J. Linguistics **34** (1998), 213–226. Printed in the United Kingdom
© 1998 Cambridge University Press

REVIEW ARTICLE

The Minimalist Program

JAN-WOUTER ZWART

NWO/University of Groningen

(Received 5 August 1997; revised 3 October 1997)

Noam Chomsky, *The Minimalist Program*. (Current Studies in Linguistics **28**.) Cambridge, MA: MIT Press, 1995. Pp. 420.

1. Introduction

The Minimalist Program, by Noam Chomsky, is a collection of four articles, 'The Theory of Principles and Parameters' (written with Howard Lasnik, 13–127), 'Some notes on Economy of Derivation and representation' (129–166), 'A Minimalist Program for linguistic theory' (167–217), and 'Categories and transformations' (219–394). The first three articles have appeared elsewhere, and are reprinted here with minor revisions.[1] The fourth was circulated in manuscript form earlier in 1995 and is commonly referred to as 'Chapter four'.[2] The volume opens with an 'Introduction' (1–11) and closes with a general bibliography and an index (395–420).

The work collected here is based on material presented by Chomsky, and discussed by participating students, faculty, and visitors, in Chomsky's fall term lecture-seminars at MIT in the period of 1986 through 1994.[3] For those who have ever wanted to attend these class lectures, but were never in the

[1] 'The theory of principles and parameters', written in 1991, appeared as chapter 24 of *Syntax: an international handbook of contemporary research* (*Volume 1*), edited by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, Berlin & New York: Walter de Gruyter, 1993 (pp. 506–569). 'Some notes on Economy of Derivation and representation' appeared in *MIT Working Papers in Linguistics* **10**, 1989 (pp. 43–74), and in *Principles and parameters in comparative grammar*, edited by Robert Freidin, Cambridge, MA: MIT Press, 1991 (pp. 417–454). 'A Minimalist Program for linguistic theory' appeared as *MIT Occasional Papers in Linguistics* **1**, 1992, and in *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, edited by Kenneth Hale & Samuel J. Keyser, Cambridge, MA: MIT Press, 1993 (pp. 1–52). The revisions appear to involve mainly referencing.

[2] 'Categories and transformations' is based on the much shorter paper 'Bare phrase structure', which appeared as *MIT Occasional Papers in Linguistics* **5**, 1994, and in *Government and binding theory and the Minimalist Program*, edited by Gert Webelhuth, Oxford: Blackwell, 1995 (pp. 385–439).

[3] Chapter 2 is based on the fall 1986 lecture-seminars, chapter 3 on the fall 1991 lecture-seminars, and chapter 4 on the fall 1993 lecture-seminars.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

position to, this is a must read. The MIT Press is to be commended for having made this collection available in such an exemplary inexpensive volume.

2. The Minimalist Program

This book is a masterpiece. Chomsky ruthlessly subjects his own work, past and present, to the conceptual guidelines of 'the minimalist program', shifting with each chapter to a higher gear. Along the road, cherished notions, concepts, principles, constraints, and descriptive tools, still riding high in the earlier two chapters, are discarded as 'conceptually unnecessary and empirically inadequate' (375). As Chomsky puts it in his Introduction (10):

> The field is changing rapidly under the impact of new empirical materials and theoretical ideas. What looks reasonable today is likely to take a different form tomorrow. That process is reflected in the material that follows. (…) Though the general framework remains, the modifications at each point are substantial. Concepts and principles regarded as fundamental in one chapter are challenged and eliminated in those that follow. These include the basic ideas of the Extended Standard Theory that were adopted in the [Principles and Parameters] approaches: D-Structure; S-Structure; government; the Projection Principle and the θ-Criterion; other conditions held to apply at D- and S-Structure; the Empty Category Principle; X-bar theory generally; the operation Move α; the split-I[nfl] hypothesis; and others. (…) Whether these steps are on the right track or not, of course, only time will tell.

It should be clear from this quote alone that anyone remotely interested in generative syntax is advised to study this volume carefully.

Reading *The Minimalist Program* from cover to cover was a fascinating ride, taking away my earlier impression of chapter 4 as containing fuzzy writing. On the contrary, chapter 4 is extremely carefully argued, sometimes verging on the rhetorical, but always clear and candid. The construction is intricate – in fact, highly suitable to a hypertext format, with all its intratextual references ('as noted earlier', 'we will return to this').

