

ABSTRACT

Title of dissertation: RELATING MOVEMENT AND ADJUNCTION
IN SYNTAX AND SEMANTICS
Timothy Hunter, Doctor of Philosophy, 2010

Dissertation directed by: Professor Amy Weinberg
Department of Linguistics

In this thesis I explore the syntactic and semantic properties of movement and adjunction in natural language, and suggest that these two phenomena are related in a novel way. In a precise sense, the basic pieces of grammatical machinery that give rise to movement, also give rise to adjunction. In the system I propose, there is no atomic movement operation and no atomic adjunction operation; the terms “movement” and “adjunction” serve only as convenient labels for certain combinations of other, primitive operations. As a result the system makes non-trivial predictions about how movement and adjunction should interact, since we do not have the freedom to stipulate arbitrary properties of movement while leaving the properties of adjunction unchanged, or vice-versa.

I focus first on the distinction between arguments and adjuncts, and propose that the differences between these two kinds of syntactic attachment can be thought of as a transparent reflection of the differing ways in which they contribute to neo-Davidsonian logical forms. The details of this proposal rely crucially on a distinctive treatment of movement, and from it I derive accurate predictions concerning the equivocal status of adjuncts as optionally included in or excluded from a maximal projection, and the possibility of counter-cyclic adjunction. The treatment of movement and adjunction as interrelated phenomena furthermore enables us to introduce a single constraint that subsumes two conditions on extraction, namely adjunct island effects and freezing effects.

The novel conceptions of movement and semantic composition that underlie these results raise questions about the system’s ability to handle semantic variable-binding. I give an unconventional but descriptively adequate account of basic quantificational phenomena, to demonstrate that this important empirical ground is not given up.

More generally, this thesis constitutes a case study in (i) deriving explanations for syntactic patterns from a restrictive, independently motivated theory of compositional semantics, and (ii) using a computationally explicit framework to rigourously investigate the primitives and consequences of our theories. The emerging picture that is suggested is one where some central facts about the syntax and semantics of natural language hang together in a way that they otherwise would not.

RELATING MOVEMENT AND ADJUNCTION IN SYNTAX AND
SEMANTICS

by

Timothy Andrew Hunter

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:
Professor Amy Weinberg (chair)
Professor Norbert Hornstein
Professor Paul Pietroski
Assistant Professor Alexander Williams
Professor Allen Stairs

© Copyright by
Timothy Hunter
2010

ACKNOWLEDGEMENTS

It should go without saying that I owe many thanks to many people.

Amy Weinberg has been tremendously helpful in guiding me through everything I have worked on over the last five years, nudging me just enough in what have always turned out to be the right directions to productively develop my interests — even in the early days when I could barely describe why I wanted anything to do with this linguistics caper, and she could have been forgiven for wondering what she had gotten herself into.

Norbert Hornstein has not just helped me with all things syntax, but also deepened and sharpened my understanding of what a theory of linguistics (or of anything else) is and how to go about constructing and analysing one. It is impossible to not be inspired by his intellectual enthusiasm.

Paul Pietroski has opened my eyes to new ways to think about logic and other formal systems, and how they relate to linguistics and cognitive science; and like Norbert, has taught me a great deal about how to do theoretical work. Much of what Paul has said to me seems to become more insightful the more I think about it.

Alexander Williams has spent many hours talking with me and meticulously reading through drafts, helping me get many important details just right. I have learnt a lot from him about the perspective that comes from a familiarity with a variety of grammatical frameworks, and I will miss his sharp dry wit which regularly reduces me to stitches.

The fact that Jeff Lidz was not part of the official advisory committee for this dissertation in a way sells short the valuable role he has played, through collaborations on other projects, in my overall training as a linguist.

Juan Uriagereka and Howard Lasnik were influential teachers in early syntax classes, and provided occasional but very helpful pieces of advice in the course of my work on this dissertation. I am also grateful for many discussions with Philip Resnik and Colin Phillips that have challenged, and led me to better understand, some of the framing assumptions on which I have come to base my work.

Being a part of the UMD linguistics graduate student community has been a fantastic experience. I won't try to list everyone I had so many productive conversations and good times with, but special mention to: my class cohort and early officemates Chris Dyer, Shiti Malhotra and Akira Omaki; my more recent officemates Alex Drummond and Terje Lohndal, and frequent interloper Dave Kush, for many fun syntax conversations (and to Terje especially for access to his impressive bookshelf); Brian Dillon, Ellen Lau, Akira Omaki and Masaya Yoshida, for many fun psycholinguistics conversations; and to Ariane Rhone for being a great friend and neighbour, if barely even half a linguist.

Outside of Maryland, Ed Stabler and Greg Kobele have both taken significant amounts of time to read my work and talk with me about it. Without their help my understanding of Minimalist Grammars would be much shallower, and this dissertation would certainly be worse for it.

Perhaps most of all I owe thanks to Mengistu Amberber, who got me hooked on linguistics when I happened to take a couple of elective classes towards the end of my undergraduate study at UNSW in Sydney. I think I must have been rather stunned at first when he suggested that I consider going halfway around the world for graduate studies in linguistics, but I am extremely glad that he did.

Finally, thanks to my parents for their constant support of these mystifyingly esoteric exploits, and to Stacey for making life so much fun in all the ways that aren't related to work.

Contents

1	Introduction and Background	1
1.1	Goals	1
1.2	Methodology	2
1.3	Structure of the thesis	4
1.4	Motivation: Why movement and adjunction?	5
1.5	The MG formalism	9
1.5.1	The original MG formalism	9
1.5.2	Reduced representations for MG expressions	15
1.5.3	Movement as re-merge	21
1.5.4	A remark on notation	25
1.6	The Conjunctivist conception of neo-Davidsonian semantics	28
1.6.1	Neo-Davidsonian logical forms	28
1.6.2	Conjunctivist semantic composition	35
1.6.3	Conjunctivist details	46
1.6.4	Potential objections	52
2	Arguments, Adjuncts and Conjunctivist Interpretation	59
2.1	Overview	59
2.2	Syntactic properties of arguments and adjuncts	61
2.2.1	Descriptive generalisations	61
2.2.2	Adjuncts in the MG formalism	66
2.3	Syntactic consequences of Conjunctivism	70
2.4	Conjunctivist interpretation of MG derivations	75
2.4.1	Getting started	76
2.4.2	Interpretation of arguments	81
2.4.3	Interpretation of adjuncts	86
2.5	Discussion	96
2.5.1	Potential objections	97
2.5.2	The role of syntactic features and semantic sorts	101
2.6	Counter-cyclic adjunction	107
2.6.1	MG implementations of counter-cyclic adjunction	108
2.6.2	Constraints on counter-cyclic adjunction	115
2.7	Conclusion	117
2.A	Appendix: Structures with <i>vP</i> shells	118

3	Adjunct Islands and Freezing Effects	126
3.1	Overview	126
3.2	Previous accounts of adjunct islands and freezing effects	127
3.2.1	Early work: non-canonical structures	127
3.2.2	The complement/non-complement distinction	130
3.2.3	Subject islands as freezing effects	134
3.3	Constraining movement	138
3.3.1	Extraction from adjuncts, as currently permitted	139
3.3.2	Prohibiting extraction from adjuncts	141
3.3.3	Freezing effects follow	145
3.4	Remnant movement	148
3.4.1	Predictions for remnant movement	148
3.4.2	English VP-fronting	149
3.4.3	German “incomplete category fronting”	153
3.4.4	Japanese scrambling	154
3.5	Conclusion	155
4	Quantification via Conjunctivist Interpretation	157
4.1	Basic semantic values	158
4.2	Assignments and assignment variants	161
4.2.1	Pronouns and demonstratives	161
4.2.2	Tarskian assignment variants	163
4.3	Syntactic details	166
4.4	Semantic details	175
4.5	Multiple quantifiers	184
4.5.1	Semantics	185
4.5.2	Syntax	187
4.6	Discussion	191
4.6.1	Quantifiers as adjuncts	192
4.6.2	The target position of quantifier-raising	194
5	Conclusion	196
	Bibliography	198

Chapter 1

Introduction and Background

1.1 Goals

In this thesis I will explore the syntactic and semantic properties of movement and adjunction in natural language, and I will suggest that these two phenomena are closely related in a novel way. In a precise sense, the basic pieces of grammatical machinery that give rise to movement, also give rise to adjunction. In the system I propose, there is no atomic movement operation and no atomic adjunction operation; the terms “movement” and “adjunction” serve only as convenient labels for certain combinations of other, primitive operations. As a result the system makes non-trivial predictions about how movement and adjunction phenomena should *interact*, since — in contrast to systems with distinct atomic operations for movement and adjunction — we do not have the freedom to stipulate arbitrary properties of movement while leaving the properties of adjunction unchanged, or vice-versa. I will:

- argue that in this way the system provides new insights into some well-known properties of movement and adjunction, among them
 - the equivocal and ambiguous status usually assigned to adjuncts in order to permit them to be optionally included in or excluded from a maximal projection (chapter 2), and

- the prohibitions on movement out of adjoined constituents and movement out of moved constituents, which are unified in the framework I propose (chapter 3);

and then

- demonstrate that the system I develop maintains the ability to handle semantic variable-binding phenomena, despite unconventional conceptions of movement and semantic composition that may make this appear difficult, by presenting an unconventional but descriptively adequate account of standard quantificational phenomena (chapter 4).

1.2 Methodology

As the statement of goals above suggests, I will treat movement and adjunction as *phenomena* to be accounted for; the corresponding empirical observations concerning these phenomena are descriptive generalisations that have been noted in the past. Research from “the GB era”¹ provides the kind of generalisations that I will take as empirically correct observations. In this methodological respect I follow the minimalist research program stemming from Chomsky (1995) and developed in much work since; see especially Hornstein and Pietroski (2009).² This application of Ockham’s razor, aiming to reduce the inventory of primitives in linguistic theory, is independent of any empirical hypotheses about the extent to which linguistic systems are simple or complex by any independently established metrics.³

¹Among many others: Chomsky (1981, 1986), Baker (1988), Cinque (1990), Haegeman (1994), Hale and Keyser (1993), Huang (1982), Kayne (1984), Lasnik and Saito (1992), Lasnik and Uriagereka (1988), Rizzi (1990).

²Among many others: Chomsky (2001, 2004), Baker (2003), Boeckx (2008), Bošković and Lasnik (2007), Chametzky (1996), Epstein and Hornstein (1999), Epstein and Seely (2002), Fox and Pesetsky (2005), Hornstein (2001, 2009), Kitahara (1997), Martin et al. (2000), Uriagereka (1998).

³See Epstein and Hornstein (1999, pp.xi–xii) on “methodological economy” versus “linguistic economy”, and Roberts (2000), Chomsky (2002, p.98) and Culicover and Jackendoff (2005, pp.89–90) and on “methodological minimalism” versus “substantive minimalism”.

As well as the minimalist methodological viewpoint, I also take from recent syntactic research two ideas that guide my analysis of movement and adjunction: first, the idea that movement is usefully thought of as merely re-merging (Kitahara 1997, Epstein et al. 1998), and second, the idea that adjuncts are in some sense more loosely attached than arguments are (Chametzky 1996, Uriagereka 1998, Hornstein and Nunes 2008). In addition, I draw on the neo-Davidsonian tradition in semantics (Parsons 1990, Schein 1993), the particular role of which I delineate more clearly shortly. In order to explore the consequences of these ideas, and the relationship between them, I express them in a variant of the Minimalist Grammar (MG) formalism (Stabler 1997, 2006), a computationally precise formulation of the central and widely-adopted aspects of the system presented in Chomsky (1995). In this more mathematically explicit setting the full ramifications of a particular adjustment, or set of adjustments, to the framework can be rigorously investigated; it is therefore a useful setting in which to pursue theoretical goals of the sort outlined above.

While this project therefore has much in common with other minimalist syntax research, the system I propose may “look different” from what these similarities might lead the reader to expect. I believe there are two main reasons for this, and I hope that a brief mention of them here will prevent these differences in appearance from obscuring the extent to which my proposal is “normal minimalism”.

The first reason is that, as just mentioned, I will be presenting everything in a more mathematically explicit manner than is typical; even at the beginning of the thesis, before I present anything particularly novel, the way in which familiar ideas are represented may appear strange at first. More importantly, however, a central theme throughout will be to argue that, with the help of this formal approach, existing ideas can be followed to conclusions that might otherwise go unnoticed, and so it should not surprise that the system I arrive at by the end of the thesis differs (not just in its degree of explicitness, but also) in some substantive ways from the conventional.

The second reason is that I will address the semantic, as well as syntactic, properties of (sentences analysed in terms of) movement and adjunction. I take a theory of linguistic competence to be a system that generates sound-meaning pairings, and correlatively the observations that I aim to explain concern (roughly-GB-era generalisations about) native speakers’ judgements of the acceptability of such pairings. I will be guided by the need to provide a syntax that supports the composition of neo-Davidsonian logical forms, according to a particular “Conjunctivist” view (Pietroski 2005, 2006), and I will argue that some well-known grammatical generalisations follow relatively naturally from the syntax that this naturally leads us to. The style of reasoning is familiarly minimalist throughout: this applies to both (i) as Pietroski (forthcoming) emphasises, the arguments for the particular conception of semantic composition that I adopt, and (ii) the attempt to construe properties of the grammar as natural consequences of its interaction with logical forms.

1.3 Structure of the thesis

In the remainder of this chapter, I discuss some reasons to think that movement and adjunction are particularly interesting phenomena to focus on in §1.4, and then outline two areas of previous work that provide the starting points for the thesis. In §1.5 I introduce the MG formalism. It can not be over-emphasised that the large majority of this section aims only to familiarise the reader with the terminology and notation that will be used to make some relatively standard assumptions explicit. In §1.6 I present the Conjunctivist approach to the semantic composition of neo-Davidsonian logical forms. In contrast to the discussion of the MG formalism, and crucially, this section does introduce a (very strong) empirical hypothesis, not merely a notation. Then the rest of the thesis, in large part, uses the MG formalism to rigorously explore the consequences of this empirical hypothesis.

In chapter 2 I address the nature of adjuncts and their relationship to arguments.

I argue that the system introduced in §1.5 permits a particular analysis of adjunction in which certain syntactic behaviour can be seen as a natural result of their role in Conjunctivist composition. The empirical target of explanation in this chapter is the observation that there are two distinct ways (not one, and not more than two) in which the constituents X and Y can combine to form a new constituent headed by X .

In chapter 3 I argue that the system developed in chapter 2 has a further beneficial syntactic consequence, namely that it permits a natural unification of two kinds of domains that are well-known to prohibit extraction: adjoined constituents (“adjunct island” effects) and moved constituents (“freezing” effects).

In chapter 4 I show how the system developed so far can be extended to handle the semantics of quantificational sentences. The main aim here is to address worries that the framework, whatever its syntactic benefits, will struggle to deal with more complex semantic phenomena. In particular, the relatively novel approach to movement and the strongly event-based approach to semantics may raise doubts over the ability of the system to handle “variable binding”. I present an account of quantification that (i) at the very least, covers the facts that more standard approaches cover, (ii) does so within the rough confines of the relatively restrictive system motivated by the aims of chapter 2, and (iii) suggests interesting lines of investigation for further research into the nature of quantifier raising and its relation to other approaches to quantification.

1.4 Motivation: Why movement and adjunction?

It is interesting to briefly consider the ways in which movement and adjunction phenomena have been treated in various grammatical frameworks. To a reasonable first approximation, different frameworks tend to have at their base a syntactic composition mechanism for expressing very local dependencies (canonical cases being perhaps dependencies between a verb and its arguments) which has roughly the generative

capacity of a context-free grammar (see Figure 1a). Although the notation differs, in each framework there is a basic mechanism for encoding statements to the effect that putting *this* kind of thing next to *that* kind of thing produces one of *this other* kind of thing. Where different frameworks diverge is in their treatment of linguistic phenomena that do not fit so well into this mould, and these phenomena — again, to a first approximation — are roughly those of movement and adjunction.

Consider first, for example, the categorial grammar tradition. Underlying many variations is the system of basic categorial grammars or “AB grammars” (Ajdukiewicz 1935, Bar-Hillel 1953), which have expressive power comparable (weakly equivalent) to that of context-free grammars (Bar-Hillel et al. 1960); see Figure 1b. In order to model dependencies spanning longer distances, this core system has been supplemented in two different ways: either by interpreting the basic combination rule as a sort of inference and exploring extensions suggested by analogy with logics, in particular hypothetical reasoning (Lambek 1958, 1961, Morrill 1994, Moortgat 1988, 1996, 1997, Carpenter 1997, Jäger 2005); or by adding other combinators as primitives alongside the basic ones (Bach 1976, 1979, Dowty 1982, 1988, Geach 1972, Ades and Steedman 1982, Steedman 1988, 2000). Both of these strategies can be thought of as roughly emulating (what is treated in the Chomskyan tradition as) movement. Adjunction is not usually treated as an independent phenomenon at all in these frameworks: what one might otherwise think of as an adjunct to category X is typically simply assigned a category of X/X or $X\backslash X$, encoding the fact that it attaches to a constituent of category X and a new constituent of the same category results. This approach gets the basic distributional facts right but is incompatible with some common assumptions about “headedness”, and makes it unclear how to account for any distinctive syntactic properties that adjuncts have since they are assumed to compose via the very same syntactic operation as all other constituents. I return to these issues in §2.2.

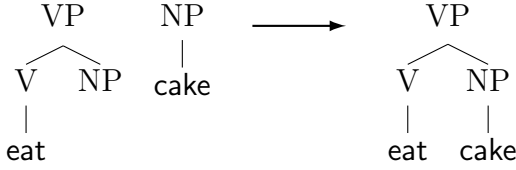
$\begin{array}{l} \text{VP} \rightarrow \text{V NP} \\ \text{V} \rightarrow \text{eat} \\ \text{NP} \rightarrow \text{cake} \end{array}$ <p>(a) Context-free string-rewrite grammar</p>	$\frac{\text{eat} : \text{VP/NP} \quad \text{cake} : \text{NP}}{\text{eat cake} : \text{VP}}$ <p>(b) Categorical grammar</p>
 <p>(c) Tree adjoining grammar</p>	$\frac{\text{eat} : =n \text{ v} \quad \text{cake} : n}{\text{eat cake} : v}$ <p>(d) Minimalist Grammar</p>

Figure 1: Various ways to model canonical local dependencies

The more recent tree-adjoining grammar (TAG) formalism (Joshi 1987, Joshi and Schabes 1997, Abeillé and Rambow 2000) can likewise be construed as a context-free base component supplemented in a distinctive way. The underlying base here is a system of elementary tree structures composable via the Substitute operation, again comparable (weakly equivalent) to context-free grammars; see Figure 1c. The full TAG formalism adds a second primitive operation, Adjoin, which interleaves the contents of two tree structures. The addition of this operation has consequences for both adjunction and movement phenomena. Most straightforwardly, it provides a mechanism for attaching optional or “unselected” (eg. adjectival and adverbial) constituents in a way that does not alter the valence of the host, as required for the basic distributional properties of adjunction. But the very same operation is also used to create long-distance dependencies that extend beyond a certain local domain, and this distinctive treatment has been argued to reveal natural explanations for some of the constraints on the formation of such long-distance dependencies (Kroch 1987, 1989, Frank 2002). The TAG formalism therefore shares with the current proposal the idea that a single addition to a context-free base system plays a role in producing the two classes of phenomena that I call movement and adjunction.

In early transformational grammars (Chomsky 1957, 1965), it was particularly explicit that movement phenomena were produced by a certain addition to a context-free base: a set of phrase-structure rules produced an underlying structure or deep structure, to which transformations were then applied. In more recent work in the framework of Chomsky (1995) this picture has changed, in that the steps of a derivation that create dependencies previously established at deep structure are interspersed with steps that create dependencies previously established via transformations. The former operation (i.e. merge) has clear similarities to the core syntactic composition operations from categorial grammars and TAGs discussed above; I will use the notation in Figure 1d, introduced more formally in the next section, for this.⁴ The latter operation (i.e. move and/or copy) is what the minimalist framework adds to this base to produce movement phenomena. The precise nature of adjunction configurations has varied. At first they were simply a result of extrinsic X-bar node labelling conventions (Chomsky 1981), but at least since Lebeaux (1988) there have been suggestions that adjunction could be thought of as the result of a distinct combinatory operation. Something along these lines would appear to be the only feasible option if we take the minimalist step of eliminating the possibility of X-bar-style labelling, as I will discuss in §2.2. Under the analysis that I will propose, adjunction is a genuinely distinct form of syntactic composition, but one that arises as a possibility as a by-product of the additions to the minimalist system already required for movement.

In the following section I introduce the particular formulation of minimalist syntax that I will take as a starting point. This introduction will cover only the merging operation that plays the core context-free structure-building role, and the ways in which this can be supplemented to model movement. The question of how to treat adjunction will then be addressed in chapter 2.

⁴For discussion of the parallels between minimalist syntax and categorial grammar, and more specifically between the merge operation in Figure 1d and the function application or slash-elimination rule in Figure 1b, see Lecomte and Retoré (2001), Retoré and Stabler (1999, 2004), Lecomte (2004), Cornell (2004).

1.5 The MG formalism

Stabler (1997) presents the Minimalist Grammar (MG) formalism as a precise formulation of many of the ideas from Chomsky (1995). MGs provide a relatively neutral setting for rigorous exploration of the consequences (as well as the more abstract computational properties) of theoretical proposals, using notation and terminology closely related to that which mainstream minimalist syntacticians are familiar with.

I introduce the MG formalism as originally presented by Stabler (1997) in §1.5.1, and an importantly insightful reformulation of this same system in §1.5.2. With a minor modification (Stabler 2006) which I describe in §1.5.3, the formalism can be adapted to provide an explicit formulation of the conception of movement as “re-merging”. Although it is this modified system that is the point of departure for the original contributions of the remaining chapters, an understanding of how that system has evolved from the others discussed here will be useful for grasping crucial intuitions later.

Readers familiar with the MG formalism may wish to skip or skim at least §1.5.1 and §1.5.2.

1.5.1 The original MG formalism

The basic building blocks from which larger expressions are constructed are lexical items. For present purposes, I will take a lexical item to be a pairing of a string and a list of features.⁵ I will write these two components separated by a colon. For example, `like:=d V` is a lexical item consisting of the string `like` and a list of two features, `=d` and `V` in that order. The order is significant: features are “checked” in a strictly left-to-right order. This particular list of features indicates that this lexical item selects a DP, and then can act as a VP. Similarly, `the:=n d` is a lexical item that selects an NP and then can act as a DP; and `book:n` is a lexical item that

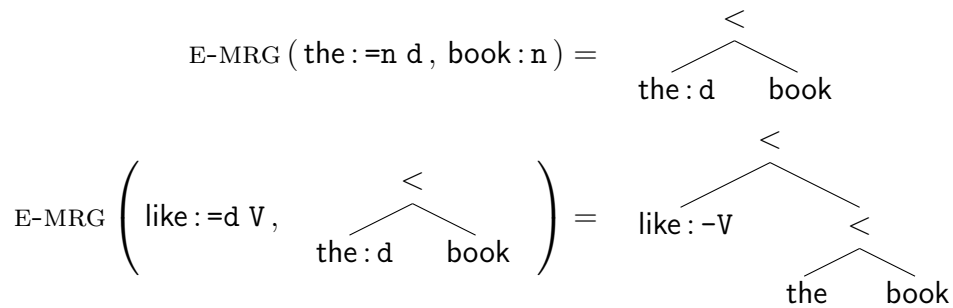


Figure 2

does not select anything, rather it is ready to act as an NP.

The two functions that Stabler (1997) introduces for constructing larger expressions out of smaller ones — and lexical items are just small expressions — are E-MRG and I-MRG.⁶ The function E-MRG takes two expressions as input and combines them into a new expression; I-MRG takes one expression as input and rearranges its contents. The three lexical items given as examples above can be combined to form a VP via two applications of E-MRG, as shown in Figure 2.

The first application of E-MRG takes the two lexical items `the : =n d` and `book : n`, and produces a tree structure representing a DP. The `=n` and `n` features are deleted or “checked”, and do not appear in the resulting expression; we call features of the form `=f` “selector” features, and those of the form `f` “selectee” features. Only the `d` feature remains, and it is in virtue of this `d` feature that the new expression is understood to be a DP. More specifically, it is because this `d` feature is on the leaf of the tree that is the *head* of the new expression: the head of the result of an application of E-MRG is the head of the input expression which had a feature of the form `=f` checked. We

⁵To understand our grammars as generating a set of string-meaning pairs, it is necessary to take lexical items to be *triples*, consisting of a string, a meaning and a list of features. For the purposes of this introduction we can limit our attention to the syntactic system understood as a generator of strings. This will be supplemented with a semantic component in chapter 2.

⁶What I call E-MRG and I-MRG are usually called MERGE and MOVE, respectively. The reason for this departure from convention will become clear below.

indicate the head of a complex expression by labelling its root with either ‘<’ or ‘>’, such that this label “points to” the expression’s head. Since it was `the:=n d` and not `book:n` that had its `=n` feature checked, the head of the resulting expression is `the:d` and not `book`. This is indicated by labelling the root of the resulting tree structure with the symbol ‘<’, “pointing to the left”. This allows us to define familiar notions such as projection, complement and specifier in the obvious ways. The trees that are produced by E-MRG (and I-MRG) are ordered: complements are always to the right of their heads (at least for the purposes of this thesis), and specifiers always to the left.⁷

The second application of E-MRG shown in Figure 2 uses the DP resulting from the first application as one of its arguments. The `d` feature on the head of this expression “matches” the `=d` feature on `like:=d V` in the same way that the `n` feature on `book:n` “matched” the `=n` feature of `the:=n d`. (A lexical item, understood as a simple expression, is the head of itself.) In the expression resulting from this second application, the `=d` and `d` features are checked as the `=n` and `n` features were before, and the “selecting” expression is again established as the head by assigning a label that “points” to it.

The other way of constructing larger expressions out of smaller ones in the MG formalism is via I-MRG. To set up an expression to which I-MRG is applicable, we will need a lexical item with a new kind of feature.⁸ The lexical item `which:=n d -wh` is much like `the:=n d`, the only difference being the additional `-wh` at the end of its feature sequence. We will see shortly that this will trigger an application of I-MRG — informally, a “movement step”. Before that can happen, however, this new lexical item plays the same role in the two applications of E-MRG in Figure 3 as `the:=n d`

⁷More formally, each pair of sister nodes is actually ordered by the linear precedence relation as well as by the “projects over” relation we denote by ‘>’; see Stabler (1997, p.69–71). Adopting a consistent specifier-head-complement order means that movement must be used to derive word orders other than SVO; Stabler (1997, pp.79–83) gives some simple examples of how to achieve this for SOV and VSO languages. Alternatively one could very easily parametrise the system to permit head-final linear order, without disrupting any properties of the formalism relevant to this thesis.

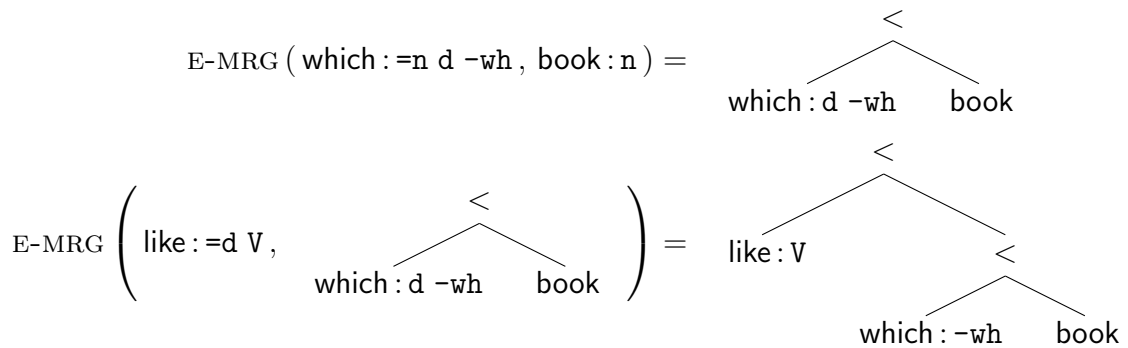


Figure 3

did in those in Figure 2.

The $-wh$ feature remains unchecked after these two applications of E-MRG. Because this feature is not on the head of the expression we have now constructed, it does not affect the immediate future of this expression, which is ready to be selected as a VP since its head bears a V feature. Leaving aside VP-internal subject positions for the moment, we will continue and build up a TP via two more applications of E-MRG, using two additional lexical items: $will : =V =d t$ (i.e. a T head which selects a VP and a DP), and $John : d$ (i.e. a DP). The completed TP will be selected by the lexical item $\epsilon : =t +wh c$, where ϵ is the empty string; E-MRG applies to these two expressions as shown in Figure 4. Note the “right-pointing” label at the root of the complex TP, as a result of the fact that the T head has selected two constituents and therefore **John** is a specifier.

We have now built an expression (specifically, a C' constituent) to which I-MRG can apply. The $+wh$ feature on the C head “looks for” the $-wh$ feature on **which** in much the same way that $=d$ features look for d features, for example; the only difference being that while $=d$ features look for d features on (the head of) some

⁸The assumption that the features checked by E-MRG (“merge”) and I-MRG (“move”) steps are different in kind will be rejected in §1.5.3.

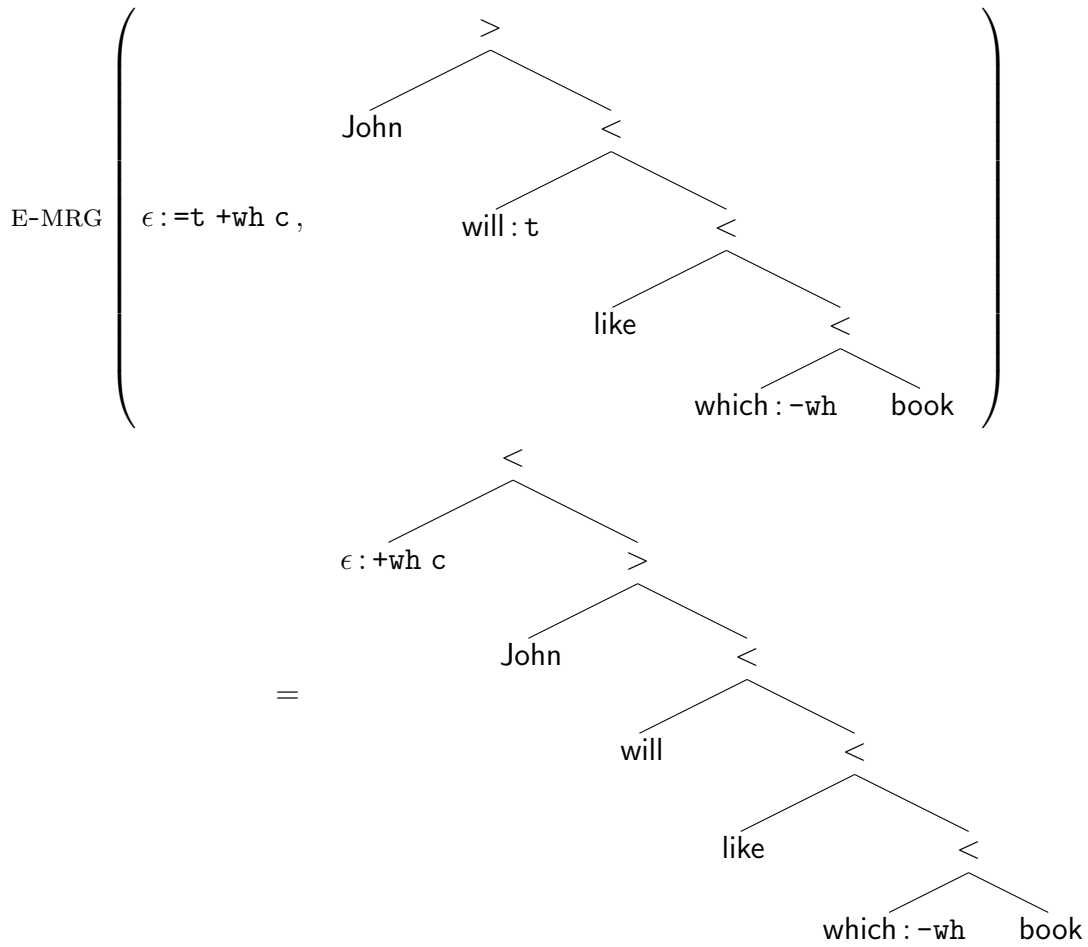


Figure 4

completely separate expression, the $+wh$ feature looks for a $-wh$ feature somewhere embedded inside the very same expression. Whereas we call $=d$ and d a “selector” and a “selectee”, we call $+wh$ and $-wh$ an “attractor” and “attractee”, respectively. When we apply the I-MRG function, the $+wh$ and $-wh$ features are checked and the subconstituent headed by the $-wh$ feature is “moved” up to become a specifier of the head that has the $+wh$ feature. This is shown in Figure 5.

