This book offers a comprehensive account of adjuncts in generative grammar, seeking to reconcile the differing ways in which they have been treated in the past by proposing a method of analysis grounded in simplification based on Simplest Merge.

The volume provides an up-to-date review of the existing literature on adjuncts and outlines their characteristic properties and the subsequent difficulties in adequately defining and treating them. The book compares previous attempts to account for adjuncts which have tended to use additional mechanisms or syntactic operations as a jumping-off point from which to propose a new way forward for analyzing them grounded in minimalist theory. Adopting an approach in the spirit of the strong minimalist thesis (SMT), Bode suggests an analysis of adjuncts that applies a minimalist approach based on theoretical simplicity, one which does not resort to extra mechanisms in capturing the empirical properties of adjuncts.

Offering a comprehensive overview of research on adjuncts and foundational minimalist principles, this book will be of particular interest to graduate students and practicing researchers interested in syntax.

Stefanie Bode is Adjunct Lecturer at the Georg-August University of Göttingen, Department of English Philology—Linguistics, Germany. Her doctoral thesis, on the verb ‘be’ in the English sentence structure, was published in 2003.
Routledge Studies in Linguistics

Time Series Analysis of Discourse
Method and Case Studies
Dennis Tay

Heart- and Soul-Like Constructs across Languages, Cultures, and Epochs
Edited by Bert Peeters

Systemic Functional Political Discourse Analysis
A Text-based Study
Eden Sum-hung Li, Percy Luen-tim Lui and Andy Ka-chun Fung

Systemic Functional Language Description: Making Meaning Matter
Edited by J.R. Martin and Y.J. Doran

Rarely Used Structures and Lesser-Studied Languages
Insights from the Margins
Emily Manetta

Externalization
Phonological Interpretations of Syntactic Objects
Yoshihito Dobashi

Approaches to the Study of Sound Structure and Speech
Interdisciplinary Work in Honour of Katarzyna Dziubalska-Kołaczyk
Agnieszka Kiełkiewicz-Janowiak, Magdalena Wrembel and Piotr Gąsiorowski

Casting a Minimalist Eye on Adjuncts
Stefanie Bode

For more information about this series, please visit: www.routledge.com/Routledge-Studies-in-Linguistics/book-series/SE0719
Casting a Minimalist Eye on Adjuncts

Stefanie Bode
Casting a minimalist eye on adjuncts / Stefanie Bode.


Typeset in Sabon
by Apex CoVantage, LLC
For Daniel
Contents

Acknowledgments viii

1 Introduction 1

2 What Are Adjuncts? 7

3 The Minimalist Framework 19
   3.1 Basic Minimalist Assumptions 19
   3.2 Merge and Structure-Building 25
   3.3 Why Are Adjuncts a Challenge? 38

4 Approaches to Adjuncts 50
   4.1 A Different Dimension 50
   4.2 Labeling and Projection 57
   4.3 A Simplification Ahead? 70

5 A New Proposal: Labeling Is Transfer 96
   5.1 Adjuncts and Labeling 96
   5.2 Adjuncts and Islands 120
   5.3 Labeling, Head Movement, and Adjunction 130
   5.4 A Final Remark 148

6 Conclusions and Outlook 162

Index 173
Acknowledgments

I thank the following copyright holders and the authors for granting permission to reprint quotes used for argumentation: Cambridge University Press, Columbia University Press, De Gruyter, Elsevier, John Benamins, Oxford University Press, Peter Lang GmbH, Routledge, John Bowers, and Balazs Surányi. I am also grateful for the helpful open access policy of many publishers and MIT’s general copyright strategy concerning reprinting smaller quotes. I am basically indebted to every author referred to, and especially to those whose works I discuss in the book. They share the deep interest in the Minimalist Program and contributed to this book. A project like this would not be possible without them. Thanks also to my publisher, Routledge, for giving me the opportunity to present my ideas on adjuncts. On a personal note, I want to thank Elysse Preposi (Editor, Linguistics), Helena Parkinson (Editorial Assistant), Eleni Steck (Production Editor), and Ramachandran Vijayaraghavan (Project Manager) for their help. I also thank two reviewers for their helpful suggestions. I thank my family for supporting me and being there for me. Eventually, I want to express my deepest gratitude to Daniel Seely, for his invaluable comments and encouragement.
1 Introduction

Why study adjuncts today? Recent minimalist discussion elaborates on the architecture of the grammar and raises such questions as: *what is part of narrow syntax or what belongs to the interfaces SEM and PHON?* There is a lot of discussion in the generative approach to syntax and, broadly speaking, minimalism splits into two opposing views. Either syntax is taken to be deterministic, leading to a crash-proof situation at the interfaces, or syntax is basically free, and its output is subject to interpretation at the interfaces. The most recent debates build upon the latter view, since applying Merge freely simplifies UG in a crucial way. Adjuncts have been under debate in the previous framework of *Government & Binding* and, strikingly, the issue has not been settled under minimalist approaches either.