A major disadvantage, however, is that these references are impossible to keep track of when reading the sections in isolation, making especially chapter 4 more arcane than it needs to be. If one tracks the references down, one will find none of them to be vacuous, and one cannot but admire the construction of the work and the genius of its author. But it would have been helpful if the publisher had insisted on making the intratextual references more precise.[4]

[4] Unfortunately, the book does not have a detailed table of contents either, and the index is rather poor. These factors conspire to make chapter 4, in particular, inaccessible to

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

The Minimalist Program deals with the perennial question of how to link sound and meaning. As in earlier work, Chomsky proposes that the language faculty involves a computational system that feeds into the two components of the mind/brain dealing with sound and meaning: the articulatory–perceptual system and the conceptual–intentional system. The computational system of human language (C_{HL}) interacts with these 'external' systems through two distinct interface levels, Phonetic Form (PF) and Logical Form (LF). Chomsky notes that what goes on at these interfaces is 'poorly understood' (3, 222); the theory of language explored here is not about the interfaces but about the computational system, of which we basically know little more in advance than that it should construct a pair of interface representations out of items taken from the lexicon (169).

The sound-meaning link is described as a derivation, taking a single array of lexical elements as its input (the NUMERATION), and the two interface representations PF and LF as its output. The two interface representations are different and one is not derived from the other (229). The point in the derivation where the computation splits is called SPELL OUT (in chapters 3–4; formerly, it was called S-structure). The derivation from Spell Out to LF is just a continuation of the derivation from the Numeration to Spell Out, with the exception that no new elements can be added from the Numeration (189). The derivation from Spell Out to PF is taken to be radically different, especially in chapter 4 (220, and passim).

In the derivation, the relations between the elements in the Numeration are made explicit by linking the various elements up in a phrase structure (by the operation MERGE). The relevant relations are both thematic (θ-role assignment, predication) and syntactic (Case assignment, agreement). The thematic relations are basic, the syntactic relations are secondary, only to be established by a second structure building operation, MOVE (222). Word order variation is the effect of languages applying the operation Move in different stages of the derivation, either before Spell Out (in OVERT SYNTAX) or in the Spell Out–LF derivation (in COVERT SYNTAX). The derivation is restricted by general conditions of economy, favoring local relations and simple structures, and prohibiting superfluous steps and superfluous symbols.

In these respects, the minimalist program is not essentially different from previous versions of the theory of generative syntax. The crude outline of the minimalist approach presented above may be applied in one form or another to all earlier stages of the generative enterprise. What is remarkable about the minimalist program, however, is its contention that this is basically all there is. It is significant, not (by comparison) for what it attempts to achieve, but for the limitations it imposes on how to achieve it. Hence the elimination of

anyone who does not start reading at the beginning and does not have a pencil ready to upgrade the index and to supplement the references with concrete page numbers.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

most of the standard concepts of the theory of Principles and Parameters, listed in the quotation above.

In the Introduction, Chomsky points out how 'the leading questions that guide the minimalist program came into focus as the principles-and-parameters (…) model took shape' (3), and he presents a brief, illuminating sketch of the development of the principles-and-parameters model out of earlier stages of generative syntax, ultimately reaching back to pre-structuralist thinking about language and mind. This emphasis on the continuity of the basic concerns of generative grammar is very useful, especially because on the face of it, the minimalist program looks so different from its predecessors.

Perhaps the notion GOVERNMENT is the most spectacular victim of the minimalist clean up operation conducted here.[5] Government is basically a local version of command, and it plays a central role in almost all modules of grammar in the principles-and-parameters approach (as is evident from chapters 1 and 2). But even in the principles-and-parameters framework government has always had an elusive character, fueling the suspicion that 'government' is not a unified concept, but a label covering highly diverse grammatical relations (see Aoun & Sportiche 1983 for early discussion of this problem). Thus, government covers the head–complement relation and the head–specifier relation, two entirely different configurations. It is defined as a relation between a HEAD and a dependent element, but also as the relation between a moved NONHEAD and its trace ('antecedent government', still described in terms of binding in Chomsky 1981:183f). Furthermore, government is used for thematic licensing (θ-role assignment) and formal licensing (Case assignment), for determining local domains for binding (governing category) and extraction (barriers, via the notion θ-government), and for deriving the distribution of PRO (via the PRO-theorem, stating that PRO must be ungoverned).

