, this describes UG and whatever else in the initial state of the language faculty accounts for language use. Second, an I-language constitutes the knowledge an individual speaker has of the language he speaks. Chomsky proposed the term to establish a scientific concept of language that escapes from the ill-defined normative, common sense notion. Unfortunately, the discussion in this section merely amalgamates the two concepts, as the following quote illustrates. 'Linguists'

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

theories of given I-languages (for example, the I-language in the mind of individuals commonly referred to as "English speakers", "French speakers", etc.) are grammars (of "English", "French", etc.)' (5–6). Under Chomsky's formulation any two speakers of English (even siblings) may have different I-languages.

The Culicover text also misrepresents the I-language/E-language distinction. Like Roberts, he misidentifies I-language as 'the character of what is in the human mind that accounts for our ability to acquire, speak, and understand language' (2). Moreover, he compounds the confusion by virtually equating I-language/E-language distinction with the linguistic competence/linguistic performance distinction. Thus competence is defined as 'the linguistic knowledge that a native speaker has, idealized away from actual situations of use; that is, I-language' and performance is identified as 'what a native speaker does, which in essence produces E-language' (2). Recall that Chomsky proposes I-language and E-language as technical notions of language developed for various scientific approaches. They lead to very different views about grammar and Universal Grammar. Moreover, Chomsky considers the notion of E-language to have no status in linguistic theory (1986a: 151). Culicover's equation thus leads to the absurd view that the concept of language (and hence of grammar and Universal Grammar) depends on whether we are investigating linguistic knowledge or the behavior that results when it is put to use. (See Smith (1999: 37–38) for additional discussion on why the two distinctions cannot be equated.)

Defining I-language as the knowledge an individual speaker has of the language she speaks is reasonable as a first approximation, but ultimately fails to convey the more precise meaning of the term. An I-language is a state of the language faculty in the mind of an individual, a mental state. It is primarily the result of acquiring a lexicon and fixing the parameters of the initial state of the language faculty (principally UG). Presumably, much of the structure of an I-language is already specified in UG, so it is crucial to clarify the relation between the initial state and the I-language attained to give a precise definition of the term.

Radford correctly identifies I-language as 'the internalized linguistic system' but does not make the important connection between an I-language and UG. What he does say about principles of UG in section 1.6 is far from enlightening. His sole example of a principle concerns the structure-dependence of rules (Chomsky 1975 and elsewhere), which he formulates as 'all grammatical operations are structure-dependent' (14). One problem with this discussion is that the notion 'grammatical operation' is introduced without the necessary explication. The discussion prior to the introduction of the principle only talks of operations that move words to the front of a string. This may not be a serious problem for a reader who already has some knowledge of transformational grammar, but given that the book is supposed to be 'suitable for true beginners' one wonders how such readers

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

are going to make sense of this principle. Chomsky's discussion compares two different formulations of the rule for forming yes/no questions, one which moves the first auxiliary in a sentence to sentence-initial position and the other which moves the auxiliary following the subject NP to that position. (Radford never makes this crucial comparison.) Only the second formulation involves an analysis of hierarchical structure (the notion NP) and therefore is structure-dependent. Because the first involves only linear order, it is structure-independent. This whole discussion makes sense within a theory of transformations that was laid to rest almost a quarter of a century ago. That is, the grammatical operation for forming yes/no questions was considered to be structure-dependent because the formulation of the transformation contained NP as a constant term in its structural description. With the advent of Move α such formulations were no longer available. What prevents the fronting of an auxiliary from a relative clause in subject position (e.g. **Is the boy who__happy is here?*) is not the formulation of Move α, but rather whatever principle of grammar blocks the movement. (See Freidin 1991, Lasnik & Uriagereka (2002) for further discussion.) This is going to be true in particular of Radford's account of head movement. So what Radford has provided as the introduction to the notion of 'principle of UG' is a poorly explicated piece of history that is irrelevant to the theory of grammar developed in the rest of the book.[7]

Furthermore, much of Radford's discussion in the introductory chapter misrepresents the concept of I-language and its ramifications. Consider in particular the definitions of grammar and UG when we take language to be I-language. According to Chomsky, 'the grammar would then be a theory of the I-language, which is the object under investigation' (1986a: 22). UG 'is construed as the theory of human I-languages, a system of conditions deriving from the human biological endowment that identifies the I-languages that are humanly accessible under normal conditions' (ibid. p. 23). Now compare this to how Radford uses these terms:

> While the concern of the descriptive linguist is to devise grammars of particular languages, the concern of the theoretical linguist is to devise a *theory of grammar*. A theory of grammar is a set of hypotheses about the nature of possible and impossible grammars of *natural* (i.e. human) languages … (4)

[7] The reason for the weak explication becomes clear in the opening of the second chapter where the 'structure dependence principle' is simply misrepresented:

> … the structure dependence principle determines that all grammatical operations in natural language are category-based (so that any word-based operation will apply to whole categories of words rather than to specific individual words) (37).

As Chomsky's discussions of the structure-dependence of rules clearly demonstrate, the distinction between word-based and category-based operations is not at issue. Whether we move the auxiliary as a word or a category will not prevent the classical deviant examples.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

and on the next page:

> So, what we mean by saying that *universality* is a criterion of adequacy for a theory of grammar is that a theory of grammar must enable us to devise a descriptively adequate grammar for every natural language: in other words, our ultimate goal is to develop a *theory of Universal Grammar*.

First it should be noted that the precise interpretation of these passages remains hidden because Radford does not state what he means by the crucial term 'grammar'. Nonetheless, the division of labor Radford assigns to the descriptive linguist vs. the theoretical linguist is revealing. Radford's formulation of a strict division between devising a grammar of an I-language and constructing a theory of human I-languages is rejected both in the literature of generative grammar from the outset and by contemporary practice. Constructing a grammar of an I-language crucially involves paying attention to UG and often contributes something new to it. What Radford calls 'a theory of grammar' is in fact UG, even though, as illustrated in the second quoted passage, he seems to be saying that they are separate. Note too that he gets the notion of universality backwards. UG is universal in the sense that it is a genetic property of the species. The idea that UG is universal because it correctly describes all natural languages is a concept that comes from the study of E-language (see Chomsky 1986a: 20). So it appears that Radford is mixing up the two concepts, even when he is supposed to be elucidating the concept of I-language.[8]

Confusion about the object under investigation, 'a theory of the I-language' as Chomsky puts it, also shows up in Radford's claims about how his text differs from others. Unlike other texts, which illustrate crosslinguistic variation by comparing English with various other languages (in the teleological normative sense), Radford uses only 'different varieties of English' (e.g. Belfast English, Jamaican Creole, Child English, Shakespearean English, etc.). 'Hence it does not make the unrealistic assumption that the reader knows (for example) Spanish, German, Arabic, Chinese and Chuckchee' (xi). Obviously, no introductory text cites examples from other languages based on this assumption.[9] Moreover, Radford's implication that the reader knows (in the relevant sense of I-language) all of these 'varieties of English' is simply ludicrous. Furthermore, the list is not at all homogeneous. While Belfast English and Jamaican Creole can be identified as I-languages of speakers from Belfast and Jamaica respectively, Shake-

[8] Radford's previous syntax text (1988) also failed to define the term 'grammar' and similarly confused the notion of universal grammar. See Freidin 1989 for discussion.