We have now completed the derivation: the only remaining unchecked feature is c on the expression’s head, encoding the fact that the expression we have derived

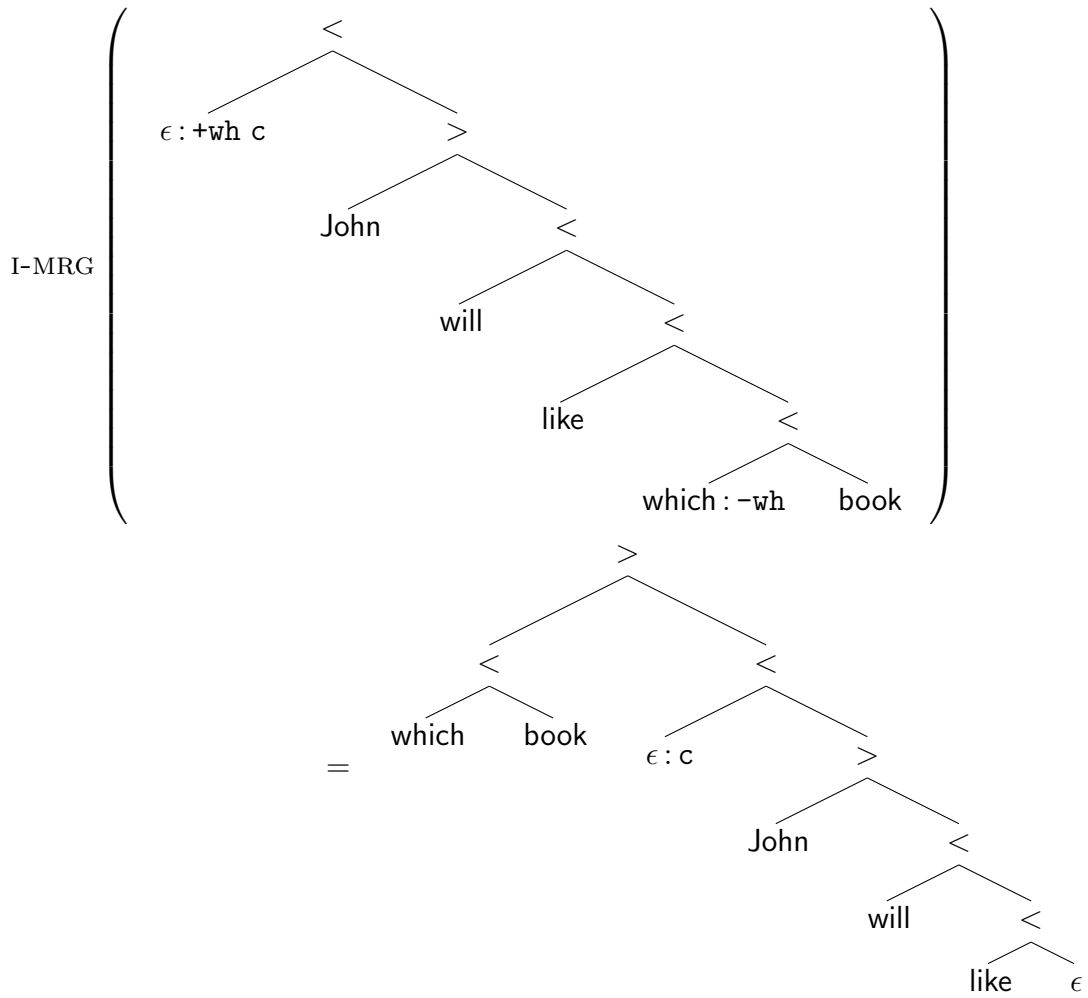


Figure 5

is a CP. The generated string is *which book John will like* — not an English sentence as it is, but I will ignore head movement for expository convenience throughout this thesis, since only phrasal movement will be significant for what follows.⁹

To be clear, the trees shown in Figure 5 are the *derived objects* of the system. We can represent an entire derivation in tree format by showing the involved lexical items at the leaves and labelling each non-leaf node with the name of the function (E-MRG or I-MRG) applied. The *derivation tree* for the derivation that ends as shown

⁹For treatments of head movement within the MG formalism see Stabler (1997, 2001).

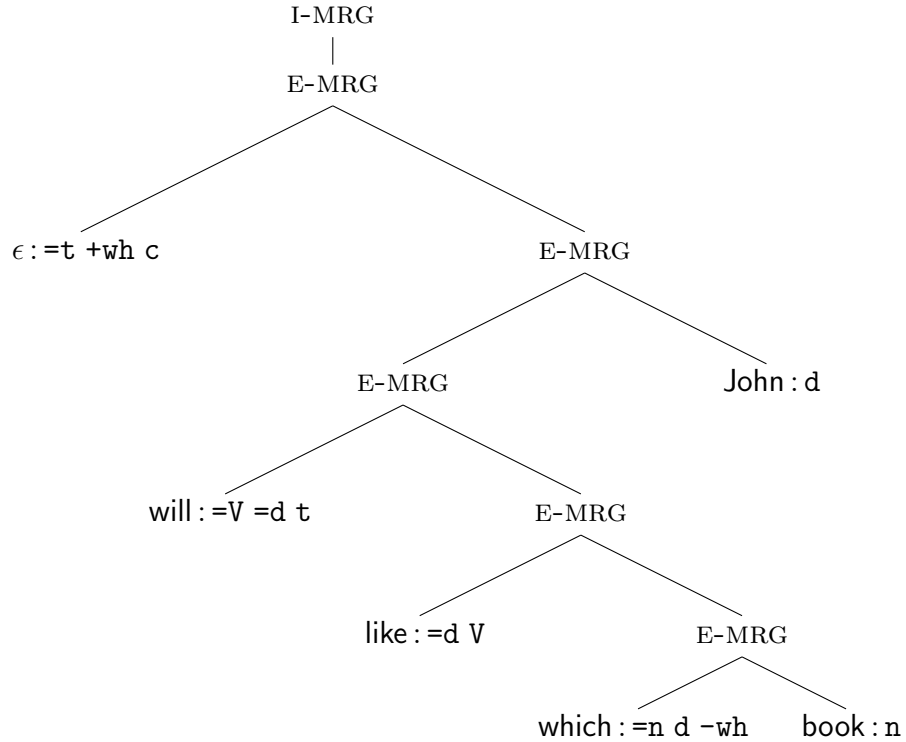


Figure 6

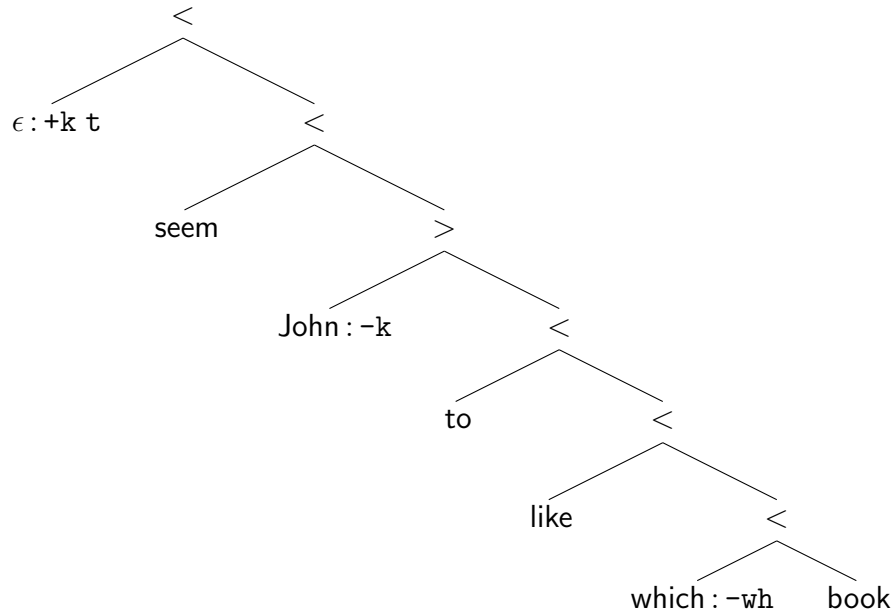
in Figure 5 is given in Figure 6.

1.5.2 Reduced representations for MG expressions

It was soon noticed (Michaelis 2001a,b) that the derived tree structures presented thus far contain lots of redundant information. We can use much simpler structures in their place while maintaining exactly the same expressive power. To illustrate this point, consider a derivation of the sentence in (1.1) using the formalism as introduced so far. At the point in the derivation immediately before I-MRG applies to raise the subject, we will have the structure shown in (1.2), where +**k** and -**k** features represent Case. (I still abstract away from head movement and VP-internal subject positions.)

(1.1) Which book does John seem to like?

(1.2)



Given the structure in (1.2), we can already be certain of some facts about any string that will be generated by any successful continuation of this derivation. For example, we can be certain that **to** and **like** will remain in their current adjacent positions, so we can be certain that any string generated from this structure will contain **to like** as a substring. We can not be certain that **which** and **book** will follow this substring; in fact, we can be certain that **which** and **book** will *not* follow this substring, because there are unchecked features on **which** that will trigger a “movement” step at some future point. We can, however, be certain that **which** and **book** will move together at this future point, and in fact will remain together for the rest of the derivation — **book** has no unchecked features that can cause it to move away from **which** — so we can be certain that **which book** will appear as a substring of whatever string is finally derived, in whatever position the **-wh** feature determines that it will move to.

Similarly, we can be certain that **John** will move out of its current position, because it has a **-k** feature that must be checked at some future point, so **John** will not immediately precede **to like** in the string that is eventually generated. Note, however, that this means that **seem** *will* immediately precede **to like**. All that intervenes be-

tween these two fragments at the moment is **John**, and we can be certain that this will move out of the way at some point. So we have discovered three substrings that are bound to appear in the string eventually generated from this structure: **seem to like**, **John** and **which book**. In the same way that **book** must stay to the immediate right of **which**, **seem** (and therefore **seem to like**) must stay to the immediate right of the T head whose phonological contribution happens to be null; therefore the eventual location of the string **seem to like**, or perhaps more clearly the string ϵ **seem to like**, will be determined by the **+k t** features currently on this T head.

Having noted all this, it follows that we can discard a large amount of redundant structure and maintain only the relevant information by replacing the representation in (1.2) with the more economical one in (1.2').

(1.2') $\langle \text{seem to like: +k t}, \{ \text{John: -k}, \text{which book: -wh} \} \rangle$

This indicates that we have constructed a T' constituent — something that will assign Case and then act as a TP — with string yield **seem to like**. We also have two other constituents, with yields **John** and **which book**, that have not yet reached their final positions. But since these are associated with **-k** and **-wh** features respectively, we have retained everything we need to know in order to correctly determine the positions they will eventually move to. Stabler (2001) writes that the idea behind these reduced representations

... can be traced back to Pollard (1984) who noticed, in effect, that when a constituent is going to move, we should not regard its yield as included in the yield of the expression that contains it. Instead, the expression from which something will move is better regarded as having multiple yields, multiple components — the “moving” components have not yet reached their final positions.

$$\frac{\langle \text{like} : =\mathbf{d} \mathbf{V}, \{\} \rangle \quad \frac{\langle \text{the} : =\mathbf{n} \mathbf{d}, \{\} \rangle \quad \langle \text{book} : \mathbf{n}, \{\} \rangle}{\langle \text{the book} : \mathbf{d}, \{\} \rangle} \text{E-MRG}}{\langle \text{like the book} : \mathbf{V}, \{\} \rangle} \text{E-MRG}$$

Figure 7: The derivational steps from Figure 2 rewritten with reduced representations

$$\frac{\langle \text{like} : =\mathbf{d} \mathbf{V}, \{\} \rangle \quad \frac{\langle \text{which} : =\mathbf{n} \mathbf{d} \text{-}\mathbf{wh}, \{\} \rangle \quad \langle \text{book} : \mathbf{n}, \{\} \rangle}{\langle \text{which book} : \mathbf{d} \text{-}\mathbf{wh}, \{\} \rangle} \text{E-MRG}}{\langle \text{like} : \mathbf{V}, \{ \text{which book} : \text{-}\mathbf{wh} \} \rangle} \text{E-MRG}$$

Figure 8: The derivational steps from Figure 3 rewritten with reduced representations

This highlights the fact that in (1.2') we do *not* regard **John** and **which book** as part of the yield of the T' constituent.

Since no relevant information is lost by adopting these reduced representations, we can reformulate the E-MRG and I-MRG functions to act on these instead (Stabler and Keenan 2003). For example, the derivational steps shown in Figure 2 can now be more compactly represented as shown in Figure 7. Every expression here has the simple form $\langle x, \{\} \rangle$, for some pairing x of a string with a feature-list. This is because the set component of an expression, intuitively speaking, encodes the “moving sub-parts” of the expression — recall that **John** : -**k** and **which book** : -**wh** are “moving sub-parts” of the T' represented in (1.2') — and nothing shown in Figure 7 has been or will be involved in any movement.¹⁰

More interesting is the more compact representation of the derivational steps from Figure 3 as Figure 8, since this includes some features that will eventually trigger movement. Note that at the second application of E-MRG in (8), **which book**

¹⁰Note that this applies not only to non-primitive expressions that happen to have no moving sub-parts, but also to lexical expressions: instead of **the** : =**n d** as a lexical expression we now have $\langle \text{the} : =\mathbf{n} \mathbf{d}, \{\} \rangle$ for example. This is not a significant complication added to the framework, just a correlate of the fact that in the original formulation with more elaborate tree-shaped objects, we treated lexical expressions as trivial (one-node) trees.

$$\frac{\langle \epsilon : =t +wh\ c, \{\} \rangle \quad \langle \text{John will like} : t, \{ \text{which book} : -wh \} \rangle}{\langle \text{John will like} : +wh\ c, \{ \text{which book} : -wh \} \rangle} \text{E-MRG}$$

$$\langle \text{which book John will like} : c, \{\} \rangle \text{I-MRG}$$

Figure 9: The derivational steps from Figure 4 and Figure 5 rewritten with reduced representations

is *not* included as part of the string associated with the V feature, unlike the book in Figure 2. Because its *-wh* feature remains unchecked, *which book : -wh* remains in the set component of the expression, waiting for a chance to check this feature at some future point. Thus the VP that we have derived has *which book : -wh* as a “moving sub-part” in just the same way that the T’ represented in (1.2’) has *John : -k* and *which book : -wh* as “moving sub-parts”.

The eventual checking of this *-wh* feature was illustrated in the original tree notation in Figure 4 and Figure 5. These same two derivational steps can be represented in the new notation as shown in Figure 9. From the point where the VP is constructed at the end of Figure 8 to the point where the C’ is constructed in the first step shown in Figure 9, *which book : -wh* remains in the set component of each generated expression as E-MRG builds up larger and larger expressions. Once the C’ is constructed, the *+wh* feature on the C head becomes available to check remaining feature of *which book : -wh*, and I-MRG applies. Since *which book* has no further features to check, it will not move any further, and so this string is included in the yield of the resulting CP, in its specifier position.

Representations much like these reduced ones will be used in the rest of this thesis. But before continuing I will pause to address some concerns that might be raised by their adoption.

First, the switch to the reduced representations discards a lot of information that would seem to be necessary in order to enforce many of the well-known “island”

constraints. For example, in adopting the reduced representations we have no record of where *which book* is “coming from” when it moves to check its *-wh* feature in Figure 9, and therefore no record of any bounding nodes or barriers or phase edges or whatever that may have been “crossed” by this movement. This is exactly the kind of problem that chapter 3 will be concerned with: I will propose to record a little bit more structure than the representations in Figure 9 do — but still much less than the original ones in Figure 4 and Figure 5 — in such a way that adjunct island effects and freezing effects fall out as instances of the same constraint.

The original MG formalism from Stabler (1997) does, however, include one basic locality constraint, and this is maintained (in fact, crucially made use of) by the more compact representations. This is the Shortest Move Condition (SMC), a constraint along the lines of Relativized Minimality (Rizzi 1990) or the Minimal Link Condition (Chomsky 1995). The SMC states that I-MRG can only apply to a tree whose head is looking to check a *+f* attracting feature, if there is *exactly one* *-f* feature further down in the tree. The expression to which I-MRG applies in Figure 5 satisfies this constraint because it contains exactly one *-wh* feature; if it contained two such features, I-MRG would not be able to apply.¹¹ In terms of the reduced representations, this means that no two strings in the set component of an expression can ever share the same first feature (if the expression is to take part in a convergent derivation). So no question arises of which of two “competing” constituents waiting to check features undergoes the movement; if such competition exists, neither movement is allowed. The SMC will not play a crucial role in this thesis, except insofar as it allows us to use the reduced representations.

Second, one might worry that the reduced representations discard information that is crucial for proper semantic interpretation; for example, that having no record

¹¹Multiple *wh*-fronting constructions therefore raise a problem, on the assumption that the two moving *wh*-phrases are looking to check the same kind of feature. See Gärtner and Michaelis (2007) for an addition to the formalism that addresses this via a “clustering” operation.

of where a *wh*-phrase is “coming from” might prevent its binding of the appropriate variable (at least, under a quantifier-raising approach to semantics). Indeed, the results of Michaelis (2001a,b), and my discussion above of “relevant information”, were concerned only with the sets of strings generated by the formalism.¹² But this worry relies on the assumption that the linking of a quantifier with its variable must “happen when” the quantifier moves, or re-merges, into its c-commanding position. The approach to quantification that I take in chapter 4, like that of Kobele (2006), avoids this problem by separating the “binder” component of a quantifier’s meaning from the “bindee” component as soon as the bindee is merged. The reduced MG representations naturally allow us to store the quantificational binder in the set component of the expression, in a manner very reminiscent of “Cooper storage” (Cooper 1983); in fact, the reduced MG representations can be roughly construed as applying a Cooper storage approach to movement, where the things stored are not quantifiers, but strings.

1.5.3 Movement as re-merge

Recall that our reduced representations have the general form $e = \langle x, \{y_1, y_2, \dots, y_n\} \rangle$, where x, y_1, \dots, y_n are all pairings of a string with a feature list. Let us call such pairings “units”. Let us call the unit x the “head” of the expression e . We currently have two functions for deriving new expressions from old ones: E-MRG, which checks features of the head of one expression against features of the head of another, and combines the two expressions into one; and I-MRG, which checks features of the head of an expression against features of one of the units

¹²In fact, I have not quite included all the information relevant for maintaining the stringsets generated by the formalism: whether an application of E-MRG places the selected constituent to the right or to the left of its head depends on whether it is the first constituent selected by that head, and this distinction is not recorded in the reduced representations I have presented. It is for this reason that Stabler and Keenan (2003) encode, for each expression, whether it is lexical or derived. I gloss over this for now because in chapter 2 I will introduce, on unrelated grounds, a distinction between internal and external arguments which can serve this purpose.

in the set component of that very same expression. Note that as things stand, a unit will get into the set component of an expression *only* as a result of checking some, but not all, of its selectee/attractee features, for example when **which book** checks its **d** feature but has its **-wh** feature remaining in Figure 8. Thus the first selectee/attractee feature of any feature list will be checked by an application of E-MRG, and any other selectee/attractee features will be checked by applications of I-MRG, drawing on the set component of an expression.¹³

The two different feature-checking operations, E-MRG and I-MRG, can be reduced to one by introducing a new operation, INS (“insert”), which adds units directly to the set component of an expression (Stabler 2006). We can then effectively do away with the E-MRG operation. What was previously achieved by an application of E-MRG(e_1, e_2) — checking a feature of the head of e_1 against a feature of the head of e_2 — can now be achieved by (i) an application of INS(e_1, e_2), which inserts all the units comprising e_2 into the set component of e_1 , and then (ii) an application of I-MRG to the resulting expression, checking a feature of its head against a feature of a unit in its (newly supplemented) set component. For example, instead of Figure 7 we now have Figure 10.

Remaining attractee features on a unit checking a selectee feature will cause a unit with the same yield to remain in the set component (just as the I-MRG operation is already required to do to allow for expressions that undergo multiple movement steps, although we have not seen examples of this).¹⁴ Thus instead of Figure 8 we have Figure 11, where **which book: -wh** is “left behind” in the set component when I-MRG applies to check the **=d** and **d** features.

All the checking of selector and selectee features in Figure 10 and Figure 11 in-

¹³Actually, to participate in a successful derivation a feature list can not have more than one selectee feature, and any attractee features on a lexical item must follow a selectee features. But I will not fuss about these details because the distinction between selectee features and attractee features will soon collapse.

¹⁴This is very similar to the “Survive Principle” of Stroik (1999, 2009). For further discussion of the relationship between “survive minimalism” and MGs see Kobele (2009).

$$\begin{array}{c}
\langle \text{the} : =n \text{ d}, \{\} \rangle \quad \langle \text{book} : n, \{\} \rangle \\
\hline
\langle \text{the} : =n \text{ d}, \{\text{book} : n\} \rangle \quad \text{INS} \\
\langle \text{like} : =d \text{ V}, \{\} \rangle \quad \hline
\langle \text{the book} : d, \{\} \rangle \quad \text{I-MRG} \\
\hline
\langle \text{like} : =d \text{ V}, \{\text{the book} : d\} \rangle \quad \text{INS} \\
\hline
\langle \text{like the book} : V, \{\} \rangle \quad \text{I-MRG}
\end{array}$$

Figure 10: The equivalents of the derivational steps from Figure 7, using the INS operation

$$\begin{array}{c}
\langle \text{which} : =n \text{ d} \text{ -wh}, \{\} \rangle \quad \langle \text{book} : n, \{\} \rangle \\
\hline
\langle \text{which} : =n \text{ d} \text{ -wh}, \{\text{book} : n\} \rangle \quad \text{INS} \\
\langle \text{like} : =d \text{ V}, \{\} \rangle \quad \hline
\langle \text{which book} : d \text{ -wh}, \{\} \rangle \quad \text{I-MRG} \\
\hline
\langle \text{like} : =d \text{ V}, \{\text{which book} : d \text{ -wh}\} \rangle \quad \text{INS} \\
\hline
\langle \text{like} : V, \{\text{which book} : \text{-wh}\} \rangle \quad \text{I-MRG}
\end{array}$$

Figure 11: The equivalents of the derivational steps from Figure 8, using the INS operation

volved a selector ($=f$) feature on the head of an expression and a selectee (f) feature on the set component of that same expression. So when the eventual checking of the -wh feature on **which book** occurs via an application of the I-MRG function, as shown in Figure 12, this is now no different from the operation that performed all the earlier feature checking in the derivation.

At this point we can note that since all feature-checking occurs under the same configuration, it is natural to eliminate the distinction between $=f/f$ feature pairs (those that were formerly checked by E-MRG) and $+f/-f$ feature pairs. I adopt the $+f/-f$ notation for all features (since it is E-MRG that we have eliminated). For example, what was previously $\text{which} : =n \text{ d} \text{ -wh}$ (“select an NP, get selected as a DP, then get attracted as a WH”) is now $\text{which} : +n \text{-d} \text{-wh}$ (“attract an NP, get attracted as a DP, then get attracted as a WH”); and we now use the word “attract” in a way

$$\begin{array}{c}
\langle \epsilon := \mathbf{t} \ \mathbf{+wh} \ c, \ \{\} \rangle \quad \langle \text{John will like} : \mathbf{t}, \ \{\text{which book} : \mathbf{-wh}\} \rangle \\
\hline
\langle \epsilon := \mathbf{t} \ \mathbf{+wh} \ c, \ \{\text{John will like} : \mathbf{t}, \ \text{which book} : \mathbf{-wh}\} \rangle \quad \text{INS} \\
\hline
\langle \text{John will like} : \mathbf{+wh} \ c, \ \{\text{which book} : \mathbf{-wh}\} \rangle \quad \text{I-MRG} \\
\hline
\langle \text{which book John will like} : \mathbf{c}, \ \{\} \rangle \quad \text{I-MRG}
\end{array}$$

Figure 12: The equivalents of the derivational steps from Figure 9, using the INS operation

that is agnostic about whether this is the first feature-checking step for one of the units involved. It is also natural to similarly simplify our naming slightly: since I will not be making any further reference to versions of MGs that use the E-MRG function, from now on I will refer to I-MRG simply as MRG. With these two minor changes of notation, the steps illustrated in Figure 11 and Figure 12 can be rewritten one more (last) time as Figure 13 and Figure 14, respectively.

The system therefore provides a way to flesh out the intuition that the features checked by movement steps do not differ in kind from those checked by merging steps (contra Chomsky 1995), and that movement is simply “re-merging” (Kitahara 1997, Epstein et al. 1998, Chomsky 2004). Consider for example some constituent that, intuitively speaking, occupies three positions in a derived tree structure; say, a DP that must check features in a theta position, a Case position, and a quantificational position. In the original MG formalism presented in the previous subsection, such an element would be affected by one application of E-MRG and then two applications of I-MRG; in our revised version, however, it is affected by one application of INS and then *three* applications of (I-)MRG.¹⁵

¹⁵ A concrete advantage of this is that it avoids redundancies that otherwise arise when we want to talk about the details of our feature-checking system. For example, Frey and Gärtner (2002) present an addition to the formalism that encodes adjunction (unrelated to what I propose in the next subsection) via “asymmetric” feature-checking. To spell this out, they are forced to add two different kinds of asymmetrically-checked features, one for “base-generated” adjuncts and one for constituents that move into adjunct positions, and correspondingly two new asymmetric feature-checking operations, the relationship between which is exactly analogous to the relationship between

$$\begin{array}{c}
\frac{\langle \text{which} : +n-d-wh, \{\} \rangle \quad \langle \text{book} : -n, \{\} \rangle}{\langle \text{which} : +n-d-wh, \{\text{book} : -n\} \rangle} \text{INS} \\
\frac{\langle \text{like} : +d-V, \{\} \rangle \quad \frac{\langle \text{which} : +n-d-wh, \{\text{book} : -n\} \rangle}{\langle \text{which book} : -d-wh, \{\} \rangle} \text{MRG}}{\langle \text{like} : +d-V, \{\text{which book} : -d-wh\} \rangle} \text{INS} \\
\frac{\langle \text{like} : +d-V, \{\text{which book} : -d-wh\} \rangle}{\langle \text{like} : -V, \{\text{which book} : -wh\} \rangle} \text{MRG}
\end{array}$$

Figure 13: The equivalents of the derivational steps from Figure 11, with the distinction between “merge” and “move” features eliminated

$$\begin{array}{c}
\frac{\langle \epsilon : +t+wh-c, \{\} \rangle \quad \langle \text{John will like} : -t, \{\text{which book} : -wh\} \rangle}{\langle \epsilon : +t+wh-c, \{\text{John will like} : -t, \text{which book} : -wh\} \rangle} \text{INS} \\
\frac{\langle \epsilon : +t+wh-c, \{\text{John will like} : -t, \text{which book} : -wh\} \rangle}{\langle \text{John will like} : +wh-c, \{\text{which book} : -wh\} \rangle} \text{MRG} \\
\frac{\langle \text{John will like} : +wh-c, \{\text{which book} : -wh\} \rangle}{\langle \text{which book John will like} : -c, \{\} \rangle} \text{MRG}
\end{array}$$

Figure 14: The equivalents of the derivational steps from Figure 12, with the distinction between “merge” and “move” features eliminated

1.5.4 A remark on notation

In this introduction to the formalism I have used two different styles of notation to represent the steps of a derivation. Applications of INS, MRG, and so on have been sometimes written in the form of a mathematical equation with the grammatical operations as functions taking certain arguments, as in (1.3), and sometimes in the form of a logical inference from certain premises, as in (1.4).

E-MRG and I-MRG. With the INS operation and a single feature-checking operation, this duplication would be avoided.

As a second example, consider the fact that as presented in Figure 7, Figure 8 and Figure 9, both E-MRG and I-MRG need to be specified with sub-cases according to whether or not the unit contributing the checked selectee/attractee feature should have its string yield concatenated with that of its selector/attractor. (If not, a unit is recorded in the set component of the resulting expression.) Assuming for simplicity that all movement is overt, in applications of E-MRG this concatenation should only take place if the selectee feature being checked is not followed by any attractee features; in applications of I-MRG, this concatenation should likewise only take place if the attractee feature being checked is not followed by any attractee features. In the revised system with just a single MRG operation, this duplication is also avoided.

$$(1.3) \quad \text{INS}(\langle \text{the: +n-d}, \{\} \rangle, \langle \text{book: -n}, \{\} \rangle) = \langle \text{the: +n-d}, \{\text{book: -n}\} \rangle$$

$$(1.4) \quad \frac{\langle \text{the: +n-d}, \{\} \rangle \quad \langle \text{book: -n}, \{\} \rangle}{\langle \text{the: +n-d}, \{\text{book: -n}\} \rangle} \text{INS}$$

It may be worth stressing that despite the difference in notation, (1.3) and (1.4) say *exactly* the same thing.¹⁶ I simply choose whatever notation is more convenient on a case-by-case basis. (And in chapter 3 I briefly adopt, for convenience, yet another format for presenting derivations.)

An advantage of the “inference notation” is that it can be used to efficiently convey the dependencies among a sequence of derivational steps, where the results of certain operations are used as inputs to other subsequent operations, as in Figure 13 for example. Note the close similarity between this notation and the derivation tree given in Figure 6 on page 15. Written using the “equation notation”, such a sequence of interrelated steps can lead to significant repetition; rewriting the steps in Figure 13 this way yields Figure 15. The expression to the right of the equals sign on each line is repeated as one of the inputs to the step on the next line.

The inference notation, however, can become impractical for longer sequences of steps (requiring too much horizontal space to fit on the page), in which cases I use the equation notation. To reduce the clutter that comes from repeating expressions in Figure 15, I assign each new expression to one of the variables ‘ e_1 ’, ‘ e_2 ’, ... so that it can be referred to on subsequent lines; the steps in Figure 15 are instead written as in

¹⁶ As an analogy, the logical law of modus ponens (‘MP’) is typically thought of as licensing an inference from two premises to a conclusion, but could equally be construed as a function that takes two formulas as arguments and returns a third formula:

$$\frac{p \rightarrow q \quad p}{q} \text{MP} \qquad \text{MP}(p \rightarrow q, p) = q$$

In the other direction, the operation of multiplication (‘MULT’) can be naturally thought of as a function that takes two numbers as arguments and returns a third. But if we consider it as the operation which generates the positive integers from the primes (in the same way that a grammatical operation generates expressions from a lexicon), it licences an inference from two premises (eg. that 2 is an integer, and that 3 is an integer) to a conclusion (that 6 is a natural number):

$$\frac{2 \quad 3}{6} \text{MULT} \qquad \text{MULT}(2, 3) = 6$$

$$\begin{aligned}
& \text{INS}(\langle \text{which} : +\text{n-d-wh}, \{\} \rangle, \langle \text{book} : -\text{n}, \{\} \rangle) = \langle \text{which} : +\text{n-d-wh}, \{\text{book} : -\text{n}\} \rangle \\
& \quad \text{MRG}(\langle \text{which} : +\text{n-d-wh}, \{\text{book} : -\text{n}\} \rangle) = \langle \text{which book} : -\text{d-wh}, \{\} \rangle \\
& \text{INS}(\langle \text{like} : +\text{d-V}, \{\} \rangle, \langle \text{which book} : -\text{d-wh}, \{\} \rangle) = \langle \text{like} : +\text{d-V}, \{\text{which book} : -\text{d-wh}\} \rangle \\
& \quad \text{MRG}(\langle \text{like} : +\text{d-V}, \{\text{which book} : -\text{d-wh}\} \rangle) = \langle \text{like} : -\text{V}, \{\text{which book} : -\text{wh}\} \rangle
\end{aligned}$$

Figure 15: The derivational steps from Figure 13 rewritten, somewhat repetitively, in “equation notation”

$$\begin{aligned}
e_1 &= \text{INS}(\langle \text{which} : +\text{n-d-wh}, \{\} \rangle, \langle \text{book} : -\text{n}, \{\} \rangle) &= \langle \text{which} : +\text{n-d-wh}, \{\text{book} : -\text{n}\} \rangle \\
e_2 &= \text{MRG}(e_1) &= \langle \text{which book} : -\text{d-wh}, \{\} \rangle \\
e_3 &= \text{INS}(\langle \text{like} : +\text{d-V}, \{\} \rangle, e_2) &= \langle \text{like} : +\text{d-V}, \{\text{which book} : -\text{d-wh}\} \rangle \\
e_4 &= \text{MRG}(e_3) &= \langle \text{like} : -\text{V}, \{\text{which book} : -\text{wh}\} \rangle
\end{aligned}$$

Figure 16: The derivational steps from Figure 13 rewritten, slightly less repetitively, in “equation notation”

Figure 16. For example, the first line illustrates the derivation of one new expression, e_1 , which then serves as an input to the derivational step on the second line which produces a new expression, e_2 .

I will switch between the format in Figure 16 and the basic inference format in Figure 13 throughout the thesis. But to repeat, Figure 16 and Figure 13 convey *exactly* the same information. The differences between them are no more significant than a change in font. What matters is the operations that are available, what they apply to and what they produce as a result — not the orthogonal issue of which notation we use to express the relationship between an operation, its operands and its result.

1.6 The Conjunctivist conception of neo-Davidsonian semantics

In this section I outline the semantic perspective that will play a crucial role in the syntactic system I develop in this thesis. In §1.6.1 I briefly review neo-Davidsonian logical forms and the motivation for them. In §1.6.2 and §1.6.3 I introduce the distinctive “Conjunctivist” proposal (Pietroski 2005, 2006) concerning *how these logical forms are composed* out of the lexical items of a sentence and the syntactic relationships among them. Before continuing I respond to some potential objections in §1.6.4. Readers familiar with the general neo-Davidsonian proposal may wish to skip or skim §1.6.1 and §1.6.4.

To be clear, the Conjunctivist proposal about modes of composition presupposes (some form of) neo-Davidsonian logical forms, but the reverse is *not* true. The results I will present in the rest of this thesis rely crucially not just on the fact that sentential logical forms are written with event variables, thematic relations and conjunction symbols as discussed in §1.6.1, but also on the specific proposal in §1.6.2 and §1.6.3 about how these logical forms are composed out of the meanings of sentences’ parts.

1.6.1 Neo-Davidsonian logical forms

Many have argued that sentences should be associated with event-based logical forms of the sort shown in (1.5) (Davidson 1967, 1985, Castañeda 1967, Parsons 1990, Schein 1993).¹⁷

(1.5) a. Brutus stabbed Caesar violently with a knife.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e) \wedge \mathbf{withknife}(e)]$$

¹⁷The predicate **withknife** used in (1.5) might be such that **withknife**(*e*) is merely a shorthand for something along the lines of **With**(*e*, **knife**) or $\exists k[\mathbf{knife}(k) \wedge \mathbf{With}(e, k)]$. Also, for the moment I am using “verb-specific” thematic relations such as **Stabber** and **Stabee** rather than more general notions such as **Agent** and **Patient** or **Subject** and **Object**. These details are not important for now but I will return to them in §1.6.2.

b. Brutus stabbed Caesar with a knife violently.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{withknife}(e) \wedge \mathbf{violent}(e)]$$

c. Brutus stabbed Caesar violently.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)]$$

d. Brutus stabbed Caesar with a knife.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{withknife}(e)]$$

e. Brutus stabbed Caesar.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$$

This derives the desired pattern of entailments. Specifically, (1.5a) and (1.5b) are logically equivalent, and each entail both (1.5c) and (1.5d), which in turn both entail (1.5e). Furthermore, (1.5c) and (1.5d) do not jointly entail (1.5a) or (1.5b), as desired: (1.5c) and (1.5d) might both be true, by being descriptions of distinct events, in which case (1.5a) and (1.5b) are false.

Using this sort of logical forms also allows us to give a simple account of the relationship between the four sentences in (1.6), and between the three sentences in (1.7) (putting aside questions of tense).