Every textbook on syntax addresses adjuncts. Adjuncts actually are such a basic phenomenon of language that they need to be part of any introduction; nevertheless, we do not find a consensus among scholars, even at the textbook level. The way in which beginners of linguistics are confronted with adjuncts is particularly telling because the degree of consensus among linguists can often be judged by considering introductions into a subject. Textbooks discuss more or less the same properties of adjuncts relating to selection, theta roles, order, optionality, islandhood, and more, but the structural proposals look rather different. For instance, Koeneman and Zeijlstra (2017) use a V’-recursion so that, in terms of X-bar theory, adjuncts are placed in a X’-sister position, making it hard to distinguish them from specifiers and, more concretely, from subjects. Adger (2003) repeats the maximal level and thus resorts to segments of XP (again using X-bar terminology). He leaves room for a subject ‘agent’ to be projected in the specifier of V, and the most crucial aspect of his analysis is that he assumes an extra syntactic operation *Adjoin* that introduces adjuncts into the derivation (as opposed to other constituents which enter the derivation via *Merge*). Both textbooks use a recursion at a specific X-bar theoretic level, which is actually impossible to formulate in a purely minimalist framework. We will come back to this point later. Similarly, Radford (2006) uses the X’-level, with the
additional complication that, in his proposal, adjuncts and arguments basically occur in the same positions. The distinction between arguments and adjuncts gets blurred and, if it is reflected in the tree diagrams at all, only works in terms of X-bar-theoretic principles formulated in the previous Government & Binding framework. Thus, it is crucial that we observe that a frequent linguistic phenomenon as adjuncts does not receive a uniform treatment even at textbook level. Hence, it is not surprising that scientific research has not come to a general solution or consensus concerning adjuncts either. The need for a theoretical implementation of adjuncts is obvious. In the beginning of his discussion of adjuncts, Adger stresses that the mechanisms by which adjuncts are implemented into phrase structure still pose major research questions (Adger 2003: 111). After decades of research into the matter, this is a surprising fact. It should invite scholars to put adjuncts on their agenda again, because the theoretical and the empirical nature of the subject needs to be addressed in concert.

Natural language makes frequent use of adjuncts. You can take any kind of linguistic data pool in any particular language and you will make this observation. In fact, it is surprising to see how little is left of a sentence if you strike out the adjuncts from an arbitrarily chosen set of sentences. This simple observation can be illustrated by looking at some everyday examples, taken from an instruction manual of a digital camera.

(1)

a. With the card’s label facing toward the camera back, insert it into the slot without touching the card’s contacts with your finger.

b. If the battery is left in the camera for a prolonged period, a small amount of power current is released, resulting in excess discharge and shorter battery life.

c. To save battery power, the camera turns off automatically after about 30 seconds of non-operation.

d. If you want to zoom, do it before focusing.

Apart from the observation that, let’s say in d., only two words are not part of adjuncts—namely, do it—we can make the following points. First, most of the adjuncts in (1) are internally structured as any other phrases are and some contain adjuncts themselves too. Second, adjuncts occur in various types of sentences: finite (see (1) b. and c.) or non-finite clauses, in declarative clauses, and interrogative clauses or imperatives (see (1) a. and d.), to name just a few. The reader can easily verify this him- or herself. Third, all adjuncts clearly add to the overall meaning conveyed by the sentences in (1). Fourth, various categories are used in (1) for adjuncts; that is, prepositional phrases, adverbs, finite sentences, infinitival sentences, and gerundial expressions, which all may occur in other, non-adjunct positions too. This can easily be verified by considering
prepositional phrases, for instance, which occur in predicative (2), argument (3), and adjunct position (4), respectively.

(2)
   a. She expected the man (to be) in the garden.
   b. There is a man in the garden.
   c. A man is in the garden.

(3)
   a. They put the chairs in the garden.
   b. He was interested in the garden.
   c. He went into the garden.

(4)
   a. She read a paper in the garden.
   b. She was working in the garden.
   c. The man in the garden is her friend.

No speaker of English has any difficulties marking the adjuncts in the preceding examples. In fact, everyone has clear intuitions about what adjuncts are. Given the observation that adjuncts occur frequently in natural language and that they do not pose any particular problems for a competent speaker when it comes to distinguish them from other phrases in a structure, it is rather striking that linguists working in the tradition of generative grammar still do not agree on how to analyze these phrases. Why are adjuncts such a controversial issue? Under standard assumptions, adjuncts occur adjoined to other phrases. Chomsky (2004) describes adjuncts as being on a separate plane. What does this mean? How do adjuncts fit into a Merge-based syntax? What are the distinctive and defining properties of adjuncts? Do we need a specific operation called Pair-Merge for adjuncts (Chomsky 2004), or do proper accounts of adjuncts refer to labeling (Hornstein & Nunes 2008) and projections of features (Zeijlstra 2019)? How can we implement adjuncts while keeping to the most minimal version of a Universal Grammar (UG)? In the ideal case, there is just a single, simple structure-building device operating on linguistic expressions, generating non-adjuncts and adjuncts likewise.

Basically, the title Casting a Minimalist Eye on Adjuncts describes the general aim of this book. The term minimalist actually has two meanings. First, it stresses the fact that the minimalist program investigates language from the perspective of what is needed conceptually to explain the human capacity for language. Under the minimalist view, the previous framework of Government & Binding, with its complex architecture, had to be radically minimized. The second meaning follows from the general scientific demand to keep a theory as minimal as possible. Casting a
minimalist eye on adjuncts thus means that we need to look at generative grammar to see how adjuncts have been handled in the past and how they are captured under recent minimal approaches. Since the advent of minimalism, scientific research essentially emphasized the need for simplification. Analyzing adjuncts by means of additional mechanisms or syntactic operations goes against the strategy of minimizing UG. This book attempts to clarify the phenomenon of adjuncts on empirical and theoretical grounds. In the spirit of the strong minimalist thesis (SMT), which reduces the computational part to simple set formation—that is to say, Simplest Merge—the ultimate goal of the book is to propose an analysis of adjuncts that does not resort to any extra mechanisms. We compare previous minimalist accounts of adjuncts in terms of simplicity and show that they complicate the grammar either by putting adjuncts in a separate dimension or by requiring featural triggers.

It is noteworthy that linguists like to use adjuncts as providing evidence for the basic property of language, namely its recursive power. Adjuncts are unbounded in principle; they can occur in a stacked fashion and prepositional adjuncts frequently are used to illustrate the phenomenon of structural ambiguity (see (5) and (6)).