[9] However, it is important when citing foreign language data (especially in introductory texts) to make sure that the syntactic analysis of examples is sufficiently clear so that the reader can easily see how the syntactic properties of the examples cited relate to the concepts of analysis under discussion. Merely citing foreign language examples without providing the requisite analysis would be a serious defect in an introductory text.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

spearean English cannot for obvious reasons. Furthermore, Child English is a virtually meaningless category if it refers to all children who are acquiring some variety of English. The state of the language faculty of these children depends crucially on what variety of English they are acquiring. The fact that there is some overlap in vocabulary and syntax between these so-called varieties is, if anything, misleading.

Exactly what concept of English McCawley is writing about is not entirely clear. He claims that 'the phenomena taken up in this book are primarily mental in nature' and that 'the rules and structures that are proposed below are hypotheses about what users of a language know and what faculties play a role in their use of language' (8), which is compatible with the I-language concept. But he also says that 'the grammars developed below can be regarded as specifying the membership of a set', defined as 'what complexes of sentence, meaning, context, and style/register the given language allows', which suggests the notion of E-language.

2. Phrase structure

Because the interpretation and generation of phrase structure representations lie at the heart of generative syntax, it is essential that an introductory text be extremely clear about the fundamental concepts of phrase structure, including 'lexical item', 'head', and 'phrase' as well as the relations between them. Of equal importance is a precise formulation of the mechanisms that construct phrase structure representations as well as whatever constrains the application and output of such mechanisms.

Although it may be uncontroversial to say that the individual items that make up the mental lexicon constitute the basic building blocks of syntax, exactly how we characterize the notion 'lexical item' is far from obvious. Culicover, Radford and Roberts all equate 'lexical item' with 'word', while McCawley only uses the term 'word'. However, as Culicover points out in footnote 2 of chapter 2, sometimes a lexical item 'may be a part of a word or an idiomatic expression consisting of more than one word' (403). While Culicover, McCawley and Roberts appear to use the term 'word' in the standard way, where, in the lexicon, it specifies idiosyncratic phonological, syntactic and semantic properties, Radford presents a bizarre interpretation. On page 44 he says that 'although a word like *interesting* can function as a verb [example omitted], it can also function as an adjective [example omitted]'. The notion 'word' is thus reduced to spelling. That this is what Radford intends becomes clear when on page 52, following a discussion of infinitival *to*, he refers to 'an alternative use of *to* as a preposition'.

In transformational generative grammar there have been three basic proposals about how lexical items enter syntactic representations. In the earliest generative grammars (for example, Chomsky 1957), lexical items are inserted into derivations via phrase structure rules. Thus a lexical category

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

like N rewrites as a noun like *clarification*. Given that the rewrite relation is generally interpreted as a constituency relation, it was assumed that the item *clarification* was a constituent of the lexical category N and thus the category N dominated the noun *clarification* in the resulting phrase marker. In the mid 1960s, lexical insertion via phrase structure rule was abandoned for a variety of reasons (see Chomsky 1965 for discussion). Instead, lexical items were inserted via a transformational substitution operation that replaced a partially specified feature matrix with the fully specified feature matrix of the lexical item itself, consisting of its phonological, syntactic and semantic features. What corresponded to the lexical category of the earlier phrase structure grammars was the categorial feature of the partially specified feature matrix. Lexical insertion via substitution simply replaced what corresponded to the lexical category with the lexical item itself, and thus there could be no constituency relation between a lexical category and the lexical item that instantiated it. (See Freidin 1992 for further discussion, where this point was made explicit, apparently for the first time.) This should have been clear for the past thirty-five years, yet even today we find texts in which lexical items are represented as constituents of lexical categories. Culicover, Radford and Roberts all make this mistake, as does Haegeman (1994), Hayes et al. (2001) and many others. McCawley, it should be noted, is an exception, along with Carnie (2002). This perspective is inescapable in bare phrase structure theory (Chomsky 1995a), the third alternative for lexical insertion, where a lexical item is simply concatenated either with another lexical item to produce a phrasal construction or with some phrasal construction so produced.

The first two alternatives rely on phrase structure rules, whereas bare phrase structure theory eliminates phrase structure rules as a grammatical device for constructing phrase structure representations. Phrase structure rules represent an essentially top-down analysis of sentence structure. In contrast, the binary merge operation of bare phrase structure theory gives a fundamentally bottom-up analysis, which captures the bottom-up concept of phrasal projection from a lexical head, the core concept of X-bar theory since its inception in the late 1960s. Moreover, Merge is a transformational operation.

It is therefore surprising to find that even though Radford adopts binary merger as the mechanism for creating phrase structure and explicitly acknowledges that phrases so formed are projections of a lexical head (88), he still represents lexical items as constituents of lexical categories.[10] It is also

[10] Note that Radford distinguishes category labels in a tree from words, calling the former NODES and the latter LEAVES of the tree. However, he distinguishes a lexical category node from the lexical item that projects it – the error in question. Even if we note that the branch of the tree connecting a lexical category node to a lexical item constitutes a different relation from one which connects a node to another node (in Radford's terminology), there is first of all no motivation for separating the lexical item and its categorical designation

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

peculiar that even though he has adopted the mechanics of bare phrase structure at the outset, he introduces bare phrase structure as a separate topic almost halfway through the book (section 9 of chapter 5 on checking). The discussion in this two-page section is sketchy, inaccurate, inconclusive, and essentially unhelpful. The inaccuracy concerns his allegation that there is a 'potential notational inconsistency' in representing lexical items as constituents of lexical categories. According to Radford, the categorial property of a word would be represented by the category label while other grammatical properties are attached to the word itself. Of course this doesn't hold because the word contains its categorial features as well as the others (cf. Chomsky 1965). Moreover, the representations he has been assuming are illegitimate because binary merge cannot produce the unary branching they involve. Culicover makes the same error in his discussion of bare phrase structure (see section 3.2 of chapter 10).

The relations between lexical item, head and phrase are perfectly clear. A lexical item constitutes the head of a phrase containing it if its categorial features are projected as the label of the phrase. In the earliest versions of X-bar theory, each level of phrasal projection of a head is categorially distinguished by a different number of bars associated with the category label. In later versions the maximal phrasal projection of a head X is labeled XP and the intermediate phrasal projections between XP and X are all labeled X′. Given the top-down phrase structure rule analysis, all phrases would have to represent XP and X′ even if they contain only the head itself. The alternative is to abandon the bar notation and interpret the categorial labels of phrases functionally. The lexical item whose category features label a phrase constitutes the head of the phrase and the largest phrase bearing the label of the head constitutes its maximal phrasal projection. Thus a bare head will also be a maximal projection.

This is essentially the perspective of bare phrase structure theory. However, Chomsky abandons the X-bar notation for a much deeper reason. On the minimalist assumption that language is a perfect system, he proposes that language should meet a condition of INCLUSIVENESS: the computation that generates structures (especially PF and LF) applies only to the elements of lexical items initially selected for the computation and is restricted to rearrangements of lexical properties – thus the computation cannot create new objects like bar-levels (or indices) because such elements are not features or properties of items in the mental lexicon. The condition of inclusiveness provides the foundation for the bare phrase structure theory adopted under

in this way. To beginning students (and maybe not just beginners) this looks like another branch of a tree. The point I am trying to make here is not just an aesthetic notational quibble. The branches given in tree representations are crucial for calculating c-command relations – for example in Kayne's analysis of the Lexical Correspondence Axiom (1994), which may not be viable if there is no dominance relation between a lexical item and its lexical category.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

the minimalist program. Specifically, 'categories are elementary constructions from properties of lexical items, satisfying the inclusiveness condition' (1995b: 249).