It is instructive to see that even within the principles-and-parameters approach, these implementations of the notion government have largely evaporated, even if the term continued to be used. Thus, the core case of government, the head–complement relation (also involved in θ-role assignment), reduces to the more primitive relation of sisterhood (recognized in chapter 3, 178). The specifier–head relation emerges in the principles-and-parameters framework as the designated formal licensing relation (for structural Case (120), including Exceptional Casemarking (122) and the 'Null Case' that licenses PRO (119)). It has nothing in common with the traditional notion of government, and the specifier–head relation indeed seems to be a contribution of generative grammar.

[5] Chomsky (1992) cited this as generative grammar's first radical break with traditional grammar.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

In the minimalist program, both local relations (head-complement and head-specifier) are derived from the notion 'minimal domain', a minimal set of positions associated with a head, divided in an internal domain (the complement) and a checking domain (the specifier(s)). Government and c-command play no role in the definition of these domains (177f., 299).

Note that the developments within the principles-and-parameters model by and large allow the elimination of government as a licensing relation, a fact not recognized at the time. It has been clear from the start that 'proper government', the requirement that traces be identified, involves two distinct requirements, one of trace identification (head government), and one of locality of movement (antecedent government). In the clear cases, the former again reduces to sisterhood, even if work by Rizzi (1990) and Cinque (1990) has brought out that other factors are involved, such as referentiality of the moved category. As for the definition of local domains (in terms of barriers), it is clear that the core case of a barrier is an XP that is not a complement of a lexical head (79) – where lexical heads potentially include functional heads associated with inflectional features of the verb (86) – again suggesting that sisterhood, not government, is involved. Apart from that, research in the principles-and-parameters framework has come to focus on conditions of 'minimality', leading to a notion of 'shortest move' that is not defined in terms of government but in terms of more general conditions of economy (90). Finally, government is involved in the definition of the local domain for binding relations only to accommodate Exceptional Casemarking constructions like *John believes* [*himself to be clever*], where the binding domain for *himself* must be larger than the 'complete functional complex' in brackets, apparently because *believes* governs *himself*. But if Exceptional Casemarking involves a licensing relation between a verb and an element in its specifier, as argued in chapter 1 (122), the relevant binding relation must be established in covert syntax in English, again making reference to government superfluous.

This little exercise goes to show that already before the minimalist program was announced, 'government' had lost much of its significance in generative grammar. More generally, the minimalist program seeks to bring out and eradicate weaknesses in the theory, lingering there in the guise of convenient terminological and notational devices.

3. Contents

As always, it is important to separate a program from its particular implementations. The book under review is open to many forms of critique, as it combines a proposal for a program with a number of concrete implementations (each in itself essentially hypothetical). The program informs the implementations, but other implementations are possible and should be pursued.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

In order to sort out the program and its implementations in *The Minimalist Program*, I will present a detailed description of its contents. My summary of the minimalist program in the previous section reflects what I believe are its essential points. Much of what follows goes beyond that.

Chapter 1, written by Chomsky and Howard Lasnik, is a general introduction to the theory of principles and parameters as understood around 1991.[6] The presentation starts from scratch, introducing the theory as 'a particular approach to classical problems of the study of language, guided by certain leading ideas that had been taking shape since the origins of modern generative grammar some 40 years ago' (13). It is probably too dense to serve as a textbook, but it is an excellent reference article for the topics covered.

There is detailed discussion of the typology of empty categories (35–43), movement and chain formation (43–50), phrase structure of the lexical and functional domain (54–64), including a proposal to replace the A/A′-distinction with a distinction between L-related and non-L-related positions (64), locality conditions on movement (78–92), binding theory (92–110), and Case theory (110–124).

In the presentation of these topics, a faithful reflection of the research agenda of the 1980s, many details prepare for the transition to the minimalist framework. The strong-lexicalist checking approach to inflectional morphology is announced several times (20, 55, 76), though not consistently (121). There is a lengthy discussion of the status of S-structure conditions (64–78), where the authors seem to waver somewhat, although it is clear that 'the preferable option [is] that conditions involving interpretation apply only at the interfaces' (73). The Minimal Link Condition is introduced as a way to derive Rizzi's Relativized Minimality effects from general principles of derivational economy (90). It is proposed that PRO has Null Case (119), paving the way for the generalization that 'structural Case (…) is the realization of a Spec–head relation' (120), anticipating the division of the sentence into a thematic licensing domain and a formal licensing domain, an internal domain and a checking domain.

Unfortunately, the discussion is sometimes marred by digressions that seem out of place in a general introduction to the principles-and-parameters theory, and have little relevance to the minimalist program as well (31–33, 57–58, 66–67, 96–100).