(5) He saw the girl [with a friend] [on Monday] [in the evening] [at the party] [in the garden].
(6) He took a picture of a man with a camera.
   a. He took a picture with a camera (not with a cell phone).
   b. A man with a camera was photographed by him.

Hence, adjuncts are pretty good candidates to exemplify that language is hierarchically structured. Therefore, it should come as a surprise that after decades of research in generative grammar, there is no consensus as to how adjuncts should be analyzed. How can it be that a phenomenon that illustrates the basic property of human language so well is still so poorly understood? The most recent research, which lays a focus on minimizing UG (see, for instance, Chomsky 2013, 2015, 2017; Epstein, Kitahara & Seely 2014, 2015a, 2015b, 2016, 2017a, 2017b; Chomsky, Gallego & Ott 2019), is promising and exciting. It is only natural, therefore, to finally include the phenomenon of adjuncts in those considerations and search for the simplest integration of these long-debated syntactic objects into minimalist analyses. Casting a minimalist eye on adjuncts while keeping track of theoretical simplicity forms the methodological strategy applied in this book. It is hoped that this book contributes to the most recent research in minimalism because it tries to fill an obvious gap, namely, to come to grips with adjuncts, which are a controversial though extremely frequent phenomenon of language, as stressed in the beginning of this introduction.

The book has two main objectives. First, we want to show that the existing accounts of adjuncts need additional mechanisms, a strategy that
is not consonant with the simplest version of minimalism. Second, we suggest a new analysis that accords with the latest minimalist reasoning. The book is structured into six chapters. After the introduction in the present chapter, Chapter Two presents a characterization of the main properties of adjuncts. Adjuncts differ from arguments in specific ways and though most readers are familiar with the basic facts, the data are included here to stress the necessity to find a minimalist analysis that captures the dichotomy. After the empirical setting we will be concerned with the minimalist framework in Chapter Three. Here we focus on basic assumptions, the structure-building operation Merge, and the question of why adjuncts pose a particular challenge to minimalist theorizing. It is worth mentioning that the book aims to address readers with different backgrounds and different degrees of expertise. Consequently, readers who are familiar with the problems posed by adjuncts and with the most recent minimalist developments might decide to skip Chapters Two and Three and move on to Chapter Four. The underlying idea of Chapters Two and Three is just to highlight central issues of the phenomenon and the framework we use before we get to specialized discussion in Chapter Four and advance a new proposal in the most current minimalist terms in Chapter Five. Chapter Four addresses specific approaches to adjuncts in minimalism. We will present the different accounts and debate the pros and cons of the respective approaches. We also attempt to clarify what an account should look like under a minimalist point of view. The different accounts are ordered and classified in sub-chapters according to their basic reasoning. First, approaches are discussed that put adjuncts into a different dimension, followed by accounts that are based on projection and labeling. Since very recent minimalist discussion elaborates on labeling and its role in a free and simple syntax, we will focus on approaches that touch upon simplification in the last sub-chapter of Chapter Four. In Chapter Five, we will finally put forth a new analysis of adjuncts referring to labeling and Transfer that is based on Simplest Merge. The gist of the analysis is that adjuncts are assembled by simple set formation just as other linguistic expressions are and the resulting structure underlies labeling conceived of as Transfer relevant to the interpretation of sets at the interfaces. We will also deal with the topic of islands here, which not only has a long tradition in generative grammar, but also relates to the phenomenon in question because adjuncts exhibit island effects cross-linguistically. The goal is to put adjunct islands into the larger context of islands and formulate a generalization in line with the most recent minimalist reasoning. In a last step, we address the issue of head movement which is related to our discussion though it is a controversial subject of its own and try to advance a potential implementation into the proposal based on labeling. A final remark on the project closes this chapter. In Chapter Six, we will summarize the most important conclusions and investigate some ideas pointing to future research.
A reviewer rightly stresses that the manner adverb actually is not needed in this case either and provides a context to the example in (4) b, such as I know what I want to say in the headline, but I haven’t worded it (yet). Hence, this basically underlines the general optionality of adjuncts.

Using a projectional system, one could also assume that only V-projections but not V-heads can be moved (Daniel Seely p.c), but a solution like this won’t work in a projection-free syntax.

In German, predicative adjectives and adjuncts in an adverbial function are formed with the same items, but attributive modifiers agree in number and case with the nominal expression to which they refer. Despite this difference in inflection, it is striking that adjectival expressions are used as adjuncts in DPs and clauses in German as well.

In the examples (18) f. and g., an upstairs reading of the respective adjuncts is strongly favored because of the affinity between how-could and when-will.

In German impersonal passives, like Gestern wurde gearbeitet (‘yesterday was worked’), the adjunct does not occupy an argument position but is internally merged to C, hence in a non-A position.

Adding phrases, as in i–ii, makes data like (21) less felicitous, again stressing the rather robust island property of adjuncts:

i. *Who did John leave the party yesterday night with?
ii. *Which song did he arrive at the station yesterday humming?

Also, one has to further check the cross-linguistic validity of (22) a. and b. In German, it is possible to extract from external and internal arguments, but not—and this is crucial—from adjuncts.

i. Von wem gefiel ihr ein Foto? (Of whom did a picture please her?)
ii. Von wem machtest du ein Foto? (Of whom did you take a picture?)
iii. *Welchen Kuchen kam er an bevor sie <welchen Kuchen> hereinbrachte? (*Which cake did he arrive before she brought in?)