Neither Culicover nor Radford mentions the condition of inclusiveness in their discussion of bare phrase structure. Moreover, Radford retains the standard X-bar notation of XP and X′ throughout his text, even though it is billed – inaccurately in this regard – as 'a minimalist approach'.

The concept of phrasal projection, which is the core of X-bar theory and essentially all that remains of X-bar theory under bare phrase structure, is superficially explicated at best in the three texts that adopt some notion of X-bar syntax. McCawley essentially rejects X-bar syntax but nonetheless uses the X-bar notation, though in an odd way. Although he uses X′ to designate a phrasal unit whose head is an instance of category X, he does not use XP for a maximal phrase whose head is X. Thus a VP is designated as V′. McCawley uses the category NP to designate a determiner plus a phrasal unit whose head is a noun (see chapter 12) – which is basically the DP analysis of nominal phrases. However, he also posits NPs without heads, where an NP dominates just a Comp′ (i.e. CP), and retains the phrasal category S which has no head in his analysis. While part of McCawley's analysis could easily be reconciled with current versions of X-bar theory, there are clearly other parts which could not.

The Radford text introduces the notion of phrasal projection in a perfunctory fashion. 'The phrase which is thereby formed [by merger] is (to use a traditional technical term) **endocentric** (i.e. headed), in that it is a **projection** of a **head** word' (88). It never explains that a head projects its categorial features as the label of the newly formed phrase. The summation in section 3.10 seems to confuse the notions of phrasal projection from X-bar theory with the notion of phrasal expansion from phrase structure grammar. 'We suggested that one way of forming a larger structure is to project (i.e. expand) a head word into a phrase by combining it with a following complement …' (116). This presentation confuses the bottom-up view of X-bar theory with the top-down view of phrase structure grammar. We talk about derivations in phrase structure grammar in which some XP expands as the string of its immediate constituents. A head does not expand into a phrase.[11]

Culicover also neglects to explicate the crucial notion of phrasal projection. In the three-page discussion of X-bar theory in chapter 2, the term 'projection' is introduced as follows: 'Nodes of the same categorial type as

[11] The glossary entry for 'projection' in Radford's book introduces a further problem with terminology. It defines INTERMEDIATE PROJECTION as 'a constituent which is larger than a word, but smaller than a phrase' (524). The entry for PHRASE says 'used to denote an expression larger than a word which is a maximal projection: see **projection**' (521). It should follow then that an intermediate projection is not a phrase, which then contradicts Radford's discussion of Merger quoted at the beginning of the paragraph.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

a head are called projections of the head' (19).[12] Example (4) on page 19 gives a schematic where VP dominates V′, which dominates V. What is missing here (and in Radford and Roberts as well) is the fact that the three distinct category labels are all equivalent to a categorial feature specification (i.e. [+v]), as was noted in the first discussion of X-bar syntax (see Chomsky 1970).

Culicover says nothing about the feature analysis of categories, whereas both Roberts and Radford cover this topic. Roberts gives the standard account in one page. Radford has more extensive discussion in sections 2.8–2.9 of the second chapter. However, here too Radford introduces terminology that could well cause confusion because it is nonstandard. He calls features of nominals like [±count] and [±plural] – inherent features in Chomsky (1965) – SUBCATEGORIAL, inviting confusion with the more standard notion of subcategorization feature (though this is not discussed in the text). What are standardly called categorial features elsewhere are designated CROSS-CATEGORIAL features in Radford – presumably because more than one part of speech can be designated by a single feature – for example, both verbs and adjectives are [+v]. The way Radford specifies the relationship between categories and features is potentially problematic.

> A *grammatical category* is a set of elements which have the same value(s) for a given set of grammatical features. (67)

Radford considers inherent as well as categorial features to be grammatical features, so his definition allows for a much wider class of syntactic categories than has been standardly assumed – for example, V_{plural} vs. $V_{singular}$. This is made explicit in chapter 5, on checking:

> What we are tacitly assuming here is that it is in the very nature of the **merger** operation by which phrases are formed that the head-features of a head are shared by all projections of the head (e.g. a phrase headed by an infinitive verb form is an infinitival verb phrase). (179)

Radford's notion of head features conflates subcategorization features with inherent features and agreement features based on specifier/head agreement. However, the utility, let alone the validity, of this analysis is never discussed and Radford generally ignores it in what follows. (See Freidin 1983 for reasons against adopting this position.) The characterization of projection given in chapter 5 on page 157 (taken from Speas 1990) does little more than to define 'maximal projection' and 'head' GIVEN A STRUCTURE IDENTIFIED AS A PROJECTION OF THE HEAD X.

Culicover's discussion of X-bar theory, like Radford's, confuses the top-down view of phrase structure grammar with the bottom-up view of X-bar

[12] The main discussion of X-bar syntax is delayed until chapter 5, for no apparent reason. The material preceding it, on arguments, government, case, binding and A-movement, does not provide information necessary to its clear explication. If anything, the opposite is true.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

syntax. Thus he claims that 'the fundamental relation of X′-theory is dominates' (18). Although it is true that domination is a fundamental relation for the theory of phrase structure in general, it is not THE fundamental relation for X-bar theory because it follows from the more basic notion of phrasal projection that is a consequence of Merge.

Roberts introduces the notion of phrasal projection as a solution to an empirical problem for a phrase structure theory that is limited to lexical categories and their phrasal counterparts, for example, N and NP only. He notes that such a theory cannot account for the fact that in a phrase like *the very beautiful girl*, the AP and the head N can constitute a single constituent (as illustrated by *one*-pronominalization).[13] Roberts then generalizes the analysis to 'every X projects an X′'(13), without specifying precisely what is meant by 'projects'. On the following page, he enunciates 'a general principle of phrase-structure rules and representations': 'Every lexical category X corresponds to a phrasal category XP.' Unfortunately, at this point we have not been told that XP is a maximal phrasal projection of X (which occurs five pages later) nor is it made clear how 'corresponds' relates to 'projects', all of which leaves open the precise way phrase structure analyses are constructed.

Of the four texts under review, only Radford's is clear about the mechanism for constructing phrase structure representations (via Merge). Roberts' discussion cited above suggests that he is assuming some variety of phrase structure rules, constrained by the postulates of X-bar theory. He does not provide any examples of such rules.[14] Although Culicover mentions phrase structure rules at the beginning of his chapter on X-bar theory, why they are dropped from discussion in the rest of the chapter remains unexplained. The main reason (prior to the advent of bare phrase structure theory) that a phrase structure rule subcomponent has been eliminated from the syntactic component is that it is considered (starting in September of 1979 (Chomsky class lectures)) to be redundant with other, more general parts of the grammar (for example, case theory and the theta role properties of lexical

[13] In contrast, Culicover's initial discussion involves no empirical examples and Radford's involves an example that does not work because it is analyzed incorrectly. Radford's example is *Government plans to privatize hospitals*, where he proposes that *plans to privatize hospitals* constitutes an intermediate phrasal N. The problem with this is that *Government plans* is a compound noun, itself the head of the nominal. Note further that the *one*-pronominalization test fails for Radford's proposed structure:

(i) *John read about a government plan to privatize hospitals and I read about an industry one.