Chapter 2 is in many ways less advanced than chapter 1.[7] It is concerned

[6] 1.1 Introduction, 13; 1.2 The lexicon, 30; 1.3 Computational system, 33 [1.3.1 General properties of derivations and representations, 33; 1.3.2 D-structure, 50; 1.3.3 Derived syntactic representations, 64]; 1.4 Modules of language, 78 [1.4.1 Government theory, 78; 1.4.2. Binding theory, 92; 1.4.3 Case theory, 110]; 1.5 Further topics, 125.

[7] 2.1 Preliminary Assumptions, 130; 2.2 Some properties of verbal inflection, 133; 2.3 A 'Least Effort' account, 138 [2.3.1 Minimizing derivations, 138; 2.3.2 The Element I, 143]; 2.4 Summary: on Economy of Derivation, 145; 2.5 The agreement system: Some

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

with the differences in verb raising between English and French (*John often kisses Mary* vs. *Jean embrasse souvent Marie*, cf. Emonds 1976, Pollock 1989), assuming the weak-lexicalist position that inflectional affixes in English are generated in INFL and lowered onto the verb stem in V.

Nevertheless, it is clear why this essay is included. The discussion yields 'varied evidence suggesting that both derivations and representations are subject to a certain form of "least effort" condition and are required to be minimal in a fairly well defined sense, with no superfluous steps in derivations and no superfluous symbols in representations' (161).

The notion of economy of derivation applied is a 'global' one: various derivations are compared, and the most economical one is selected as the only one allowed. In some cases, where two derivations are equally economical, we get optional word order variation (as in French *souvent paraître triste* and *paraître souvent triste*) (144). Chomsky notes, however, that economy 'cannot simply reduce to the matter of counting steps in a derivation' (139). He proposes that language specific steps, such as *do*-insertion, are more costly than universal ones, a notion giving way to its converse – Move more costly than Merge – in chapter 4 (346).

Chomsky also speculates that other cases of optionality might have to be assigned to a stylistic component of the mapping of S-structure to PF (reminiscent of Chomsky & Lasnik 1977: 431). He adds: 'This may well be too strong a conclusion, raising a problem for the entire approach' (146). Nevertheless, the notion of 'syntax in the phonological component' returns in the final sections of chapter 4.

Economy of representation is derived from the visibility requirement of Full Interpretation (151). In this context, the question of what are legitimate objects at the LF-interface is discussed (153–154), and the English expletive *there* is analyzed as 'an LF-affix, with its associate adjoining to it at LF' (155).

Chomsky in this article adopts Pollock's Split-INFL hypothesis (to be abandoned later in section 4.10), introducing the two agreement projections AgrSP and AgrOP (147). Chomsky immediately adds that 'AgrP plays no role at LF' (141), which makes it possible that the trace of Agr is deleted, once Agr has been raised or lowered. Elsewhere, Chomsky states that agreement and Case are 'checked at LF since they have LF consequences having to do with visibility (the Case Filter) and the Chain Condition' (155), leaving us to doubt the uninterpretability of Agr.[8]

---

speculations, 146; 2.6 Economy of representation, 150 [2.6.1 Operators and variables, 151; 2.6.2 Legitimate LF elements, 153; 2.6.3 FI and expletives, 154; 2.6.4 Further questions concerning LF Raising, 157]; 2.7 Some conclusions on language design, 161.

[8] The Chain Condition: every argument chain must be headed by a Case position and must terminate in a θ-position (46). Case and agreement are two sides of the same coin in chapter 1: 'Case is a relation of XP to H, H an X^0 head that assigns or checks the Case of XP. Where the feature appears in both XP and H, we call the relation "agreement"; where it

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

To my knowledge, this is the first time Chomsky makes 'a distinction between deletable and nondeletable elements on the basis of their LF-role' (145). This distinction plays an increasingly important role in the final two chapters, leading eventually to a phrase structure without agreement nodes and with multiple specifiers.

Chapter 3 is a brief exposé of the new minimalist approach, treated more fully in chapter 4.[9] It introduces some of the basic terminology of the minimalist program (converging vs. crashing derivations (171), the principles of Greed (201) and Procrastinate (198), minimal domain, checking domain, internal domain (178), equidistance (184); other terms and notions like merger (226), numeration (225), attract (297), feature movement (262), label (243), and term (247) are not introduced until chapter 4).