(1.6) a. Brutus stabbed Caesar.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$$

b. Cassius saw a stabbing.

$$\exists e \exists e'[\mathbf{seeing}(e) \wedge \mathbf{See-er}(e, \mathbf{cassius}) \wedge \mathbf{See-ee}(e, e') \wedge \mathbf{stabbing}(e')]$$

c. Cassius saw a stabbing of Caesar by Brutus.

$$\exists e \exists e'[\mathbf{seeing}(e) \wedge \mathbf{See-er}(e, \mathbf{cassius}) \wedge \mathbf{See-ee}(e, e') \wedge \mathbf{stabbing}(e') \wedge \mathbf{Stabber}(e', \mathbf{b}) \wedge \mathbf{Stabee}(e', \mathbf{c})]$$

d. Cassius saw Brutus stab Caesar.

$$\exists e \exists e'[\mathbf{seeing}(e) \wedge \mathbf{See-er}(e, \mathbf{cassius}) \wedge \mathbf{See-ee}(e, e') \wedge \mathbf{stabbing}(e') \wedge \mathbf{Stabber}(e', \mathbf{b}) \wedge \mathbf{Stabee}(e', \mathbf{c})]$$

- (1.7) a. Brutus stabbed Caesar.
 $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$
- b. Cassius slept.
 $\exists e[\mathbf{sleeping}(e) \wedge \mathbf{Sleeper}(e, \mathbf{cassius})]$
- c. Cassius slept after Brutus stabbed Caesar.
 $\exists e\exists e'[\mathbf{sleeping}(e) \wedge \mathbf{Sleeper}(e, \mathbf{cassius}) \wedge \mathbf{After}(e, e') \wedge \mathbf{stabbing}(e') \wedge \mathbf{Stabber}(e', \mathbf{b}) \wedge \mathbf{Stabee}(e', \mathbf{c})]$

Finally, it also permits an elegant account of the entailment from (1.8a) to (1.8b), if we allow ourselves some more intricate structure in (at least) such cases; I return to this issue in §2.A.¹⁸

- (1.8) a. Brutus boiled the water.
 $\exists e\exists e'[\mathbf{Boiler}(e, \mathbf{b}) \wedge \mathbf{Result}(e, e') \wedge \mathbf{boiling}(e') \wedge \mathbf{Boilee}(e', \mathbf{water})]$
- b. The water boiled.
 $\exists e[\mathbf{boiling}(e) \wedge \mathbf{Boilee}(e, \mathbf{water})]$

Many of these examples can be treated equally well using a single polyadic predicate that relates an event to its participants (eg. $\mathbf{stabbing}(e, \mathbf{b}, \mathbf{c})$) instead of introducing a separate conjunct for each thematic relation. But there is significant, if (perhaps unavoidably) subtle, evidence that sentential logical forms require this “thematic separation”. In other words we require, in our language of logical forms, (open) propositional constituents which express, say, that an event is a stabbing and has a certain patient without expressing anything about any agent of the event, or that an event has a certain agent without expressing anything about what sort of event it is or about any patient of the event.

¹⁸There is debate about the precise properties of the relation called **Result** in (1.8a), but these details are not crucial for the moment. I return to the issue briefly in footnote 29; see Pietroski (2003b, 2005) for further discussion.

The relationship between (1.9a) and a “missing participant” variant such as (1.9b) *can* be accounted for without abandoning the monolithic three-place predicate **stabbing**(e, x, y), via existential quantification as shown in (1.10).

- (1.9) a. Brutus stabbed Caesar
 b. Caesar was stabbed

- (1.10) a. $\exists e[\mathbf{stabbing}(e, \mathbf{b}, \mathbf{c})]$
 b. $\exists e\exists x[\mathbf{stabbing}(e, x, \mathbf{c})]$

Parsons (1990, pp.96–99) argues against this analysis, in favour of the thematically separated (1.11).

- (1.11) a. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$
 b. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabee}(e, \mathbf{c})]$

Under the non-separated approach in (1.10), it is not possible for a logical form to refer to a stabbing event in a way that does not entail that a corresponding agent exists. Parsons argues that although in the real world all stabbing events have agents, as (1.10b) entails, similar sentences can be used to describe scenarios where no such agent exists, and so (1.10b) is inappropriate. In particular, he gives the example of reporting a dream as in (1.12).

- (1.12) In a dream last night, I was stabbed, although in fact nobody had stabbed me, and I wasn’t stabbed with anything.

Parsons (1990, p.98) presents this as “a report of an incoherent dream, one in which, say, I am bewildered by the fact that I have been stabbed but not by anyone or anything”.¹⁹ Such a report is difficult to accommodate without thematic separation: (1.10b) wrongly asserts the existence of an agent, for example, and (1.13) yields a contradiction. But with a thematically separated account the report is straightforwardly analysed as in (1.14).

$$(1.13) \quad \exists e[\exists x[\mathbf{stabbing}(e, x, \mathbf{i})] \wedge \neg\exists x[\mathbf{stabbing}(e, x, \mathbf{i})]]$$

$$(1.14) \quad \exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabbee}(e, \mathbf{i}) \wedge \neg\exists x[\mathbf{Stabber}(e, x)]]$$

Parsons acknowledges that he presents this argument “without having a general criterion of how we apportion truths into those due to meanings and those due to knowledge of the world”, and so the facts are, of course, open to doubts; but any data showing that a verb meaning does not in general entail the existence of a participant in a certain role, would be evidence for thematic separation. This observation also highlights a crucial difference between (1.10a) and (1.11a) that the next two arguments exploit: any subset of the three conjuncts in (1.11a) can together form a propositional logical constituent.

Herburger (2000) presents an account of focus, whose application to certain simple sentences relies crucially on the ability to separate the three conjuncts in (1.11a). The essence of Herburger’s proposal is that non-focussed (intuitively presupposed) material is expressed as a restriction on event quantification. Thus (1.15a) will be analysed as either (1.15b) or (1.15c), depending on whether one adopts thematic separation or not. (I follow Herburger and others in representing semantic focus with uppercase.)

(1.15) a. Brutus stabbed Caesar VIOLENTLY

b. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabbee}(e, \mathbf{c})][\mathbf{stabbing}(e) \wedge$
 $\mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabbee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)]$

c. $\exists e[\mathbf{stabbing}(e, \mathbf{b}, \mathbf{c})][\mathbf{stabbing}(e, \mathbf{b}, \mathbf{c}) \wedge \mathbf{violent}(e)]$

Intuitively, (1.15b) and (1.15c) say roughly that, among (real or hypothetical) stabblings of Caesar by Brutus, there was a violent one. Consider now focussing not an adverb as in (1.15a), but an argument as in (1.16a) or the verb as in (1.17a) Here

¹⁹Importantly, it is not to be understood as merely “a report that, according to the dream, I *had* been stabbed by somebody, but that the stabbing had taken place earlier than the events in the dream, and so I did not actually experience (in the dream) the stabbing.”

only a thematically separated approach permits an analogous analysis.

(1.16) a. Brutus stabbed CAESAR violently

b. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{violent}(e)] [\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)]$

(1.17) a. Brutus STABBED Caesar violently

b. $\exists e[\mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)] [\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)]$

With a monolithic three-place predicate as in (1.15c), there is no way to omit the patient specification for (1.16a) or the event-type specification for (1.17a) from the restriction of the quantifier.

One of the strongest arguments for thematic separation comes from the semantics of plurals. Consider the reading of (1.18a) shown in (1.18b) (adapted from Schein 1993, p.57), under which the sentence is true if, say: games g_1 and g_2 taught quarterback q_1 one play each, game g_2 taught quarterback q_2 two plays, and game g_3 taught quarterback q_3 two plays (and there are no other quarterbacks). Of the existential quantifier over games ($\exists G$) and the universal quantifier over quarterbacks ($\forall q$), neither has the other in its scope. The unseparated relation $\mathbf{teaching}(e, G, q, P)$ used in (1.18c) does not let us leave the relationship between the three games G and the various quarterback-teachings undetermined in the way that (1.18b) does; any time we have two quantifiers each binding a participant, this sort of unseparated relation will force us to put one of these quantifiers in the scope of the other. However the compositional details of such sentences turn out, it seems hard to avoid the need for two-place relations (such as **Agent** and **Theme**) that hold between an event and one of its participants in the complete logical forms; see Schein (1993, ch.4,5) for more discussion of this point.

- (1.18) a. Three video games taught every quarterback two new plays.
- b. $\exists e [\exists G [\mathbf{three-games}(G)] [\forall x [\mathbf{Agent}(e, x) \leftrightarrow Gx]] \wedge$
 $\forall q [\mathbf{quarterback}(q)] [\exists e' [e' \sqsubseteq e] [\mathbf{teaching}(e') \wedge$
 $\forall r [\mathbf{Beneficiary}(e', r) \leftrightarrow r = q] \wedge$
 $\exists P [\mathbf{two-plays}(P)] [\forall x [\mathbf{Theme}(e', x) \leftrightarrow Px]]$
 $]]$
- c. $\exists e [\exists G [\mathbf{three-games}(G)] [$
 $\forall q [\mathbf{quarterback}(q)] [$
 $\exists P [\mathbf{two-plays}(P)] [\mathbf{teaching}(e, G, q, P)]$
 $]$
 $]]$
-

Taking it as given now that we would like to end up with “thematically separated” logical forms of the sort shown in (1.19), a distinct question remains open about exactly how we should understand them to be composed from the meanings of sentences’ parts. In the rest of this section I will present arguments, drawn mainly from Pietroski (2005, 2006), for a particular answer to this further question, on which the syntactic proposals in the rest of this thesis will depend. Specifically, it will be important that these logical forms are *not* assumed to be constructed from lexical meanings like (1.20) where the thematic relations, while expressed in separate conjuncts, are all introduced by the verb (as in Landman 2000, Higginbotham 1985).

$$(1.19) \quad \exists e [\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$$

$$(1.20) \quad \llbracket \mathbf{stabbed} \rrbracket = \lambda x \lambda y \lambda e. [\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, y) \wedge \mathbf{Stabee}(e, x)]$$

I now turn to presenting the justification for not adopting such lexical meanings.²⁰

²⁰There is perhaps already good reason to be suspicious of lexical meanings like (1.20), given the arguments of Schein and Parsons for thematic separation in logical forms. With lexical meanings

1.6.2 Conjunctivist semantic composition

The basic premise of the Conjunctivist approach to semantic composition (Pietroski 2005, 2006) is that whenever two constituents combine, they each contribute some monadic predicate, and that the meaning of the newly-formed constituent is the result of conjoining the two contributed predicates. Thus the logical form of (1.21) is thought to be the result of conjoining the four predicates in (1.22), each contributed by exactly one of the words in the sentence.

(1.21) Brutus stabbed Caesar violently.

$$\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c}) \wedge \mathbf{violent}(e)]$$

- (1.22)
- | | | | |
|----|---|----|--|
| a. | $\lambda e.\mathbf{Stabber}(e, \mathbf{b})$ | c. | $\lambda e.\mathbf{Stabee}(e, \mathbf{c})$ |
| b. | $\lambda e.\mathbf{stabbing}(e)$ | d. | $\lambda e.\mathbf{violent}(e)$ |

In the case of ‘Brutus’ and ‘Caesar’ — but not in the case of ‘violently’ — the predicate contributed in (1.22) appears to differ from what could plausibly be considered the lexical meaning of the word. A central idea in later chapters will be that the shift from (something like) \mathbf{c} to (something like) $\lambda e.\mathbf{Stabee}(e, \mathbf{c})$ (see Carlson 1984, Krifka 1992) requires that ‘Caesar’ establish a certain “rich” syntactic relationship with ‘stab’; and since no such shift is required in the case of ‘violently’, a less “rich” syntactic relationship with ‘stab’ suffices.

The motivation for the Conjunctivist approach lies in the fact that the more standard approaches based on function-application (eg. Klein and Sag 1985, Heim and Kratzer 1998) leave unexplained certain generalisations about the meanings that

like (1.20) it is unclear, to say the least, how the **Stabber** relation could end up absent from the logical form in (1.11b), how the **Stabber** and **Stabee** relations could end up applied to distinct event variables in Schein’s (1.18b), and how the thematic conjuncts could end up “far enough apart” to allow scopal independence in (1.18b). But in the interest of keeping distinct (i) questions of the logical forms that should be associated with complete sentences and (ii) questions of how these logical forms are composed from the meanings of the sentences’ parts, I will (for now) proceed on the conservative assumption that the discussion of thematic separation so far bears only on questions of the first sort.

natural languages lexicalise. While the lambda calculus is very likely *sufficient* to describe the way meanings of parts of sentences are composed, it is less clear that it is *necessary*. Conjunctionism proposes a less powerful alternative with the aim of explaining these generalisations. This is the *independent* motivation for the Conjunctionist approach; I aim to show in subsequent chapters that it also contributes to an insightful explanation of some otherwise unexplained syntactic generalisations.

1.6.2.1 Pure function application

To begin, consider the alternative of deriving the logical form in (1.21) using a function-application approach. Assuming that the verb combines with both its arguments before the adverb attaches (the most favourable syntactic assumption for this approach), this would lead us to the lexical meaning in (1.23a) for ‘stab’, and to the lexical meaning in (1.24a) for ‘violently’. But the lexical verb meaning in (1.23a) seems to leave open the possibility of other logical connectives in place of the conjunctions. We would expect to find verbs in the world’s languages with meanings like the others in (1.23), but we do not. Similarly, the adverb meaning in (1.24a) would lead us to expect to find lexicalized meanings using other logical connectives such as those in (1.24), but we do not.

- (1.23) a. $\lambda x \lambda y \lambda e. \mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, y) \wedge \mathbf{Stabbee}(e, x)$
 b. $\lambda x \lambda y \lambda e. \mathbf{stabbing}(e) \vee \mathbf{Stabber}(e, y) \vee \mathbf{Stabbee}(e, x)$
 c. $\lambda x \lambda y \lambda e. \mathbf{stabbing}(e) \vee \mathbf{Stabber}(e, y) \rightarrow \mathbf{Stabbee}(e, x)$
- (1.24) a. $\lambda P \lambda e. P(e) \wedge \mathbf{violent}(e)$
 b. $\lambda P \lambda e. P(e) \vee \mathbf{violent}(e)$
 c. $\lambda P \lambda e. P(e) \rightarrow \mathbf{violent}(e)$

The claim that there is a generalisation being missed in (1.23a) relies on the premise that thematic relations are expressed individually in logical forms, as argued for

in §1.6.1. Of course for any n -place predicate p , one *can* play the purely formal trick of decomposing p into artificial pieces such that $p(x_1, x_2, \dots, x_n)$ is rewritten as $\exists e[P(e) \wedge R_1(e, x_1) \wedge \dots \wedge R_n(e, x_n)]$. If this were all that were happening, the claim that a generalisation is being missed would be unfounded. But the evidence for thematic separation given in §1.6.1 is evidence that this is *not* all that is happening: we need a one-place predicate **stabbing** and two two-place predicates **Stabber** and **Stabee**. Given these pieces the question arises of which logical operator will connect them, and the answer seems to be that they are always conjoined.

Note that the two meanings in (1.23a) and (1.24a) fail to suggest any distinction between the relationship that ‘stab’ bears to ‘Caesar’ and the relationship that ‘violently’ bears to ‘Brutus stabbed Caesar’: in each case the former denotes a function which is applied to the denotation of the latter. If nothing else, the headedness of the resulting constituents does not seem to agree with this: the result of combining ‘stab’ and ‘Caesar’ is headed by ‘stab’ (the semantic function), but the result of combining ‘violently’ and ‘Brutus stabbed Caesar’ is headed by (the head of) ‘Brutus stabbed Caesar’ (the semantic argument).²¹ Of course one *can* nonetheless maintain that these two pairs of constituents indeed combine via a common semantic operation (i.e. function application), and assume that their differing syntactic relationships have another cause; but I will suggest that the Conjunctivist alternative forces us to a difference in semantic composition that can lead to insightful explanations of syntactic facts.

1.6.2.2 Function application with adjustment in certain configurations

In order to account for the absence of the unattested meanings in (1.24), one might propose that the lexical meaning of ‘violently’ is simply $\lambda e.\mathbf{violent}(e)$. The precise manner in which this composes with the verb (or verb phrase) meaning will need

²¹I return to this point in §2.2.2.

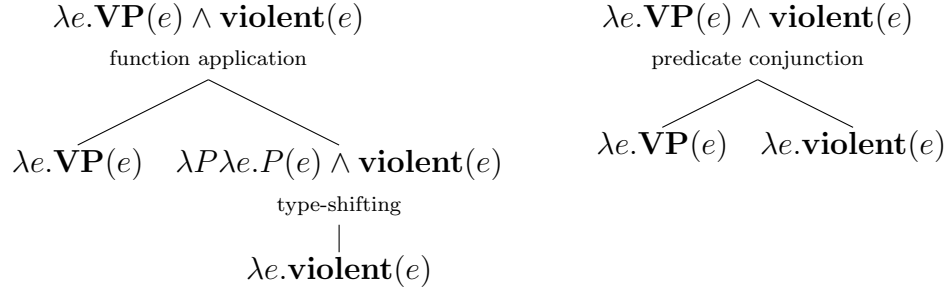


Figure 17

to be adjusted somewhat: it may be either (i) “type-shifted” to the function in (1.24a) before combining via function-application (along the lines of Partee and Rooth 1983, Partee 1986), or (ii) combined via a distinct compositional axiom of predicate conjunction (Higginbotham 1985, Larson and Segal 1995, Heim and Kratzer 1998, Chung and Ladusaw 2004); see Figure 17.

Such a system also allows adjectives to have simple lexical meanings (eg. $\lambda x.\mathbf{red}(x)$) and yet appear both in predicative positions (eg. ‘The ball is red’) via simple function application, and in modifier positions (eg. ‘red ball’) via one of the conjunction-introducing variations in Figure 17. One might furthermore suppose that in this system the choice between simple function application and a conjunction-introducing alternative is correlated with the syntactic distinction between arguments and adjuncts; generally, adjuncts will be constituents that require adjustment in order to be composable in this system, while arguments require no such adjustment, as illustrated for the composition of ‘stabbed Caesar violently’ in Figure 18 (abstracting away from external arguments for the moment, and adopting the type-shifting alternative from Figure 17 for concreteness). Note that we end up with a meaning containing two conjunction symbols, one of which was introduced lexically, the other of which was introduced via a compositional principle.

Most importantly, the underlying idea here is to sometimes *adjust the way in which a constituent contributes meaning to a sentence on the basis of its syntactic*

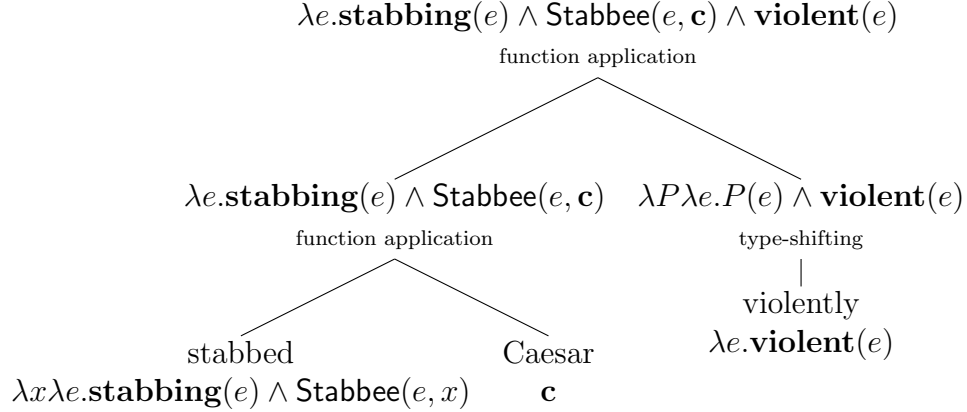


Figure 18

configuration. Conjunctivism uses a strategy of this same familiar sort, the only differences being in *what sort of adjustment* is applied and *which syntactic configurations* prompt it.

1.6.2.3 Conjunction with adjustment in certain configurations

Supplementing function application with a conjunction-introducing mechanism may eliminate the possibility of the unattested adverb meanings in (1.24) (though I return to this below), but the problem of verb meanings in (1.23) remains. It is hard to avoid noticing that the logical connective introduced via “special” means in Figure 17 is also the one that we need to explain the prevalence of in verb meanings, i.e. conjunction. One natural possibility to consider, then, is that conjunction is not introduced via “special” means at all, and that it is the unmarked case. This will require that a syntactically-conditioned adjustment is made to the meanings of arguments, since these are not a perfect fit in a conjunction-oriented system, just as adjuncts required a syntactically-conditioned adjustment because of their imperfect fit in a function-application-oriented system.

The result is that a sort of type-shifting is hypothesised for arguments (instead of for adjuncts) that marks them with a certain thematic relation, as shown in Figure 19;

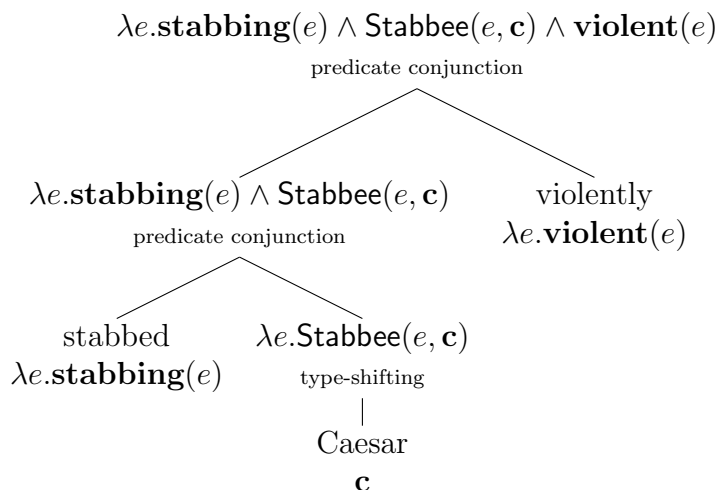


Figure 19

compare with Figure 18.²²

This does of course raise the question of *which* thematic relation an argument is marked with, and how that choice is made in general. I will assume that the predicate an argument contributes is determined not only by the fact that it is in *an* argument position, but furthermore by *which* argument position of the relevant head — roughly, internal or external, though perhaps there are more than two options — it is in, along the lines of the Uniformity of Theta Assignment Hypothesis (UTAH) (Baker 1988, 1997).²³ Accordingly, we can think of the shift in Figure 19 as producing not $\lambda e.\mathbf{Stabbee}(e, \mathbf{c})$ but rather $\lambda e.\mathbf{Patient}(e, \mathbf{c})$ or $\lambda e.\mathbf{Internal}(e, \mathbf{c})$. I will return to these details, but more importantly, the adjustment applied will certainly depend on the fact that ‘Caesar’ in Figure 19 is an *internal* argument (rather than an external argument). This means that the adjustment can not take the form of a genuinely unary rule as Figure 19 suggests. But the syntax I will propose in chapter 2 to support this compositional system accommodates, and indeed makes crucial use of, the fact

²²This is related to the idea from Kratzer (1996) that the external argument should be “severed” from the verb; the difference is that while Kratzer severs only external arguments and leaves internal arguments to combine basically as in Figure 18, in Figure 19 *all* arguments are severed from the verb.

that an argument’s adjustment relies on its position relative to other arguments.

A system where thematic relations are introduced roughly as in Figure 19 makes it possible to derive logical forms that contain, say, ‘**stabbing**’ but not ‘**Stabber**’. Suppose that we accepted (say, on the basis of Schein’s argument) that complete logical forms must contain thematic conjuncts such as ‘**Stabber**(e, x)’; but suppose that we were *not* swayed by Parsons’ argument that the existence of a stabbing event does not entail the existence of a stabber. Then we would be free to consider the analysis in (1.25) for the two sentences in (1.9), repeated here.

- (1.9) a. Brutus stabbed Caesar
 b. Caesar was stabbed

- (1.25) a. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$
 b. $\exists e\exists x[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, x) \wedge \mathbf{Stabee}(e, \mathbf{c})]$

An arguably simpler solution, however, would be to drop the assumption that the ‘**Stabber**’ relation originates as part of the lexical meaning of the verb itself, and instead suppose that all thematic relations are introduced in the manner shown in Figure 19. Then since no stabber is expressed in (1.9b), the ‘**Stabber**’ relation will simply remain absent from the final logical form, as shown in (1.11) (also repeated), rather than being introduced by the verb meaning and then nullified by existential closure.

- (1.11) a. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabee}(e, \mathbf{c})]$
 b. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabee}(e, \mathbf{c})]$

With thematic relations treated as shown in Figure 19, “the omission of arguments does not require that extra semantic operations be defined to ‘plug’ the unbound

²³One could even argue that (something along the lines of) UTAH is *predicted by* the Conjunctivist theory that we have been independently led to by the desire to eliminate the problems noted with respect to (1.23) and (1.24), for there is seemingly, by hypothesis, nowhere else for thematic relations to come from.

argument positions” (Carlson 1984, p.262). So Parsons’ is not the only argument for (1.11).

In cases such as passivisation, it might be possible to reduce the theoretical burden of requiring this semantic “plugging” operation by associating it with the syntactic operation that modifies the verb’s subcategorisation frame. In other cases, however, it is less obvious that there is an independently-required syntactic operation on which to piggyback. Nominalisations can occur with or without participants expressed, for example, as evidenced by (1.26), but there is no clear reason — other than the desire to maintain the lexical meanings with thematic relations included — to postulate an operation in (1.26b) that covertly plugs a position overtly filled by ‘Caesar’ in (1.26a).

- (1.26) a. Stabbing Caesar is dangerous
 b. Stabbing is dangerous

Another consequence of the approach to thematic relations illustrated in Figure 19 is the possibility that a verb that requires a certain argument in some syntactic contexts, might not require it in other contexts. Such flexibility would be surprising under the assumption that all thematic relations are part of a verb’s lexical meaning, as in Figure 18. Williams (2008, pp.11–12) presents evidence of exactly this sort from Mandarin: while the verb ‘qiē’ (‘cut’) requires that its patient is expressed when used in simple clauses, as illustrated by (1.27), this requirement disappears when it is used as part of a resultative construction, as illustrated by the acceptability of (1.28).²⁴

- (1.27) a. Lǎo Wèi qiē -le zhúsún
 Lao Wei cut -PFV bamboo shoots
 Lao Wei cut bamboo shoots
 b. *Lǎo Wèi qiē -le
 Lao Wei cut -PFV
 Lao Wei cut
 c. *Lǎo Wèi zài qiē
 Lao Wei PROG cut

Lao Wei is cutting

- (1.28) Tā hái qiē dùn -le nǐde cǎidāo
3S also cut dull -LE your food.knife
He also made your cleaver dull by cutting

See Williams (2005, 2008) for elaboration of this point, and for illuminating discussion of the subtle relationship between “projective” theories (where thematic relations are introduced by verbs) and “non-projective” theories (where they are introduced by syntactic context) of verbal valence.

Finally, if thematic relations are introduced by certain syntactic configurations and not projected from lexical verb meanings, the task of giving a compositional derivation of sentential logical forms like Schein’s (1.18b) seems at least manageable.²⁵ With thematic relations originating in the verb meanings, this task is extremely problematic (as foreshadowed in footnote 20). The lexical meaning in (1.20) gives us no way to produce thematic conjuncts like **Stabber**(e, \mathbf{b}) that involve distinct event variables, but this is exactly what (1.18b) requires: note that in (1.18b), **Agent** is applied to e whereas **Theme** is applied to e' . Similarly, (1.20) leaves us no way to have one quantifier scoping over the **Stabber** specification and another quantifier scoping over the **Stabee** specification, with no scope relation between the two quantifiers. With thematic relations introduced “at a distance” from the verb, however, this becomes at least plausible: Figure 20 shows a sketch of how the logical form in (1.18b) might be derived, abstracting away from many details. Whatever the syntactic details of the introduction of the higher event variable, the crucial point is that the specification of **Agent** is contained in a constituent that (i) only modifies this higher event variable, and (ii) contains a quantifier that has the **Agent** specification, but nothing else, in its scope. This would not be possible if every thematic role specification originated

²⁴The sentences marked unacceptable with the given interpretations in (1.27) can be analysed as including a silent object pronoun, and therefore understood with a distinct meaning, namely that Lao Wei cuts or is cutting something contextually salient. But Williams crucially shows that (1.28) does not contain such a silent pronoun.

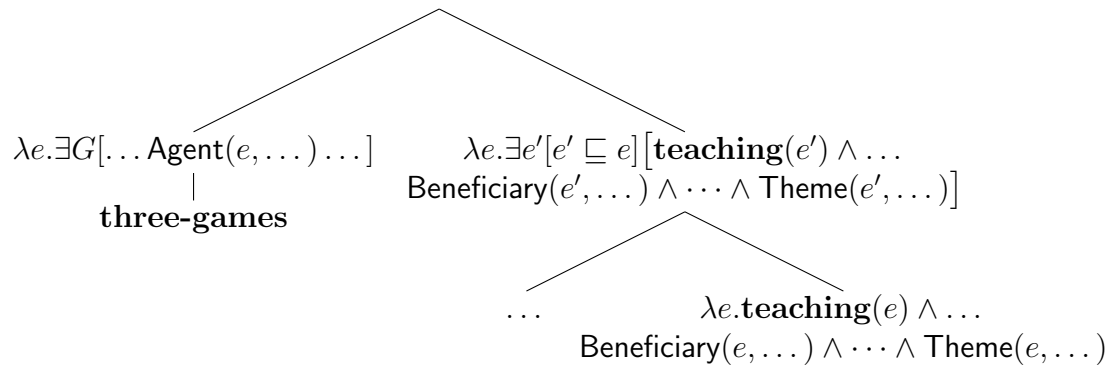


Figure 20

from the lexical meaning of a single terminal node: mere “lexical separation” is not enough (Schein 1993, p.10).

1.6.2.4 Comparisons and discussion

It is worth emphasising that this Conjunctivist idea, of using structural configurations to alter the way a constituent contributes meaning, is no more ad hoc than adding a mechanism of predicate conjunction like either of those in Figure 17, used in certain syntactic circumstances, to a system otherwise based on function application.²⁶

In fact, to the extent that the two theories — function application supplemented with syntactically-conditioned predicate conjunction, and predicate conjunction supplemented with syntactically-conditioned thematic relations — differ at all, surely the latter is the more minimal. The former theory adds predicate conjunction to a system that already overgenerates lexical meanings: a theory with meanings such as (1.24a) is descriptively adequate, just explanatorily toothless. But adding a rule to the system cannot *reduce* its generative capacity, and so cannot increase its explanatory capacity. It is not at all clear that adding a predicate conjunction mechanism

²⁵And likewise, if we are convinced by Parsons’ argument that stabbings need not have stabbers, and therefore convinced that (1.10) and (1.25) must be rejected in favour of (1.11) even before issues of syntax are considered, we will very likely be forced to adopt something like Figure 19 when they are.

²⁶Pietroski (2005, pp.67–75) discusses this comparison in detail.

can explain the *lack* of the unattested meanings in (1.24). Supplementing function application with predicate conjunction results in two ways to encode the attested adverb meanings, and one way to encode the unattested ones. This is not obviously an improvement — and for all this, it still says nothing about the lack of the unattested verb meanings in (1.23). When Conjunctivism supplements predicate conjunction with thematic relations, it expands its descriptive coverage to, it seems, just about the right degree. All and only the attested meanings in (1.23) and (1.24) are predicted, and the attested verb meaning *requires* that we allow syntactic configurations to affect semantic contributions (specifically, to impart thematic relations).

Note also that given neo-Davidsonian logical forms, taking predicate conjunction to be the basic semantic operation means generalising to the unbounded case of syntactic combination: syntactically, any number of constituents can attach to a verb the way ‘violently’ does. We invoke some additional machinery to deal with the much more restricted case of arguments: syntactically, only a bounded number of constituents (typically one or two, occasionally three or perhaps four) can attach the way ‘Caesar’ does. It seems plausible that arguments, unlike adjuncts, should be severely limited in number like this if they indeed require some additional “help” in order to be semantically composable.

To be clear, I have not yet said anything about what the particular syntactic configuration is that provides arguments with their thematic relations. Formulating a precise hypothesis about this is a main aim of chapter 2. For now the point is just that we have independent semantic justification for the idea. Note, however, that a theory that permits semantic composition via function-application is much less likely to *force* us to the conclusion that there should be such differences in kind of syntactic configuration, because given plausible meanings for α and for a complex expression $[\alpha \beta]$ it is rarely difficult to find *some* function that maps the former to the latter and call this the meaning of β ; recall (1.24a), for example.

1.6.3 Conjunctivist details

Having provided the broad motivation for the Conjunctivist approach to the composition of neo-Davidsonian logical forms, in this subsection I briefly fill out some further details of the way it can be put into practice.