[14] However, in section 2.6 on checking theory, Roberts does allude to the perspective of Chomsky 1993 that excludes phrase structure rules, pointing out that 'lexical insertion can take place at several points in the derivation' (99). Unfortunately, there is no further clarification of exactly what this means, so his presentation ends up being opaque.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

items). That is, phrase structure rules stipulate what follows from lexical information plus UG.[15]

Culicover does discuss a mechanism for constructing phrase structure representations in the three-page section on bare phrase structure in the last chapter, but the discussion is fundamentally flawed. It identifies the merger operation as the Generalized Transformation (GT) 'whose function is to map sets of one or more phrase markers into single phrase markers' (354). This is a mischaracterization to the extent that GT always applies to pairs of phrase markers, in contrast to singulary transformations, which only apply to a single phrase marker. The example Culicover uses involves a verb which combines with an NP to form a VP, where he gives the phrase marker for the verb *cook* as the unary branching structure VP dominating V′ dominating V dominating *cook*. However, it is precisely the point of bare phrase structure that unary branchings like this do not occur. Moreover, under bare phrase structure a head (that is the lexical item itself) will also be a maximal phrasal projection when it has no complements, specifier or adjuncts. This property creates nontrivial problems for formal structural definitions that distinguish specifier, complement and adjunct, of the sort discussed in the Culicover, Radford and Roberts texts.

[15] It is worth noting here that the methodological assumption that redundancy constitutes a defect of the theory that should be eliminated has motivated a lot of important research within the principles and parameters framework over the past quarter century and has been discussed explicitly in the literature. Culicover, while claiming to be concerned with the 'conceptual and methodological foundations of the theory' (vi), misses it completely.

Culicover's initial discussion of conceptual and methodological foundations occurs in the brief four pages of section 3 of the first chapter. The first part of this section discusses the issue of idealization with respect to native speaker competence and then 'autonomy'. The one paragraph on native speaker competence tells us that we must ignore differences among individual speakers, thereby constructing an idealized account of what a native speaker knows about her language. This is essentially the line taken in Chomsky (1965). It makes some sense for a notion of E-language, but not for the concept of I-language. As Culicover gives no examples of how this sort of idealization would operate in practice, it is going to be difficult, if not impossible, for the student to see how this methodological assumption should be put into practice. The same problem holds for Culicover's discussion of what he calls 'autonomy'. He defines autonomy as 'the notion that a linguistic theory can be organized into such components' as syntax, semantics, morphology, and phonetics/phonology (11), which is not the way the notion has been discussed in the literature (see Chomsky 1957, 1975). Culicover states as a methodological assumption the view that 'the components are strictly autonomous, in the sense that the principles and primitives that make up one component of grammar are unique to that component' (11). The empirical ramifications of this are not discussed, nor is this supposedly fundamental methodological assumption mentioned again.

The other assumptions in section 3.2 of chapter 1, which he claims 'are essential to an understanding of the development of and the motivation for theories within the [P&P] framework' (12) and which are unique to Culicover's text, are so vaguely formulated that how they apply remains unclear. Moreover, the initial empirical examples cited as support for these assumptions do not work, as Culicover himself points out in the first case.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

3. Transformations

The exact nature of the transformational operations required for a grammar depends to a large extent on the way that phrase markers are constructed. Consider, for example, the derivation of the sentence in (1).

(1) Mary expects that John will be elected.

Given phrase structure rules, lexical insertion will be via substitution and the NP *John* will undergo movement from the object position of *elected* to the Spec-IP position of the complement clause, also via a substitution operation – in this case replacing an empty NP in that position with the NP *John*. Because phrase structure rules operate in top-down fashion, the phrase-marker of the entire sentence is constructed before the movement operation applies. In contrast, if there are no phrase structure rules and (1) is constructed by Merge, then no substitution operation is needed for either lexical insertion or NP movement. Merge is basically the adjunction operation that concatenates two syntactic objects to form a single new object bearing the label of one of the objects so concatenated. In the case of movement, the adjunction operation concatenates the object of the verb *elected* with the IP in (2a), thereby creating the Spec-IP position at the left edge of the extended IP in (2b).

(2) (a) $[_{I}$ will $[_{V}$ be elected John]]
(b) $[_{I}$ John $[_{I}$ will $[_{V}$ be elected John]]

Thus the derivation never involves an empty NP position created by a phrase structure rule. Moreover, in principle, it couldn't because a derivation in which an empty NP was merged with IP and then filled with a lexically realized NP would be disallowed by economy considerations: the creation of the empty NP position would be a superfluous step in the derivation.

This is a fundamental point about the very nature of transformational analysis. Unfortunately, neither Radford, who essentially adopts the bare phrase structure theory, nor Culicover, who discusses it in the final chapter, mentions it.

Equally fundamental is the specification of the precise nature of the transformational operations utilized in analysis. Regrettably, none of the books under review addresses this important issue. McCawley warns us that the 'rules expressing the relationship between the underlying structures and the corresponding surface structures … are commonly referred to as **transformations**, in view of the metaphor in which they are spoken of as "operations" that "convert" the underlying structures into "derived" structures' (18). However, his refusal to provide explicit formulations for the transformational rules he assumes (see page xviii) simply ensures that his analyses will themselves be based on the kind of metaphor he is warning us against. McCawley's implied criticism misses the point that transformations are

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

operations in a formal system of grammar. Metaphorical interpretations of formal systems have no bearing on the validity of such systems, whether we are talking about grammar, logic or mathematics.

Transformational analysis is primarily determined by the elementary transformational operations utilized. Thus in Chomsky's *LSLT* theory of 1955, grammatical transformations (like the rule for forming passives) were constructed from a set of elementary operations, so that one grammatical transformation could perform several elementary operations at once. This view of grammatical transformations was abandoned by the mid 1970s when it became clear that that complexity of transformations could be reduced by relying on general principles of grammar. This led ultimately to the formulation of a single movement rule, Move α, whose behavior was governed by lexical properties and UG constraints on rules and representations. Thus the elementary operations available from UG constituted the grammatical transformations of the theory.

McCawley's text retains the complex grammatical transformations of early transformational grammar. His passive rule, for example, involves both subject postposing and object preposing as well as the insertion of both the passive auxiliary and the passive *by*. His rule of subject-to-subject raising involves replacing a subject NP which dominates only an infinitival clause (and is therefore headless) with the subject NP of the clause it dominates as well as extraposing the remainder of the infinitival clause to become a sister of the matrix verb. McCawley simply ignores the generalization of NP-preposing and raising as Move NP (see Chomsky 1976; also Freidin 1994 for discussion). He also misses the generalization of *wh*-movement for interrogatives and relative clauses by distinguishing two different rules, *Wh*-movement and Relative *Wh*-movement. As a result, the syntactic analysis in the text, essentially the kind that was prevalent in the late 1960s, seems woefully outdated. It is not that McCawley presents any specific argument against the current view of what constitutes a grammatical transformation; rather, he just ignores the issue.