In many respects, chapter 3 is superseded by chapter 4. For example, X-bar theory is still regarded as an independent module in chapter 3 (189), whereas in chapter 4 the properties of X-bar theory are fully derived from the structure building process (246).

Chapter 3 remains interesting, however, for its discussion of the status of the levels of representation D-Structure (3.3) and S-Structure (3.4).

D-Structure is eliminated in the sense that there is no base component applying rewrite rules to generate an empty structure which is to be fleshed out later by 'all at once' lexical insertion (187). Instead, structures are created by combining elements drawn from the lexicon, and there is no stage in the process at which we can stop and say: this is D-Structure.[10] The approach leads to a unification of insertion (merge) and movement, which differs from merge only in that the element to be merged is contained in the target of merger.[11]

The question of the existence of S-Structure conditions is discussed with more confidence than in chapter 1. Differences between languages at S-Structure are reduced to requirements at the PF-interface. Certain features that are visible but not interpretable at PF must be eliminated by the feature checking operation (a function of movement). Hence there is no need for language particular S-Structure conditions in order to describe word order

---

appears only on XP, we call it "Case"' (119). Essentially the same definitions are given on p. 175 of chapter 3, but there Case is not assigned by Agr but by T or V, VIA Agr. Case is then dissociated from agreement given that 'agreement appears without Case (as in NP–AP agreement) and Case appears without agreement (as in transitive expletives [where the expletive gets Case, but does not agree with the verb])' (175).

[9] 3.1 Some general considerations, 167; 3.2 Fundamental relations: X-bar theory, 172; 3.3 Beyond the interface levels: D-structure, 186; 3.4 Beyond the interface levels: S-structure, 191; 3.5 Extensions of the Minimalist Program, 200.

[10] On the other hand, the way in which elements from the lexicon are combined is not fully free, in the sense that a verb is first combined with its internal argument, etc.

[11] Merge and Move have predecessors in the generalized and singulary transformations of early generative grammar. The former gave way to the system of recursive rewrite rules in the 1960s, obscuring the formal similarity of the two processes.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

variation (198–199). Chomsky in addition discusses various types of evidence for S-Structure conditions, showing in each case that the relevant condition CAN or MUST apply at LF (193), or even that it MUST NOT apply at S-Structure (208).

Chapter 4 is a fuller exposition of the minimalist program, again building everything up from scratch, starting with a review of the guiding ideas of the minimalist program, but taking it much further than chapter 3, reaching at times rather surprising conclusions.[12]

Section 4.1 outlines the basic concerns of the minimalist program, stressing the derivational character of the grammar.

Section 4.2 discusses the Lexicon (it contains collections of features) and the way sentences are built up by combining elements from an array of lexical choices (the numeration) via the operations Merge and Move. Overt movement is forced, not by a PF visibility condition, as in chapter 3, but by the stipulation that certain features ('strong' features) must be eliminated at once (233).[13]

Section 4.3 is about phrase structure, and contains a proposal to replace the 'informal notation' of phrase structure in tree diagrams with a notation in terms of sets of the type $\{\gamma,\{\alpha,\beta\}\}$, where α and β are two constituents (terms) which together form a new object with label γ (indicating the relevant properties of the new constituent, derived from either α or β).

Section 4.4 lists the three basic requirements on movement: the antecedent must c-command its trace, antecedent and trace must have the same phrase structure status, and movement must be a Last Resort operation, driven by feature checking requirements (253). The elements that move are not lexical items (i.e. collections of phonological, formal, and semantic features) but just the formal features that need to be checked (262). In overt syntax, the remaining features are moved along with the formal features, to ensure PF

[12] 4.1 The Minimalist Program, 219; 4.2 The cognitive system of the language faculty, 225 [4.2.1 The computational component, 225; 4.2.2 The lexicon, 235]; 4.3 Phrase structure in a minimalist framework, 241; 4.4 The operation Move, 249 [4.4.1 Movement and economy, 249; 4.4.2 Projection of target, 256; 4.4.3 Last Resort: some problems, 261; 4.4.4 Move F, 261; 4.4.5 Covert raising, 272]; 4.5 Interpretability and its consequences, 276 [4.5.1 Types of features, 277; 4.5.2 Checking theory, 279; 4.5.3 Expletives, 286; 4.5.4 Clause type, 289; 4.5.5 The Minimal Link Condition, 294; 4.5.6 Attract/Move, 297]; 4.6 Movement and theta theory, 312; 4.7 Properties of the transformational component, 316 [4.7.1 Why Move?, 316; 4.7.2 Departures from the Best Case, 317; 4.7.3 XP-Adjunction and the architecture of linguistic theory, 324; 4.7.4 Other improprieties, 326; 4.7.5 Adjuncts and shells, 329]; 4.8 Order, 334; 4.9 Expletives and economy, 340; 4.10 Functional categories and formal features, 348 [4.10.1 The status of Agr, 349; 4.10.2 Core concepts reconsidered, 355; 4.10.3 Empirical expectations on Minimalist assumptions, 367]; 4.11 Summary, 378.