First, if the theory is to apply to semantic composition not just in the verbal domain but more generally, then the distinction we have noted between modifying an event variable directly (as ‘violently’ does) and indirectly (as ‘Caesar’ does) must extend to other categories as well. Relative clauses and adjectives are obvious examples of adjuncts in the nominal domain with a clear conjunctive interpretation; all the comments made above with respect to the conjunctive interpretation of adjuncts carry over to adjectives, and the question of whether the lexical meaning of ‘red’ should be $\lambda P \lambda x. P(x) \wedge \mathbf{red}(x)$ or $\lambda x. \mathbf{red}(x)$. Thus the nominal expression in (1.29) is associated with the conjunction of two monadic *predicates of individuals*.

$$(1.29) \quad \text{Brutus [who is tall]} \\ \lambda x. \mathbf{b}(x) \wedge \mathbf{tall}(x)$$

If expressions of every category are to be associated with monadic predicates, our understanding of the adjustment that an argument’s meaning undergoes must be modified slightly to include existential closure. Since every semantic value is now a monadic predicate, we may as well adopt notation that treats predicates as “first class” entities; so we define a predicate conjunction operator, written ‘&’. Sentential logical forms are then written as shown in (1.31).

$$(1.30) \quad P \ \& \ Q = \lambda x. P(x) \wedge Q(x)$$

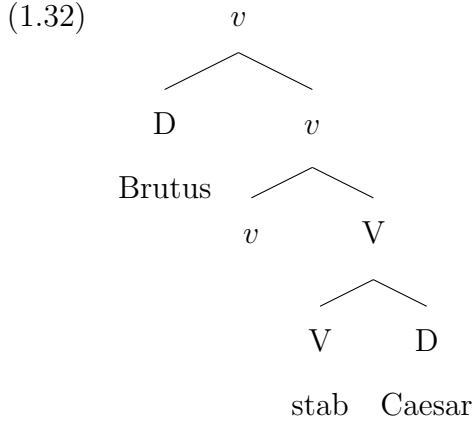
$$(1.31) \quad \text{Brutus [who is tall] stabbed Caesar.}$$

$$\exists e[\mathbf{stabbing}(e) \wedge \exists x[(\mathbf{b} \ \& \ \mathbf{tall})(x) \wedge \mathbf{Stabber}(e, x)] \wedge \exists y[\mathbf{c}(y) \wedge \mathbf{Stabee}(e, y)]]$$

The general significance of being interpreted in an argument position of a verb is no longer to shift an individual x to a predicate of events $\lambda e. R(e, x)$ for some binary

relation R , but rather to shift a predicate of individuals p to a predicate of events $\lambda e.\exists x[p(x) \wedge R(e, x)]$. But the general points about the strategy of associating semantic effects with particular syntactic configurations in the last subsection are not affected by this change.²⁷

A second detail to clarify is the syntactic position of subjects. If we adopt the standard assumption that arguments' thematic relations are actually established in the positions shown in (1.32), where the relationship between v and 'stab' is the same as that between 'stab' and 'Caesar', then we are led to suppose that the VP and the v P denote the predicates Φ and Ψ in (1.33), respectively. I assume that the v head denotes some trivial or near-trivial event predicate (perhaps satisfied by all and only events that have agents) which I write as ' \boldsymbol{v} '.



$$\begin{aligned}
 (1.33) \quad \Phi &= \lambda e.\mathbf{stabbing}(e) \wedge \exists y[\mathbf{c}(y) \wedge \mathbf{Internal}(e, y)] \\
 \Psi &= \lambda e.\boldsymbol{v}(e) \wedge \exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)] \wedge \exists e'[\Phi(e') \wedge \mathbf{Internal}(e, e')]
 \end{aligned}$$

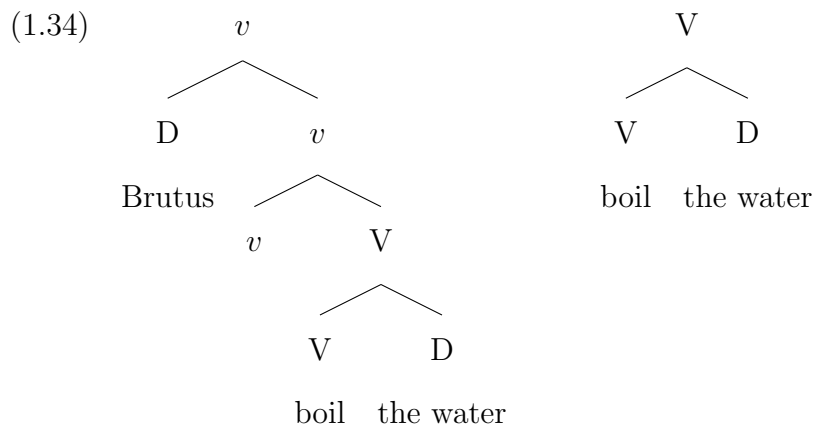
²⁷If a selected argument such as 'Caesar' denotes a monadic predicate as suggested here, then this *systematic* existential closure would be another unexplained coincidence in a system based on function-application. In just the same way as it is mysterious why only one of the possibilities in (1.23) and (1.24) is attested, it would be mysterious why only the first of the following is attested (holding the outer connectives constant as conjunction):

- (i) $\lambda P\lambda Q\lambda e.\mathbf{stabbing}(e) \wedge \exists x[P(x) \wedge \mathbf{Stabbee}(e, x)] \wedge \exists y[Q(y) \wedge \mathbf{Stabber}(e, y)]$
- (ii) $\lambda P\lambda Q\lambda e.\mathbf{stabbing}(e) \wedge \forall x[P(x) \rightarrow \mathbf{Stabbee}(e, x)] \wedge \forall y[Q(y) \rightarrow \mathbf{Stabber}(e, y)]$
- (iii) $\lambda P\lambda Q\lambda e.\mathbf{stabbing}(e) \wedge \exists x[P(x) \wedge \mathbf{Stabbee}(e, x)] \wedge \forall y[P(y) \wedge Q(y) \rightarrow \mathbf{Stabber}(e, y)]$

In general, the lack of meanings where two distinct constituents contribute to the one top-level conjunct, as P and Q do in the last of the above, can be seen as further evidence that each constituent somehow contributes exactly one conjunct.

Now that we are considering argument-like relations more generally rather than just those between verbs and nominals, I have switched to the generic **Internal** and **External** relations in place of specific thematic relations such as **Stabbee** and **Stabber**; recall the role of UTAH outlined above. “Metaphorically . . . verbs ‘infuse’ [**Internal** and **External**] (which by themselves have purely formal significance) with specific thematic significance” (Pietroski 2005, p.53), subject to constraints (or preferences or “defaults”) that (something like) UTAH brings.²⁸ If bearing the **Internal** relation to (a satisfier of the predicate denoted by) ‘stab’ amounts to being a thing stabbed, bearing the **Internal** relation to (a satisfier of the predicate denoted by) *v* can amount to being some kind of “result”.²⁹ From this the entailment from (1.8a) to (1.8b), repeated below, follows immediately, on the standard assumption that the subject of (1.8b) is underlyingly an internal argument (Burzio 1986, Belletti 1988, Baker 1997). The structures in (1.34) yield the predicates Φ and Ψ in (1.35) for (1.8a) and (1.8b), and so the relevant truth conditions are expressed by $\exists e[\Psi(e)]$ and $\exists e[\Phi(e)]$, the former of which entails the latter.

- (1.8) a. Brutus boiled the water.
 b. The water boiled.



²⁸I address some objections to this in §1.6.4. See also Pietroski (2005, pp.51–55,200–209) on UTAH and the **Internal** and **External** relations.

$$(1.35) \quad \Psi = \lambda e. \mathbf{v}(e) \wedge \exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)] \wedge \exists e'[\Phi(e') \wedge \mathbf{Internal}(e, e')] \\ \Phi = \lambda e. \mathbf{boiling}(e) \wedge \exists y[\mathbf{water}(y) \wedge \mathbf{Internal}(e, y)]$$

It is now helpful to adopt some further notational conveniences that will make this kind of semantic expressions much more readable. Applying the operator ‘**int**’ to a predicate P transforms it into the predicate contributed by a constituent denoting P which is in an internal argument position, and likewise ‘**ext**’ for external argument positions. These are defined explicitly in (1.36). For example, while \mathbf{c} is a predicate of individuals, $\mathbf{int}(\mathbf{c})$ is a predicate of events; specifically, the predicate satisfied by an event iff its internal “participant” satisfies \mathbf{c} . So $\mathbf{stabbing} \ \& \ \mathbf{int}(\mathbf{c})$, for example, is the predicate satisfied by all and only stabbings of Caesar.

$$(1.36) \quad \mathbf{int}(P) = \lambda e. \exists x[P(x) \wedge \mathbf{Internal}(e, x)] \\ \mathbf{ext}(P) = \lambda e. \exists x[P(x) \wedge \mathbf{External}(e, x)]$$

Then the two predicates in (1.35) can be rewritten as in (1.37), corresponding transparently to the two structures in (1.34).

$$(1.37) \quad \Psi = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\Phi) = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})) \\ \Phi = \mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})$$

This notation, and its clear relation to syntactic structure, may arouse suspicion. What is really happening in the translation from the trees in (1.34) to (1.37)? If these expressions are making reference to seemingly syntactic relationships like being internal and external arguments, are they really “saying anything about semantics” at all? For comparison, consider what the same structures would be translated into under a more standard approach based on function application: the two expressions in (1.38).

²⁹The well-known observation that the relevant “resulting” relation seems to differ from verb to verb can be accounted for by supposing that the head movement of the lexical verb to v has some semantic consequences; see Pietroski (2003b, 2005).

$$(1.38) \quad \Psi' = v(\mathbf{boil}(\mathbf{water}))(\mathbf{b})$$

$$\Phi' = \mathbf{boil}(\mathbf{water})$$

These are no more enlightening than their counterparts in (1.37). In fact *they are less so*: the expressions in (1.37) account for an entailment that the expressions in (1.38) do not, because the basic Conjunctivist compositional axioms, unlike function application, bring in some logical vocabulary, specifically conjunction and existential closure.³⁰ No translations of the kind in (1.38) can *ever* derive any entailments among sentences in and of themselves; entailments can only follow when lexical items are assigned meanings with logical content, eg. **every** = $\lambda P \lambda Q. \forall x [P(x) \rightarrow Q(x)]$ or **and** = $\lambda P \lambda Q. P \wedge Q$.

Comparing semantic translations of the sort in (1.37) with those of the sort in (1.38) can also reinforce the point, made in §1.6.2, that the Conjunctivist view generalises to the unbounded case of syntactic composition. Recall the basic idea that whereas many standard approaches take function application as the fundamental composition operation and adjust the semantic contribution of adjuncts, Conjunctivism takes predicate conjunction as the fundamental operation and adjusts the semantic contribution of arguments. To facilitate the comparison, let us temporarily adopt a new notation for function application: instead of writing ‘ $f(x)$ ’ for the result of applying a function f to x , suppose we write ‘ $f \$ x$ ’. With this adjustment the two proposed fundamental semantic composition operations now have equally “noisy” notations: ‘ $\phi \& \psi$ ’ and ‘ $f \$ x$ ’.³¹ Similarly, to parallel the ‘int’ and ‘ext’ operators adopted by Conjunctivism to express syntactically-conditioned adjustments, let us temporarily introduce an operator ‘shift’ to express the kind of syntactically-conditioned conjunction-introducing adjustment that one might add to a system based

³⁰Alternatively, one might think of what is happening as follows: rather than providing a translation from phrase markers into first-order logic, with int and ext just shorthands, we are providing a translation from phrase markers into a specialised logic. This specialised logic has one binary operator $\&$ and two unary operators int and ext; with inferences licenced (i) from $\phi \& \psi$ to ϕ , (ii) from $\phi \& \psi$ to ψ , (iii) from int(ϕ) to ϕ , and (iv) from ext(ϕ) to ϕ . (Note that if we have proven both ϕ and ψ , we can not infer $\phi \& \psi$.)

on function application, as shown in Figure 17 on page 38.

Consider now (1.39), a sentence with one canonical argument and one canonical adjunct. In the function-oriented alternative, it is the adjunct, ‘quietly’, which has its semantic contribution adjusted (converted from a property to a property-modifier, roughly); in the notation just introduced, the semantic composition of this sentence’s meaning can be expressed as shown in (1.40a). (I abstract away from complications about whether the adverb attaches below the subject, in which case ‘\$’ should be understood to be left-associative, or attaches above the subject as sketched in (1.23) and (1.24) above, in which case ‘\$’ should be understood to be right-associative.) On the Conjunctivist view, it is the argument, ‘Brutus’, which undergoes semantic adjustment, as shown in (1.40b).

(1.39) Brutus slept quietly

- (1.40) a. **shift(quiet) \$ sleep \$ b**
b. **quiet & sleeping & ext(b)**

These two expressions make clear the symmetric nature of the relationship between the two proposals under discussion: what one does to certain constituents, the other does to certain other constituents. A difference is that the Conjunctivist approach uses its unmarked composition operation for the unbounded case of syntactic composition (adjuncts), whereas the function-oriented view uses its unmarked composition operation for the bounded case (arguments). In the general case, a sentence can have an unboundedly large number of adjuncts (say, n of them, $\alpha_1, \alpha_2, \dots, \alpha_n$) but only a small bounded number of arguments (say, 2 of them, β_1 and β_2). The relevant representations for the two approaches considered here are shown in (1.41).

³¹It is convenient for mathematicians and computer scientists to use the “non-noisy” ‘ $f(x)$ ’ (or ‘ fx ’) notation for function application because of the generality and universality of this operation. For cognitive scientists interested in circumscribing the ways in which humans do and do not compose meanings, however, it is far from obvious that this generality and universality is something we should like to hide.

- (1.41) a. $\text{shift}(\alpha_1) \$ \text{shift}(\alpha_2) \$ \dots \$ \text{shift}(\alpha_n) \$ (\text{verb} \$ \beta_1) \$ \beta_2$
 b. $\alpha_1 \& \alpha_2 \& \dots \& \alpha_n \& \text{verb} \& \text{int}(\beta_1) \& \text{ext}(\beta_2)$

Thus if function application is our basic composition operation, the general case requires an unbounded number of adjustments; whereas if conjunction is our basic operation, the general case requires a small bounded number of adjustments.³² As mentioned above, the severely limited capacity of the language faculty to accommodate arguments plausibly relates to this sense in which arguments require some additional “help” in order to be semantically composable. “In an event-based semantics, arguments — not adjuncts — are the interpretive oddballs” (Hornstein and Nunes 2008, pp.70–71).

1.6.4 Potential objections

Here I discuss two well-known objections to assumptions that I have made in this section: first, there is the issue of adverbs (and adjectives) with “non-conjunctive” meanings, and second, concerns over whether the notions of “being an internal argument” and “being an external argument” have some meaningful content independent of particular heads that things might be arguments of. My aim here is not to settle these issues (which have been discussed at length elsewhere) beyond any doubt, only to outline some reasons for believing that the assumptions I have adopted can survive these challenges, in one way or another. Note also that to the extent that these semantic assumptions play a role in providing insightful accounts of various syntactic phenomena in the rest of this thesis, the overall evidence in their favour grows, quite apart from anything I say here.

Cases of adjectival and adverbial modification like those shown in (1.42) appear at first to be problematic for a theory based heavily on conjunction. The sentences

³²Of course the fact that the sentences speakers *use* may not typically contain many more adjuncts than arguments is beside the point, since it is the limits of the system generating these expressions that we are investigating.

in (1.42a) appear to violate the pattern of entailments illustrated above in (1.5); and an alleged crook is not (necessarily) a crook, a fake diamond is not a diamond, and a big ant is not (obviously) big.

- (1.42) a. Brutus stabbed Caesar allegedly/apparently/nearly/virtually.
b. alleged crook
c. fake diamond
d. big ant

First it should be noted that the adjectival cases are problematic for *any* theory that systematically associates a phrase like ‘red ball’ with the conjunction of two predicates $\lambda x.\mathbf{red}(x)$ and $\lambda x.\mathbf{ball}(x)$, whether this is achieved via type-shifting or via a predicate-conjunction rule; recall Figure 17. So while a solution to this puzzle is certainly required, it is required by many theories that do not rely so prominently on conjunction, including many that do not make any use of event variables (eg. Heim and Kratzer 1998). Whatever solution is applied to ‘alleged crook’ will quite plausibly carry over to ‘allegedly’, with event variables in place of individual variables.

It seems likely that the correct account of these “non-conjunctive” cases will involve *some* kind of underlying conjunction. While ‘fake diamond’ is not simply the obvious conjunction, it is very plausibly the conjunction of a certain kind of “fake-ness” and a certain kind of “diamond-ness”, perhaps with one (or both) of the conjuncts relativized to the other in some way. There is evidence that non-conjunctive modifiers differ syntactically from strictly conjunctive ones (Cinque 1994, Bouchard 1998, Valois 2005); while these differences are too subtle to figure in the coarse-grained discussion of adjunct syntax here, it would make sense in the Conjunctivist setting to hypothesise that some different syntactic structure is contributing to these adjusted interpretations. Alternatively, one might maintain the Conjunctivist hypothesis by making some auxiliary assumptions about exactly what the conjunctively-constructed predicate is a predicate of: Larson (1998) argues for an analysis of ‘beautiful dancer’

where ‘beautiful’ conjunctively modifies an event variable, and McNally and Boleda (2004) make a similar proposal where it is properties of kinds that are conjoined. Given such possibilities, weakening our theory to allow adverbial meanings like the radically non-conjunctive ones in (1.24) seems to be going too far, even if more must be said about exactly which predicates are conjoined in certain cases.

In fact, taking adjectives to be arbitrary property-modifiers as in (1.24) is not only worryingly permissive, but also too restrictive in other ways: Kamp and Partee (1995), following Kamp (1975), suggest that many non-conjunctive adjective interpretations are context-dependent “with the most relevant aspect of context a comparison class which is often, *but not exclusively*, provided by the noun of the adjective–noun construction” (emphasis added) (Kamp and Partee 1995, p.142). While the correct interpretation of ‘big ant’ does indeed require relativising the interpretation of ‘big’ to (something like) the class of ants, the relevant comparison class need not be expressed by the noun that the adjective modifies, as illustrated by the differing standards of height suggested in (1.43) (their example (9)) despite the fact that the same noun (‘snowman’) is being modified in both cases.

- (1.43) a. My two-year-old son built a really tall snowman yesterday.
b. The D.U. fraternity brothers built a really tall snowman last weekend.

Even given the freedom to have the standard of height vary as a function of the noun that ‘tall’ modifies, the contrast in (1.43) would have to be left to context-dependence. It therefore seems unproblematic to suppose that the same context-dependence will account for cases where the relevant comparison class happens to be expressed more locally, such as ‘big ant’.

The second contentious assumption is that there is some meaningful content to the notions “internal argument” and “external argument”. Some preliminary motivation for this idea comes from the observation that the structural positions in which an event’s participants are expressed are, to a large extent, predictable from the roles

the participants play. This is most clear in the case of two-participant verbs involving an agent and a patient/theme.

Virtually every such verb expresses the agent of the event as its subject as shown in (2a); there are no verbs that follow the pattern in (2b), where the theme of the event is expressed as the subject.

- (2) a. John hit/built/found/pushed/bought/cleaned/broke/described
the table.
- b. *The table plit/puilt/vound/fushed/pought/bleaned/proke/tescribed
John.

(Baker 1997)

This observation prompts the well-known Uniformity of Theta Assignment Hypothesis (UTAH) (Baker 1988, 1997), according to which the thematic role borne by a participant determines the syntactic position it will be expressed in. To the extent that such a generalisation is correct, it reveals another “accidental” lack of lexical meanings, for implementations of neo-Davidsonianism based on function-application, of the sort illustrated in (1.23) and (1.24) with respect to logical connectives; specifically, while verbs in many languages exist with the meaning in (1.44a), none are known with the meaning in (1.44b).³³

- (1.44) a. $\lambda x \lambda y \lambda e. \mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, y) \wedge \mathbf{Stabee}(e, x)$
- b. $\lambda x \lambda y \lambda e. \mathbf{stabbing}(e) \wedge \mathbf{Stabee}(e, y) \wedge \mathbf{Stabber}(e, x)$

Besides the typological facts, additional evidence that (something like) UTAH holds — perhaps as a violable “soft” constraint or default preference — comes from the finding that infants’ hypotheses about the meanings of novel verbs appear to be constrained by the structural positions in which participants are expressed (Gleitman

1990, Naigles 1990, Gropen et al. 1991, Fisher et al. 1994, Kako and Wagner 2001).

Although it seems broadly correct that some such generalisation holds, a number of related questions remain open. It is far from clear exactly what a general thematic relation such as **Agent** could be taken to mean: is there a precise sense in which John does “the same thing” (Schein 2002, p.267) in each of Baker’s (2a) sentences above? UTAH, as originally stated, implies that the answer must be “yes” — at least modulo any differences between underlying structure and surface form that can be independently supported, as have been proposed for unaccusative verbs for example (Burzio 1986, Belletti 1988, Levin and Rappaport Hovav 1995).

Dowty (1991) is sceptical that this strong claim can be maintained, and proposes an alternative in terms of “proto-roles” instead. On this view it is not true that John’s role in the hitting is the same as his role in the building, for example, in Baker’s (2a). What is true, however, is that John’s role in the hitting is the most “agent-like” of the roles played by that event’s participants, and that his role in the building is the most “agent-like” of the roles played by that event’s participants; for Dowty this suffices to establish that ‘John’ will be the external argument in both cases. The success of this proposal depends, of course, on an adequate definition of “agent-like” and “patient-like”, and I will not address these questions here. If such a picture turns out to be correct, the sense in which the notions of internal and external argument are significant will be different from that suggested by Baker, but will plausibly be consistent with the assumptions I rely on here.

Even Dowty’s weakened view becomes problematic, however, when we consider

³³A parallel remark can be made for the fact that natural language determiners are universally conservative (Barwise and Cooper 1981, Higginbotham and May 1981, Keenan and Stavi 1986). Specifically, if there is no general significance to being an internal or external argument to a determiner, then it is a mystery why so many languages have, for example, a determiner (eg. ‘every’) with the meaning $\lambda P\lambda Q.\forall x[P(x) \rightarrow Q(x)]$, and none have a determiner with the meaning $\lambda P\lambda Q.\forall x[Q(x) \rightarrow P(x)]$. This generalisation seems more robust than UTAH (at least when we look beyond verbs with only agent and patient/theme participants), and experimental evidence has shown that children will obey this constraint when hypothesising meanings for novel determiners (Hunter and Conroy 2009).

examples like the ‘buy’/‘sell’ pair in (1.45).

- (1.45) a. John sold a book to Mary
b. Mary bought a book from John

At least intuitively, it would appear that these two sentences describe the same event, and therefore that the role John plays in (the event described by) (1.45a) is the same as the role he plays in (the event described by) (1.45b). Whatever definitions of agent-like and patient-like one proposes, if John plays the most agent-like role in (the event described by) (1.45a) one would expect the same to be true of (1.45b), and yet ‘John’ does not occupy the same syntactic position in the two sentences. These kinds of facts suggest that either (i) the significance of being an external argument (for example) must be relativised not only to the other participants in the event, as Dowty suggests, but also to the verb used to describe the event,³⁴ or (ii) events (in the relevant sense of the word) are more fine-grained than intuition would have us believe, such that the two sentences in (1.45) do *not* in fact describe the same event; see Schein (2002).

Working out the details of either of these alternatives is not straightforward, and requires answers to many subtle questions well beyond the scope of this thesis. I will use generic **Internal** and **External** relations in the rest of this thesis without further justification. By avoiding **Agent** and **Patient** specifically I aim to leave open the pos-

³⁴This option seems to amount to completely denying that there is any independent significance of external or internal argument-hood. Strictly speaking, however, it may be possible to shoehorn even this view into working with logical forms expressed in terms of **Internal** and **External**, if one really wanted to. Suppose a verb with the UTAH-incompatible meaning in (1.44b) were discovered. There still must be some way to identify “which argument goes where”, i.e. which argument takes the place of the x variable and which the place of the y variable. Usually this is thought of as following straightforwardly from the assumption that a lexical meaning is a *curried* function and thus can semantically compose when not all of its arguments are present. If one instead stipulated that lexical meanings were *uncurried* functions (taking a single *ordered pair* as an argument: $\lambda\langle x, y \rangle \dots$) then **Internal** and **External** would simply be labels indicating what should be the first coordinate of the ordered pair and what should be the second; then whether a verb’s meaning assigns the agent role to the first or second coordinate of the pair would be purely a matter of lexical semantics. I do not mean to suggest that this stipulative uncurrying of functions should be seriously considered purely so that **Internal** and **External** can become significant notions, but the thought experiment demonstrates, I think, that adopting these relations is not obviously as strong a claim as it might initially appear.

sibility that accounts of the semantic significance of syntactic positions more flexible than UTAH itself, such as that of Dowty (1991), may be compatible with my proposals. I will, however, definitely be committed to the claim that whatever significance comes with being an external argument (for example) comes as a result of being in a particular *syntactic configuration* with the verb, not from a rule delineating possible and impossible *lexical* meanings (allowing (1.44a) but disallowing (1.44b), for example) in some pre-syntactic sense.

Chapter 2

Arguments, Adjuncts and Conjunctivist Interpretation

2.1 Overview

The goal of this chapter is to extend the MG formalism as presented in §1.5 to permit adjunction. I will argue that the adjustments I propose are independently motivated by theories of the composition of neo-Davidsonian logical forms, and provide a natural account of some of the basic syntactic differences between arguments and adjuncts. In the resulting system, the syntactic behaviour of arguments and adjuncts can therefore be construed as reflection of their differing contributions to event-based logical forms. The basic intuition to be exploited is that the mode of semantic composition used by adjuncts is simpler than that used by arguments, allowing a degree of syntactic freedom for adjuncts that is unavailable for arguments. The particular way in which adjuncts are more “loosely” attached than arguments is made possible by the same redistribution of structure-building labour (using INS and (I-)MRG, instead of E-MRG and I-MRG) that lets us see movement as mere re-merging, as discussed in §1.5.3.

The pattern of the facts to be explained (discussed in more detail in §2.2) can be seen in (2.1) and (2.2). We would like to know what it is about the two verb phrases, ‘sleep quietly’ and ‘stab Caesar’, that permits ‘quietly’ to be either included in or excluded from the constituent targeted for fronting/clefting in (2.1), but does not permit ‘Caesar’ the same flexibility in (2.2).³⁵

- (2.1) a. Brutus [_{VP} slept quietly].
 b. Sleep quietly, (is what) Brutus did.
 c. Sleep, (is what) Brutus did quietly.
- (2.2) a. Brutus [_{VP} stabbed Caesar].
 b. Stab Caesar, (is what) Brutus did.
 c. * Stab, (is what) Brutus did Caesar.

My solution to this puzzle draws on the idea from neo-Davidsonian semantics that an argument bears a more complex relation to the head it modifies than an adjunct does. On this view, (2.1a) and (2.2a) are associated with the logical forms in (2.3a) and (2.3b), respectively. The crucial point is that while ‘quietly’ modifies the event variable directly in (2.3a), ‘Caesar’ does so only indirectly, via a thematic relation, in (2.3b).

- (2.3) a. $\exists e[\mathbf{sleeping}(e) \wedge \mathbf{Sleeper}(e, \mathbf{b}) \wedge \mathbf{quiet}(e)]$
 b. $\exists e[\mathbf{stabbing}(e) \wedge \mathbf{Stabber}(e, \mathbf{b}) \wedge \mathbf{Stabee}(e, \mathbf{c})]$

This calls for an account of the syntax-semantics interface which produces these distinct kinds of semantic contributions. I will show that syntactic structures motivated

³⁵Rather than taking (2.1) and (2.2) to reveal a difference about what can count as a VP for the purposes of VP-fronting, one might pursue an alternative explanation where ‘quietly’ is right-extraposed in (2.1c) to allow fronting of the remnant [_{VP} sleep ____]. Such an alternative would indeed avoid the need to explain why some (but not all) constituents of VPs are subject to fronting, but would instead need to explain why there is no analogous derivation of (2.2c) involving extraposition of ‘Caesar’. Since the extraposition of the object itself appears to be possible, as illustrated by the possibility of heavy NP shift in (i), the source of the ungrammaticality would presumably need to be the remnant movement of the VP. Such remnant movement, of a VP from which an object has been extraposed, does indeed seem to be disallowed, as illustrated in (ii).

- (i) Brutus stabbed violently [every last Belgian warlord he encountered].
 (ii) * Stab violently, Brutus did [every last Belgian warload he encountered].

But then one is faced with explaining a condition on which remnant VPs can and cannot be fronted, that is sensitive to the position (descriptively: argument vs. adjunct) of the VP-internal gap.

I adopt the simpler assumption that whatever rules out (ii) (one plausible candidate would be the constraint on remnant movement I derive in chapter 3) also rules out the purported extraposition derivation of (2.1c), and take the contrast between (2.1) and (2.2) to reveal a fact about VP constituency.

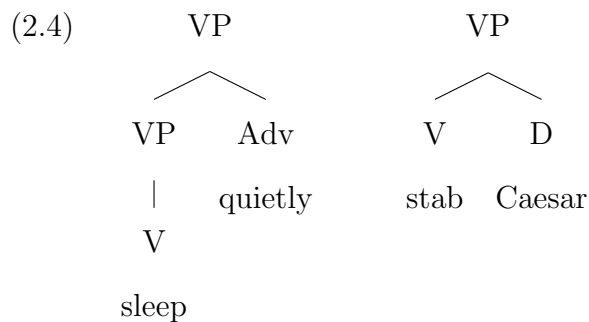
by the need to produce logical forms like those in (2.3) naturally account for syntactic distinctions such as the one illustrated in (2.1) and (2.2). This general approach is inspired by intuitions from Hornstein and Nunes (2008).

The rest of the chapter is organised as follows. In §2.2 I discuss in more detail the facts I aim to explain, sketched in (2.1) and (2.2), and look at how the relevant properties have been stipulated, but never adequately explained, in previous theories. In §2.3 I identify some non-trivial implications of the Conjunctivist hypothesis for our understanding of how semantic interpretation relates to syntactic structure, and in §2.4 I show that when the MG formalism is modified to accommodate the Conjunctivist hypothesis it permits a natural explanation for the facts discussed in §2.2. With some relatively conservative further assumptions, the system also reveals a simple explanation of the phenomenon of “counter-cyclic adjunction”, which I present in §2.6. Thus an independently-motivated hypothesis about how meanings are composed provides an insightful account of some otherwise puzzling syntactic facts.

2.2 Syntactic properties of arguments and adjuncts

2.2.1 Descriptive generalisations

The distinction between (2.1a) and (2.2a) has often been encoded by associating the relevant constituents with structures like those in (2.4), and requiring that the fronting/clefting operation illustrated in (2.1) and (2.2) apply to a node labelled VP.



This allows the relevant generalizations to be stated. But *prima facie* we simply have two VPs, each composed of a head verb and one other word, so we may still wonder *why* they differ in the way described by (2.4).³⁶

I have been deliberately agnostic about the details of the fronting/clefting operation that is applied in (2.1) and (2.2), and will continue to be, because the relevant facts about constituency are orthogonal to them. The very same pattern of judgments can be observed in the context of an ‘it’-cleft in (2.5) and VP-ellipsis and ‘do so’-substitution in (2.6). The relevant fragments are underlined.

- (2.5) a. i. Brutus [_{VP} slept quietly].
 ii. It is slept quietly that Brutus did.
 iii. It is slept that Brutus did quietly.
- b. i. Brutus [_{VP} stabbed Caesar].
 ii. It is stab Caesar that Brutus did.
 iii. * It is stab that Brutus did Caesar.
- (2.6) a. Brutus [_{VP} slept quietly] ...
 i. ... and Antony will (do so) (too). (slept quietly ellided)
 ii. ... and Antony will (do so) noisily. (slept ellided)
- b. Brutus [_{VP} stabbed Caesar] ...
 i. ... and Antony will (do so) (too). (stab Caesar ellided)
 ii. * ... and Antony will (do so) Cassius. (stab ellided)

My aim is not to provide new insight into the details of any of these particular operations, but rather to explain the fact that certain constituents seem to be “accessible”

³⁶I also assume that the two phrases ‘stab Caesar’ and ‘slept quietly’ are both projections of a lexical head verb to which ‘Caesar’ and ‘quietly’ attach in different ways, in line with traditional conceptions of adjunction rather than the “functional specifier” approach (Cinque 1999). My aim is to show how an independently-motivated theory of semantics should lead us to *expect* two modes of attachment *with these particular properties* to be exhibited by natural languages; not to entirely eliminate one, as Cinque does and as the categorial grammar *X/X* approach does. See §2.2.2 and §2.5.1 for some further discussion.

to the syntax in a sense that is independent of the mechanics of any particular operation that might apply to them.

This pattern of accessible and inaccessible constituents is not specific to the structure of verb phrases. Probing the structure of nominal phrases using ‘one’-substitution in (2.7) reveals the same pattern, with ‘tall’ acting parallel to ‘quietly’ and ‘of physics’ parallel to ‘Caesar’.

- (2.7) a. I taught this tall student ...
- i. ... and you taught that one. (tall student replaced)
 - ii. ... and you taught that short one. (student replaced)
- b. I taught this student of physics ...
- i. ... and you taught that one. (student of physics replaced)
 - ii. * ... and you taught that one of chemistry. (student replaced)

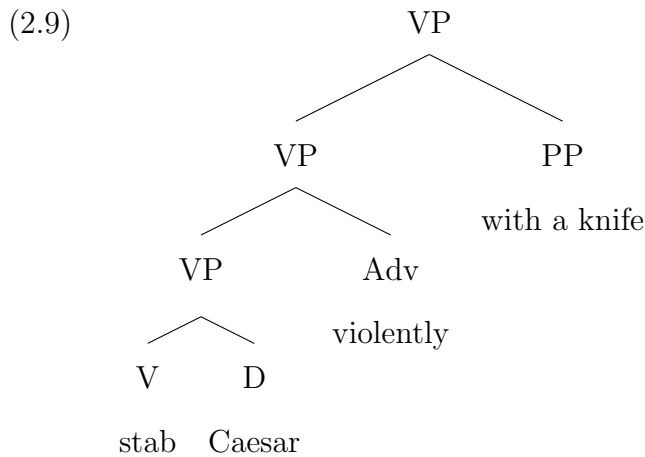
Descriptively, fragments like ‘tall’ and ‘quietly’ can be either included in or excluded from a phrase when it is targeted for some syntactic operation, whereas fragments like ‘Caesar’ and ‘of physics’ must be included. It is customary to label the former kind of fragment “adjuncts” and the latter kind “arguments”, but these labels are often assigned precisely on the basis of this syntactic behaviour.³⁷ Of course, it is possible that this is the best we can do, and the fact that there are these two ways for constituents *X* and *Y* to combine into a new constituent headed by *X* can not be made to follow naturally from anything else; but my claim here will be that we can, indeed, do better.

When a phrase containing both arguments and adjuncts is targeted by some syntactic operation, the generalisations noted above interact in the expected way: the argument(s) must be obligatorily included in the phrase, and each adjunct can

³⁷The data presented here are frequently given as *diagnostics* for argument-hood and adjunct-hood — that is, for determining which of the structures in (2.4) a phrase should be assigned — in syntax textbooks (Haegeman 1994, pp.88–91; Carnie 2007, pp.169–171).

independently be either included or excluded. This pattern is shown in (2.8) using the simple fronting construction; the reader can verify that the various constructions used above pattern together in these more complex cases as well.

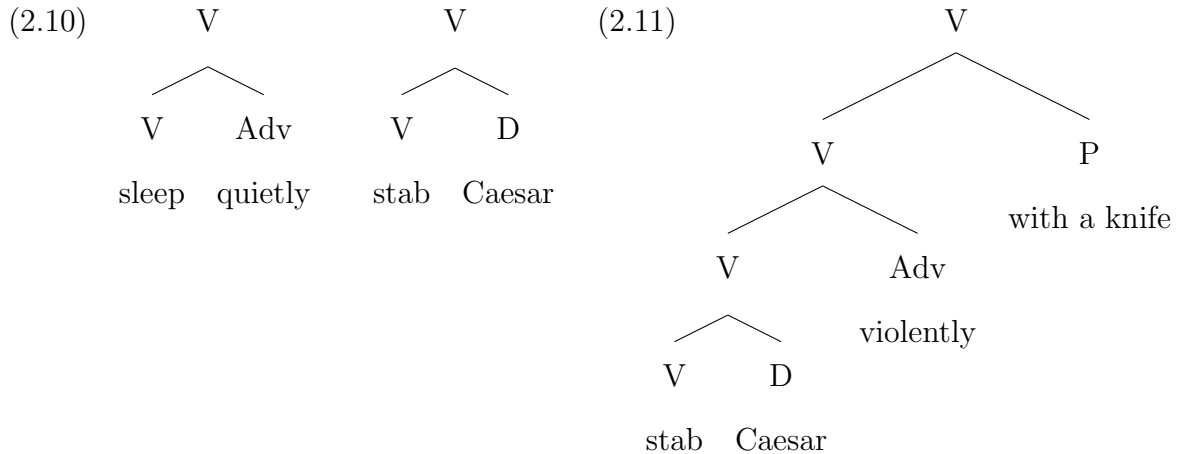
- (2.8) a. Brutus [_{VP} stabbed Caesar violently with a knife].
 b. Stab Caesar violently with a knife, (is what) Brutus did.
 c. Stab Caesar violently, (is what) Brutus did with a knife.
 d. Stab Caesar, (is what) Brutus did violently with a knife.
 e. *Stab, (is what) Brutus did Caesar violently with a knife.