Culicover, Radford and Roberts all cover movement phenomena in terms of the standard distinctions between head and phrasal movements. However, their terminology for phrasal movement is, for such a central topic, strangely disjoint. Culicover follows the standard terminological distinction of A- vs. A′-movement. Radford refers to A-movement and operator movement (which is nonstandard in spite of his claim to the contrary).[16] Roberts distinguishes DP-movement and *wh*-movement, although he does explicate the A vs. A′ distinction in the chapter on binding and has references to A-

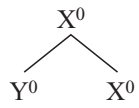
[16] Note that the chapter on operator movement is devoted almost exclusively to interrogative *wh*-phrases. Relative clause *wh*-phrases are introduced only in one exercise. The generalization about *wh*-movement is too important to relegate to an exercise that could be overlooked.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

and A′-movement in the index. However, the pages the index refers to do not mention the terms themselves and some precede the point in the text where the distinction is explicated.

Given that the notion of grammatical transformation reduces to that of elementary transformational operation, an essential question about movement is whether it results from substitution or adjunction. Although Culicover, Radford and Roberts state that head movement is by adjunction, creating a structure as in (3) (cf. Chomsky 1993), this comes as a clarification to the discussion in which head movement has been represented as an arrow between two head positions, suggesting substitution.

(3)



Culicover and Roberts return to the arrow between two head positions to illustrate head movement, whereas Radford continues to represent adjunction structures.

The adjunction structure in (3) is not unproblematic. Roberts notes that such adjunction structures create a potential difficulty for the c-command condition on trace binding he postulates. If the X^0 sister of Y^0 c-commands Y^0, then Y^0 will not c-command its trace. Therefore we must invoke the category/segment distinction by claiming that the X^0 sister of Y^0 is a segment of X^0, the other segment dominating Y^0, and that c-command applies only to categories, not segments. Radford also formulates a c-command condition on trace binding, but does not discuss the problem that his adjunction analysis for head movement creates.

There is, however, a potentially more serious problem with the adjunction analysis of head movement. If, as discussed above, a lexical item is itself the head of its phrasal projection, then representations like (3) are simply illegitimate because lexical items do not have syntactic constituents. Adjunction between two heads can only create a phrase – which is precisely what we don't want.

One alternative would be to perform head movement by adjoining a head to a root. For example, the question in (4) might be derived by adjoining the modal *will* to the IP in (5).

(4) Will you arrive at 5:00?
(5) [$_{I}$ you will arrive at 5:00]

Yet supposing that *will* contains the categorial features for I, it is not obvious how it winds up projecting the categorial features for C, given that (4) is a CP that contains an IP. This could be resolved if instead the modal substitutes for a phonetically null head that projects CP. There is, of course, still the issue of how the head of I can move to C without violating the

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

nondistinctness condition on substitutions, which the adjunction analysis avoids.

As for phrasal movement, Culicover, Radford and Roberts all represent such movements with an arrow between the launching and landing sites – presumably a substitution operation, though none of them is explicit about what elementary operation is involved. As noted above, if phrase structure is created by phrase structure rules, then phrasal movement should result from substitution. In Radford, where phrase structure is transformationally derived via Merge, phrasal movement could only occur via adjunction. Yet, Radford's discussion of movement, supposedly along the line of Chomsky (1995b), is grossly inaccurate. In *The Minimalist Program* Chomsky assumes that only formal features move, so that the movement rule is Move Feature, not Move Category. The displacement of phrasal categories is carried out by some sort of general pied-piping operation. Later in chapter 4, he proposes an alternative rule of Attract Feature, which has somewhat different properties. In chapter 7 (page 293) Radford mentions feature movement by attraction, claiming incorrectly that 'attraction is more economical than movement' on the grounds that movement applies to categories.

4. Conditions on rules and representations

Conditions on rules and representations have been an extremely lively topic in generative grammar beginning with Chomsky's proposal of the A-over-A principle (Chomsky 1964). Since the advent of the principles and parameters framework in the late 1970s, these conditions have been at the core of syntactic analysis in generative grammar. Exactly what sorts of principles are necessary as well as their precise formulations remain important research questions today.

As the four books under review illustrate, what principles should be presented in an introductory textbook and how they should be developed is open to considerable interpretation. For example, the Ross (1967, 1984) island constraints (the Complex NP Constraint (CNPC), the Coordinate Structure Constraint (CSC), the Left Branch Condition (LBC), the Sentential Subject Constraint (SSC – not to be confused with the Specified Subject Condition, which will not be discussed here) and the Right Roof Constraint (RRC)) figure prominently in McCawley's text. In contrast, they are covered in the Culicover and Roberts texts primarily to point out how some of them reduce to the Subjacency Condition. Radford does not mention the Ross island constraints or the Subjacency Condition.[17]

[17] The only context in which the notion of island comes up is in a discussion of pied piping concerning *wh*-extraction out of a subject, where Radford states that subjects are islands. He then cites the Condition on Extraction Domains (Huang 1982) as saying that only complements permit extraction. No formal statement of the CED is given, perhaps because

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

Culicover's coverage of the Ross constraints consists of giving formulations for the CNPC, the LBC, the SSC, and the CSC (in a footnote), without really explicating how they apply. The CNPC and SSC come up again when Culicover proposes that Subjacency captures both under a single condition. The example chosen to illustrate the CNPC case is a relative clause on an object, given as (6) ((51) on page 196).

(6) $[_{CP}$ C[+wh] $[_{IP}$ you will hire $[_{NP}$ a man $[_{CP}$ who$_j$ $[_{IP}$ t$_j$ can fix what$_i$]]]]]

Culicover says that extraction of *what* crosses the NP and the lower IP, the latter being forced by the fact that Spec-CP is already filled by *who*. He then suggests that it becomes a problem if the *wh*-phrase in subject position does not have to move, thereby providing an escape hatch for *what*. But this is simply wrong, even if the subject *wh*-phrase does not move, because *what* would still have to cross the NP and the matrix IP in violation of Subjacency.

Culicover then raises a problem for the Subjacency analysis of *wh*-movement out of sentential subjects, citing (7) ((53a) on page 197, arrows between traces and *who* excluded).

(7) *$[_{CP}$ who$_i$ would $[_{IP}$ $[_{CP}$ t′$_i$ that John saw t$_i$] surprise Susan]]

In (7) the extraction of *who* from the Spec of the CP in subject position to the matrix Spec-CP crosses only one IP.[18] However, Culicover's analysis in (7) is not obviously correct, especially in light of Koster (1976) (cited in the bibliography) where apparent sentential subjects are analyzed as constituents outside of IP, possibly in Spec-CP. If this is correct, then the prohibition against *wh*-movement out of a 'sentential subject' follows automatically because (7) could not be generated.[19] Culicover later shows that a revision of

---

a formal statement would have to include the notion of government, which is supposed to be illegitimate given the assumptions of the minimalist program. Radford's citation here demonstrates a certain unfortunate laxness about the material he is supposed to be explicating – that is, the minimalist program. We can see that this is not an isolated occurrence in Radford just by considering the material that precedes the discussion of extraction out of subjects. Here Radford invokes 'the preposition-stranding constraint' (in boldface and thus on a par with chain uniformity, enlightened self-interest (Lasnik's version of Last Resort) and other conditions that actually occur in the literature on minimalism), a constraint that exists only in Radford's text. This so-called constraint handles only one example (**Whom were you talking to?*). No formulation is provided, so there is no explanation of how a previous example (*Who were you talking to?*) (35a) is not also prohibited.

[18] Culicover claims that this problem forced Chomsky (1970) to propose a separate Subject Condition. This discussion is not quite accurate. Chomsky 1970 had to propose a Subject Condition because CP (that is, S′) rather than IP (that is, S) was taken to be the relevant bounding category. In Chomsky (1977: 112–113), IP is adopted and the Subject Condition is dropped.