[13] Strength is defined not as a property of another feature, but as a feature in itself, to be checked by a categorial feature (232). Later on, Chomsky refers to the strong feature as 'strong F, where F is categorial' (277). 'Strong F' is always a feature of a (functional) head, and is −Interpretable (282).

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

interpretability. In covert syntax, movement is always feature movement, i.e. head movement (271). Expletive replacement therefore does not adjoin the associate to the expletive at LF, as in chapter 2, but adjoins the features of the associate to INFL (273).

Section 4.5 revises the earlier idea of chapter 3 that feature checking entails feature elimination. Only those features are eliminated (erased) that are not interpretable at the LF interface.[14] +Interpretable features remain active even after feature checking, and may trigger an additional movement (280). However, Chomsky allows certain −Interpretable features to escape erasure after checking, so that they, too, may trigger additional movements (281). This is the mechanism by which multiple specifiers are allowed (354). A second discussion of expletive constructions follows, in which expletives of the type of *there* are defined as containing only categorial features, no Case features or ϕ-features (287). Finally, movement is described as ATTRACTION of the relevant features by a functional head (the target), eliminating the principle Greed of chapter 3 (297); the definitions of minimal domains and the notion 'shortness' (relevant to the equidistance principle) are adapted accordingly (299).

Section 4.6 restates the complementarity of θ-role assignment and formal licensing (checking) (312). The VP-internal subject hypothesis is adopted, but in a slightly different guise: unergative verbs contain two VP-shells, the higher of which (vP) hosts the external argument in its specifier position (315).

Section 4.7 returns to the theory of movement. Chomsky follows Kayne (1994) in severely restricting the possibilities of adjunction (323), removing the adjunction position from the checking domain (except for head adjunction). Cases where XP-adjunction seems to occur may fall outside of the computation from Numeration to LF (i.e. they may be part of the derivation from Spell Out to PF) (325). Adverbs are base-adjoined to X′ (330).

Section 4.8 addresses Kayne's (1994) Linear Correspondence Axiom, according to which linear ordering is a function of hierarchical ordering. The results of the LCA are captured in the bare phrase structure theory adopted here, with one exception: the order of the most deeply embedded head and its complement. Chomsky proposes that the most deeply embedded complement must always be a trace (337). Then if the LCA is a condition of the phonological component (dealing with OVERT word order), the most deeply embedded complement can be taken out of the equation (it will be invisible to the LCA) (340).

[14] Chomsky generally ignores interpretability at PF when it comes to elimination or erasure of features. Since Chomsky assumes that morphology is regulated in the course of the derivation from Spell Out to PF (319), a problem arises when a −Interpretable feature that has an overt reflex in morphology is checked in overt syntax. Erasure of this feature raises the question of how Morphology can do its job (see 385, note 50).

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

Section 4.9 contains the third discussion of expletives. The question is addressed why economy conditions apparently favor merger of an expletive in the embedded clause over raising of the embedded subject in *there seems* [*t to be someone in the room*], whereas raising of the embedded subject is preferred to merger of *we* in the subject position of the embedded clause in *we expected* [*someone to be t in the room*]. The solution is that θ-theory, prohibiting insertion of *we* in the lower clause in the Exceptional Casemarking construction, overrides the economy condition of 'merge over move', leading to the conclusion that an argument without a θ-role is not a legitimate object, causing the derivation to crash (347).

Section 4.10 subjects 'functional categories to the same minimalist critique' (348–9). Chomsky admits to speculating rather freely here (349). Arguing that Agr is present for theory-internal reasons only (being uninterpretable at LF), Chomsky proposes that the familiar phrase structure of chapters 1–3, with multiple heads, each with a single specifier, be replaced by a phrase structure with fewer functional heads, each potentially having multiple specifiers (354). The subject then does not move to a specifier of AgrS, but to an outer specifier of T. Similarly, the object, when it shifts, does not move to the specifier of AgrO, but to a specifier of *v* (the head of the higher VP-shell).[15] The proposal is applied to transitive expletive constructions (as in Icelandic), where the expletive and the associate occupy multiple specifier positions of T (instead of Spec,AgrS and Spec,T, as in Jonas 1996). Since the finite verb in Icelandic appears between the expletive and the associate, Chomsky proposes that verb second may take place in the phonological component (368).