The labelling convention in (2.4), and the requirement that (all and only) nodes labelled VP are available for syntactic manipulation, can be extended to capture the facts in (2.8) as illustrated in (2.9). But this remains more a description of the facts than an explanation of them. What motivation is there for treating some projections of the verb as a VP and others only as a V (recall that ‘sleep’ alone is a VP in (2.4)), other than the observed facts about what can be targeted and what can’t?

As Hornstein and Nunes (2008) point out, this question comes sharply into focus when we consider the recent trend towards a *relational* conception of X’ levels (Muysken 1982, Fukui 1986, Speas 1986, Chomsky 1995). On this “Bare Phrase Structure” (BPS) view we do not permit ourselves the use of non-branching nodes, or labels that provide any information beyond identifying the head constituent of a

complex one. Thus the syntactic structures in (2.4) are instead represented as in (2.10), and the larger one in (2.9) as in (2.11).



This spare conception of syntactic structure allows us to state all the necessary generalisations when only arguments — those fragments which, by definition, must obligatorily be included in the relevant phrase — are present, by defining a maximal projection as the highest projection of a given head. This correctly distinguishes between the two nodes labelled V in the structure of ‘stab Caesar’ in (2.10), only the higher one being maximal and therefore available for syntactic operations. But it incorrectly predicts that ‘sleep’ should be non-maximal and therefore unavailable, in just the same way that ‘stab’ is; and also incorrectly predicts that ‘stab Caesar’ and ‘stab Caesar violently’ should be non-maximal and unavailable in (2.11). The pattern of facts outlined above is a serious problem for this minimalist conception of phrase structure.³⁸ But whether one generally subscribes to BPS or not, the distinction in labelling encoded in (2.4) and (2.9) remains a stipulation that we should want to eliminate from our theory; BPS is just a methodological line of thought which very clearly brings this to our attention.

³⁸It is perhaps worth noting that the account of adjuncts in the original proposals of BPS (Chomsky 1994, 1995) is extremely problematic. Chomsky suggests a system where the upper two nodes labelled V in (2.11) would have the ordered pair ⟨V,V⟩ as their labels, identifying the V ‘stab’ as the head without rendering ‘stab Caesar’ non-maximal by sharing a label with it. But if labels such

To summarise, we have identified a sense in which constituents can be “accessible” for syntactic manipulation in a way that is independent of any particular construction or operation. My aim is to find a natural explanation for which constituents are accessible in this sense and which are not. The pattern to explain is that some non-head constituents, such as ‘Caesar’ in (2.2a), contribute crucially to making a projection maximal and therefore accessible; whereas others, such as ‘quietly’ in (2.1a), leave the “maximality” of a projection unchanged.

2.2.2 Adjuncts in the MG formalism

The difficulty in accounting for the distinction between arguments and adjuncts in any explanatory way carries over to the MG instantiation of minimalist syntax introduced in §1.5. To begin, let us assume, standardly, that the phrase ‘stab Caesar’ is constructed as shown in (2.12).

$$(2.12) \quad \frac{\frac{\langle \text{stab} : +d-V, \{\} \rangle \quad \langle \text{Caesar} : -d, \{\} \rangle}{\langle \text{stab} : +d-V, \{ \text{Caesar} : -d \} \rangle} \text{INS}}{\langle \text{stab Caesar} : -V, \{\} \rangle} \text{MRG}$$

The question then arises of how the phrase ‘sleep quietly’ is to be constructed. One possibility that can be immediately discounted is that it is constructed in essentially the same way as ‘stab Caesar’. Supposing that the only difference is in the category of feature checked — say, ‘a’ instead of ‘d’ — this possibility is shown in (2.13).

$$(2.13) \quad \frac{\frac{\langle \text{sleep} : +a-V, \{\} \rangle \quad \langle \text{quietly} : -a, \{\} \rangle}{\langle \text{sleep} : +a-V, \{ \text{quietly} : -a \} \rangle} \text{INS}}{\langle \text{sleep quietly} : -V, \{\} \rangle} \text{MRG}$$

It is worth briefly considering exactly what is right and what is wrong about this as $\langle V, V \rangle$ are permitted, then it is trivial to re-formulate the X-bar system of the GB era in these terms by using $\langle X, X \rangle$ in place of X' , $\langle X, X, X \rangle$ in place of X'' , and so on. Whatever the force of the Inclusiveness Condition that motivates BPS, it is generally thought to rule out the intrinsic X-bar levels used in earlier theories; and if it rules out that, it rules out Chomsky’s account of adjunction.

hypothesis. What is correct is that ‘stab Caesar’ and ‘sleep quietly’ have the same feature sequence (i.e. they are “of the same type”), and that their respective heads are ‘stab’ and ‘sleep’. The first problem with the hypothesis in (2.13) is that it *requires* that ‘sleep’ combine with exactly one thing of the same category as ‘quietly’, whereas we would like this to be optional and iterable. We would be forced to say while ‘sleep’ has features $+a-V$ in this case, it would have $-V$ when no adverb is present, $+a+a-V$ when two adverbs are present, $+a+a+a-V$ when three are present, and so on. The second problem with (2.13) is simply that it makes no structural distinction between ‘stab Caesar’ and ‘sleep quietly’, making it impossible to even state (let alone provide an insightful account of) the generalisations noted in §2.2.1.

An alternative is to suppose that instead of ‘sleep’ selecting ‘quietly’ as shown in (2.13), ‘quietly’ selects ‘sleep’. This is the direct equivalent of the standard treatment in categorial grammar, which would assign ‘quietly’ a category such as $VP \setminus VP$. This immediately runs into problems, since it establishes ‘sleep’ as the complement of ‘quietly’, as shown in (2.14).

$$(2.14) \quad \frac{\frac{\langle \text{quietly} : +V-V, \{\} \rangle \quad \langle \text{sleep} : -V, \{\} \rangle}{\langle \text{quietly} : +V-V, \{\text{sleep} : -V\} \rangle} \text{INS}}{\langle \text{quietly sleep} : -V, \{\} \rangle} \text{MRG}$$

This incorrectly establishes ‘quietly’ as the head of the phrase, and therefore derives the incorrect word order. (It is true that ‘quietly sleep’ is a possible verb phrase in English, but this account predicts that it should be the only possible word order in the same way that ‘stab Caesar’ is.) Even granting a stipulation that corrects for word order, the fact that the adverb is the head of the phrase remains. While (2.14) ensures that the *category* of the constituent to which ‘quietly’ attaches remains unchanged, the significance of a phrase’s head goes beyond just contributing its category.

The two approaches illustrated in (2.13) and (2.14) both have major problems, but these are the only options we have for the structure of ‘sleep quietly’ within the

MG formalism as it stands. The restrictiveness of the formalism exactly mirrors that observed in the discussion of BPS and the trees in (2.10): we have no way to establish that ‘sleep’ is the head of ‘sleep quietly’ and ‘stab’ is the head of ‘stab Caesar’ without assigning the two phrases identical structure. The system does not permit it at the moment, but we need there to be two distinct ways in which constituents X and Y can combine such that X provides the head of the new expression.

Frey and Gärtner (2002) present an addition to the MG formalism which satisfies this need. They introduce a new combinatory operation and a new kind of feature that is checked by this new operation. Returning temporarily to the tree-shaped MG expressions used in §1.5.1, this new operation, ADJOIN, works as shown in (2.15).

$$(2.15) \quad \text{ADJOIN}(\text{sleep} : -V, \text{quietly} : \approx V) = \begin{array}{c} < \\ \diagdown \quad \diagup \\ \text{sleep} : -V \quad \text{quietly} \end{array}$$

The $\approx V$ feature encodes the fact that ‘quietly’ must combine with something bearing a $-V$ feature; this much it has in common with a $+V$ feature. The ADJOIN function that checks a $\approx V$ feature, however, is defined such that it differs from the MRG function (that checks a $+V$ feature) in two ways: first, the $-V$ feature on the other involved constituent is not checked, only the $\approx V$ feature; and second, the constituent bearing the $-V$ feature contributes the head of the newly-formed constituent. As such, the $-V$ feature remains unchecked after ADJOIN applies in (2.15), and the head of the phrase is ‘sleep’.³⁹

³⁹Considering the Frey and Gärtner (2002) approach also makes clear the advantages of the INS/MRG system over the more standard version with separate E-MRG and I-MRG operations. In the INS/MRG system it follows immediately, once we have added the machinery to allow adjuncts to appear at once place in the derived structure, that they can appear at multiple places as well, by assigning feature sequences of the form $\approx f \approx g$. And of course we also have the possibility of sequences of the form $-f \approx g$, as befits some scrambled elements, and $\approx f -g$, as befits words like ‘why’ (more specifically, perhaps $\approx V -wh$). In contrast, Frey and Gärtner (2002) are forced to add *two* new operations that check adjunction features, the relationship between which is analogous to the relationship between E-MRG and I-MRG.

By adding new primitive operations and features like this one can, of course, formulate a descriptively adequate system. The stipulated differences between ADJOIN and MRG avoid the problems observed in (2.13) and (2.14), and one can easily add further stipulations about how this new machinery works: for example, Frey and Gärtner add an encoding of the adjunct island constraint by forbidding adjunction of constituents containing unchecked movement-triggering features, and Gärtner and Michaelis (2003) add the possibility of ADJOIN applying counter-cyclically as proposed by Lebeaux (1988, 2000). But these additions to the formalism, like the stipulative labelling conventions illustrated in (2.4) and (2.9), are more a restatement of the descriptive generalisations than an explanation of them. It remains an open question why it is *these* particular combinatory operations that our formalism needs. If such “one-sided” feature-checking operations are allowed, why is there no one-sided feature-checking operation where the constituent whose features *are checked* is the head? Why should it be the one-sided feature-checking operation ADJOIN that can apply counter-cyclically and prohibits extraction, rather than MRG? Must the ability to apply counter-cyclically correlate with a prohibition on extraction? Must any or all of these correlate with the possibility of being stranded in the manner discussed in §2.2.1? We could just as easily define an operation which has one of these properties and not the others. In short there seems to be no obvious reason why it is the ADJOIN operation, as defined, that is the one operation that needs to be added to the system alongside MRG.

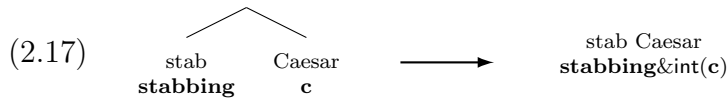
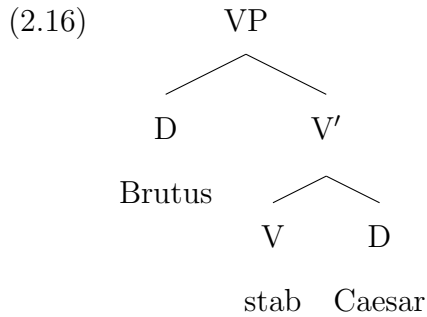
In §2.4 I will propose a different account of adjunction. Of course, this will necessarily involve adding *something* to the MG formalism, since we have seen in (2.13) and (2.14) that there is simply no room in the current formalism for the facts. My aim, however, is to develop a supplemented formalism which yields the essential properties of adjunction that Frey and Gärtner (2002) and Gärtner and Michaelis (2003) achieve by stipulation with the ADJOIN operation, and yet is independently motivated by a

particular view of the syntax-semantics interface that Conjunctivism forces upon us.

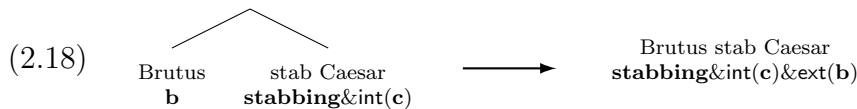
2.3 Syntactic consequences of Conjunctivism

Adopting the Conjunctivist view of semantic interpretation outlined in §1.6 has certain consequences for how we think of the relationship between syntactic derivations and semantic composition. Since we reject the common conception of semantic composition based on “saturation” or “valency”, syntactic relationships must take up some of the slack by playing a part in identifying how meanings of a sentence’s constituents are composed (eg. identifying whether a verb’s argument expresses an agent or a patient). The changes to our conception of syntactic derivations that are forced by this independently-motivated theory of semantics will yield a natural explanation for the puzzles posed in §2.2 about the behaviour of arguments and adjuncts. I will sketch these syntactic consequences with reference to informal tree diagrams here, and then carry the result over to the MG formalism in §2.4.

Suppose that semantic interpretation occurs at every merging step of a syntactic derivation, and that we compute the semantic value of a newly-formed constituent on the basis of only the semantic values of its immediate constituents (Bach 1976); so there is no need to explicitly represent the internal structure of a constituent that has been assigned an interpretation. Consider now the interpretation of the structure informally illustrated in (2.16). (For ease of exposition I leave aside the more articulated *vP* structures discussed in §1.6.3; these are dealt with in §2.A.) Semantic interpretation at the first merging step could proceed as shown in (2.17).



This first step suggests the general rule that when we combine two constituents denoting Φ and Ψ , where Φ is (the denotation of) the one that projects, the resulting constituent denotes $\Phi \& \text{int}(\Psi)$. But this is clearly not general enough: the next step is to combine the result of (2.17) with ‘Brutus’, and the resulting new constituent *should* be constructed using the *ext* operation, as shown in (2.18).



So the strictly incremental interpretation considered here is not tenable. The essence of the problem is that we “lose track of” how many arguments a head has already combined with, and therefore cannot correctly decide whether to apply *int* or *ext*. But this problem does not arise if we suppose that the “sliding window” through which our interpretative rules look is slightly larger. Specifically, it suffices to avoid interpretation of structures in which a head has combined with some, but not all, of its arguments — where, crucially, we understand an ‘argument’ to be defined as a constituent that requires semantic adjustment of the sort provided by *int* and *ext*. Then we would not perform the interpretive step shown in (2.17) — because at that stage ‘stab’ had combined with only one of its two arguments — and instead would maintain the syntactic structure until the entire VP was complete, only then interpreting (and discarding structure), as shown in Figure 21.⁴⁰

⁴⁰Recall footnote 12 (on page 21): I have been implicitly assuming a distinction between internal and external arguments in all derivations so far for the purposes of phonological, rather than

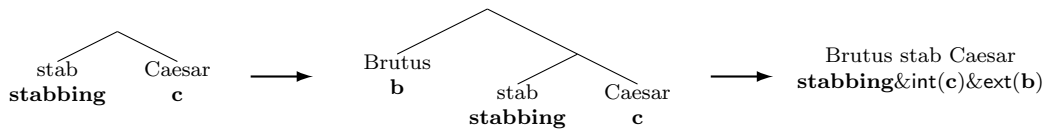


Figure 21

To put this another way, we delay interpretation so that the relationships that are not, on the Conjunctivist view, encoded in lexical semantics, can instead be encoded via syntactic relations. (Note that this delay of interpretation, which I will claim turns out to be desirable, would *not* be forced by adopting more standard theories of semantic composition based on function application.) But while this work of establishing particular syntactic relations is crucial in the case of arguments, it is unnecessary in the case of adjuncts. The reason we must establish internal and external argument configurations is that **c**, for example, is not conjoinable with **stabbing**; **int(c)**, however, is (and the necessity of distinguishing it from **ext(c)** is what forces the delay). But an event predicate like **violent** does not need this “help” in order to combine conjunctively, and so ‘violently’ is less restricted than ‘Caesar’ is in the way it can participate in the syntactic derivation.

To illustrate more clearly what sort of flexibility adjuncts like ‘violently’ have, we must consider the composition of the VP with the rest of the TP projection. I will have more to say about the semantics of the TP projection in chapter 4, but for now let us just make some simple assumptions in order to proceed. I assume that the occurrence of the subject in the specifier position of TP for Case reasons

semantic, interpretation: external arguments have been placed to the left of the head, and internal arguments to the right. One might therefore worry that the proposed delay in interpretation is actually independently necessary for phonological purposes, and that all the results that follow from this could be explained on similar logic without invoking Conjunctivism. Linearisation concerns alone, however, will not always delay interpretation until the end of a maximal projection. First, in head-final languages all arguments appear to the left, so incremental interpretation at every MRG step will permit correct linearisation. Second, even in languages like English it is generally assumed that all arguments after the first one (i.e. all specifiers) are linearized on the left, which would allow incremental interpretation at all points except the very first intermediate projection.

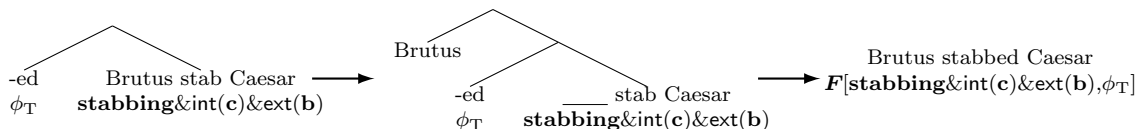


Figure 22

is semantically vacuous. As a result, the meaning computed at the end of the TP projection is a function of (only) the VP’s meaning and the T head’s meaning; for now I will simply write ‘ $F[\phi_{VP}, \phi_T]$ ’ for the semantic interpretation of a TP containing a T head with meaning ϕ_T and a VP with meaning ϕ_{VP} .⁴¹ Then we can illustrate the composition and interpretation of the TP intuitively as in Figure 22.

Now, consider the semantic composition of an adverb such as ‘violently’ under these assumptions. The desired effect is to add a conjunct to the event predicate constructed in the VP, and so the meaning computed at the completion of the TP projection should be the one shown in (2.19).

- (2.19) a. Brutus stabbed Caesar violently
 b. $F[\text{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b}) \ \& \ \mathbf{violent}, \phi_T]$

One relatively straightforward way to go about constructing this meaning is to suppose that the additional conjunct, **violent**, is involved in build-and-interpret sequence shown in Figure 21 — where the details of this “involvement” remain to be made clear, but it must be somehow different from the way in which the two arguments’ meanings are involved. Then the interpretation of the TP will have exactly the same form as Figure 22, just with a slightly “larger” meaning being contributed

⁴¹Note that ‘ F ’ is *not* a symbol of the language in which meanings are expressed in the way that ‘&’ and ‘int’ and ‘ext’ are, just a place-holder for whatever turns out to be the appropriate “two-hole context” when we consider the proper interpretation of the TP projection.

by the VP. This possibility can be represented as in (2.20).

$$\begin{aligned}
 (2.20) \quad \phi_{VP} &= \mathbf{stabbing} \ \& \ \mathbf{int(c)} \ \& \ \mathbf{ext(b)} \ \& \ \mathbf{violent} \\
 \phi_{TP} &= \mathbf{F}[\phi_{VP}, \phi_T] \\
 &= \mathbf{F}[\mathbf{stabbing} \ \& \ \mathbf{int(c)} \ \& \ \mathbf{ext(b)} \ \& \ \mathbf{violent}, \phi_T]
 \end{aligned}$$

Under the analysis I will present, the flexibility that an adjunct like ‘violently’ has will be to contribute to the final meaning in (2.19) either as indicated in (2.20) *or* via another alternative. The second alternative is that **violent** is added as a conjunct to **stabbing & int(c) & ext(b)** not at the interpretive step illustrated in Figure 21 but rather at the interpretive step illustrated in Figure 22. To illustrate the way in which this second possibility differs from the first, consider the way in which (2.21) differs from (2.20).

$$\begin{aligned}
 (2.21) \quad \phi_{VP} &= \mathbf{stabbing} \ \& \ \mathbf{int(c)} \ \& \ \mathbf{ext(b)} \\
 \phi_{TP} &= \mathbf{F}[\phi_{VP} \ \& \ \mathbf{violent}, \phi_T] \\
 &= \mathbf{F}[\mathbf{stabbing} \ \& \ \mathbf{int(c)} \ \& \ \mathbf{ext(b)} \ \& \ \mathbf{violent}, \phi_T]
 \end{aligned}$$

Note that this does not eliminate or affect any accepted differences between adjuncts that “modify the VP” and those that “modify the TP”. The proposal is *not* simply that ‘violently’ is flexible between these two options. If ‘violently’ were to modify the TP, the resulting meaning would be $\mathbf{F}[\phi_{VP}, \phi_T] \ \& \ \mathbf{violent}$. Instead, I am proposing that there are two different ways in which semantic modification of the VP can be achieved.

Crucially, this flexibility that ‘violently’ has is *not* available for arguments. An argument like ‘Caesar’ does *not* have the possibility of adding its conjunct at the interpretive step shown in Figure 22 because the particular conjunct contributed by ‘Caesar’ is dependent on its relative position within the structure of the VP. At the

interpretive step in Figure 22, the internal structure of the VP has been lost, and so there is no longer the possibility of establishing internal or external argument relationships. An adjunct like ‘violently’, however, doesn’t need to be placed in any particular structural position inside the VP relative to anything else; a less intricate relation, the details of which will be made precise, will suffice. Roughly speaking: it is not possible to appear as an internal or external argument of the VP in Figure 22, but it is possible to appear as an adjunct to it.

The contrast in (2.1) and (2.2), repeated here, will thus be explained by a sort of structural ambiguity, though *not one that produces any difference in meaning* (recall that the same meaning is composed in both (2.20) and (2.21)). There are two distinct derivations of (2.1a), one of which makes ‘sleep quietly’ available for manipulation, the other just ‘sleep’; but there is only one derivation of (2.2a), which makes ‘stab Caesar’ available.

- (2.1) a. Brutus [_{VP} slept quietly].
 b. Sleep quietly, (is what) Brutus did.
 c. Sleep, (is what) Brutus did quietly.
- (2.2) a. Brutus [_{VP} stabbed Caesar].
 b. Stab Caesar, (is what) Brutus did.
 c. * Stab, (is what) Brutus did Caesar.

Having sketched the relevant intuitions, I now turn to fleshing out these ideas in the MG formalism.

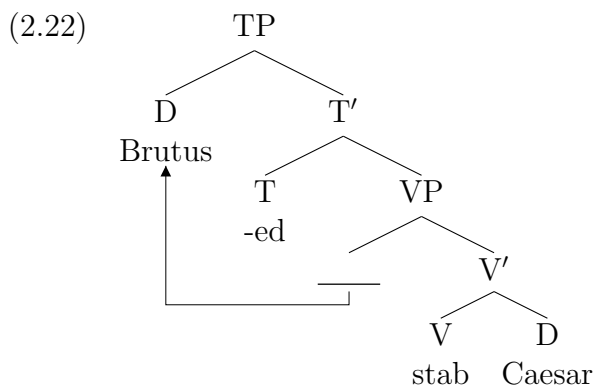
2.4 Conjunctivist interpretation of MG derivations

In this section I will develop a slightly modified version of the MG formalism introduced in §1.5, motivated by the requirement that semantic interpretation proceeds according to the Conjunctivist proposal introduced in §1.6, such that the puzzling

pattern of facts discussed in §2.2 is naturally accounted for. After briefly introducing the relevant stage-setting syntactic assumptions in §2.4.1, I present modifications to deal with arguments in §2.4.2 and then minimal further modifications to deal with adjuncts in §2.4.3.

2.4.1 Getting started

To begin, Figure 23 shows the derivation of the TP part of a simple transitive sentence. Where it improves clarity, I represent a phonetic gap with ‘____’. The informal tree corresponding to the derivation in Figure 23 is given in (2.22).⁴²



Suppose for concreteness that the informal structure in (2.23), for some functional head X, can serve as a model of the structure of the VP-fronting sentence ‘Stab Caesar, Brutus did’. A more thorough account of this construction will almost certainly include further projections in between X and TP, but that will not affect our discussion here. As mentioned in §2.2, under investigation is the general sense in which some constituents are accessible for syntactic manipulation (in some way that transcends particular constructions/transformations) whereas others are not. Setting ourselves

⁴²Although I have not given an account of head movement here, I write the last two lines of the derivation in Figure 23 showing its effects for readability. A more “honest” representation of what I have derived would show the following instead:

$$\begin{aligned}
 e_6 = \text{MRG}(e_5) &= \langle \text{-ed} \text{ ____ } \text{stab Caesar} : +\mathbf{k}-\mathbf{t}, \{ \text{Brutus} : -\mathbf{k} \} \rangle \\
 e_7 = \text{MRG}(e_6) &= \langle \text{Brutus -ed} \text{ ____ } \text{stab Caesar} : -\mathbf{t}, \{ \} \rangle
 \end{aligned}$$

Lexicon:

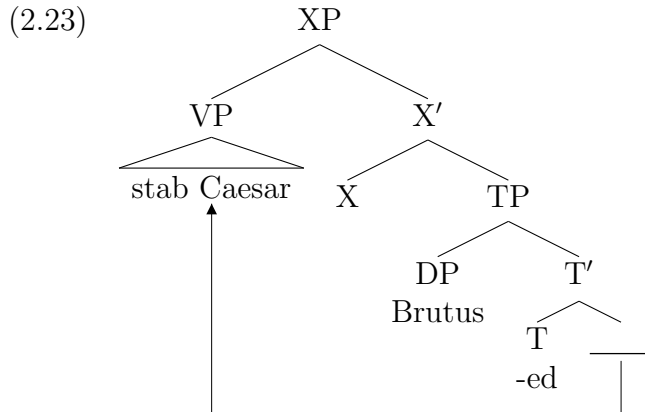
$\ell_{\text{Brutus}} = \langle \text{Brutus: -d-k}, \{\} \rangle$	$\ell_{\text{stab}} = \langle \text{stab: +d+d-V}, \{\} \rangle$
$\ell_{\text{Caesar}} = \langle \text{Caesar: -d}, \{\} \rangle$	$\ell_{\text{T}} = \langle \text{-ed: +V+k-t}, \{\} \rangle$

Derivation:

$e_1 = \text{INS}(\ell_{\text{stab}}, \ell_{\text{Caesar}})$	$= \langle \text{stab: +d+d-V}, \{ \text{Caesar: -d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{stab Caesar: +d-V}, \{\} \rangle$
$e_3 = \text{INS}(e_2, \ell_{\text{Brutus}})$	$= \langle \text{stab Caesar: +d-V}, \{ \text{Brutus: -d-k} \} \rangle$
$e_4 = \text{MRG}(e_3)$	$= \langle \text{___ stab Caesar: -V}, \{ \text{Brutus: -k} \} \rangle$
$e_5 = \text{INS}(\ell_{\text{T}}, e_4)$	$= \langle \text{-ed: +V+k-t}, \{ \text{___ stab Caesar: -V}, \text{Brutus: -k} \} \rangle$
$e_6 = \text{MRG}(e_5)$	$= \langle \text{stabbed Caesar: +k-t}, \{ \text{Brutus: -k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \text{Brutus stabbed Caesar: -t}, \{\} \rangle$

Figure 23

the modest goal of moving ‘stab Caesar’ to the specifier of X is enough to investigate this general notion of accessibility.



This structure is derived as shown in Figure 24. The derivation involves two lexical items not seen in Figure 23. First, there is a variant of the ℓ_{stab} lexical expression from before, call it ℓ'_{stab} , which has an additional $-f$ feature; this is the feature that is checked when the VP merges as the specifier of the functional head X. The other new lexical item, ℓ_X , represents this X head; its feature sequence $+t+f-x$ indicates that

Lexicon:

$\ell_{\text{Brutus}} = \langle \text{Brutus} : -\text{d}-\text{k}, \{\} \rangle$	$\ell'_{\text{stab}} = \langle \text{stab} : +\text{d}+\text{d}-\text{V}-\text{f}, \{\} \rangle$
$\ell_{\text{Caesar}} = \langle \text{Caesar} : -\text{d}, \{\} \rangle$	$\ell_{\text{T}} = \langle -\text{ed} : +\text{V}+\text{k}-\text{t}, \{\} \rangle$
	$\ell_{\text{X}} = \langle \epsilon_{\text{X}} : +\text{t}+\text{f}-\text{x}, \{\} \rangle$

Derivation:

$e_1 = \text{INS}(\ell'_{\text{stab}}, \ell_{\text{Caesar}})$	$= \langle \text{stab} : +\text{d}+\text{d}-\text{V}-\text{f}, \{ \text{Caesar} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{stab Caesar} : +\text{d}-\text{V}-\text{f}, \{\} \rangle$
$e_3 = \text{INS}(e_2, \ell_{\text{Brutus}})$	$= \langle \text{stab Caesar} : +\text{d}-\text{V}-\text{f}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_4 = \text{MRG}(e_3)$	$= \langle ___ \text{stab Caesar} : -\text{V}-\text{f}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_5 = \text{INS}(\ell_{\text{T}}, e_4)$	$= \langle -\text{ed} : +\text{V}+\text{k}-\text{t}, \{ \text{stab Caesar} : -\text{V}-\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_6 = \text{MRG}(e_5)$	$= \langle -\text{ed} ___ : +\text{k}-\text{t}, \{ \text{stab Caesar} : -\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \text{Brutus -ed} ___ : -\text{t}, \{ \text{stab Caesar} : -\text{f} \} \rangle$
$e_8 = \text{INS}(e_7, \ell_{\text{X}})$	$= \langle \epsilon_{\text{X}} : +\text{t}+\text{f}-\text{x}, \{ \text{Brutus -ed} ___ : -\text{t}, \text{stab Caesar} : -\text{f} \} \rangle$
$e_9 = \text{MRG}(e_8)$	$= \langle \epsilon_{\text{X}} \text{ Brutus -ed} ___ : +\text{f}-\text{x}, \{ \text{stab Caesar} : -\text{f} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle \text{stab Caesar } \epsilon_{\text{X}} \text{ Brutus -ed} ___ : -\text{x}, \{\} \rangle$

Figure 24

it combines with a TP complement and then with something bearing a $-\text{f}$ feature.⁴³ (Although it is not represented in (2.23), the subject moves out of the VP before the fronting occurs.)

⁴³Representing the $-\text{f}$ features explicitly and presenting the expressions listed at the top of Figure 24 as a “lexicon” may give the impression that there is some inelegant redundancy, with ‘stab’ effectively “represented twice”. But on the assumption that the need for feature-checking drives every merging step, these features must be present, though presumably it is uncontroversial that at some level — though perhaps not a level relevant to the mechanics of derivations — features like the $-\text{V}$ on ‘stab’ are different in kind from features like the $-\text{f}$, only the former being an intrinsic property that will be there every time ‘stab’ is used in any derivation. Chomsky (1995, p.231) discusses this distinction between “intrinsic features” and “optional features”. The expressions that the derivation “begins with” must be able to bear such optional features, as shown in Figure 24, though no doubt we would like ℓ_{stab} and ℓ'_{stab} to be in some sense variants of the one thing. In general, one can assume that the optional features are nondeterministically added as a lexical item enters the derivation. The current framework’s distinction between units and expressions permits a perhaps more elegant option: a “lexicon” is a set of *units*, and the set of “lexical expressions” corresponding to a lexicon U contains, for each $u \in U$, both $\langle u, \{\} \rangle$ and $\langle u', \{\} \rangle$, where u' is the unit just like u but with $-\text{f}$ added at the end of its feature sequence. Then the lists at the top of Figure 23 and Figure 24 would show the “lexical expressions” relevant for their respective derivations, derived from a “lexicon” in which the only record of ‘stab’ is the unit $\text{stab} : +\text{d}+\text{d}-\text{V}$. However these details are

We would now like to supplement the MG formalism with semantics. As things stand, these grammars derive strings without associated meanings. The obvious modification to make to include semantic interpretation is to say that a unit is not just a string with an associated sequence of features, but rather a string-meaning pair with an associated sequence of features. So whereas we have until now thought of $\text{stab} : +d+d-V$ as a unit, we will now take $\langle \text{stab}, \text{stabbing} \rangle : +d+d-V$ to be a unit instead, where **stabbing** is a monadic event predicate as discussed in §1.6. It will be convenient to write these units as $\left[\begin{smallmatrix} \text{stab} \\ \text{stabbing} \end{smallmatrix} \right] : +d+d-V$.

The next step is to say how these meanings are composed in the course of MG derivations. When MRG applies and two units are combined into one new unit, what meaning should the new unit have? This will lead us to the idea of slightly delayed interpretation as discussed in §2.3, but the logic outlined there can now be rehearsed in a more precise framework.

Consider the first two MRG steps from the derivation in Figure 23, rewritten here as (2.24).

$$\begin{array}{l}
 (2.24) \quad e_2 = \text{MRG} (\langle \text{stab} : +d+d-V, \{ \text{Caesar} : -d \} \rangle) \\
 \quad \quad = \langle \text{stab Caesar} : +d-V, \{ \} \rangle \\
 e_4 = \text{MRG} (\langle \text{stab Caesar} : +d-V, \{ \text{Brutus} : -d-k \} \rangle) \\
 \quad \quad = \langle \text{_____ stab Caesar} : -v, \{ \text{Brutus} : -k \} \rangle
 \end{array}$$

The desired semantic effects of these two steps are to add the $\text{int}(\mathbf{c})$ and $\text{ext}(\mathbf{b})$ conjuncts to the “head predicate” **stabbing**. (Although the subject does not contribute to the phonological yield of the VP, it does undergo thematic interpretation here.) So, writing the meaning components as well as string components of units now, we would *like* these two steps to proceed as shown in (2.25).

spelled out though, the fact that I am explicitly representing features like $-f$ should not be mistaken for an indication that there is any more redundancy here than is implicit elsewhere.

$$\begin{aligned}
(2.25) \quad e_2 &= \text{MRG} \left(\left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right] : +\mathbf{d}+\mathbf{d}-\mathbf{V}, \left\{ \left[\begin{array}{c} \text{Caesar} \\ \mathbf{c} \end{array} \right] : -\mathbf{d} \right\} \right\rangle \right) \\
&= \left\langle \left[\begin{array}{c} \text{stab Caesar} \\ \text{stabbing\&int(c)} \end{array} \right] : +\mathbf{d}-\mathbf{V}, \left\{ \right\} \right\rangle \\
e_4 &= \text{MRG} \left(\left\langle \left[\begin{array}{c} \text{stab Caesar} \\ \text{stabbing\&int(c)} \end{array} \right] : +\mathbf{d}-\mathbf{V}, \left\{ \left[\begin{array}{c} \text{Brutus} \\ \mathbf{b} \end{array} \right] : -\mathbf{d}-\mathbf{k} \right\} \right\rangle \right) \\
&= \left\langle \left[\begin{array}{c} \text{stab Caesar} \\ \text{stabbing\&int(c)\&ext(b)} \end{array} \right] : -\mathbf{V}, \left\{ \left[\begin{array}{c} \text{Brutus} \\ \mathbf{b} \end{array} \right] : -\mathbf{k} \right\} \right\rangle
\end{aligned}$$

These two steps reflect what was intuitively represented in (2.17) and (2.18). But as discussed there, it is not possible to define a single rule for the semantic effects of the MRG operation that will give the desired results. Sometimes we need ‘int’ to be applied to the selected meaning, and sometimes ‘ext’ — but the expression to which MRG applies in the second step of (2.25), checking the features of ‘Brutus’, has no more internal structure than the one to which it applies in the first step. The first step seems to require the general rule in (2.26) but the second step seems to require (2.27) instead.

$$(2.26) \quad \text{MRG} \left(\left\langle \left[\begin{array}{c} \dots \\ \phi \end{array} \right] : +f \dots, \left\{ \left[\begin{array}{c} \dots \\ \psi \end{array} \right] : -f, \dots \right\} \right\rangle \right) = \left\langle \left[\begin{array}{c} \dots \\ \phi \&\text{int}(\psi) \end{array} \right] : \dots, \left\{ \dots \right\} \right\rangle$$

$$(2.27) \quad \text{MRG} \left(\left\langle \left[\begin{array}{c} \dots \\ \phi \end{array} \right] : +f \dots, \left\{ \left[\begin{array}{c} \dots \\ \psi \end{array} \right] : -f, \dots \right\} \right\rangle \right) = \left\langle \left[\begin{array}{c} \dots \\ \phi \&\text{ext}(\psi) \end{array} \right] : \dots, \left\{ \dots \right\} \right\rangle$$

Since the neither of the two expressions to which MRG applies in (2.25) has any more or less structure than the other, we “lose track of” how many arguments a head has already combined with, and therefore cannot correctly decide whether to apply int or ext.