[19] However, this analysis raises another problem – namely, a 'sentential subject' that constitutes an indirect question.

(i) *[which books John knows Susan borrowed] surprised Bill
Cf. the well formed *that John knows which books Susan borrowed surprised Bill.*

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

the Subjacency Condition under the *Barriers* framework will prohibit extraction out of sentential subjects on his analysis. In his discussion of how the *Barriers* version of Subjacency subsumes the CNPC, Culicover fails to point out that this formulation only gets the relative clause cases but not the N-complement cases in a complement position. As Chomsky (1986b: 35) notes, both the complex NP and its complement CP are L-marked and therefore are not blocking categories. Thus NP cannot inherit barrierhood from CP as it does in the relative clause cases.[20]

Both the Culicover and Roberts texts recapitulate the history of the field regarding constraints on movement. They start with the Ross island constraints and the *wh*-island constraint, show how some of these might be unified under Subjacency, introduce *that*-trace phenomena and the Empty Category Principle, and show how further unification was attempted using the notion of barrier to government, ending up with a discussion of Relativized Minimality. Roberts is more systematic and careful with details. For example, he points out where the *Barriers* formulation of Subjacency fails to capture the N-complement case of the CNPC and the simple case of the LBC.

McCawley's discussion of Subjacency is primarily aimed at what he perceives as a failure to reduce the Ross island constraints, in particular the Complex NP Constraint and the Coordinate Structure Constraint, to Chomsky's 1970s versions. Although he cites Chomsky 1986b in his references, McCawley does not discuss the *Barriers* formulation of Subjacency where Chomsky makes a similar point. Furthermore, given the last resort analysis of movement, where the motivation for successive cyclic *wh*-movement is no longer obvious, McCawley's complaint that successive cyclic *wh*-movement is unmotivated because the intervening positions lack a Q complementizer may be either uncontroversial or wrong, given the overt effects of successive cyclic movement we find in other languages, e.g. Spanish and Irish, see Torrego (1984) and McCloskey (2001). Even so, there is no reason to assume that the Ross island constraints constitute any kind of benchmark for a theory of conditions on movement.

In addition to the Ross island constraints, McCawley employs the syntactic cycle, starting in the second chapter and then again in chapters 6

[20] This kind of inattention to detail occurs elsewhere in Culicover's text. For example, in discussing whether a nominal subject is a DP or an NP, Culicover cites (i)–(ii) as evidence against DP and in favor of NP.

(i) Mary(*'s) owns a car.
(ii) *The left.

((i) should not be confused with the construction [$_{NP}$ Mary's N], where N is silent.) Yet, these examples are irrelevant to the issue. Both are prohibited for other quite basic reasons. **Mary's owns a car* violates Case restrictions. Either *Mary* is marked for both genitive and nominative Case, in violation of Case Uniqueness, or genitive Case is not licensed in the configuration that *Mary's* occurs in. (ii) is a clear θ-criterion violation – the θ-role of *left* is not assigned to any argument.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

(on the interaction of Raising & Equi-NP Deletion and Raising & Passive), 11 (on the cyclic nature of pronominalization (Ross 1969)) and 14 (on the interaction of Tag Formation and Negative Raising). He mentions the concept of strict cyclicity (without crediting Chomsky 1973), referring to it a few times in passing, mostly regarding Quantifier Lowering. McCawley motivates the cycle with the kind of A > B > A rule ordering arguments that were in fashion in the late 1960s. These arguments are notoriously easy to defeat because they rely crucially on the formulation of complex transformations that have long been abandoned. Moreover, as Kimball (1972) demonstrated, even if we retain such complex transformations, we can achieve that same output by applying rules in a linear fashion where a single rule applies to all cyclic domains before another does, thereby defeating the A > B > A ordering argument.

McCawley gives another argument for cyclic application based on extrinsic ordering. He cites (8) as the outcome of applying Passive after *Wh*-movement.

(8) *How many jokes do you think Bob Hope was stolen from by Henry Kissinger?

His analysis assumes that once the *wh*-phrase *how many jokes* has been moved from the object position of *stolen*, no trace of the *wh*-phrase remains, making it possible for Passive to prepose the PP object to subject position. Given trace theory, the argument collapses. The trace of the *wh*-phrase could easily be fronted instead of the overt object of *from* (i.e. *Bob Hope*), giving a valid result (i.e. *How many jokes do you think were stolen from Bob Hope by Henry Kissinger?*). McCawley is aware of this, as his footnote to the passage indicates, but suggests that with the cyclic principle we do not need to postulate empty categories to account for the example. The question then is which is more fundamental, traces or the cycle?

As Chomsky has said many times (for example, in Chomsky 2000), traces are the null hypothesis, not something that requires additional stipulation. As noted in Chomsky (1980), traces follow automatically from a prohibition against the compounding of elementary operations. Only deletion can remove a trace; therefore substitution and adjunction always leave traces.[21]

[21] Culicover stipulates this unnecessarily by positing a Trace Principle ('All movements leave traces' (181)), which overlaps, as he notes, with the effects of the Projection Principle and the Extended Projection Principle. Radford also treats traces as elements that require justification. He cites the fact that a c-command condition on trace binding can predict that movement will be upward rather than downward as THEORETICAL evidence for traces, whereas the blocking of contraction by the presence of traces he considers EMPIRICAL evidence. Putting aside the fact that both kinds of evidence are equally empirical, these are both weak arguments based on various unstated assumptions, compared to the more fundamental considerations about elementary operations. Roberts does not discuss the motivation for traces. He just assumes that movement leaves behind a trace and uses traces in the standard way.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

Whether the cyclic principle is an axiom of the theory or its effects are derivable from other, more fundamental principles of the theory is controversial. (For extensive discussion on why it should not be considered an axiom, see Freidin 1978, 1999.)

The transformational cycle has been around since its initial formulation in 1964 lectures (Chomsky 1966) and has recently been revived by Chomsky as a principle of UG for minimalist analysis (see Freidin 1999 for discussion). It is therefore surprising that Radford, who claims to be developing the 'key concepts' of the minimalist program (in particular Chomsky 1995b), says nothing about it. Culicover is also silent about the strict cycle – strangely so, because the *wh*-island violations he discusses in motivating the Subjacency condition can be derived counter-cyclically without violating Subjacency, as discussed in Chomsky (1973), which proposed the Strict Cycle Condition to eliminate the possibility. Culicover formulates Subjacency as a condition on the application of rules, which allows the counter-cyclic derivation. He doesn't consider the alternative, making Subjacency a condition on representations (as discussed in Freidin 1978, 1992, 1995 and Browning 1991), which would subsume the effect of the SCC for *wh*-islands.

Roberts discusses the SCC, but his formulation in (9) ((34) on page 195) is bizarre because it makes no mention of the notion cyclic domain.

(9) Nothing can move from a position c-commanded by an intermediate trace.[22]

Nonetheless, this formulation does help rule out the derivation of *wh*-island violations. Applying (9), the only way to derive these constructions would be to violate Subjacency, as Roberts notes. Thus consider (10).

(10) $[_{CP}$ who$_1$ do $[_{IP}$ you wonder $[_{CP}$ what$_2$ $[_{IP}$ t$_1$ saw t$_2$]]]]

Given (9), if *who* moves to the intermediate Spec-CP and then to the matrix Spec-CP, then its trace in Spec-CP will block the movement of *what*. This

[22] This is a rather striking example of a textbook writer reformulating a standard constraint in a nonstandard way. Roberts' formulation of Subjacency is closer to what is standard, but is also somewhat odd. He says that (i) is a simplified version of the original formulation in Chomsky (1973).