Structural ambiguities in languages also provide strong evidence for hierarchical decisions as opposed to linear ones. Take a look at a German example that can be easily translated into English showing that ambiguities resolve by relying on hierarchical order and that this is observable in any language. The italicized adjunct in i. can either refer to the man or to the event of following.

i. Er verfolgt einen Mann in einem Taxi. (He follows a man in a taxi.)

Thanks go to the reviewer who reminded of the important role of Speas (1990) in the discussion of projection.

In NS, syntactic relations (among existing terms) cannot be altered throughout the derivation (Epstein, Kitahara & Seely 2012: 256).

We will not discuss solutions based on parsing, such as Phillips (2003), who assumes that structures are built incrementally top-down from left to right because they cannot be reconciled with the derivational bottom-up conception of (Simplest) Merge.

If aspectual adverbs are interpreted at a higher position, which presupposes that this higher level is already present in the child’s language, this actually is no real argument in favor of a specifier head licensing but also can be explained by referring to interpretive domains.

A reviewer objects that (48) a. is predicted to be ungrammatical in a later paper since locative PPs would be located lower in the clause. However,
Cinque’s account in the framework of cartography requires a complex feature system that is not in line with a free syntax and SMT.

7. See Bobaljik (1999) for a detailed discussion.

8. All accounts agree on the necessity to derive the interpretation of adjuncts, but they clearly differ in the relation that holds between syntax and semantics.

1. The NTC in the sense of the Law of Conservation of Relations (LCR) (see Epstein, Kitahara & Seely 2012: 256) would normally block any form of counter-cyclic Merge, which changes the syntactic objects already formed. Late-Merge actually represents a form of counter-cyclic Merge.

2. If Late-Merge exists, one would have to allow it to apply recursively too, which would lead to a massive complication. Moreover, the question arises of whether it would be possible to rule out Internal Late-Merge as a further natural option, which would basically rule in island violations. See Sportiche (2019a, 2019b) for details.

3. There are scholars who argue against differences between adjuncts and arguments with respect to reconstruction (cf. Georgi, Salzmann & Wierzba 2018 and the works cited therein). If reconstruction takes place in any case, and all copies remain in place, this would provide further evidence against Late-Merge.

4. The analysis of Takahashi and Hulsey (2009), based on Late-Merge of a complement known as “wholesale Late-Merge,” will not be discussed here. What is relevant to our discussion is that Late-Merge of the NP complement of NP is conditioned in this case. However, a conditioning or restriction on Merge goes against the conception of free Merge anyway. The same basically applies to Stepanov’s label inspection which is clearly not compatible with free Merge either.

5. A reviewer points out that there are languages (the reviewer mentions Finnish and Korean) in which adjectives, relative clauses, various PPs, and some manner adverbs are case-marked. He also objects that reconstruction is sometimes possible, and adjuncts are not always islands. First, the properties under (8) are an approximation. Second, modifying elements that are case-marked also exist in German, embedded into an argument-DP. They are not independently case-marked but agree with the D-N, and we probably deal with agreement at PHON. Third, we agree that the reconstruction data are tricky; still, as far as wh-movement is concerned, there is an observable distinction between arguments and adjuncts. As for islands, this issue will be addressed in Chapter Five.

6. Chomsky (1995: 244) discusses three options for label construction. If projection takes place when X and Y merge, the intersection or the union of X and Y, or properties of either X or Y, may label the product.

7. Given BPS, projection is taken to be a relational concept. A projection is maximal if it no longer projects. Hence, adjuncts cannot preserve a (bar-) level information.

8. Moreover, any (unmoved) adjunct would syntactically be ambiguous between concatenated and concatenated plus labeled. Since syntactic outputs get interpreted at SEM, one would, however, expect this ambiguity to correlate with some distinction in the respective interpretation. Still, there seems to be just a single interpretation available for a sentence like John met Mary at the party.

9. This would actually capture what Hornstein (2009: 109) calls the “double personality” of adjuncts, namely, that they seem to be simultaneously inside and outside a given phrase. Consequently, they can participate in VP-ellipsis or be stranded.

10. However, it is worth noticing that the idea of extended projections dating back to Grimshaw (1991 [2005]) is in fact considered by Bowers (2018),
who elaborates on it as bottom-up selection yielding clausal and nominal projections in his approach. As noted previously, Bowers (2001, 2018) also analyzes adjuncts as selecting their “host.” Thus, the idea of an extended projection would presumably be applicable to Zeijlstra’s approach. Adverbs and prepositional phrases with the V uV specification would thus be able to select aspectual verbal projections, too. In (i), one must argue that recently bearing the features V uV uPerf selects perfective have composed of the features V Perf uV. Have in turn would select the verbal participle. Note, however, that under this view, the need for more features, or more sub-features would arise, rendering the system of triggered Merge more complicated.

i.

a. He has been reading the book recently.

b. [V] [uV] [Perf] [uT] (has)

c. [V] [u Perf] [uV] [Prog] (been)

d. [V] [uProg] (reading)

e. [V] [uV] [uPerf] (recently)

It must also be guaranteed that the verbal inflections are expressed on the correct items (at PHON). For example, as can be seen in i.c. and e. the adverb has the same features as the auxiliary be (with the exception of the assumed [Prog] (sub-)feature), still, it is the auxiliary that will be inflected for the participle feature in the end and not the adverb. We might argue that since, according to Zeijlstra, -ly values the Pred of the root recent as V, there could be no second affix due to a ban on two endings, but this does not help in cases where the adverb lacks -ly, for instance, fast.

11. In a similar vein, Panagiotidis (2015: 116) assumes that there is a functional head, a supercategory called [perspective] that categorizes roots. Zeijlstra (2017: 319) differs in that he abandons the categorizer and assigns the supercategory, contentive, now [pred], directly to roots.