4. Discussion

In the final summary (Section 4.11), Chomsky stresses that 'these seem to be the right KINDS of moves, though doubtless not yet the right ones' (379, emphasis mine). As always, it would be a mistake to confuse the minimalist program with a particular IMPLEMENTATION of the minimalist program. In the remaining pages, I would like to highlight a number of aspects of the particular implementation under review that may merit further discussion.

First, note that the minimalist program is presented as inherently derivational, putting to rest a discussion that had been lingering since the introduction of trace theory. The developments can be traced throughout the pages of this volume. The tentative stance of chapters 1 and 2, reminiscent of *Lectures on government and binding* (cf. 49, 133), is dropped in the opening remarks of chapters 3 (169) and 4 (219). The derivational character of the computational system is stressed time and again, never more so than in the

[15] Chomsky suggests that the specifier position for the object may actually be an inner specifier, c-commanded by the external argument in its base position (358).

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

construction of phrases via the operations Merge and Move.[16] The minimal domain can only be determined derivationally (179) and so can the feature content (the label) of any constituent (243). The issue returns in a brief discussion of Optimality Theory (380, notes 4 and 5), where Chomsky maintains that 'crucial properties appear to hold not of output–input pairs but of intermediate stages, so that no output–input condition is formulable'.[17]

Much more tentative appears to be the adoption of the concept of 'global economy', by which alternative derivations from a given numeration are compared (yielding a reference set of derivations), the most economical one being selected as the only one allowed (201, 220). Chomsky recognizes the unattractiveness of global economy (as opposed to local economy, by which at each step in the derivation the most economical one is selected, cf. Collins 1997) and attempts to curb its effects of computational intractability (227, 348). Nevertheless, global economy continues to play a certain role throughout chapter 4, perhaps unnecessarily so.[18]

Another distinction that might be dispensed with in a minimalist approach is the distinction between substitution and adjunction. The operations are formally indistinguishable as two instantiations of Merge (189, 191, 382 note 18).[19] The distinction is made by defining adjunction as a merger that yields a two-segmented category, described by Chomsky as an ordered pair (248). Like Kayne (1994) Chomsky argues that adjunction to XP is an option with a very limited range (323). Unlike Kayne, Chomsky proposes to relegate a number of processes that seem to involve adjunction to a 'stylistic' component of the grammar, part of the derivation from Spell Out to PF (324–5).[20] At the same time, the possibilities of substitution are broadened by allowing multiple specifiers, again deviating from the more constrained

[16] See Epstein (1995) for considerable strengthening of this point. Epstein argues that the definition of grammatical relations such as c-command can only escape stipulation in a derivational approach. On the other hand, as a reviewer points out, it may be that the derivational character of the minimalist program is just a technical aspect of Chomsky's presentation, and that a representational version of minimalism is equally feasible (cf. Brody 1995). My personal feeling is that the derivational character of the grammar is much more central to the minimalist approach than it was to earlier generative approaches.

[17] A concrete example is provided in the discussion of head movement of a lexical head containing an incorporated lower head (49–50, 214 n. 20, 223–224). After head movement, the trace of the lower head is too far removed from its incorporated antecedent, although locality conditions are satisfied at each STEP. (The earlier discussion in chapter 2, p. 143, appears to reach a different conclusion.)

[18] See Johnson & Lappin (1997) for the computational complexity of global economy. Note that Johnson & Lappin present global economy as a key notion of the minimalist program (1997:275), while recognizing at the same time that it is essentially anti-minimalist (1997:285). This suggests that their critique is not directed at the minimalist program, but at Chomsky's implementation of it.

[19] The distinction does not play a role in other work that is minimalist in spirit, such as Kayne (1994) and Zwart (1997).

[20] The two-segment category would not be interpretable at LF (320).