(i) In the following structure, α and β cannot be related by movement:
$\ldots \alpha \ldots [_{BC} \ldots [_{BC} \ldots \beta \ldots] \ldots]$
where α and β are separated by more than one blocking category BC. (193)

Since 'blocking category' is a concept developed for the *Barriers* analysis in Chomsky 1986b, this formulation and its explication are anachronistic, as are the references to DP and IP on the following page. McCawley and Radford avoid these problems by giving informal descriptions of constraints, which makes them difficult if not impossible to evaluate (see the quote from Chomsky 1957 discussed at the beginning of this article). Culicover's formulations, in contrast, are closer to what is found in the technical literature.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

leaves two possible derivations, both of which violate Subjacency because *who* will move directly from Spec-IP to the matrix Spec-CP. Whether the movement occurs before or after *what* has moved is irrelevant. However, notice that (9) allows the counter-cyclic movement of *what* – something that the standard formulation of the SCC disallows.

Radford, in marked contrast to McCawley, Culicover and Roberts, has virtually nothing to say about island phenomena. The chapter on *wh*-movement discusses only four examples involving interclausal *wh*-movement (his (22a), (24), (55) and (58b) in chapter 7).

(11) Who do you think will say what?
(12) *What do you think who will say?
(13) How do you think that he is feeling?
(14) *How do you wonder whether he is feeling?

In (11) and (12) he assumes that *think* takes an IP complement, so (12) would not be a *wh*-island violation. Rather, (12) violates what Radford calls the shortest movement principle, which he never formally defines. In (13) *how* is assumed to move successive cyclically through an intermediate Spec-CP, forced by the shortest movement principle. However, this conflicts with Last Resort, where movement is driven only by feature checking. Regarding (14) Radford claims that *whether* is an adverb in Spec-CP, which blocks the successive cyclic movement of *how*, thereby creating a shortest movement violation. The problem with this analysis is that *whether* clauses are only weak islands for *wh*-objects. Thus (15a) is not as deviant as (15b).

(15) (a) ?What did you wonder whether Bill gave to Mary?
(b) *What did you wonder to whom Bill gave?

If *what* is moving successive cyclically in (15a), then why not in (14)? The answer depends on the precise analysis given to a wide range of cases, which is not to be found in Radford.[23]

Although Radford claims to be providing a minimalist approach to English syntax, he actually presents very little idea of what the minimalist program is really about, aside from some of its engineering aspects. He defines minimalism as 'the requirement to minimize the theoretical and descriptive apparatus used to describe language' (6). First, this is just normal science and was ever thus in generative grammar. But more important, it comes nowhere close to the heart of the minimalist program, which is based

[23] Whether there is a convincing minimalist analysis of all so-called island phenomena to replace the Subjacency Condition, which may be too stipulative to count as a minimalist principle, remains to be seen. However, *wh*-island phenomena, a central topic in syntax for decades, are susceptible to analysis in minimalist terms (i.e. minimal link and shortest move). Regrettably, Radford ignores this important topic.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

on two assumptions: i) that language is a 'perfect system' and ii) that the computational system for human language is optimal in some interesting sense. Regarding how language might be a perfect system, Chomsky is quite clear. He says that language should meet a condition of INCLUSIVENESS, mentioned above. (See Chomsky 1995b and Freidin 1997 for discussion.) Given the centrality of inclusiveness (see comments above on bare phrase structure), the fact that it is never mentioned in Radford reveals how superficial this 'minimalist approach' really is. Instead, we find in Radford a lot of feature mechanics, some of it, like specifier features, of his own devising, but surprisingly missing one of the major innovations in Chomsky's 1995b minimalist analysis, the distillation of movement to a rule of Move/Attract Feature. The end result is a presentation of minimalism that focuses on current mechanics, which may change and indeed have changed (e.g. the abandonment of feature movement in favor of long distance agreement in Chomsky (2000; 2001a, b)), and leaves out the basic ideas that motivated the minimalist program in the first place.

Culicover's brief section on the minimalist program in the last chapter also fails to mention the two fundamental assumptions of the minimalist program or the inclusiveness condition. Instead, he says that 'the MP takes as its point of departure the hypothesis that D-structure, S-structure, the government relation, the X′ schema, and referential indices play no essential role in syntactic theory and must be dispensed with' (347). This is, however, somewhat of a mischaracterization. As noted above, it follows from inclusiveness that indices and bar-levels are unavailable. Given bare phrase structure, which Culicover discusses briefly and inadequately in a later subsection, there is no way to construct canonical D-structures. Because the system of principles employed includes output conditions only at the interface levels PF and LF, there is nothing to define a level like S-structure as significant. As for the government relation, the motivation for excluding it might simply be that it is unnecessary. Consider, for example, the ECP. It overlaps with Last Resort, which has stronger conceptual motivation as an economy condition. Presumably a computational system that is optimal (in almost any sense) would avoid overlapping conditions. Culicover's discussion of economy conditions is limited to Greed/Last Resort and Procrastinate, omitting the equally fundamental Full Interpretation (Chomsky 1986a), which determines whether a derivation crashes or converges at the interface levels. Because of these various omissions, his characterization of the minimalist program in this section provides a somewhat distorted picture.

Although Roberts' text lists the four papers that comprise Chomsky (1995b) in his references and the term 'minimalism' has eight entries in the subject index, it contains no extended discussion of the minimalist program. In particular, there is no mention of inclusiveness, economy conditions, or feature movement. In the chapter on Case and Agreement, Roberts provides an eleven-page section on checking theory based on Chomsky (1993), but

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

cautions that it 'is very much at the forefront of research, and many aspects of what I've presented here are likely to undergo revision in the near future' (106).

Roberts' comment raises the basic issue for a syntax textbook writer: what version of generative grammar should be explicated? The four books under review illustrate at least three different answers. Radford's text embraces the latest theoretical developments, avoiding any mention of what came before. There are two linked dangers with such an approach. First, as Roberts notes, the latest ideas are highly susceptible to change – in part because they are under the most intense scrutiny. Second, because the ideas are new, they may not be well understood and therefore may be resistant to a clear and useful explication. One alternative is to present a version of syntactic theory that seems more established and less subject to change, ignoring completely the current ideas. This seems to be a reasonable characterization of McCawley's text. Regrettably, McCawley's notion of what theoretical ideas have been well established doesn't extend much beyond the early 1970s. The third alternative, exemplified by the Culicover and Roberts texts, is to recount some version of the history of ideas for the field, including some of the most recent work. This provides students with the relevant background for understanding work in the field and with some insight into how ideas change. While this is an important aspect of a linguist's overall education, it is probably not the optimal way to INTRODUCE students to the study of syntax, where they are looking for a single set of tools for doing syntactic analysis.