12. The necessity of additional features becomes evident if we look at the following examples:

i. John numbered the examples.

ii. the number of cookies

Since Zeijlstra does not assume a categorizer, but assigns the Pred feature to roots directly, number has to be valued as Pred:V in i. and Pred:N in ii. Consequently, T must have a V feature, and D an N feature that values the respective root. Now, we also need to add the respective selectional features of T and D which are needed in order to guarantee that Merge has a purpose. Does T have uV, V and D uN, V? Note that uV, V are the same features adjuncts shall be assigned. If T also selects a subject-DP, it will have the same features as a preposition, that is, V, uV, uD—the only difference being tense.

13. It is worth mentioning that Zeijlstra (2019) starts his reasoning by pointing out that the first LA (Chomsky 1995, 2000) would rest on an unmotivated assumption that the selector projects. His analysis would actually give a reason why this seems to be the case. Though matching under sisterhood and percolation of features result in the desired labeling, it has to be stressed that he actually only substitutes the “selector projects” assumption by two other stipulations. The first one is that Merge must always have a purpose (his rule 2), and the second one is that there is a specific, obligatory percolation mechanism (his rule 1). Uninterpretable features always percolate up until they find a matching feature in a sister, whereas interpretable ones would never percolate beyond their maximal projection. Note also that the latter
part of the percolation mechanism does not define projection as a relative aspect of structure-building as is normally the case since the beginning of minimalism and BPS.

14. Arguing for a PRO as the subject of the infinitival adjunct mediating the binding options so that an adjunct internal solution is possible will not help, because PRO itself needs to be interpreted. In (59) a.–b., PRO refers back to the subject of the matrix clause, and in (59) c.–d., it is controlled by the object of the main clause. Both phrases occur outside the adjunct.

15. Gallego’s proposal implies that adjuncts have to be traced back to two quite distinct situations:
   i. unlabeled syntactic objects
   ii. double, —triple, etc., peaks created by means of Parallel Merge


17. Daniel Seely (p.c.) notes that Pair-Merge is a potentially serious problem for the evolutionary argument about Merge, since two types of Merge significantly weaken the evolutionary argument.

18. Stockwell’s account builds on the idea that children acquire labeling in syntax by means of parallel derivations, leading the way to crash-proof grammar. Taking optional labeling as a preliminary hypothesis, the child exposed to a set of (AP, NP), for instance, would assume that the label could either be AP or NP, but, since the derivation goes on by merging a determiner, the label leading to an ultimately crash-proof result, would be NP.

1. Note that SOs still present (like copies) but inaccessible to syntax are different than inaccessible meaning removed from the syntax as in the sense of Narita (Daniel Seely p.c.).

2. We remain agnostic as to whether nominal expressions such as proper names involve a D-node. The possibility of the John I know point into that direction. What is relevant to our discussion is that we deal with XP in this case and not with a single head.

3. A reviewer objects that, given Rizzi’s conception of criteria, both C and the moved XP could bear [+Foc]. It would be quite easy to posit a feature, and something must trigger the movement, for example, license Internal Merge. First, the reviewer overlooks the point that we are searching for a solution in terms of Simplest Merge. Under this simpler conception, there are no triggers determining the computation. External and Internal Merge being two sides of the same coin, apply freely. Second, postulating a lexical feature for various distinct categories remains stipulative and considering Minimal Search, it is questionable how deep it can search into a set. Given that heads are atoms, it is counter-intuitive to search beyond head level. Anyway, a free syntax and a labeling algorithm that is unequivocally based on heads is much simpler than a deterministic syntax building on heads and features.

4. Note that since, labeling applies freely, you also could label both sets at once before you merge them, but then they would both be inaccessible to any further Merge operation. Hence, we only look at the labeling possibilities after Merge. There is no labeling before Merge, which follows from the intrinsic ordering of syntax taking place before semantics can apply. However, one has to investigate the option of labeling both sets at once after merge. What would be left in the syntax in this case is the unlabeled set containing the by-now labeled and transferred sets. This unlabeled set is the only thing further operations of Merge can be applied to, and whatever gets merged to it would
serve as a labeler. So, in essence the result would be truly exocentric sets of the form i. or ii. at SEM.

i. \[ \ldots Z \{ \ldots \{ Z \{ XP, YP \} \} \ldots \} \]

ii. \[ \ldots \{ Z WP \} \{ \ldots \{ Z \{ XP, YP \} \} \ldots \} \]

This possibility has to be investigated further, but it is questionable whether sets like those would be interpretable. If not, they are ruled out at SEM.

5. Notice that this does not mean that labeling is driven by interpretation. The system is not crash-proof. Some derivations yield an interpretable result, others do not.

6. Note that sentences like the following are not ruled out by labeling:

   i. *Which dog do you wonder <which dog> C John likes <which dog>?

   ii. *John thought which dog that John likes <which dog>

Epstein, Kitahara and Seely (2015: 230) argue that the first example is ruled out because the embedded question can only be interpreted properly if there is a WH-Q C-Q label. Since we do not assume that there is labeling by means of feature sharing, we cannot resort to their solution. Instead, one might think about a solution in terms of the PHON interface here requiring a WH-Q C-Q agreement and ruling out WH-Q C-decl. as a violation of this agreement.


8. Complex copies such as welches Buch (‘which book’) must not be pronounced in i. here but, as ii shows, this seems to be a phonological restriction because the simple, expletive was (‘what’) can occur instead:

   i. *Welches Buch meinst du welches Buch er gelesen hat? (‘Which book think you which book has he read’)

   ii. Was meinst du welches Buch er gelesen hat? (‘What think you which book has he read’)

9. The lowest copy at SEM is of course always semantically interpreted. External Merge necessarily has an impact on interpretation. Still, it makes sense to consider further effects at intermediate positions, provided that a phase head is involved.