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

approach of Kayne (1994). The suspicion is never laid to rest that substitution, as a concept distinct from adjunction, is a relic of earlier stages in the theory, where certain positions were given in advance by phrase structure rules, only to be filled up (or not) by movement rules of a particular type (A-movement).[21]

Multiple specifiers are introduced as 'permitted in principle on minimalist assumptions about phrase structure theory' (285, 245). But of course multiple specifiers were not excluded in principle either when specifiers were defined by the rewrite rules of earlier stages of the theory. The question to ask is what considerations prompted us earlier to restrict the number of specifiers to at most one, and whether these considerations are still in force. This question is not addressed, unfortunately. The mechanism by which multiple specifiers are ultimately allowed (suspension of erasure of checked −Interpretable features, 354, and tolerance of unforced violations of Procrastinate, 375) strikes this reviewer as remarkably inelegant from the perspective of the minimalist program.

The ultimate consequence of the multiple specifier hypothesis as employed here in the description of transitive expletive constructions in Icelandic is that the Germanic verb second phenomena are the result of 'phonological operations that are extraneous to the [Numeration to LF] computation' (368). Here I would have liked to see the caveat of chapter 1 ('This may well be too strong a conclusion, raising a problem for the entire approach' 146) repeated. The 'entire approach' is guided by the idea that the computation serves interpretability at the interfaces only. Verb movement admittedly does not seem to affect interpretation at LF. On the other hand, the little we know about the phonological component suggests that the operations taking place there have the sole effect of turning X^0 elements into strings of phonemes (through the Morphology component, for instance). Verb movement seems pretty remote from that kind of operation. The problem may be that verb second is taken to be an operation whose only purpose is to achieve a certain linear order, taking the notion 'second' to be a linear rather than a structural notion (i.e. some relatively high functional head position). In fact, the verb second effect is an epiphenomenon, to be described and explained in terms of syntactic theory, guided by the general ideas underlying the minimalist program.[22]

This last point underlines the fact that there can be no such thing as 'canonical minimalism', beyond the general framework outlined in this book. There is no need to stipulate that specific implementations, whether found here or elsewhere, are just attempts to achieve the ultimate goal of the

[21] Note that Chomsky's (248) assumption that specifiers are distinct in properties from adjuncts does not necessarily imply a STRUCTURAL difference between the processes of substitution and adjunction.

[22] See Zwart (1997) for a recent analysis of verb second within the minimalist framework.

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press

minimalist program: finding out to what extent language can be described as a 'perfect' system.

The Minimalist Program is a major contribution in that it lays the groundwork for a new stage in syntactic theorizing, in combination with a detailed presentation of the kind of questions that can be asked, and the kind of results that may be obtained. If Chomsky is right in estimating that 'a rich and exciting future lies ahead for the study of language and related disciplines' (9), this book is bound to feature as a milestone in that development.

REFERENCES

Aoun, J. & Sportiche, D. (1983). On the formal theory of Government. *The Linguistic Review* **2**. 211–236.
Brody, M. (1995). *Lexico-Logical Form. A radically minimalist theory*. Cambridge, MA: MIT Press.
Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, N. (1992). Economy considerations in language design. Guest lecture at the 8th Comparative Germanic Syntax Workshop, Tromsø, 21 November.
Chomsky, N. & Lasnik, H. (1977). Filters and control. *Linguistic Inquiry* **8**. 425–504.
Cinque, G. (1990). *Types of A′-dependencies*. Cambridge, MA: MIT Press.
Collins, C. (1997). *Local economy*. Cambridge, MA: MIT Press.
Emonds, J. (1976). *A transformational approach to English syntax: root, structure-preserving, and local transformations*. New York: Academic Press.
Epstein, S. D. (1995). Un-principled syntax and the derivation of syntactic relations. Ms., Harvard University.
Johnson, D. & Lappin, S. (1997). A critique of the Minimalist Program. *Linguistics and Philosophy* **20**. 273–333.
Jonas, D. (1996). Clause structure, expletives and verb movement. In Abraham, W., Epstein, S. D., Thráinsson, H. & Zwart, C. J. W. (eds.) *Minimal ideas: syntactic studies in the minimalist framework*. Amsterdam: John Benjamins. 167–188.
Kayne, R. S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* **20**. 365–424.
Rizzi, L. (1990). *Relativized minimality*. Cambridge, MA: MIT Press.
Zwart, C. J. W. (1997). *Morphosyntax of verb movement: a minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer Academic Publishers.

Author's address: NWO/University of Groningen,
P.O. Box 716,
NL-9700 *AS Groningen,*
The Netherlands.
E-mail: zwart@let.rug.nl

https://doi.org/10.1017/S0022226797006889 Published online by Cambridge University Press