One may be tempted to think that we should therefore go back to the “un-reduced” MG representations from §1.5.1, so that the semantic correlate of MRG can be sensitive to the entire structure tree structure of expressions. This would, of course, be sufficient to resolve this problem, but it is not necessary: it suffices to delay interpretation just slightly, specifically until the end of each maximal projection, as shown in Figure 21. This makes visible all the necessary syntactic configurations, since it is only relative positions *within* a single maximal projection that the semantic compo-

sition rules must be sensitive to. In §2.4.2 I propose some modifications to the MG formalism which implement this idea.

2.4.2 Interpretation of arguments

The basic idea is to split the work currently done by MRG into two pieces. As things stand, an application of MRG has two effects: checking/deleting of features, and interpretation of the resulting structure/relations. We can separate the work into these two parts and have a function for each. I re-use the name MRG for the new function that checks features, and use the name SPL for the function that interprets structure that has already been established.

The general form of an MG expression must also be adjusted: whereas up until now the general form of an expression has been (2.28a), where x and the y_i are units, we now include in addition an ordered list of arguments, as shown in (2.28b).

$$(2.28) \quad \begin{array}{l} \text{a. } \langle x, \{y_1, y_2, \dots, y_n\} \rangle \\ \text{b. } \langle x, a_1, a_2, \dots, a_m, \{y_1, y_2, \dots, y_n\} \rangle \end{array}$$

Here a_1 is the complement of x , and the other a_i are specifiers. (I permit arbitrarily many specifiers, although I will not make use of more than two). When MRG applies to an expression of the form in (2.28b), the result has one of the y_i (not fully composed with the head x to form a new unit, but rather) appended to the list of arguments; when SPL applies to such an expression, x and the a_i are composed into a new unit.

For example, whereas previously MRG worked as shown in (2.24), repeated here, now MRG *only* checks features — and intuitively, “builds tree structure” — without performing any semantic composition. The effect of the new MRG operation are illustrated in (2.29).

$$\begin{aligned}
(2.24) \quad e_2 &= \text{MRG} (\langle \langle \text{stab} : +d+d-V, \{ \text{Caesar} : -d \} \rangle \rangle) \\
&= \langle \text{stab Caesar} : +d-V, \{ \} \rangle \\
e_4 &= \text{MRG} (\langle \langle \text{stab Caesar} : +d-V, \{ \text{Brutus} : -d-k \} \rangle \rangle) \\
&= \langle \text{stab Caesar} : -v, \{ \text{Brutus} : -k \} \rangle
\end{aligned}$$

$$\begin{aligned}
(2.29) \quad e_2 &= \text{MRG} (\langle \langle [\text{stab}_{\text{stabbing}}] : +d+d-V, \{ [\text{Caesar}_c] : -d \} \rangle \rangle) \\
&= \langle [\text{stab}_{\text{stabbing}}] : +d-V, [\text{Caesar}_c] : d, \{ \} \rangle \\
e_4 &= \text{MRG} (\langle \langle [\text{stab}_{\text{stabbing}}] : +d-V, [\text{Caesar}_c] : d, \{ [\text{Brutus}_b] : -d-k \} \rangle \rangle) \\
&= \langle [\text{stab}_{\text{stabbing}}] : -V, [\text{Caesar}_c] : d, [\text{Brutus}_b] : d, \{ [\text{Brutus}_b] : -k \} \rangle
\end{aligned}$$

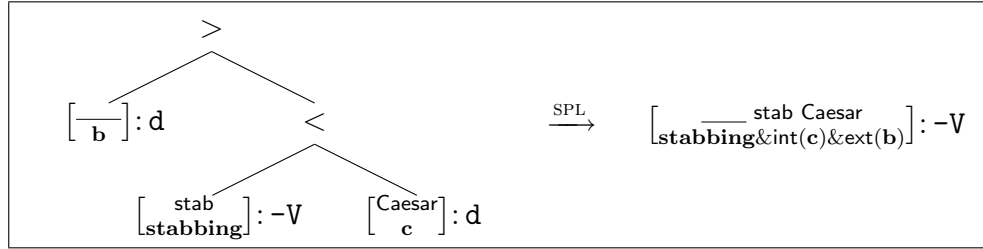
Note that in the expression e_2 that is produced by the first application of MRG, $+d$ and $-d$ features have been checked, but no composition or interpretation has occurred: $[\text{stab}_{\text{stabbing}}]$ and $[\text{Caesar}_c]$ are still separate, though their configuration in e_2 encodes the fact that a head-complement relationship has been established between them, as a result of the feature checking. We record with $[\text{Caesar}_c]$ the *category* d of the features that were checked, but the d annotation is just an inert record of this, not an attractor or attractee feature that can drive further structure-building; the use of this annotation will become clear when we deal with adjuncts.⁴⁴

The three units that make up the VP projection are semantically (and phonologically) composed, *as appropriate for the structural relations established among them*, when SPL is applied to the expression e_4 , as shown in Figure 25a. Intuitively, we can think of e_4 , the expression to which SPL is applied, as a representation of a small “window” of tree structure, which is then interpreted by SPL and transformed into an unstructured unit, as shown in Figure 25b; compare with Figure 21. In these small tree⁴⁵ structures, the head (indicated by the ‘<’ and ‘>’ labels) is the only unit in the tree with features present; these features are the ones inherited by the unit created

⁴⁴These inert category annotations, eg. d on $[\text{Caesar}_c]$ in (2.29), should not be confused with the *features* used in the original MG formalism discussed in §1.5.1 before merge steps and move steps were unified into (I-)MRG, that were likewise not prefixed with ‘+’ or ‘-’ (or ‘=’).

$$\begin{aligned}
\text{SPL}(e_4) &= \text{SPL} \left(\left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right] : -V, \left[\begin{array}{c} \text{Caesar} \\ c \end{array} \right] : d, \left[\begin{array}{c} \text{---} \\ b \end{array} \right] : d, \{ \left[\text{Brutus} \right] : -k \} \right\rangle \right) \\
&= \left\langle \left[\begin{array}{c} \text{---} \text{stab Caesar} \\ \text{stabbing\&int(c)\&ext(b)} \end{array} \right] : -V, \{ \left[\text{Brutus} \right] : -k \} \right\rangle
\end{aligned}$$

(a)



(b)

Figure 25: One formal and one intuitive illustration of the application of SPL to interpret the VP projection

upon application of SPL (i.e. the $-V$ in this example).

More generally, rather than either of the two incorrect rules for semantic composition in (2.26) and (2.27), we now have the SPL rule in (2.30) for interpreting argument relations in a Conjectivist manner (restricting attention to the two-argument case, for simplicity).

$$\begin{aligned}
(2.30) \quad \text{SPL} \left(\left\langle \left[\begin{array}{c} s_0 \\ \phi \end{array} \right] : -f \dots, \left[\begin{array}{c} s_1 \\ \psi_1 \end{array} \right], \left[\begin{array}{c} s_2 \\ \psi_2 \end{array} \right], \{ \dots \} \right\rangle \right) \\
= \left\langle \left[\begin{array}{c} s_2 s_0 s_1 \\ \phi \&\text{int}(\psi_1) \&\text{ext}(\psi_2) \end{array} \right] : -f \dots, \{ \dots \} \right\rangle
\end{aligned}$$

With these changes in place, a simple VP-fronting sentence (previously derived in Figure 24) is now derived as shown in Figure 26. The lexicon is unchanged; it is as in

⁴⁵Of course technically these are *not* tree structures, but merely lists, as the true representations in (2.29) and Figure 25a make clear. To illustrate the correspondence with standard conceptions of phrase structure, such a list $\langle x, a_1, a_2, \dots, a_n \rangle$ can be taken to represent the tree structure headed by x , with a_1 its complement, a_2 its first specifier and so on. But the new expression form $\langle x, a_1, a_2, \dots, a_m, \{y_1, y_2, \dots, y_n\} \rangle$ does not let us represent arbitrary trees, only strictly right-branching ones with a certain headedness pattern. This limitation forces us to apply SPL at the end of each maximal projection.

$e_1 = \text{INS}(\ell'_{\text{stab}}, \ell_{\text{Caesar}})$	$= \langle \text{stab} : +d+d-V-f, \{ \text{Caesar} : -d \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{stab} : +d-V-f, \text{Caesar} : d, \{ \} \rangle$
$e_3 = \text{INS}(e_2, \ell_{\text{Brutus}})$	$= \langle \text{stab} : +d-V-f, \text{Caesar} : d, \{ \text{Brutus} : -d-k \} \rangle$
$e_4 = \text{MRG}(e_3)$	$= \langle \text{stab} : -V-f, \text{Caesar} : d, \text{---} : d, \{ \text{Brutus} : -k \} \rangle$
$e_5 = \text{SPL}(e_4)$	$= \langle \left[\text{---}^{\text{stab Caesar}}_{\text{stabbing\&int(c)\&ext(b)}} \right] : -V-f, \{ \text{Brutus} : -k \} \rangle$
$e_6 = \text{INS}(\ell_{\text{T}}, e_5)$	$= \langle -ed : +V+k-t, \{ \text{stab Caesar} : -V-f, \text{Brutus} : -k \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle -ed : +k-t, \text{---} : V, \{ \text{stab Caesar} : -f, \text{Brutus} : -k \} \rangle$
$e_8 = \text{MRG}(e_7)$	$= \langle -ed : -t, \text{---} : V, \text{Brutus} : k, \{ \text{stab Caesar} : -f \} \rangle$
$e_9 = \text{SPL}(e_8)$	$= \langle \left[\text{Brutus -ed ---}^{\text{Brutus -ed}}_{\text{F[stabbing\&int(c)\&ext(b),\phi_{\text{T}}]}} \right] : -t, \{ \text{stab Caesar} : -f \} \rangle$
$e_{10} = \text{INS}(e_9, \ell_{\text{X}})$	$= \langle \epsilon_x : +t+f-x, \{ \text{Brutus -ed ---} : -t, \text{stab Caesar} : -f \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle \epsilon_x : +f-x, \text{Brutus -ed ---} : t, \{ \text{stab Caesar} : -f \} \rangle$
$e_{12} = \text{MRG}(e_{11})$	$= \langle \epsilon_x : -x, \text{Brutus -ed ---} : t, \text{stab Caesar} : f, \{ \} \rangle$
$e_{13} = \text{SPL}(e_{12})$	$= \langle \text{stab Caesar } \epsilon_x \text{ Brutus -ed ---} : -x, \{ \} \rangle$

Figure 26

Figure 24. The semantic components are included only in the results of the significant applications of SPL to reduce clutter. I have nothing to say about the semantic effects of the fronting operation.

I will use the term “phase” for a (maximal) sequence of derivational steps that does not include an application of SPL.⁴⁶ Then the INS and MRG steps that produced the expression e_4 in Figure 26, for example, amount to one phase; phases are in fact exactly maximal projections. An intuition that emerges is that a phase serves to establish the various conjuncts that are to be applied to a particular variable (in this case, an event variable): the VP phase “sets up” the three units shown in the tree in

Figure 25b in such a way that they will each contribute one of the three predicates in (2.31); recall (1.22) on page 35. The conjunction of these is the meaning assigned to the new unit produced by the application of SPL that ends the phase.

$$(2.31) \quad \mathbf{stabbing} = \lambda e.\mathbf{stabbing}(e)$$

$$\mathbf{int}(\mathbf{c}) = \lambda e.\exists x[\mathbf{c}(x) \wedge \mathbf{Internal}(e, x)]$$

$$\mathbf{ext}(\mathbf{b}) = \lambda e.\exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)]$$

This intuition will be useful in understanding the proposed treatment of adjuncts that follows. I will show that it permits adjuncts a kind of freedom that arguments do not have, as mentioned at the outset. It may not be clear at this point that arguments are in any sense “unusually restricted” in their syntactic behaviour, since all we have shown is a system where they are obligatorily included in a moved phrase, as was the case in even the unmodified MG framework; the addition of SPL steps did not interact in any interesting way with the VP-fronting mechanisms in Figure 26. But the changes I have introduced make way for a system where it naturally emerges that adjuncts have a degree of freedom that arguments do not.

⁴⁶Note, however, that the result of applying SPL to an expression is just another expression, which can then participate in a derivation in exactly the same way that other expressions whose derivation does not involve SPL can. There is no sense in which, as a result of applying SPL, anything “goes anywhere” or is “sent away”. All three operations (INS, MRG and SPL) are just functions mapping tuples of expressions (which have certain phonological, syntactic and semantic properties) to new expressions (which have certain other phonological, syntactic and semantic properties).

Readers who would like to reserve the term “phase” for, say, *vP* and CP, or whatever nodes turn out to have the properties often thought to be shared by these (Chomsky 2001, Fox and Pesetsky 2005), are free to choose their own term and substitute it as appropriate; the same goes for my use of the term “spell out”. The notion of phase which I adopt is grounded in certain assumptions about interpretation at the interfaces, in the style of Uriagereka (1998) but with respect to semantic rather than phonological interpretation, and does not currently bear at all on issues of syntactic locality (although it will play an *indirect* role in my account of certain island effects in chapter 3). Boeckx and Grohmann (2007) provide a useful overview of various interpretations of the word “phase”, perhaps none of which will turn out to be equivalent to my use of the term.

2.4.3 Interpretation of adjuncts

Let us use features of the form $*f$ to indicate whatever distinctive syntactic properties adjuncts have. Introducing some such new kind of feature seems unavoidable: as discussed in §2.2.2, the current mechanisms simply do not provide enough degrees of freedom. My aim will be to stipulate as little as possible about how these $*f$ features work. Specifically I will stipulate that bearing the feature $*f$ correlates with having a meaning that is conjoinable with (the meanings of) things that bear the corresponding $-f$ feature. So, for example, ‘violently’ can bear a $*V$ feature since it is associated with the predicate of events **violent**, which is directly conjoinable with the meaning of ‘stab’, which bears a $-V$ feature. On the assumption that this $*V$ does nothing more than indicate that ‘violently’ has an event predicate to contribute, I will now look at the syntactic behaviour that we would naturally expect to result from bearing such a feature, and show that it matches the patterns observed in §2.2.

Consider now how we might derive a sentence like (2.32), including an adjunct.

(2.32) Brutus stabbed Caesar violently.

$F[\text{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b}) \ \& \ \mathbf{violent}, \phi_T]$

Suppose that the derivation begins as shown in Figure 27. These steps will be, it seems sensible to suppose, at least part of the phase in which the conjuncts that are to be added to **stabbing** are “set up”. The interpretation we would like to end up with, when SPL applies at the end of this phase, is **stabbing** & int(**c**) & ext(**b**) & **violent**. Now, having reached e_5 , we can ask: what more must be done to “set up” ‘violently’ to contribute an appropriate conjunct to the predicate produced by this phase? The natural answer seems to be “nothing”. In the case of ‘Caesar’ and (the unpronounced occurrence of) ‘Brutus’, some extra information must be encoded about the particular conjunct that it will contribute — because conjuncts are all that *can* be contributed — and this is achieved by merging it into its argument configuration. But ‘violently’

$$\begin{aligned}
e_1 &= \text{INS}(\ell_{\text{stab}}, \ell_{\text{Caesar}}) \\
&= \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right]: +\text{d}+\text{d}-\text{V}, \left\{ \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right]: -\text{d} \right\} \right\rangle \\
e_2 &= \text{MRG}(e_1) \\
&= \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right]: +\text{d}-\text{V}, \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right]: \text{d}, \left\{ \right\} \right\rangle \\
e_3 &= \text{INS}(e_2, \ell_{\text{Brutus}}) \\
&= \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right]: +\text{d}-\text{V}, \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right]: \text{d}, \left\{ \left[\begin{array}{c} \text{Brutus} \\ \text{b} \end{array} \right]: -\text{d}-\text{k} \right\} \right\rangle \\
e_4 &= \text{MRG}(e_3) \\
&= \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right]: -\text{V}, \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right]: \text{d}, \left[\begin{array}{c} \text{---} \\ \text{b} \end{array} \right]: \text{d}, \left\{ \left[\begin{array}{c} \text{Brutus} \\ \text{---} \end{array} \right]: -\text{k} \right\} \right\rangle \\
e_5 &= \text{INS}(e_4, \ell_{\text{violently}}) \\
&= \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right]: -\text{V}, \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right]: \text{d}, \left[\begin{array}{c} \text{---} \\ \text{b} \end{array} \right]: \text{d}, \left\{ \left[\begin{array}{c} \text{violently} \\ \text{violent} \end{array} \right]: *-\text{V}, \left[\begin{array}{c} \text{Brutus} \\ \text{---} \end{array} \right]: -\text{k} \right\} \right\rangle
\end{aligned}$$

Figure 27

comes with a meaning that is ready-made for conjunction, and requires no such extra help or adjustments. To completely answer the question of how ‘violently’ is to be interpreted, it suffices to single out a bearer of a $-\text{V}$ feature (and this is *not* true of ‘Caesar’ and ‘Brutus’). As soon as ‘violently’ is inserted, producing e_5 in Figure 27, $\text{stab}: -\text{V}$ is appropriately singled out; SPL needs no more information than is already present.⁴⁷

The general claim therefore is that adjuncts are constituents that are *only inserted*, and never merged. Merging corresponds precisely to the kind of “type shifting” that Conjunctivist interpretation says that we need for arguments. This distinction between insertion and merging is reminiscent of the distinction Hornstein and Nunes

⁴⁷Note that although in Figure 27 ‘violently’ has been inserted after ‘Caesar’ has merged into its argument position, it could equally be inserted before ‘Caesar’ is, or even after ‘Caesar’ is inserted but before it is merged. Any of these orders of operations will produce the same expression (e_5) as a final result.

$$\begin{array}{l}
e_5 = \left\langle \left[\begin{array}{c} \text{stab} \\ \text{stabbing} \end{array} \right] : -V, \left[\begin{array}{c} \text{Caesar} \\ \text{c} \end{array} \right] : d, \left[\begin{array}{c} \text{---} \\ \text{b} \end{array} \right] : d, \left\{ \left[\begin{array}{c} \text{violently} \\ \text{violent} \end{array} \right] : *V, \left[\begin{array}{c} \text{Brutus} \\ \text{b} \end{array} \right] : -k \right\} \right\rangle \\
e_6 = \text{SPL}(e_5) \\
= \left\langle \left[\begin{array}{c} \text{--- stab Caesar violently} \\ \text{stabbing\&int(c)\&ext(b)\&violent} \end{array} \right] : -V, \left\{ \left[\begin{array}{c} \text{Brutus} \\ \text{b} \end{array} \right] : -k \right\} \right\rangle
\end{array}$$

Figure 28

(2008) suggest between “concatenation” and labelling.⁴⁸ The improvement that the account presented here offers over that of Hornstein and Nunes is that it avoids *ever* placing adjuncts in an argument-like configuration. For Hornstein and Nunes, labelling effectively coincides with being included with a constituent for further syntactic operations (via their A-over-A principle), which neatly derives the possibility of adjuncts being excluded by being left unlabeled, but does not as neatly explain why adjuncts can also optionally be included in the relevant constituent.

The derivation begun in Figure 27 therefore continues with an immediate application of SPL, as shown in Figure 28. The $*V$ feature is deleted in the process: there is no “inverse” feature that goes with it in the way $+f$ goes with $-f$, so this is similar to the one-sided or “asymmetric” feature checking suggested for adjuncts by Frey and Gärtner (2002). The new unit resulting from this application of SPL, with its $-V$ feature, participates in the rest of the derivation in just the same way that the adjunctless ‘stab Caesar’ unit would. This will derive the desired result in (2.32).

Similarly, if we begin with an additional “fronting feature” $-f$ on ‘stab’, we will derive a unit $\left[\begin{array}{c} \text{stab Caesar violently} \\ \text{stabbing\&int(c)\&violent} \end{array} \right] : -V-f$ that will result in ‘stab Caesar violently’

⁴⁸Chametzky (1996) likewise suggests that adjunction corresponds to unlabelled structures. The current implementation can also be thought of as a formalisation of the intuition that adjuncts are “in a different dimension”; or even more directly, the idea that “An adjunct is simply activated on a derivational phase, without connecting to the phrase-marker . . . adjuncts can be activated by simply ‘being there’” (Lasnik and Uriagereka 2005, pp.254–255) and “from a neo-Davidsonian semantic perspective, the adjunct is indeed ‘just there’ as a mere conjunct” (Uriagereka and Pietroski 2002, p.279). See also Hinzen (2009).

$e_1 = \text{INS}(\ell'_{\text{stab}}, \ell_{\text{Caesar}})$	$= \langle \text{stab} : +\text{d}+\text{d}-\text{V}-\text{f}, \{ \text{Caesar} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{stab} : +\text{d}-\text{V}-\text{f}, \text{Caesar} : \text{d}, \{ \} \rangle$
$e_3 = \text{INS}(e_2, \ell_{\text{Brutus}})$	$= \langle \text{stab} : +\text{d}-\text{V}-\text{f}, \text{Caesar} : \text{d}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_4 = \text{MRG}(e_3)$	$= \langle \text{stab} : -\text{V}-\text{f}, \text{Caesar} : \text{d}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_5 = \text{INS}(e_4, \ell_{\text{violently}})$	$= \langle \text{stab} : -\text{V}-\text{f}, \text{Caesar} : \text{d}, \text{---} : \text{d}, \{ \text{violently} : *V, \text{Brutus} : -\text{k} \} \rangle$
$e_6 = \text{SPL}(e_5)$	$= \langle \text{---} \text{stab Caesar violently} \text{---} \text{stabbing\&int(c)\&ext(b)\&violent} : -\text{V}-\text{f}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_7 = \text{INS}(\ell_{\text{T}}, e_6)$	$= \langle -\text{ed} : +\text{V}+\text{k}-\text{t}, \{ \text{stab Caesar violently} : -\text{V}-\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_8 = \text{MRG}(e_7)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \text{---} : V, \{ \text{stab Caesar violently} : -\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_9 = \text{MRG}(e_8)$	$= \langle -\text{ed} : -\text{t}, \text{---} : V, \text{Brutus} : \text{k}, \{ \text{stab Caesar violently} : -\text{f} \} \rangle$
$e_{10} = \text{SPL}(e_9)$	$= \langle \text{---} \text{Brutus -ed ---} \text{---} \text{stabbing\&int(c)\&ext(b)\&violent, \phi_{\text{T}}} : -\text{t}, \{ \text{stab Caesar violently} : -\text{f} \} \rangle$
$e_{11} = \text{INS}(e_{10}, \ell_{\text{X}})$	$= \langle \epsilon_x : +\text{t}+\text{f}-\text{x}, \{ \text{Brutus -ed ---} : -\text{t}, \text{stab Caesar violently} : -\text{f} \} \rangle$
$e_{12} = \text{MRG}(e_{11})$	$= \langle \epsilon_x : +\text{f}-\text{x}, \text{Brutus -ed ---} : \text{t}, \{ \text{stab Caesar violently} : -\text{f} \} \rangle$
$e_{13} = \text{MRG}(e_{12})$	$= \langle \epsilon_x : -\text{x}, \text{Brutus -ed ---} : \text{t}, \text{stab Caesar violently} : \text{f}, \{ \} \rangle$
$e_{14} = \text{SPL}(e_{13})$	$= \langle \text{stab Caesar violently } \epsilon_x \text{ Brutus -ed ---} : -\text{x}, \{ \} \rangle$

Figure 29

being fronted, since this unit will participate in the rest of the derivation in the same way that the ‘stab Caesar’ unit did in Figure 26. This is shown in Figure 29.

This illustrates the possibility of including an adjunct with the phrase it modifies, when that phrase is targeted for some syntactic operation. But as we have seen, it is also possible to “strand” an adjunct: specifically, it is possible to target ‘stab Caesar’ without its adjunct ‘violently’. The sense in which adjuncts are “more easily accommodated” than arguments, due to their simple conjunctive interpretation, makes possible an explanation of this fact.

The basic property of the *f features is that an element bearing them will con-

$$\begin{aligned}
e_1 &= \left\langle \left[\text{stab Caesar} \right]_{\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})} : -\mathbf{V}, \{ [\text{Brutus}] : -\mathbf{k} \} \right\rangle \\
e_2 &= \text{INS}(\ell_T, e_1) \\
&= \left\langle \left[\text{ed} \right]_{\phi_T} : +\mathbf{V}+\mathbf{k}-\mathbf{t}, \left\{ \left[\text{stab Caesar} \right]_{\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})} : -\mathbf{V}, [\text{Brutus}] : -\mathbf{k} \right\} \right\rangle \\
e_3 &= \text{MRG}(e_2) \\
&= \left\langle \left[\text{ed} \right]_{\phi_T} : +\mathbf{k}-\mathbf{t}, \left[\text{stab Caesar} \right]_{\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})} : \mathbf{V}, \{ [\text{Brutus}] : -\mathbf{k} \} \right\rangle \\
e_4 &= \text{MRG}(e_3) \\
&= \left\langle \left[\text{ed} \right]_{\phi_T} : -\mathbf{t}, \left[\text{stab Caesar} \right]_{\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})} : \mathbf{V}, [\text{Brutus}] : \mathbf{k}, \{ \} \right\rangle \\
e_5 &= \text{INS}(e_4, \ell_{\text{violently}}) \\
&= \left\langle \left[\text{ed} \right]_{\phi_T} : -\mathbf{t}, \left[\text{stab Caesar} \right]_{\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})} : \mathbf{V}, [\text{Brutus}] : \mathbf{k}, \{ [\text{violently}] : * \mathbf{V} \} \right\rangle \\
e_6 &= \text{SPL}(e_5) \\
&= \left\langle \left[\text{Brutus stabbed Caesar violently} \right]_{\mathbf{F}[\text{stabbing}\&\text{int}(\text{c})\&\text{ext}(\text{b})\&\text{violent}, \phi_T]} : -\mathbf{t}, \{ \} \right\rangle
\end{aligned}$$

Figure 30

junctively modify something of the corresponding type: ‘violently’ bears $*\mathbf{V}$ so it will modify ‘stab’ which bears $-\mathbf{V}$. In the derivations in Figure 27, Figure 28 and Figure 29, this is achieved by inserting ‘violently’ into the phase of which ‘stab’ is the head. But another way to achieve the same result (i.e. conjunction of **violent** and **stabbing**) is to insert ‘violently’ into the phase of which (a projection of ‘stab’, i.e. a unit that has inherited its $-\mathbf{V}$ feature) is a complement. In this case, the derivation of the TP phase would proceed as shown in Figure 30; note that this begins with an expression e_1 that has been constructed at spellout of the VP phase without the adjunct yet playing any role.

Recall that the occurrence of the subject in the specifier position of TP for Case is taken to be semantically vacuous. That **violent** is conjoined with **stabbing**&int(c)&

$\text{ext}(\mathbf{b})$ when SPL applies to e_5 — rather than being conjoined as a top-level conjunct to produce $\mathbf{F}[\text{stabbing} \dots, \phi_T] \& \mathbf{violent}$ — is determined by the fact that its $*V$ feature matches the V category of the phase’s complement. Intuitively speaking, $\mathbf{violent}$ is conjoined with the VP’s meaning “before” the VP’s meaning is further composed with the rest of the TP phase. Note that the meaning we derive for the TP is identical to what we derived in Figure 27, when the adjunct was inserted in the VP phase. The distinction is not between two different meanings that may be derived, but rather between two different points in the derivation at which it is possible for $\mathbf{violent}$ to be conjoined with stabbing .

This choice between inserting the adjunct in the VP phase and inserting it in the TP phase was sketched in (2.20) and (2.21) in §2.3, repeated here.

$$\begin{aligned}
 (2.20) \quad \phi_{VP} &= \text{stabbing} \& \text{int}(\mathbf{c}) \& \text{ext}(\mathbf{b}) \& \mathbf{violent} \\
 \phi_{TP} &= \mathbf{F}[\phi_{VP}, \phi_T] \\
 &= \mathbf{F}[\text{stabbing} \& \text{int}(\mathbf{c}) \& \text{ext}(\mathbf{b}) \& \mathbf{violent}, \phi_T]
 \end{aligned}$$

$$\begin{aligned}
 (2.21) \quad \phi_{VP} &= \text{stabbing} \& \text{int}(\mathbf{c}) \& \text{ext}(\mathbf{b}) \\
 \phi_{TP} &= \mathbf{F}[\phi_{VP} \& \mathbf{violent}, \phi_T] \\
 &= \mathbf{F}[\text{stabbing} \& \text{int}(\mathbf{c}) \& \text{ext}(\mathbf{b}) \& \mathbf{violent}, \phi_T]
 \end{aligned}$$

The outline in (2.20) shows $\mathbf{violent}$ being included “early”, in the VP phase, as illustrated in Figure 28 and Figure 29. The alternative in (2.21) shows $\mathbf{violent}$ being included “late”, in the TP phase, as in Figure 30. The availability of this later option has nothing to do with any particular properties of the TP projection: the general point is that if an adjunct has a meaning that conjunctively modifies the head of a certain projection XP, it can be inserted either in the XP phase (along the lines of (2.20)) or in the phase immediately above XP (i.e. the phase in which the XP

$e_1 = \text{INS}(\ell'_{\text{stab}}, \ell_{\text{Caesar}})$	$= \langle \text{stab} : +\text{d}+\text{d}-\text{V}-\text{f}, \{ \text{Caesar} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{stab} : +\text{d}-\text{V}-\text{f}, \text{Caesar} : \text{d}, \{ \} \rangle$
$e_3 = \text{INS}(e_2, \ell_{\text{Brutus}})$	$= \langle \text{stab} : +\text{d}-\text{V}-\text{f}, \text{Caesar} : \text{d}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_4 = \text{MRG}(e_3)$	$= \langle \text{stab} : -\text{V}-\text{f}, \text{Caesar} : \text{d}, \text{___} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_5 = \text{SPL}(e_4)$	$= \langle \text{[stabbing\&int(c)\&ext(b)]} : -\text{V}-\text{f}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_6 = \text{INS}(\ell_{\text{T}}, e_5)$	$= \langle -\text{ed} : +\text{V}+\text{k}-\text{t}, \{ \text{stab Caesar} : -\text{V}-\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \text{___} : \text{V}, \{ \text{stab Caesar} : -\text{f}, \text{Brutus} : -\text{k} \} \rangle$
$e_8 = \text{MRG}(e_7)$	$= \langle -\text{ed} : -\text{t}, \text{___} : \text{V}, \text{Brutus} : \text{k}, \{ \text{stab Caesar} : -\text{f} \} \rangle$
$e_9 = \text{INS}(e_8, \ell_{\text{violently}})$	$= \langle -\text{ed} : -\text{t}, \text{___} : \text{V}, \text{Brutus} : \text{k}, \{ \text{violently} : *-\text{V}, \text{stab Caesar} : -\text{f} \} \rangle$
$e_{10} = \text{SPL}(e_9)$	$= \langle \text{[Brutus -ed ___violently]} : -\text{t}, \{ \text{stab Caesar} : -\text{f} \} \rangle$
$e_{11} = \text{INS}(e_{10}, \ell_{\text{X}})$	$= \langle \epsilon_{\text{X}} : +\text{t}+\text{f}-\text{x}, \{ \text{Brutus -ed ___violently} : -\text{t}, \text{stab Caesar} : -\text{f} \} \rangle$
$e_{12} = \text{MRG}(e_{11})$	$= \langle \epsilon_{\text{X}} : +\text{f}-\text{x}, \text{Brutus -ed ___violently} : \text{t}, \{ \text{stab Caesar} : -\text{f} \} \rangle$
$e_{13} = \text{MRG}(e_{12})$	$= \langle \epsilon_{\text{X}} : -\text{x}, \text{Brutus -ed ___violently} : \text{t}, \text{stab Caesar} : \text{f}, \{ \} \rangle$
$e_{14} = \text{SPL}(e_{13})$	$= \langle \text{stab Caesar } \epsilon_{\text{X}} \text{ Brutus -ed ___violently} : -\text{x}, \{ \} \rangle$

Figure 31

participates as a non-head) (along the lines of (2.21)).⁴⁹

The option of inserting ‘violently’ only after the VP phase has been completed produces the possibility of targeting ‘stab Caesar’ and stranding its adjunct ‘violently’ when the $-\text{f}$ feature appears on ‘stab’. This is illustrated in Figure 31, which uses the same lexical expressions as Figure 29 does but results in fronting a different fragment.

⁴⁹Note that the idea is not that an adjunct that semantically modifies (the head of) XP can be inserted at *any* point in the derivation after the XP has been completed. The only late option (at least for now; see §2.6) is to insert it in the phase *immediately above* the XP phase, i.e. in the phase of which XP is a complement. In Figure 30 it is crucial that the VP constituent ‘___stab Caesar’ is still a distinct unit during the TP phase, so that ‘violently’ can conjoin with it. But after the application of SPL at the end of Figure 30 this VP constituent no longer exists as a distinct unit, so inserting ‘violently’ in the next phase up after TP (say, the CP phase) is not an option.

When SPL applies to e_9 , ‘violently’ is understood to be attached to the (as it happens, unpronounced) VP that is in the complement position, because its $*V$ feature matches the V category, so it is integrated into the new unit produced as the head of e_{10} . This has no effect on the `stab Caesar:-f` unit that is waiting to remerge later. Intuitively, the occurrence of ‘stab Caesar’ that plays a V role split away from the occurrence that plays a f role when MRG was applied to e_6 to produce e_7 ; when ‘violently’ is introduced shortly thereafter, it interacts only with the occurrence of ‘stab Caesar’ that plays a V role, and the occurrence that plays a f role is unaffected.