10. Of course, this raises questions about v as a phase head. If we want to resort to the duality of semantics, it should be one. But do we find similar effects attested at PHON for v? Also, one has to investigate A-movement further. Under the assumption that only C is a phase head, there would only be two copies in John T seems to John like her. Moreover, the role of v in connection with roots must be further examined in this connection because roots only get categorized if they are embedded in structure. Consider i. and ii. The root record either occurs in a verbal or in a nominal environment. This observation implies that roots have no category and cannot label. The label has effects on the interpretation at SEM and PHON (compare the difference in meaning and stress patterns of i. and ii. here).

   i. to record a song

   ii. the record

11. Note that VP-fronting just depends on which set gets copied (see below). The respective labeling can still be derived accordingly, since the PP-adjunct in either case must be labeled to resolve the conflict of {{read the book} {in the garden}}.

   i. {Read the book in the garden} John did <read the book in the garden>.
ii. {Read the book} John did <read the book> in the garden.

12. As Daniel Seely (personal communication) crucially observes, there is no locality violation here. What goes wrong with this derivation has nothing (directly) to do with movement from an adjunct. Rather, what goes wrong (ultimately) is that T does not select PP (it is an interpretive problem at the interface in the analysis suggested).

13. This also accounts for the status of subjects as islands. In *[Who] did [a picture of [who]] T [a picture of [who]] please Mary? a copy would contain a copy, which makes it impossible to resolve the conflict. Copies need to be labeled (= transferred) simultaneously. To label higher copies (who/C-related) before lower copies (a picture of/T-related) would violate cyclicity. If the wh-copies are labeled first, this leaves the conflict XP (subject), YP unresolved. If the subject copies are labeled (= transferred) first in accord with cyclicity, this makes who inside the subject invisible. The wh-element in the CP would thus be taken as a repetition, which violates full interpretation, too. Because of cyclicity, it is not possible first to label the wh-copies and next to label the “subject” copies.

14. In German, the grammaticality of data like Von wem gefällt dir das Foto <von wem> (Of whom did a picture <of whom> please you) seems to go against the analysis proposed in the text. But further data (see i. here) support the view that in German, the PP von wem is actually an adjunct outside the subject.

   i. Von wem gefällt es dir? (Of whom do you like it)

15. In an example like who did you fall asleep [before meeting who], we assumed so far that extraction from the before-adjunct fails because with the conflict being resolved by labelling the adjunct, the wh-phrase gets inaccessible. Alternatively, in the other derivations, we argued that the conflict is either resolved by assigning the “wrong” label (crash at the interfaces T-PP), or the conflict remains unresolved (no label/no Transfer) which yields a kind of non-generability. Daniel Seely (p.c.) rightly asks how we can have any intuitions regarding the sentence if it is not generable. In particular, he points out that we can interpret it as for which x, you fell asleep before meeting x and that we don’t get just noise at the eardrum. Since the analysis is not framed as part of a crash-proof grammar and since we indeed have this judgment, we would expect there to be a fifth derivation fulfilling the following criteria in order to explain Daniel Seely’s observation:

   i. The wh-phrases must be analyzable as copies yielding the reading in ii.
   ii. As for which X, you fall asleep before meeting X.
   iii. The example cannot be judged as perfect but violates some condition at the interfaces.

There is in fact a derivation with these properties. Suppose we have the set {before meeting who} and we merge it with the head ‘asleep,’ which would also label the set. Hence, this would be a derivation without the conflict {XP, YP} because we would simply merge the before-set as a complement. The derivation goes on and we can finally label the wh-copies simultaneously. All sets can be labeled/transferred and we can assign the intended interpretation to the structure, but it is degraded because ‘asleep’ does not select anything. Hence, the only derivation which allows extraction yields a selectional violation (crashes at SEM) which corresponds to our judgment. So there is no derivation left that would rule in extraction from adjuncts as perfect.
16. It should be stressed here that an analysis of the that-t-Filter (see i.), or ECP-violations, in terms of labeling as in Chomsky (2015) is not available. Since we can observe cross-linguistic variation in those cases too, it would be tempting to resort to a PHON-solution here as well. It is at least marginal in German, for instance, which favors V/2 in those cases (see ii. and iii).

   i. *Who do you think who that <who> T <who> comes to the party?
   ii. ??Wer denkst du <wer> dass <wer> T <wer> zu der Party kommt?
   iii. Wer denkst du <wer> kommt <wer> T <wer> zu der Party <kommt>

17. Another pool of data normally addressed in terms of agreement is the so-called there-construction in i. here. We would suggest that the associate is actually no subject but an object or predicative related to the copula. This would capture the fact that the associate is no island and the phrase following it is, in fact, an adjunct, hence, an island. This is confirmed by the data in ii. and iii., respectively.

   i. There are [pictures of John] on the table.
   ii. Who are there pictures of <who> on the table?
   iii. *Which table are there pictures of John on <which table>?

We can also cover data such as the following in iv. by analyzing the bracketed material as an adjunct. The sentence in v. is not grammatical because the example *There is likely a man is not interpretable in the first place.

   iv. There is a man [likely to be in the garden].
   v. *There is likely a man [to be in the garden].

18. Recall that another consequence of the proposal is that labels are assigned by inspecting the head only. We just deal with heads and non-heads in syntax. Heads label sets: there is no need for searching deeper into an atom or go beyond the head level to look for features.