My own predilection is to start with the current understanding and use it to rationally reconstruct the theory of grammar (see Freidin 1992). Whether this can be done in the current state of the field where the minimalist program has radically altered our perception of so many basic things, raising questions but not providing answers (Chomsky 1998: 119–120), remains to be determined. Whether it can be done in a way that beginners will understand is not obvious. The notion of language as a perfect system presupposes that one understands how language is a system, part of the goal of an introductory syntax. Further, how such a system could be formulated in a variety of ways, some less perfect than others (assuming that the concept of 'perfect grammatical system' is clear – hardly an innocuous assumption), is surely beyond what can be expected of beginning students. Nonetheless, if the criteria for a perfect grammatical system can be formulated explicitly, as is the case with inclusiveness, then even beginners could learn to apply it. Likewise for the related topic of the optimal character of the computational system. Notions like the economy of representations (viz. Full Interpretation) and derivations (viz. Minimal Link and Last Resort), as well as the nonredundancy of the computational system (i.e. no overlapping conditions), can be successfully presented in introductory syntax provided the empirical basis for these concepts is carefully explicated.

In the end though, it is a pragmatic question of what a teacher feels

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

confident that beginning students can understand, rather than the scientific question of what may be the best theory.

REFERENCES

Browning, M. (1991). Bounding conditions on representations. *Linguistic Inquiry* **22**. 541–562.
Campos, H. & Kemchinsky, P. (eds.) (1995). *Evolution and revolution in linguistic theory: studies in honor of Carlos P. Otero*. Washington, DC: Georgetown University Press.
Carnie, A. (2002). *Syntax*. Oxford: Blackwell.
Chomsky, N. (1955). *The logical structure of linguistic theory*. Ms., Harvard University & MIT. [Revised 1956 version published in part in 1975, New York: Plenum, and in 1985, Chicago: University of Chicago Press.]
Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
Chomsky, N. (1964). The logical basis of linguistic theory. In Lunt, Horace G. (ed.), *Proceedings of the Ninth International Congress of Linguists*. The Hague: Mouton. 914–978.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, N. (1966). *Topics in the theory of generative grammar*. The Hague: Mouton.
Chomsky, N. (1970). Remarks on nominalizations. In Jacobs, R. A. & Rosenbaum, P. S. (eds.), *Readings in English transformational grammar*. Waltham, MA: Ginn & Co. 184–221.
Chomsky, N. (1973). Conditions on transformations. In Anderson, S. R. & Kiparsky, P. (eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston. 232–286.
Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
Chomsky, N. (1976). Conditions on rules of grammar. *Linguistic Analysis* **2**. 303–351.
Chomsky, N. (1977). On *wh*-movement. In Culicover, P. W., Wasow, T. & Akmajian, A. (eds.), *Formal syntax*. New York: Academic Press. 71–132.
Chomsky, N. (1980). On binding. *Linguistic Inquiry* **11**. 1–46.
Chomsky, N. (1986a). *Knowledge of language*. New York: Praeger.
Chomsky, N. (1986b). *Barriers*. Cambridge, MA: MIT Press.
Chomsky, N. (1990). On formalization and formal linguistics. *Natural Language and Linguistic Theory* **8**. 143–147.
Chomsky, N. (1993). A Minimalist Program for linguistic theory. In Hale, K. & Keyser, S. J. (eds.), *The view from Building 20*. Cambridge, MA: MIT Press. 1–52.
Chomsky, N. (1995a). Bare phrase structure. In Campos, H. & Kemchinsky, P. (eds.), 51–109.
Chomsky, N. (1995b). *The Minimalist Program*. Cambridge, MA: MIT Press.
Chomsky, N. (1998). Some observations on economy in generative grammar. In Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M. & Pesetsky, D. (eds.), *Is the best good enough? Optimality and competition in syntax*. Cambridge, MA: MIT Press. 115–127.
Chomsky, N. (2000). Minimalist inquiries. In Martin, R., Michaels, D. & Uriagereka, J. (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press. 89–155.
Chomsky, N. (2001a). Derivation by phase. In Kenstowicz, M. (ed.), *Ken Hale: a life in language*. Cambridge, MA: MIT Press. 1–52.
Chomsky, N. (2001b). Beyond explanatory adequacy. Ms., MIT.
Freidin, R. (1978). Cyclicity and the theory of grammar. *Linguistic Inquiry* **9**. 519–549.
Freidin, R. (1983). X-bar theory and the analysis of infinitivals. *Linguistic Inquiry* **14**. 713–722.
Freidin, R. (1989). Review of A. Radford's *Transformational grammar: a first course*. *Journal of Linguistics* **25**. 509–518.
Freidin, R. (1991). Linguistic theory and language acquisition: a note on structure dependence. *Behavioral and Brain Sciences* **14**. 618–619.
Freidin, R. (1992). *Foundations of generative syntax*. Cambridge, MA: MIT Press.
Freidin, R. (1994). Conceptual shifts in the science of grammar: 1951–92. In Otero, C. P. (ed.), *Noam Chomsky: critical assessments* (1, tome 2). London: Routledge. 653–690.
Freidin, R. (1995). Superiority, subjacency, and economy. In Campos, H. & Kempchinsky, P. (eds.), 138–167.
Freidin, R. (1997). Review of N. Chomsky's *The Minimalist Program*. *Language* **73**. 571–582.
Freidin, R. (1999). Cyclicity and minimalism. In Epstein, S. D. & Hornstein, N (eds.), *Working minimalism*. Cambridge, MA: MIT Press. 95–126.

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press

Haegeman, L. (1994). *Introduction to government and binding theory* (2nd edn.). Oxford: Blackwell.
Hayes, B., Curtiss, S., Szabolcsi, A., Stowell, T., Stabler, E., Sportiche, D., Koopman, H., Keating, P., Munro, P., Hyams, N., Steriade, D. & Fromkin, V. (2001). *Linguistics: an introduction to linguistic theory*. Oxford: Blackwell.
Huang, C.-T. J. (1982). *Logical relations in Chinese and the theory of grammar*. Ph.D. dissertation, MIT.
Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kimball, J. P. (1972). Cyclic and linear grammars. In Kimball, J. P. (ed.), *Syntax and semantics* (vol. 1). New York: Seminar Press. 63–80.
Koster, J. (1976). Why subject sentences don't exist. In Keyser, S. J. (ed.), *Recent transformational studies in European languages*. Cambridge, MA: MIT Press. 53–64.
Lasnik, H. & Uriagereka, J. (2002). The poverty of the challenge. *The Linguistic Review* **19**.
Lightfoot, D. (1991). Review of J. D. McCawley's *The syntactic phenomena of English*, vols. 1 & 2. *Journal of Linguistics* **27**. 267–274.
McCloskey, J. (2001). The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* **37**. 67–100.
Radford, A. (1988). *Transformational grammar: a first course*. Cambridge: Cambridge University Press.
Ross, J. R. (1967). *Constraints on variables in syntax*. Ph.D. dissertation, MIT.
Ross, J. R. (1969). On the cyclic nature of English pronominalization. In Reibel, D. S. & Schane, S. (eds.), *Modern studies in English*. Englewood Cliffs, NJ: Prentice-Hall. 187–200.
Ross, J. R. (1984). *Infinite syntax*. Norwood, NJ: Ablex.
Smith, N. V. (1999). *Chomsky: ideas and ideals*. Cambridge: Cambridge University Press.
Speas, M. J. (1990). *Phrase structure in natural language*. Dordrecht: Kluwer Academic Publishers.
Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry* **15**. 103–129.

Author's address: Program in Linguistics, Clio Hall, Princeton University, Princeton, NJ 08544, U.S.A.
E-mail: freidin@princeton.edu

https://doi.org/10.1017/S0022226702001408 Published online by Cambridge University Press