What this all amounts to saying is that the untransformed sentence ‘Brutus stabbed Caesar violently’ is actually ambiguous between two distinct syntactic derivations — the one where ‘violently’ is inserted into the VP phase, as in Figure 27 and Figure 28, and the one where ‘violently’ is inserted into the TP phase, as in Figure 30 — each of which makes available a different VP constituent that can be targeted. The distinction between these two possibilities, purely with respect to semantic composition, was illustrated in (2.20) and (2.21); the same idea, with some more syntactic details, is illustrated in Figure 32 and Figure 33. The head-argument relations established in each phase are represented by tree structures, and these syntactic relationships are interpreted by SPL, as illustrated in a simpler case with no adjuncts in Figure 25b. In addition to this tree structure, a phase may include “on the side” adjuncts with a feature that matches the category of any of the pieces of the tree structure: the $*V$ feature matches the head of the VP in Figure 32 and the complement of the TP in Figure 33, as indicated by the dashed lines. If ‘stab’ bears a $-f$ feature, a derivation analogous to Figure 32 will result in the fronting of ‘stab Caesar violently’ as in Figure 29, and a derivation analogous to Figure 33 will result in the fronting of ‘stab Caesar’, standing ‘violently’, as in Figure 31.

The idea is not that ‘violently’ can be inserted in *any* phase later than the VP phase. What makes the TP phase an option is that the VP that ‘violently’ needs to

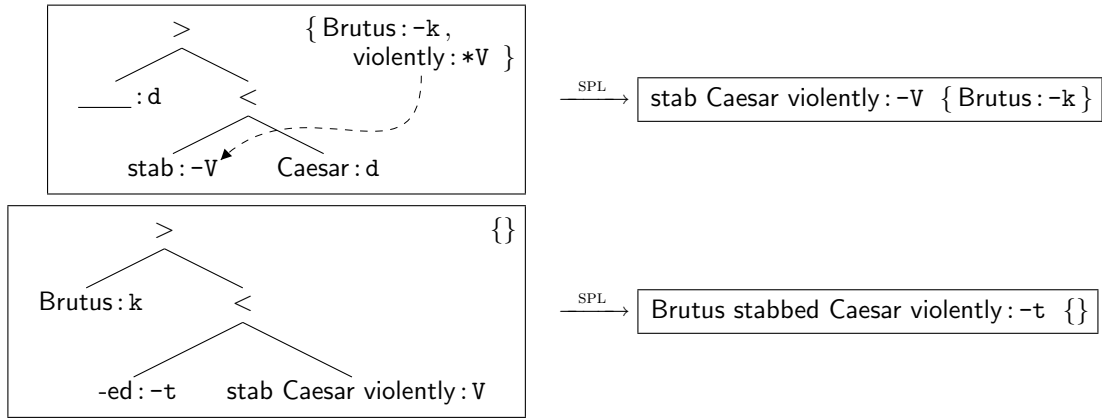


Figure 32: Illustration of a derivation with the VP-modifying adjunct ‘violently’ inserted during the VP phase, as given formally in Figure 27 and Figure 28. This permits ‘stab Caesar violently’ to act as a constituent, as shown in Figure 29. Semantic composition proceeds as shown in (2.20).

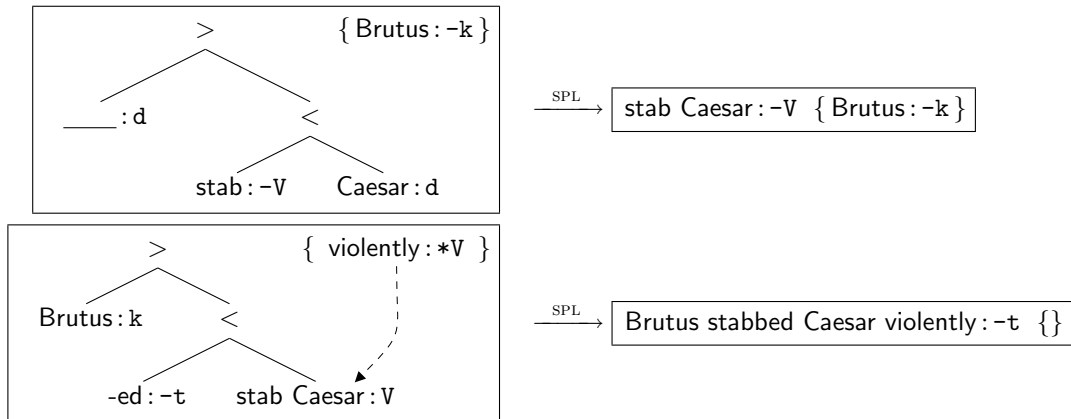
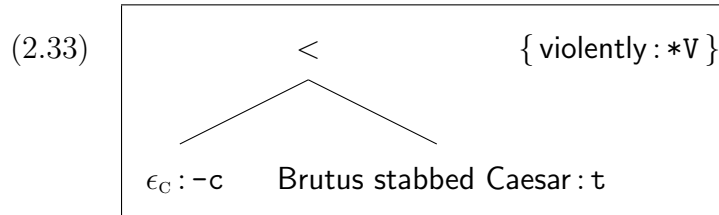


Figure 33: Illustration of a derivation with the VP-modifying adjunct ‘violently’ inserted during the TP phase, as given formally in Figure 30. This permits ‘stab Caesar’ to act as a constituent, as shown in Figure 31. Semantic composition proceeds as shown in (2.21).

modify is still “present” or “visible”. The next phase above TP — say, CP, which we can assume, at least for this sentence, just selects a TP complement and no specifier — would not be an option, because the VP is not present in the relevant sense. If ‘violently’ were inserted during the CP phase, the structure produced will be as shown in (2.33). Neither the $-c$ feature on the phase’s head, nor the t annotation on the phase’s complement, matches the $*V$ feature.



Crucially, the two-way ambiguity demonstrated for the attachment of the adjunct ‘violently’ *is not available for arguments* such as ‘Caesar’. Therefore there are *not* two distinct derivations of ‘Brutus stabbed Caesar’, one of which makes available ‘stab Caesar’ and one of which makes available ‘stab’ alone. This is a consequence of the fact that arguments require a more specific configuration than adjuncts do. In order to be properly interpretable, ‘Caesar’ must establish itself as an argument of ‘stab’. It must interact with ‘stab’ in a certain way that is only possible in the phase of which ‘stab’ is the head: only the VP phase permits establishing argument relations with ‘stab’.⁵⁰ The adjunct ‘violently’ wants nothing to do with such complications, and only requires that it appear in some phase in which something it can modify is participating, either as an argument or as a head; as we have seen, an adjunct can therefore be inserted in the phase headed by the element that it conjunctively modifies *or* in the phase immediately above it, where the element it modifies is an argument. An argument like ‘Caesar’ does *not* have this option of “later insertion”; so it is not possible to target ‘stab’ without bringing along ‘Caesar’, in the way that one can target ‘stab Caesar’ without bringing along ‘violently’ and target ‘sleep’ without bringing along ‘quietly’. This captures the basic pattern observed in (2.1) and (2.2), repeated here.

- (2.1) a. Brutus slept quietly. (ambiguous)
 b. Sleep quietly, (is what) Brutus did. ('quietly' inserted "early")
 c. Sleep, (is what) Brutus did quietly. ('quietly' inserted "late")
- (2.2) a. Brutus stabbed Caesar. (unambiguous)
 b. Stab Caesar, (is what) Brutus did. ('Caesar' inserted "early")
 c. * Stab, (is what) Brutus did Caesar.

When more than one adjunct is present, as in (2.8) reproduced here, the choice of when each adjunct is inserted is clearly independent: any that are included in the fronted constituent are inserted in the lower/earlier phase, and any that are stranded are inserted in the higher/later phase. And it follows naturally that the need for 'Caesar' to necessarily be inserted in the lower of the two phases discussed here is not affected by the presence or absence of any adjuncts.

- (2.8) a. Brutus stabbed Caesar violently with a knife.
 b. Stab Caesar violently with a knife, (is what) Brutus did.
 c. Stab Caesar violently, (is what) Brutus did with a knife.
 d. Stab Caesar, (is what) Brutus did violently with a knife.
 e. * Stab, (is what) Brutus did Caesar violently with a knife.

2.5 Discussion

In this section I discuss some potential objections to the proposals of the previous section in §2.5.1, and consider more closely the nature of the relationship between

⁵⁰In the format of (2.20) and (2.21): there is no derivation of 'Brutus stabbed Caesar' where the semantics is constructed as follows:

$$\begin{aligned}
 \phi_{VP} &= \mathbf{stabbing} \ \& \ \mathbf{ext}(\mathbf{b}) \\
 \phi_{TP} &= \mathbf{F}[\phi_{VP} \ \& \ \mathbf{int}(\mathbf{c}), \phi_T] \\
 &= \mathbf{F}[\mathbf{stabbing} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\mathbf{c}), \phi_T]
 \end{aligned}$$

syntax and semantics that it leads us to in §2.5.2. In the next section I return to empirical consequences of the account.

2.5.1 Potential objections

One may wonder whether an account that introduces a new kind of feature $*f$ that adjuncts (and only adjuncts) bear can provide any less stipulative an account than others, such as the labelling convention from §2.2.1 or the MG-based Frey and Gärtner (2002) account from §2.2.2. It is true that the account I have offered is *not* as reductive as categorial grammar accounts of adjuncts using X/X types, or equivalently the MG encoding of this idea using $+x-x$ feature sequences, because these genuinely invoke only the one composition operation for both arguments and “adjuncts”. But this kind of proposal, while arguably the most attractive account in terms of parsimony, does not seem to be viable, as discussed in §2.2.2. What I have proposed above is less reductionist than the empirically-problematic X/X approach but also less stipulative than the Frey and Gärtner (2002) approach. The introduction of $*f$ features (or something very much like them) is inevitable once we reject the entirely reductionist X/X approach, but from the point of deciding that such features *exist*, there remain more and less stipulative ways to arrive at their properties and effects.

Frey and Gärtner (2002) account for adjunction by introducing mechanisms that share little or nothing of substance with the mechanisms that treat argument configurations. Having done this, the interactions between the new mode of combining expressions and the existing MRG operation that establishes argument configurations can only be stipulated; as discussed in §2.2.2, introducing a completely distinct ADJOIN operation leaves open questions about why the system might include *that* operation and not others.

I acknowledge that an adjunct is not *just* a certain kind of argument-taker (one whose domain and range are equal, in categorial terms), but this leaves open the

possibility of providing *some* insight into why the various properties that adjuncts have pattern together. Given a particular constituent whose semantic component is thought to compose via bare conjunction, and given some pre-existing syntactic operations, there can be patterns of syntactic behaviour that naturally accommodate this semantic composition and others that do not. My aim in §2.4.3 was to stipulate as little as possible about the syntactic behaviour of *f features, and ask how the existing syntactic operations could accommodate them on the independently justified assumption that adjuncts denote pure conjuncts. This borrows from categorial grammars the intuition that syntactic interactions should be a transparent reflection of semantic interactions (while rejecting the semantic system that is standardly adopted in that framework): the aim was to ask how the existing syntactic operations could “accommodate” something with a bare conjunct as its meaning, such as ‘violently’. A system with INS and MRG as the only two structure-building operations, such that movement is re-merging, implies that there is a state that a constituent is in immediately before it is first merged (when it has been only inserted); this state seems sufficient.⁵¹

One might object that it is not clear that my treatment of adjuncts shares anything “of substance” with the existing syntactic mechanisms. The logic behind this objection is that a member of the set component of an expression is understood to be an adjunct if it begins with an appropriate *f feature, and understood to be a “moving sub-part” otherwise; everything I have said could be said just as easily if expressions included *two* separate sets, one for adjuncts to the current phase, and one for moving sub-parts. To further unify the behaviour of adjuncts and arguments, one would want evidence that adjuncts and “moving things” behave the same way for some purposes. In chapter 3 I argue that this is true, specifically, that movement out of either is prohibited: movement out of an adjunct is a violation of the well-known

⁵¹Norbert Hornstein (p.c.) notes a stronger possible interpretation of what has been shown: that a system where move is re-merge *predicts* that something with the behaviour of adjuncts exists.

adjunct island constraint, while movement out of a “moving thing” induces what has been called a “freezing effect” (Wexler and Culicover 1981, Corver 2005).

Finally, one might also question whether it is absolutely necessary to delay interpretation in the way that I introduced in §2.4.2. It is certainly true that Conjunctivist interpretation could not proceed at every merging step of a derivation using the original MG representations, as discussed in §2.4.1. This means that some sort of more detailed representation is required; but there are multiple possibilities, not only the one I chose in §2.4.2.

One alternative would be to include in each expression a “counter” indicating how many arguments the current head has already combined with.⁵² Then when MRG applies to an expression with 0 as the value of its counter, the *int* operator is introduced (and the counter incremented), and when it applies to an expression with 1 as the value of its counter, the *ext* operator is introduced (and the counter incremented) (and so on, if we permit heads that take more than two arguments — nothing important hinges on this). This way, the derivation of a VP would proceed as shown in (2.34) (compare with the solution I adopted in (2.29) and Figure 25).

$$\begin{aligned}
 (2.34) \quad e_1 &= \text{MRG} \left(\left\langle \left\langle \begin{array}{c} \text{stab} \\ \text{[stabbing]} \end{array} : +\text{d}+\text{d}-\text{V}, \quad 0, \quad \left\{ \begin{array}{c} \text{[Caesar]} \\ \text{c} \end{array} : -\text{d} \right\} \right\rangle \right) \\
 &= \left\langle \begin{array}{c} \text{stab Caesar} \\ \text{[stabbing\&int(c)]} \end{array} : +\text{d}-\text{V}, \quad 1, \quad \left\{ \right\} \right\rangle \\
 e_2 &= \text{INS} (e_1, \ell_{\text{Brutus}}) \\
 &= \left\langle \begin{array}{c} \text{stab Caesar} \\ \text{[stabbing\&int(c)]} \end{array} : +\text{d}-\text{V}, \quad 1, \quad \left\{ \begin{array}{c} \text{[Brutus]} \\ \text{b} \end{array} : -\text{d}-\text{k} \right\} \right\rangle \\
 e_3 &= \text{MRG} (e_2) \\
 &= \left\langle \begin{array}{c} \text{stab Caesar} \\ \text{[stabbing\&int(c)\&ext(b)]} \end{array} : -\text{V}, \quad 2, \quad \left\{ \begin{array}{c} \text{[Brutus]} \\ \text{b} \end{array} : -\text{k} \right\} \right\rangle
 \end{aligned}$$

Recall that the original framework does not maintain enough information about the structure we have built to distinguish between cases where *int* should apply and cases where *ext* should apply. The counter-based solution solves this by recording

⁵²This suggestion comes from Greg Kobele, p.c.

the relevant “amount of” structure, rather than recording the relevant portion of structure itself, as the solution I adopted does. Both solutions are adequate for the interpretation of arguments.

Only the “structure-storing” solution, however, permits the account of adjunction I adopted in §2.4.3, because it forces us to separate interpretation from the checking of +f/-f features. If new semantic values are only constructed at applications of MRG, as in (2.34), then there is no way for adjuncts to be only inserted into the derivation and yet contribute to interpretation (without adding a specialized adjunct-interpretation operation). So the separation of interpretation from argument-structure building is crucial for the proposed treatment of adjunction.

One might worry, nonetheless, that the choice of the structure-storing solution over the counter-storing solution is an ad hoc one motivated only by the facts about the behaviour of adjuncts. For this objection to be valid, it must be clear that, all else being equal, the counter-storing solution is *more* theoretically parsimonious than the structure-storing one; if they are equally parsimonious (or the comparison is unclear) then there is no shame in letting the empirical facts guide us, and if the counter-storing solution is less so then the point is moot.

I think that to the extent that the comparison is clear, a case can be made that the structure-storing solution is the more frugal. First, it is consistent with the intuition behind the Inclusiveness Condition (Chomsky 1995, p.228), that “no new objects are *added* in the course of computation apart from *rearrangements* of lexical properties” (emphasis added), though making this appealing intuition precise is extremely difficult. Second, while it may appear that the structure-storing solution requires “an extra operation” (namely SPL), I think this is misleading because MRG and SPL together do roughly the work of MRG alone in the counter-storing solution. Redistributing existing ideas among formal devices should not be considered theoretically costly; indeed, it is a standard strategy for increasing the empirical ground covered

by existing ideas.

Finally, one more consequence of the structure-storing solution, which does not follow from the counter-storing solution, is interesting to note: it follows immediately that “intermediate projections” can not be targeted for syntactic operations. The reason is that SPL will never be applied to a structure in which a head has taken some but not all of its arguments, and so no unit corresponding to such a partial structure will ever be formed, as would be required for a conventional X' node to be manipulated.⁵³

2.5.2 The role of syntactic features and semantic sorts

The picture I have presented raises a question about the relationship between syntactic and semantic properties of expressions: exactly *how* transparent is the relationship between semantic composition and syntactic interactions? If the lexical meaning of ‘violently’, for example, is an event predicate, then is it furthermore necessary, one might wonder, to mark it as an adjunct with a syntactic $*V$ feature. Put differently: syntactic features generally serve to encode the syntactic interactions that an expression can or must enter into, but if syntactic interactions reflect semantic ones as I have argued, what work remains to be done by syntactic features?

When we consider the details it is clear that there are distinctions that must still be encoded by syntactic features. Some instances can be seen without even looking beyond the simplistic sentences I have restricted attention to in this chapter, and others emerge when we consider some more complex cases.

Suppose that lexical semantic meanings are sorted, such that \mathbf{b} , for example, is

⁵³This requires some qualification and clarification. Nothing I have said rules out derivations where SPL does apply to X' nodes, but doing so would not result in the desired meaning. We could apply SPL after each of the merging steps in a derivation of a structure like $[_{\phi P} \psi_2 [_{\phi'} \phi \psi_1]]$, but the result would be $\phi \& \text{int}(\psi_1) \& \text{int}(\psi_2)$ rather than $\phi \& \text{int}(\psi_1) \& \text{ext}(\psi_2)$. I assume that such a logical form, with two expressions of an internal participant, is not well-formed, for if it were we would expect ‘Brutus stabbed Caesar’ to have a second reading, derived via an additional application of SPL, with $\text{int}(\mathbf{b})$ as a part of its logical form.

marked as a predicate of individuals, **violent** is marked as a predicate of events, and so on. We can represent this by adding subscripts: e for predicates of individuals, v for predicates of events. Then the two core kinds of interactions that I have discussed in this chapter can be represented, relatively abstractly, as in (2.35).

$$(2.35) \quad \begin{array}{l} \text{a. } \left[\begin{array}{c} \text{sleep} \\ \text{sleeping}_v \end{array} \right]: -V \quad \left[\begin{array}{c} \text{quietly} \\ \text{quiet}_v \end{array} \right]: *V \quad \longrightarrow \quad \left[\begin{array}{c} \text{sleep quietly} \\ \text{sleeping}_v \& \text{quiet}_v \end{array} \right]: -V \\ \text{b. } \left[\begin{array}{c} \text{stab} \\ \text{stabbing}_v \end{array} \right]: +d-V \quad \left[\begin{array}{c} \text{Caesar} \\ c_e \end{array} \right]: -d \quad \longrightarrow \quad \left[\begin{array}{c} \text{stab Caesar} \\ \text{stabbing}_v \& \text{int}(c_e) \end{array} \right]: -V \end{array}$$

Here we can already see that semantic sorts do not uniquely determine syntactic feature sequences. First, although **sleeping_v** and **quiet_v** are both predicates of events, one is marked with $-V$ and the other with $*V$. So even given that the lexical meaning of a certain word is a monadic event predicate, there is a further, seemingly arbitrary, distinction to encode: informally, whether the word acts syntactically as a verb or as an adverb. Second, note that **stabbing_v** is of the same sort again and yet its feature sequence is neither $*V$, nor simply $-V$, but $+d-V$. In other words, even given that a word with an event predicate as its meaning is a verb rather than an adverb, there are further facts that syntactic features serve to encode about the verb’s transitivity.⁵⁴

My aim has not been to derive insightful explanations for brute distributional facts of this kind. It has been to show why such brute distributional facts *correlate* with more subtle syntactic properties, such as the possibility of “stranding” under movement (and, as discussed in the next section, the possibility of counter-cyclic attachment). It may be helpful to briefly consider what would happen if syntactic features were omitted, and the sort annotations on meanings *were* the only thing that dictated how word and phrases combined: we simply say that when two constituents combine, they combine in the manner I have discussed for “adjuncts” if their meanings are of the same sort, and they combine in the manner I have discussed for “arguments” otherwise. As noted above, we would lose track of certain important information, such

⁵⁴For simplicity I assume here that external arguments are introduced outside the VP projection, as in the appendix but contra the assumptions in the body of the chapter.

as the distinction between verbs and adverbs and the transitivity of verbs, and would therefore vastly overgenerate. Some examples that would be erroneously derivable are given in (2.36).

- (2.36) a. * Brutus quietlied sleep (‘quiet’ as a verb, ‘sleep’ as an adverb)
quiet_v & ext(b_e) & sleeping_v
- b. * Brutus stabbed (‘stab’ as intransitive)
stabbing_v & ext(b_e)
- c. * Brutus slept Caesar stab (‘slept’ as transitive, ‘stab’ as an adverb)
sleeping_v & ext(b_e) & int(c_e) & stabbing_v

But crucially, to the extent that ‘sleep’ can be used “as an adverb”, it will be possible to strand it in exactly those cases; and to the extent that ‘Caesar’ can be used as the object of ‘sleep’, it will similarly follow that it does not have this flexibility. If we furthermore drop the assumption that the “backbone” of the sentence must have a meaning of sort *v*, then it will also be possible to derive such examples as in (2.37), but the correlations between the various properties of the two modes of attachment will remain.

- (2.37) a. * Stab brutused
b_e & ext(stabbing_v)
- b. * Sleep brutused stab caesar
b_e & ext(sleeping_v) & int(stabbing_v) & c_e

So while syntactic features are crucial for ruling out such cases, the work they do is independent of any explanations of the sort I have tried to provide for the aforementioned correlations.

Furthermore, the mode of combination illustrated in (2.35b) is not *restricted* to cases where two predicates of differing sorts are composed, as illustrated by (2.38) (mentioned briefly in §1.6.1).

(2.38) I saw Brutus stab Caesar

$\text{seeing}_v \ \& \ \text{ext}(\mathbf{i}_e) \ \& \ \text{int}(\text{stabbing}_v \ \& \ \text{int}(\mathbf{c}_e) \ \& \ \text{ext}(\mathbf{b}_e))$

Assuming that ‘see’ takes (something like) a bare VP as its complement here, this is an instance where two event predicates combine not purely conjunctively as in (2.35a), but rather in exactly the way that is expected for “arguments”, namely (2.35b). This can be straightforwardly accounted for by assigning ‘see’ the feature sequence +V-V. The pattern in (2.35b) is compatible with mismatching sorts, since I assume essentially that *int* and *ext* can “fix” any mismatches, but it does not *require* a mismatch between sorts; but the pattern in (2.35a) *requires matching* sorts, since the two predicates are merely conjoined. Applying the “argument” pattern in (2.35b), where there happens to be no sort mismatch, does not disrupt the derived generalisations: ‘Brutus stab Caesar’ in (2.38) does show the characteristic properties of being an argument, despite having an event predicate as its meaning. This highlights an important point: in my account of the differences between the two VPs ‘stab Caesar’ and ‘sleep quietly’, the crucial property of ‘quietly’ is not that it has as its meaning an event predicate, but rather that its meaning is *merely conjoined* with that of the verb.

To recap, we have observed cases that can be analysed as instances of the pattern in (2.35b), showing syntactic properties associated with “argument-hood” and semantic properties consistent with the *int* and *ext* operators. Some of these cases involve a sort mismatch, and others do not. We have also observed cases that can be analysed as instance of the pattern in (2.35a), showing syntactic properties associated with “adjunct-hood” and semantic properties consistent with mere conjunction. None of these cases involve a sort mismatch, since I assume that mismatching sorts are not conjoinable.

Certain prepositional phrases (PPs) are potentially problematic for this picture. First I outline an account of some PPs that *do* fit in with the story so far (namely,

as adjuncts), for example ‘in Boston’ in (2.39).

(2.39) Brutus slept in Boston

In classic Davidsonian terms, this PP contributes a conjunct along the lines of $\lambda e.\text{Location}(e, \mathbf{boston})$. This conjunct is analogous to those contributed by thematically-marked arguments, on the neo-Davidsonian view. Pursuing the intuition that the **Location** relation comes from the preposition ‘in’, we can hypothesise that while it is a certain *syntactic configuration* that supplements the interpretation of ‘Brutus’ in order to provide an event predicate, it is the preposition that likewise supplements ‘Boston’ (Pietroski 2005, pp.55–56). For present purposes I will write ‘ shift_{in} ’ for the relevant “shifting” operator that is analogous to **int** and **ext**. Then the formula to be associated with (2.39) will be (2.40). Note that since shift_{in} crucially originates “inside” the PP, the composition of the PP with the verb will be an instance of the pattern in (2.35a), as shown in (2.41); and this seems to be a good result, since such PPs by and large display the characteristic properties of adjuncts.

(2.40) $\text{sleeping}_v \ \& \ \text{ext}(\mathbf{b}_e) \ \& \ \text{shift}_{\text{in}}(\mathbf{boston})$

(2.41) $\left[\begin{array}{c} \text{sleep} \\ \text{sleeping}_v \end{array} \right] : -V \quad \left[\begin{array}{c} \text{in Boston} \\ \text{shift}_{\text{in}}(\mathbf{boston}) \end{array} \right] : *V \quad \longrightarrow \quad \left[\begin{array}{c} \text{sleep in Boston} \\ \text{sleeping}_v \ \& \ \text{shift}_{\text{in}}(\mathbf{boston}) \end{array} \right] : -V$

There remain important details to be worked out, but this sketch of a treatment of PPs that fit relatively neatly into the story so far is enough to illustrate the problems that certain others pose.

Consider the role of the PP ‘to Caesar’ in (2.42). Assuming that ‘to’ plays a bridging role analogous to ‘in’ for the purposes of semantic composition, we would expect this sentence to be associated with the formula shown.

(2.42) Brutus gave a book to Caesar in Boston yesterday

$\text{giving}_v \ \& \ \text{int}(\mathbf{a-book}_e) \ \& \ \text{ext}(\mathbf{b}_e) \ \& \ \text{shift}_{\text{to}}(\mathbf{c}) \ \& \ \text{shift}_{\text{in}}(\mathbf{boston}) \ \& \ \text{yesterday}$

Thus we would expect ‘to Caesar’ to play a role analogous to ‘in Boston’ syntactically as well as semantically, composing according to the pattern in (2.41) and (2.35a). In

at least some syntactic respects, however, ‘to Caesar’ behaves more like ‘Brutus’ and ‘a book’ (i.e. “as an argument”): it is obligatory (see (2.43)) and its position relative to other constituents is fairly fixed (see (2.44)), for example, consistent with (2.35b) rather than (2.35a).

- (2.43) a. ? Brutus gave to Caesar in Boston yesterday (omit ‘a book’)
 b. ? Brutus gave a book in Boston yesterday (omit ‘to Caesar’)
 c. Brutus gave a book to Caesar yesterday (omit ‘in Boston’)
 d. Brutus gave a book to Caesar in Boston (omit ‘yesterday’)
- (2.44) a. ? Brutus gave to Caesar in Boston yesterday a book
 b. ? Brutus gave a book in Boston yesterday to Caesar
 c. Brutus gave a book to Caesar yesterday in Boston

In this sense these PPs are problematic for the theory I have outlined: they seem to show the syntactic properties of combining as in (2.35b), but the semantic properties of combining as in (2.35b).

Note, however, that this problem stems from the observation that a plausible meaning for the PP ‘to Caesar’ appears as a bare conjunct in the formula in (2.42). The fact that this plausible meaning for the PP is an event predicate rather than an individual predicate does not in and of itself clash with the fact that it behaves syntactically like ‘Brutus’ and ‘a book’ (rather than like ‘in Boston’), as (2.38) shows. This reveals an important point about the sense in which I aim to “derive syntactic properties from semantic ones”: the semantic premises on which the explanation rests are not that a certain constituent has a meaning of a certain sort, but rather that a certain pair of constituents semantically compose in a certain way. The sorts may constrain, but do not in general determine, the mode of composition.

This leaves open the question of why certain pairs of constituents are not able to combine in all of the ways that are compatible with their meanings’ sorts: for

example, why *can't* we understand 'I saw in Boston' in a way where 'in Boston' is analogous to 'Brutus stab Caesar' in (2.38), to mean that I saw an event that was in Boston? I have no explanation for why not all such options are available. But nonetheless — with the exception of cases like (2.42), as discussed — it is generally true that *when* a constituent's semantic composition is plausibly analysed via the int and ext operators, it displays the syntactic properties following from treatment via +f and -f features and the MRG function in the theory I have proposed, and *when* it is plausibly analysed as bare conjunction, it displays those following from treatment via *f features and (only) the INS function.

2.6 Counter-cyclic adjunction

Another curious property of adjunction, not obviously related to the others discussed so far, is that it appears to be able to apply “counter-cyclically”. In this section I will show that the implementation of adjunction that I have proposed, with the help of some reasonable auxiliary assumptions, permits a simple explanation of the facts that originally motivated counter-cyclic adjunction; and furthermore, correctly predicts certain constraints on counter-cyclic adjunction that other accounts must stipulate.

In GB-era theories where counter-cyclic adjunction was first proposed, the idea was that adjuncts could be present at s-structure without being present at d-structure (Lebeaux (1988, 2000)); in more recent minimalist terms, this amounts to saying that adjunction need not obey the extension condition.⁵⁵ Evidence for this idea includes the contrast between (2.45a), where 'that Mary stole a book' is a complement of 'claim', and (2.45b), where 'that Mary made' is adjoined to 'claim'.⁵⁶

- (2.45) a. * Which claim [that Mary_i stole a book] did she_i deny?
 b. Which claim [that Mary_i made] did she_i deny?

⁵⁵Chomsky (1995, pp.204–205) discusses Lebeaux's proposal in a minimalist setting.

The unacceptability of (2.45a) is straightforwardly accounted for as a Condition C violation, since ‘Mary’, in its base position, is c-commanded (and thus bound) by ‘she’. The question then arises of why an analogous Condition C violation does not appear in (2.45b). Lebeaux suggests that the adjunct ‘that Mary made’ is only added to the structure *after* the *wh*-movement has taken place; thus there is no point in the derivation (or, there is no copy of ‘Mary’) that can induce a Condition C violation. This late-attachment is not possible for arguments, since (in GB-era terminology) they must be present by d-structure (and thus, throughout the derivation) to satisfy selectional requirements.

2.6.1 MG implementations of counter-cyclic adjunction

Gärtner and Michaelis (2003) present a simple encoding of Lebeaux’s idea in the MG formalism, based on the ADJOIN operation proposed by Frey and Gärtner (2002). Recall from §2.2.2 that this operation checks a feature on an adjunct and leaves the features of the adjoined-to constituent (which projects) unchanged, as illustrated in (2.15), repeated here.

$$(2.15) \quad \text{ADJOIN}(\text{sleep} : -V, \text{quietly} : \approx V) = \begin{array}{c} < \\ \diagdown \quad \diagup \\ \text{sleep} : -V \quad \text{quietly} \end{array}$$

Gärtner and Michaelis use the full tree-shaped representations discussed in §1.5.1, rather than the reduced variants I have adopted. This is crucial because it permits the features checked by structure-building operations in the derivational past to remain visible, so that adjunction sites of appropriate categories can be identified. To distinguish checked from unchecked features, Gärtner and Michaelis include a symbol ‘#’ in sequences of features, separating checked features (on its left) from unchecked

⁵⁶The contrast in (2.45) was first noted by Freidin (1986), but has been questioned by Lasnik (1998).

features (on its right). Thus when a feature is checked by some operation, the feature is not *deleted* from the representation, but rather moved leftward across the ‘#’ symbol, from the sublist of unchecked features to that of checked features. This treatment of the +d and -d features can be seen in the simple application of MRG in (2.46), for example; and the application of ADJOIN in (2.15) will be written as in (2.15’).

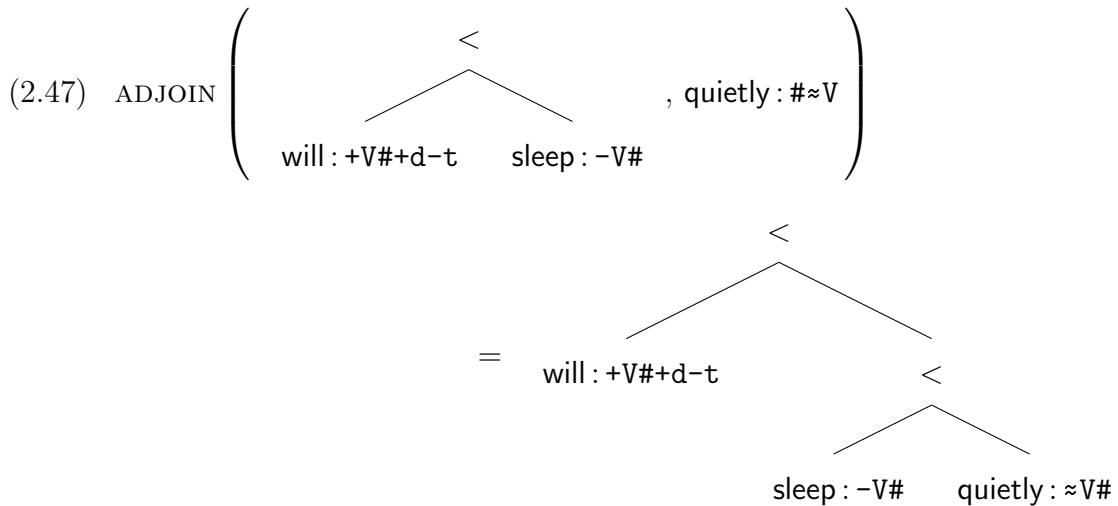
$$(2.46) \quad \text{MRG}(\text{stab} : \# + d - V, \text{Caesar} : \# - d) = \begin{array}{c} < \\ \diagdown \quad \diagup \\ \text{stab} : + d \# - V \quad \text{Caesar} : - d \# \end{array}$$

$$(2.15') \quad \text{ADJOIN}(\text{sleep} : \# - V, \text{quietly} : \# \approx V) = \begin{array}{c} < \\ \diagdown \quad \diagup \\ \text{sleep} : \# - V \quad \text{quietly} : \approx V \# \end{array}$$

Note that on the left hand side of (2.46) the ‘#’ symbol appears at the far left of both feature sequences, because *all* these lexical items’ features are unchecked there. The only feature unchecked on the right hand side of (2.46) is the -V, as expected. To “convert” these enriched feature sequences to their equivalents in the version of the formalism used elsewhere in this thesis, we would just ignore the ‘#’ symbol and all features to the left of it.