19. It would actually be necessary to investigate the role of roots more closely. It seems obvious that they point to conceptual content and that the combination of roots and categories in sets make available complex thoughts that enable human beings to move away from the here and the now. In general, this yields the capacity for mind-dependent, infinite thought mediated by recursive syntax. A full discussion of roots though relevant goes beyond the scope of this book.

20. Instead of roots and categories one could simply posit non-categories and categories. Linguistic atomic expressions can thus be divided into non-categories and categories. Only the latter can label sets. Crucially, meaning is conveyed in structure. Consider i. and ii.

   i. He has a dream. D-n\dream
   ii. He dreams of a better world. T-v-\dream

Basically, roots (non-categories) have no properties in isolation. They are simple units without internal structure. They occur obligatorily at the bottom of a syntactic derivation (“first Merge”), always embedded under categories. Any property is encoded in the structure (or set). Hence, meaning in general is conveyed in sets which results in a flexibility that cannot be bound to a single item. Data like iii. and iv. illustrate this.

   iii. John shouted.
   iv. John shouted the mouse out of the room.
Because meaning in this sense is always composed, human thought can only be mind-dependent. There are no elements referring to the external world.

21. The relation can be causal or temporal, what they share is the single event character, where one event necessarily either causes the other, or precedes the other. Truswell concludes that the Single Event Condition applies to complements and bare Present Participial Adjuncts likewise. It conditions A-bar-dependencies since they are legitimate only if the minimal constituent containing head and foot of the chain describes a single event (Truswell 2011: 24).

22. Notice that German does not have complements of the —ing-type. The prediction would thus be that German does not have transparent, participial adjuncts. This seems to be correct because only the full adjunct can be fronted. Extraction from the adjunct is banned:

   i. Welches Lied singend, kommt er an? (which song singing arrives he prt.?)
   ii. *Welches Lied kommt er singend an? (*Which song comes he singing prt.?)

23. Similar considerations might hold for double objects since, as Johnson observes, either the direct object or the indirect object behaves as an island, blocking either focus projection or movement (compare Johnson 2003: 213). We will have little to say about double objects here. Double objects would be necessary to investigate in detail in terms of labeling and the proposal outlined.

24. Rizzi (1990: 4), illustrates wh-islands with data like the following:

   i. ??Which puzzle do you wonder how John could solve t t?
   ii. *Which student do you wonder how t could solve the puzzle t?
   iii. *How do you wonder which puzzle John could solve t t?
   He analyzes the asymmetry in terms of his Relativized Minimality condition plus binding (via referential indices). Lasnik and Saito (1984) suggest an account based on lexical versus antecedent-government which is no longer applicable in minimalism since government has been completely eliminated. The same applies to the accounts of Chomsky (1986) and Lasnik and Saito (1984). Still, the asymmetry must be explained. Consider iv. and v. further underpinning its relevance (see Lasnik & Saito 1992: 174–175):

   iv. *Why do you wonder who left?
   v. ??What do you wonder who wrote?

Lasnik (2018) presents subtle data suggesting that binding of subjects into the wh-domains influences the acceptability of wh-islands. He puts forth an interesting view based on feature valuation. However, the strong violation of adjunct extraction as opposed to objects as in the data vi-vii (see Lasnik 2018: 69) cannot be captured this way:

   vi. What problem did you figure out [how [PRO to solve t t]]?
   vii. *How, did you figure out [which problem, [PRO to solve t t]]?

Lasnik puts this question aside for future investigation (Lasnik 2018: 69).

25. Note that we can apply this account of islands to complex-NP islands as in i. as well since relative clauses are standardly analyzed as adjuncts:

   i. *Which man did they meet [the girl who likes <which man>]

26. Thanks go to a reviewer for mentioning Sprouse and Hornstein (2013).

27. We refer the reader to the framework of distributed morphology (see for instance, Halle & Marantz 1993; Embick & Noyer 2007; Embick 2015; Harley 2014), the exo-skeletal account of Borer (2003, 2013, 2014) or a
derivational solution to roots suggested by de Belder and Craenenbroeck (2015).

28. Recall that we suggested that categories and non-categories (roots) are atoms of the computation which get instantiated at the point of externalization. The sets created in the computation can only be labeled by categories.

29. There are, of course, also examples demonstrating that a complement can be merged involving the same root but distinct categories as i and ii show:

   i. This is similar to his analysis.
   ii. There is some similarity to his analysis.

30. Recall that we assume that the atoms of human language are categories (such as C, T, D, v, a, n) on the one hand and non-categories (roots) on the other hand.

31. That labeling takes place in cyclic fashion is assumed for independent reasons. Subject copies have to be labeled prior to copies which have an instance merged to higher C, for instance, wh-phrases. Remember that if higher wh-phrases would be transferred before the lower DP-subject copies in i., subject-islands would be ruled in. Since it applies cyclically, the subject copies are labeled and transferred first. Consequently, there cannot be a wh-copy contained in a subject-copy. Since the derivation is bottom-up, it is no strange assumption that labeling takes place in a cyclic fashion, too, though apart from this it is free in principle and underlies interpretation at SEM.

   i. *Who does [a picture of who] T <a picture of who> please John?

32. Note that “cyclic” refers to the level of the respective sets. So, for instance, DP and root-v copies are actually sets at the same level since their respective copies are contained in sister sets. Reversing the order of 1–2 would not violate cyclicity, but would simply yield no felicitous result at SEM. However, labeling of wh-copies before the DP-copies would violate cyclicity.

33. See Dékány (2018) for an excellent, critical overview of approaches to head movement.

34. A syntactic approach to head movement is further complicated by reasons of locality. Carstens, Hornstein and Seely (2016) discuss a problem of equidistance arising with both T-to-C and v-to-T because the D-head of the external argument should be available too. Therefore the question arises as to what guarantees that it is v or T that move, but not D, which should be found by Minimal Search, too. The authors discuss a possible solution by assuming feature-inheritance or feature-copying. Features and triggers would, of course, complicate the picture.

35. A good candidate for a language-particular rule applying at PHON is do-support. It is necessary in English (see i.) but not in standard German (see ii.), to insert do in certain contexts though there is some minor variation in German. Some non-standard dialects allow a supportive element tun (see iii.).

   i. Does he work a lot?
   ii. Arbeitet er viel?
   iii. Tut er viel arbeiten?

36. Positional effects might of course also be traced back to alternative copy-pronouncement if Internal Merge has taken place in syntax.

37. A suggestion might appeal to features and heads at PHON. If Agree which is highly parametric in nature applies at externalization and the linear order of heads is also treated here, the following patterns might be explicable.

   i. Who C <who> T <who> v-saw John?
   ii. Who C John T John v-see <who>?
ii. Where C John T John v-see Mary <where>?

iv. Who C <who> T Neg <who> v-see John?

In ii., there is one uninterrupted or adjacent sequence of phi and Q features and do-support does not apply. C in ii and iii. share the Q feature with the wh-phrase, but not the phi feature with John/T. Hence, at C the sequence is interrupted, which needs to be repaired by insertion of a form of do. In negated sentences, the Neg head has the same interrupting effect because it does not share any of those features. Consequently, do needs to be inserted (see iv.).

Comparing English and German, we need to refer to the specific heads (categories) and to the positions on the sequence. In English, only the highest head with a Q-feature needs to be pronounced by a category (not including the lexical root). In German, the highest head must be overt in any case and this includes categories with roots. Thus, we either have V/2 or an overt complementizer dass (that).

i. Gestern C-arbeitete er (Yesterday worked he/He worked yesterday).

ii. C-dass er gestern arbeitete (that he yesterday worked/that he worked yesterday).

This would mean that, in principle, any head of the sequence C-T-v-root could be overt, but that it is an idiosyncratic decision of particular languages at which position the respective element bearing specific features is expressed phonologically.

38. At the end of Chapter Four, we thought about an evolutionary scenario based on the following steps i. atoms: root, categories ii. recursive Merge: SO, SO $\rightarrow$ labels (=Transfer) $\rightarrow$ SEM (uniformity) iii. Map: SEM, PHON features $\rightarrow$ Externalization iv. Agree: feature values: F, Fv (diversity). The idea behind this is that cross-linguistic variation is associated with the externalized PHON side only. We assume that all languages have categories like C, T, v, D, n, and so on, and those categories can label = transfer. Semantic features like (non-)finiteness, (in)definiteness, q(uestion), and so on can be externalized. PHON pronounces SEM. “Nominal” forms (like case) and “verbal” forms (any inflections) would under this view be PHON-members. The same applies to agreement (in feature values) and copy pronouncing. These properties are supposed to be highly language-particular in nature. Consider some examples that would be violations at PHON in English under this view:

i. *Him likes she. (*Acc < Nom)

ii. *They tried to liked her. (*to Past-ed

iii. *They likes him. (*phi-pl-phi-sing)

iv. *They wonder Bill C Mary likes. (*non-Q-C-Q)

39. In general, PHON deals with linearization and with realization. In the same vein, the idea of sequences in the case of do-support should be understood.

40. Narita (2012) suggests an account in terms of phases to explain the complementarity of head and phrasal movement: T/*TP, DP, *D).

41. At PHON, case forms are probably related to the respective heads Tfin—nominative and vtrans accusative.

42. The cyclicity of labeling and the respective order needs to be scrutinized as well. The conflicts are said to be resolved immediately after External Merge or immediately after Internal Merge. The latter (“after movement”) necessarily implies a delay because copies need to be labeled simultaneously. The examples with head movement would point into the direction that labeling in this case is in fact delayed further—we cannot, for instance, label the
PerfP in the example we discussed in this section before we reach the top of the derivation because this would render the PerfP invisible and inaccessible to syntactic operations. Neither the DP subject(-copies) nor the wh-copies would thus be accessible. This is clearly a false prediction as the reader will have already recognized.

43. De Belder and Craenenbroeck (2015) even assume that first Merge joins a categorizer and an empty slot that is replaced by root material at PHON. It also is impossible, under this view, to assume that this operation is triggered by features.

1. Of course, there are ways to capture the semantics of adjuncts. They might be taken as parts of the predicates (by means of predicate composition) or the rules of interpretation may analyze adjuncts as predicates themselves requiring a “subject.” For instance, the adjunct in i. might semantically function as a predicate, which makes the meaning similar to the meaning assigned to ii.

   i. He left the room [very quickly].
   ii. [His leaving of the room] was [very quick].

2. In some more recent lectures in 2019, Chomsky also figured out the role of MERGE and the workspace.


Chomsky, N. 2015b. Problems of Projections: Extensions. In di Domenico, E.,


Ishii, T. 2016. Transfer as Self Pair-Merge. School of Arts and Letters, Meiji University, Bungeo Kenkyuu (Studies in Literature) 129: 5–33.


Seely, T. D. 2006. Merge, Derivational C-Command and Subcategorization in a

Sportiche, D. 2019a. Neglect (or Doing Away with Late Merger and Countercyclicity). Available at Lingbuzz/002775.


Babaljik, J. D. 1995. In Terms of Merge: Copy and Head Movement. *MIT Work-


Hornstein, N. and Nunes, J. 2008. Adjunction, Labeling, and Bare Phrase Struc-
ture. *Biolinguistics* 2: 57–86.
Sportiche, D. 2019a. Neglect (or Doing Away with Late Merger and Countercyclicity). Available at Lingbuzz/002775.