With this adjustment Gärtner and Michaelis can encode Lebeaux’s counter-cyclic adjunction idea: the ADJOIN operation can attach an adjunct bearing a $\approx f$ feature not only to an expression bearing an unchecked -f feature,⁵⁷ but also to an expression bearing a *checked* -f feature. The first of these options is cyclic adjunction obeying the extension condition, as illustrated in (2.15’); the second is counter-cyclic adjunction,

an example of which is given in (2.47).

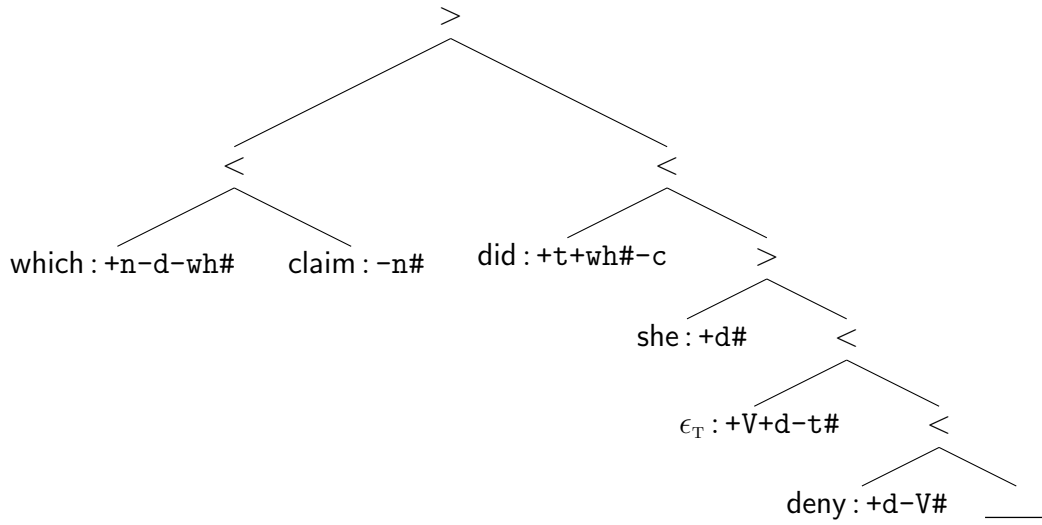


Notice that it is the *checked* $-V$ feature on ‘sleep’ that licenses this application of ADJOIN.

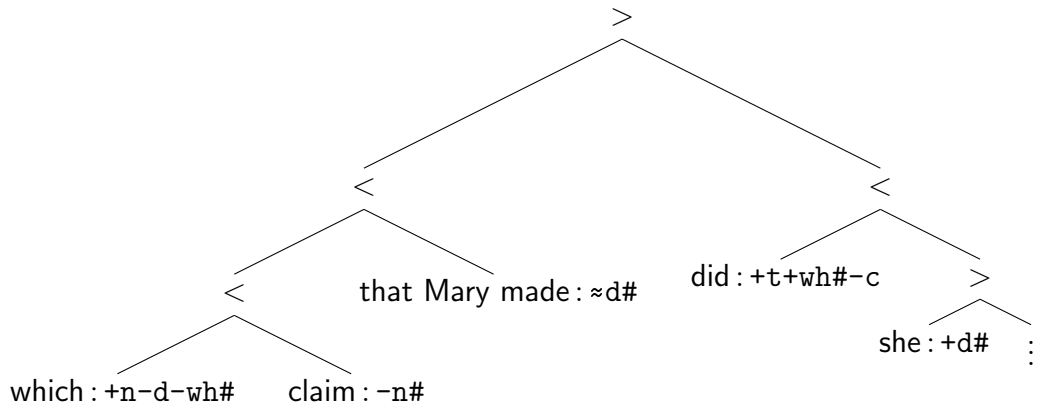
Using this mechanism it is possible to derive (2.45b) without inducing a Condition C violation, as Lebeaux suggested. First a simple question without any relative clause is derived, as shown in Figure 34a; then ADJOIN applies counter-cyclically, as in (2.47), attaching the relative clause to ‘which claim’ (the latter already having moved to the specifier of CP position), as shown in Figure 34b. I abstract away from the internal details of the relative clause, and simply follow Gärtner and Michaelis in assuming that it is an adjunct to the DP ‘which claim’,⁵⁸ represented by *that Mary made : #≈d*. Notice that just as the checked (i.e. left of ‘#’) $-V$ feature licensed ADJOIN in (2.47), the checked $-d$ feature on ‘which’ licenses adjunction to the maximal projection headed by ‘which’ to produce Figure 34b.

⁵⁷More precisely, an unchecked $-f$ feature immediately to the right of the ‘#’, i.e. the $-f$ feature must be the first unchecked feature. The counter-cyclic option, however, permits adjunction to a $-f$ feature anywhere to the left of the ‘#’.

⁵⁸This analysis is argued for by Bach and Cooper (1978). It departs from both the semantically-inspired approach (Partee 1975) where the relative clause would adjoin to ‘book’ itself (and the two denoted properties conjoined/intersected), and the “promotion” analysis (Vergnaud 1974, Kayne 1994) where the relative clause is the complement of the DP and ‘book’ raises internally to the relative clause. Chomsky (1975, pp.96–101) argues against assuming, on semantic grounds, that Partee’s structure must be correct in surface syntax; Culicover and Jackendoff (2005, pp.140–141)



(a) Immediately before the counter-cyclic adjunction step



(b) Immediately after the counter-cyclic adjunction step

Figure 34: Counter-cyclic adjunction as proposed by Gärtner and Michaelis (2003).

In a system where adjunction is achieved via a distinguished operation unrelated to other structure-building functions, there is of course no obstacle to stipulating that adjunction, alone, can apply counter-cyclically as Gärtner and Michaelis do; recall the discussion in §2.2.2. To the extent that a system of the sort I have proposed above, where adjunction is not a theoretical primitive, can account for the fact that counter-cyclic application is a possibility only for adjunction, it should be preferred. I now turn to demonstrating that my implementation of adjunction can indeed account naturally for this fact. The basic idea is as follows: if an adjunct is to be understood to modify some phrase XP, the proposal in §2.4.3 was that the adjunct can be inserted either in the phase where XP is constructed, or in the phase immediately above, say YP, in which XP is also present as a complement; but if XP is furthermore present in another phase at some point later in the derivation, say ZP, it will be possible to insert the adjunct there too.

Recall that in order to allow adjuncts to be stranded (i.e. remain unmoved when the phrase that they modify undergoes movement), I included an inert annotation indicating the category of a checked feature so that an adjunct that modifies VP can be inserted during the TP phase. Thus in the application of SPL shown in Figure 35 (illustrated earlier in Figure 33 on page 94), the adjunct can be interpreted as modifying the VP despite the fact that this is not the head of the current phase. The *V* and *k* annotations in Figure 35 are already a limited version of Gärtner and Michaelis’s proposal to keep checked features visible: when a feature of the form *-f* is checked, it effectively remains visible until the next application of SPL.

Suppose now that we were to adopt Gärtner and Michaelis’s proposal more fully, maintaining all checked features and separating them from unchecked features using the ‘#’ symbol. Then the unit produced at the end of the VP phase will be **stab Caesar: +d+d#-V** . When the *-V* feature is checked upon merging into the com-

also discuss some problems related to coordination that arise for Partee’s approach.

$$\begin{aligned} \text{SPL}(\langle -\text{ed} : -\text{t}, \text{ ____ stab Caesar} : \text{V}, \text{ Brutus} : \text{k}, \{ \text{violently} : * \text{V} \} \rangle) \\ = \langle \text{Brutus stabbed Caesar violently} : -\text{t}, \{ \} \rangle \end{aligned}$$

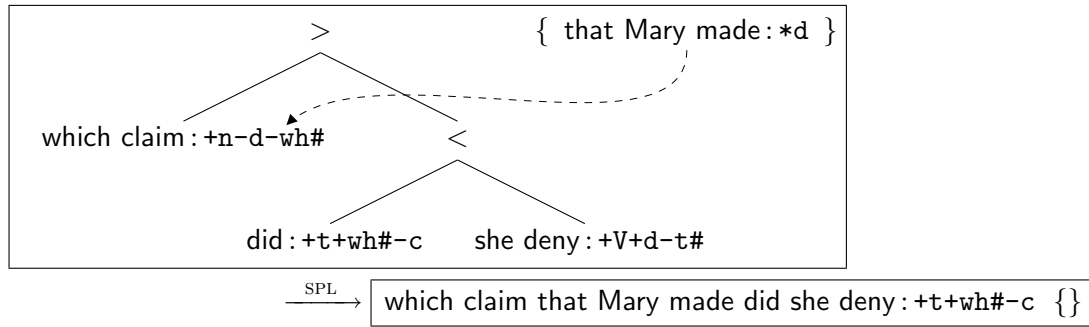
Figure 35

$$\begin{aligned} \text{SPL}(\langle -\text{ed} : +\text{V}+\text{k}\#\text{-t}, \text{ ____ stab Caesar} : +\text{d}+\text{d}-\text{V}\#, \text{ Brutus} : -\text{d}-\text{k}\#, \{ \text{violently} : \# * \text{V} \} \rangle) \\ = \langle \text{Brutus stabbed Caesar violently} : +\text{V}+\text{k}\#\text{-t}, \{ \} \rangle \end{aligned}$$

Figure 36

plement of T position, we record this in what is now the normal way by moving it to the other side of the ‘#’ symbol, yielding the feature sequence +d+d-V#. On this view the application of SPL in Figure 35 would be represented instead as in Figure 36. Here the -V feature to the left of the ‘#’ symbol is what identifies ‘stab Caesar’ as the target of adjunction.

When the “insert one phase later” option for adjuncts that I have proposed is reconceived in this way, it turns out that the possibility of counter-cyclic adjunction follows by exactly the same mechanisms. The idea suggested by Figure 36 is that when a unit bearing a feature of the form *f is in the set component of an expression to which SPL is applied, it can be interpreted as modifying either (i) the head of the expression, if it bears a -f feature that is unchecked (i.e. right of ‘#’), or (ii) an argument (of the head), if it bears a -f feature that is checked (i.e. left of ‘#’).



$$\begin{aligned} \text{SPL}(\langle \text{did: +t+wh\#-c, she deny: +V+d-t\#, which claim: +n-d-wh\#, \{ that Mary made: *d \} \rangle) \\ = \langle \text{which claim that Mary made did she deny: +t+wh\#-c, \{ \} \rangle \end{aligned}$$

Figure 37: “Counter-cyclic” adjunction, illustrated intuitively (top) and more formally (underneath). Note that the mechanisms that permit adjunction of the relative clause ‘that Mary made’ here are no different from those that permitted adjunction of ‘violently’ in Figure 33.

With this in mind consider the expression in (2.48); this corresponds to the tree in Figure 34a, representing the point in the derivation immediately before the counter-cyclic adjunction occurs.

$$(2.48) \quad \langle \text{did: +t+wh\#-c, she deny: +V+d-t\#, which claim: +n-d-wh\#, \{ \} \rangle$$

The crucial point is that the checked $-d$ feature on ‘which claim’ is visible here, and available to license adjunction, in exactly the same way that the checked $-V$ feature on ‘stab Caesar’ is in Figure 36. Thus the counter-cyclic adjunction of the relative clause is achieved by inserting it into this CP phase, as shown in Figure 37; the $*d$ feature on the relative clause identifies ‘which claim’ as its target.

To put the general idea another way, an adjunct that modifies XP can be inserted either (i) in the phase where XP is constructed (i.e. where the X head combines with its arguments) as in Figure 32, or (ii) in any other phase where XP is present,

either because XP has (first-)merged there as in Figure 33 or because XP has *re*-merged there as in Figure 37. The option of inserting the adjunct in the phase where XP is first-merged produces the “stranding” of adjuncts (in cases where XP moves away from this position, eg. the VP-fronting considered in §2.4.3), and the option of inserting the adjunct in the phase where XP is re-merged produced counter-cyclic adjunction of the sort proposed by Lebeaux.

2.6.2 Constraints on counter-cyclic adjunction

As well as covering the same basic facts in (2.45) that Gärtner and Michaelis’s formulation of counter-cyclic adjunction covers, the implementation I propose here correctly predicts a surprising condition that appears to constrain counter-cyclic adjunction.

If we allow adjuncts to attach arbitrarily late in all cases, we would predict that the relative clause ‘that Mary stole’ could be attached late to ‘book’ in (2.49) to avoid the Condition C violation, in the same way that ‘that Mary made’ was to ‘claim’ in (2.45b) (repeated below). But the Condition C violation is not avoided, so counter-cyclic adjunction here seems to be disallowed.

- (2.45) a. * Which claim [that Mary_i stole a book] did she_i deny?
 b. Which claim [that Mary_i made] did she_i deny?

- (2.49) * Which claim [that John found a book [that Mary_i stole]] did she_i deny?

The unacceptability of (2.49) is unexpected on Gärtner and Michaelis’s account: it should be possible to counter-cyclically adjoin ‘that Mary stole’ to ‘a book’ after the *wh*-movement, just as it was possible to adjoin ‘that Mary made’ to ‘which claim’ late in the derivation of (2.45b). The height of the target site ‘which claim’ in the structure in Figure 34a was not important to the licensing of the counter-cyclic adjunction, so the fact that ‘a book’ is embedded in the complement of ‘claim’ in (2.49) should not have any effect on the possibility of counter-cyclic adjunction. The facts, however, indicate that counter-cyclic adjunction to this embedded position is impossible.

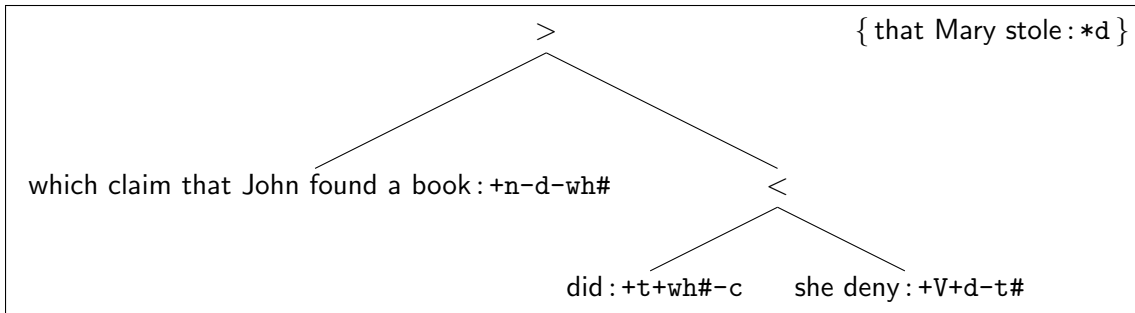


Figure 38: The relative clause ‘that Mary stole’ is unable to modify ‘a book’ here.

On the account of adjunction that I have proposed, the impossibility of counter-cyclic adjunction here follows naturally. After merging the complement and specifier of the CP phase into which we would need to insert to late-adjoined relative clause, we will have the expression in Figure 38. (Note that ‘that John found a book’ is a complement clause and therefore cannot be added counter-cyclically, just as ‘that Mary stole a book’ cannot be in (2.45a).)

Now, recall that the question to be answered is why a relative clause *can* adjoin to ‘which claim’ given the CP structure in Figure 37, but *cannot* adjoin to ‘a book’ here in Figure 38. The reason is simply that ‘a book’ is not “present” in Figure 38 in the way that ‘which claim’ is in Figure 37: a $-d$ feature originating from ‘which claim’ is visible in Figure 37, but no $-d$ feature originating from ‘a book’ is visible in Figure 38, precisely because ‘a book’ is embedded within the phrase that has re-merged into the specifier position of CP.

Note that there *is* a $-d$ feature visible in Figure 38, but it has originated from the ‘which claim ...’ phrase just as the one in (2.48) has. So while applying SPL to the expression in Figure 38 *would* have the result of adjoining ‘that Mary stole’, this relative clause would semantically modify ‘which claim ...’ exactly as in Figure 37, not ‘a book’.⁵⁹

⁵⁹Figure 38 shows no dashed line connecting ‘that Mary stole’ to the tree structure, not because the relative clause fails to be interpreted at this phase, but rather because it fails to yield the

The account of adjunction that I have proposed therefore not only dovetails with Lebeaux’s original idea of counter-cyclic adjunction (and the existing MG implementation of it by Gärtner and Michaelis), but predicts certain constraints on its applicability that appear to be empirically accurate. It seems that the constraint would be difficult to state elegantly — let alone derive — in a system like the one Gärtner and Michaelis propose, where adjunction is implemented as a theoretical primitive.

2.7 Conclusion

I have argued that some of the central syntactic differences between arguments and adjuncts are best thought of as transparent reflections of the differences between the ways they contribute to neo-Davidsonian logical forms. The account relies on (i) a restrictive theory of the modes of semantic composition that natural languages use, independently motivated by facts about the meanings that languages do and do not lexicalize, and (ii) a possibility opened up by a syntactic framework that formalizes the independently appealing idea that there should be a single feature-checking operation that is invoked at both “merge” and “move” steps. The resulting proposal provides a natural explanation for why arguments and adjuncts show the well-known differences in syntactic behaviour that they do.

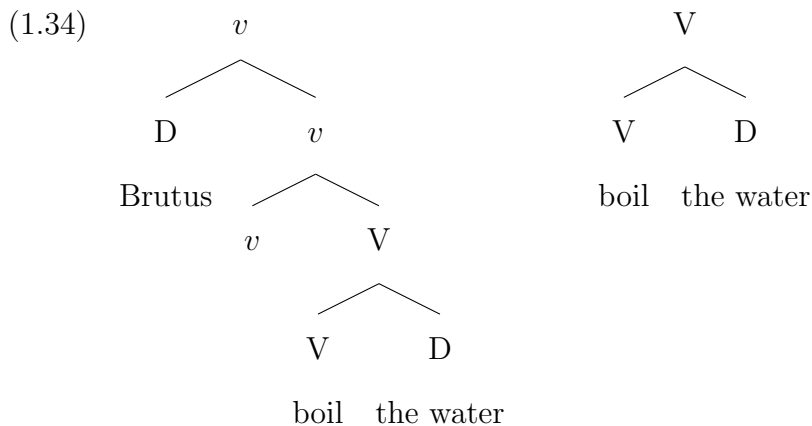
interpretation under discussion, as indicated by the square brackets in (2.49).

2.A Appendix: Structures with *v*P shells

In this appendix I will work through the details of how the implementation of adjunction proposed in §2.4 can be applied to some slightly more complex structures, namely split *v*P structures. This will (i) demonstrate that the account is not dependent on any particular properties of the simplified clause structure I assumed for ease of exposition in the body of the chapter, and (ii) emphasise that the question of which phase an adjunct is inserted into is not equivalent to that of which variable it conjunctively modifies.

I will take as examples verbs that participate in the causative/unaccusative alternation, since these are cases where there is a relatively clear semantic hypothesis about the relationship between the *v*P projection and the VP projection, as discussed briefly in §1.6.3.⁶⁰ The idea is that the implicational relationship between the sentences in (1.8) can be captured by associating them with the two structures in (1.34), so that our Conjunctivist axioms produce the two predicates in (1.35). These are rewritten in (1.37) in their more abbreviated form.

- (1.8) a. Brutus boiled the water.
 b. The water boiled.



- (1.35) $\Psi = \lambda e.v(e) \wedge \exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)] \wedge \exists e'[\Phi(e') \wedge \mathbf{Internal}(e, e')]$
 $\Phi = \lambda e.\mathbf{boiling}(e) \wedge \exists y[\mathbf{water}(y) \wedge \mathbf{Internal}(e, y)]$

$$(1.37) \quad \Psi = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\Phi) = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})) \\ \Phi = \mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})$$

Notice that in Ψ there are two event variables, one corresponding to the VP phase and one to the *v*P phase. Therefore when an adverb is added to this structure and contributes a monadic event predicate, there are two possible meanings that might result: one where the adverb's event predicate modifies the VP phase's variable, and one where it modifies the *v*P phase's variable. These two semantic possibilities, for some adverbial event predicate P , are illustrated in (2.50) and (2.51) respectively.

$$(2.50) \quad \Psi' = \lambda e. \mathbf{v}(e) \wedge \exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)] \\ \wedge \exists e'[\mathbf{boiling}(e') \ \& \ \exists y[\mathbf{water}(y) \wedge \mathbf{Internal}(e', y)] \wedge P(e')] \\ \wedge \mathbf{Internal}(e, e')] \\ = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})) \ \& \ P$$

$$(2.51) \quad \Psi'' = \lambda e. \mathbf{v}(e) \wedge \exists x[\mathbf{b}(x) \wedge \mathbf{External}(e, x)] \\ \wedge \exists e'[\mathbf{boiling}(e') \ \& \ \exists y[\mathbf{water}(y) \wedge \mathbf{Internal}(e', y)] \\ \wedge \mathbf{Internal}(e, e')] \wedge P(e) \\ = \mathbf{v} \ \& \ \mathbf{ext}(\mathbf{b}) \ \& \ \mathbf{int}(\mathbf{boiling} \ \& \ \mathbf{int}(\mathbf{water})) \ \& \ P$$

Notice now that the implication relation from the causative/transitive including an adverb to the corresponding unaccusative/intransitive (still containing the adverb) only arises if the adverb modifies the VP phase's event variable, since we assume that this is the only event variable in the unaccusative/intransitive sentence. In other words, the implication only arises if the causative/transitive is constructed as shown in (2.50). Therefore to the extent that the implication holds in (2.52) but not in

⁶⁰I leave open the significant question of how structures containing both *v*P and VP projections are interpreted when there is not obviously a two-event semantic analysis, as in the case of many “standard” transitive verbs (eg. ‘stab’). If there is genuinely only one event variable in these sentences, as assumed in the body of this chapter, we will be forced to either reject the strict one-variable-per-projection assumption, or deny that *v*P is present (contra standard assumptions).

(2.53), we would be justified in supposing that adverbs (and adjuncts more generally) can indeed modify either of these two variables: ‘loudly’ modifies the VP phase’s variable in (2.52) (on the relevant reading, at least), and ‘deliberately’ modifies the *v*P phase’s variable in (2.53).⁶¹

(2.52) a. Brutus boiled the water loudly

b. The water boiled loudly

(2.53) a. Brutus boiled the water deliberately

b. The water boiled deliberately

Then the event predicates corresponding to (the *v*P constituent in) (2.52a) and (2.53a) will be (2.54) and (2.55), respectively.

(2.54) $v \ \& \ \text{ext}(\mathbf{b}) \ \& \ \text{int}(\mathbf{boiling} \ \& \ \text{int}(\mathbf{water}) \ \& \ \mathbf{loud})$

(2.55) $v \ \& \ \text{ext}(\mathbf{b}) \ \& \ \text{int}(\mathbf{boiling} \ \& \ \text{int}(\mathbf{water})) \ \& \ \mathbf{delib}$

The flexibility to be *inserted* into the derivation “early” or “late”, which the account of adjunction in §2.4 made use of, is unrelated to this distinction between *semantically modifying* the higher or lower event variable. This should be clear because the two possibilities for *insertion* into the derivation did not produce any difference in the semantic result in §2.4. Each of the adjuncts in (2.52) and (2.53) has the flexibility to be inserted into the derivation in either of two phases: ‘loudly’ can modify the VP phase’s variable by being inserted in the VP phase or the *v*P phase, and ‘deliberately’ can modify the *v*P phase’s variable by being inserted in the *v*P phase or the TP phase. Generally: if an adjunct has a meaning that conjunctively modifies the head of a certain projection XP, it can be inserted either in the XP phase or in the phase immediately above XP.

Below I give derivations illustrating each of these four options: two possible insertion points for each of the two possible targets of modification. Each derivation

⁶¹I leave aside questions of whether/how a particular adverb can modify different variables in different sentences.

proceeds up to the completion of the TP phase, and to reduce clutter I show units' semantic components only at the end of each phase. Adapting the use of the '**F**' placeholder slightly, I now assume that the meaning resulting from composing a T head meaning ϕ_T with its complement *v*P meaning ϕ_{vP} is **F** $[\phi_{vP}, \phi_T]$.

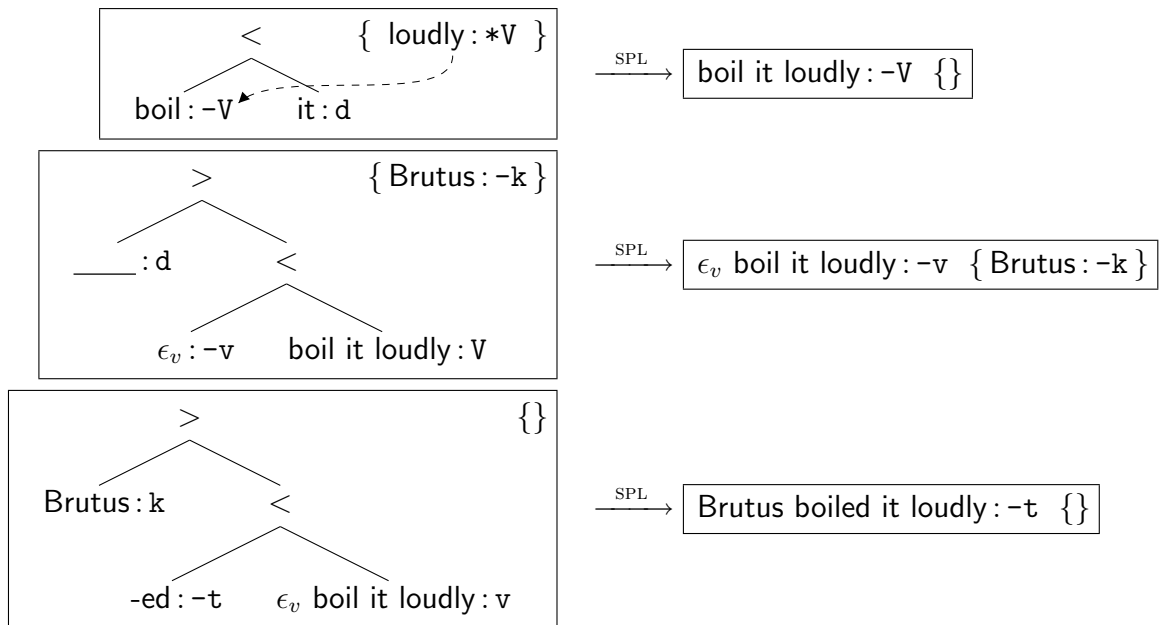
Purely for reasons of (horizontal) space, I use 'it' instead of 'the water'. For now, the meaning of 'it' I write simply as '**it**' (though a more thorough treatment of indexicals and pronouns is given in chapter 4). I also abbreviate 'deliberately' with **delib**.

The relevant lexical items are given in (2.56). Notice the differing adjunction features on 'loudly' and 'deliberately', because of their differing targets of semantic modification. Some other straightforward changes from the lexical items used in the body of the chapter are also necessary because of the addition of the *v*P layer: a *v* lexical item is added, of course, which selects a VP complement and then a DP specifier to be assigned an external theta role; and correspondingly, 'boil' now only selects one DP, and the T head selects a *v*P, rather than VP, complement.

	$l_{\text{Brutus}} = \langle \text{Brutus} : -\mathbf{d-k}, \{\} \rangle$	$l_{\text{boil}} = \langle \text{boil} : +\mathbf{d-V}, \{\} \rangle$
(2.56)	$l_{\text{it}} = \langle \text{it} : -\mathbf{d}, \{\} \rangle$	$l_v = \langle \epsilon_v : +\mathbf{V+d-v}, \{\} \rangle$
	$l_{\text{loudly}} = \langle \text{loudly} : *V, \{\} \rangle$	$l_T = \langle -\mathbf{ed} : +\mathbf{v+k-t}, \{\} \rangle$
	$l_{\text{delib}} = \langle \text{delib} : *v, \{\} \rangle$	

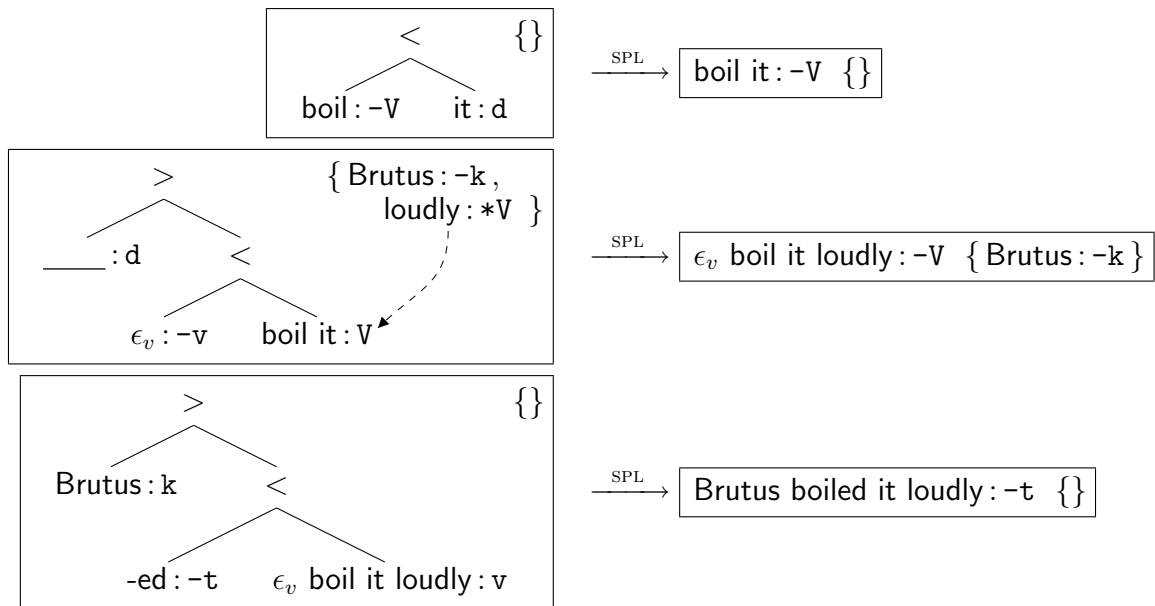
VP-modifying adjunct inserted in the VP phase

$e_1 = \text{INS}(\ell_{\text{boil}}, \ell_{\text{it}})$	$= \langle \text{boil} : +\text{d}-\text{V}, \{ \text{it} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{boil} : -\text{V}, \text{it} : \text{d}, \{ \} \rangle$
$e'_2 = \text{INS}(e_2, \ell_{\text{loudly}})$	$= \langle \text{boil} : -\text{V}, \text{it} : \text{d}, \{ \text{loudly} : *V \} \rangle$
$e_3 = \text{SPL}(e'_2)$	$= \langle \text{[boiling\&int(it)\&loud]} : -\text{V}, \{ \} \rangle$
$e_4 = \text{INS}(\ell_v, e_3)$	$= \langle \epsilon_v : +\text{V}+\text{d}-\text{v}, \{ \text{boil it loudly} : -\text{V} \} \rangle$
$e_5 = \text{MRG}(e_4)$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it loudly} : \text{V}, \{ \} \rangle$
$e_6 = \text{INS}(e_5, \ell_{\text{Brutus}})$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it loudly} : \text{V}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \epsilon_v : -\text{v}, \text{boil it loudly} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_8 = \text{SPL}(e_7)$	$= \langle \text{[v\&int(boiling\&int(it)\&loud)\&ext(b)]} : -\text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_9 = \text{INS}(\ell_T, e_8)$	$= \langle -\text{ed} : +\text{v}+\text{k}-\text{t}, \{ \epsilon_v \text{ boil it loudly} : -\text{v}, \text{Brutus} : -\text{k} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \epsilon_v \text{ boil it loudly} : \text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle -\text{ed} : -\text{t}, \epsilon_v \text{ boil it loudly} : \text{v}, \text{Brutus} : \text{k}, \{ \} \rangle$
$e_{12} = \text{SPL}(e_{11})$	$= \langle \text{[F[v\&int(boiling\&int(it)\&loud)\&ext(b),\phi_T]}] : -\text{t}, \{ \} \rangle$



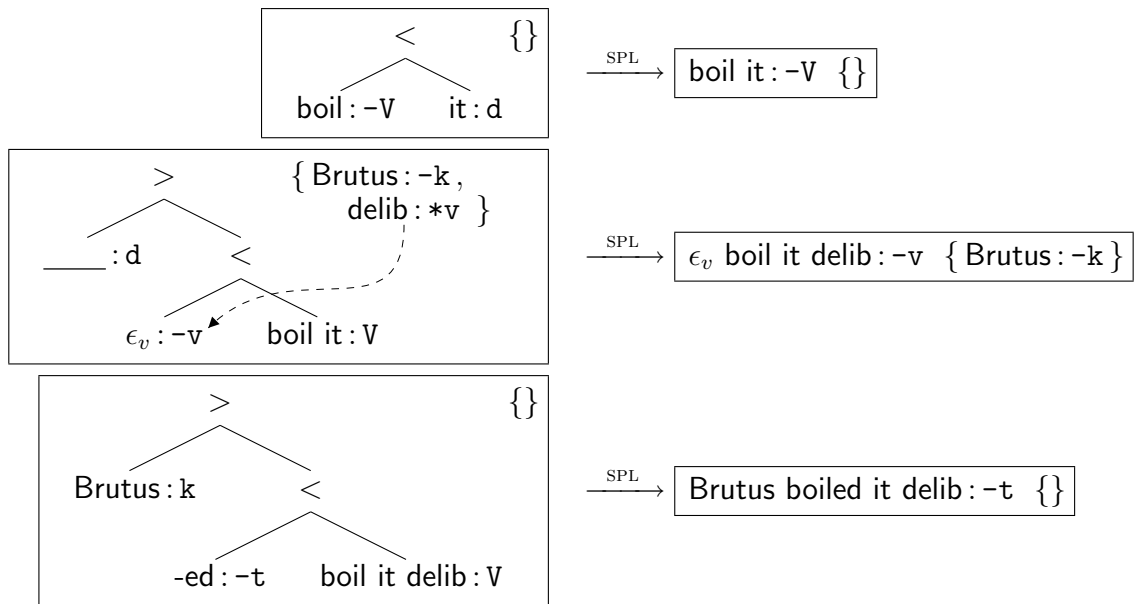
VP-modifying adjunct inserted in the *vP* phase

$e_1 = \text{INS}(l_{\text{boil}}, l_{\text{it}})$	$= \langle \text{boil} : +\text{d}-\text{V}, \{ \text{it} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{boil} : -\text{V}, \text{it} : \text{d}, \{ \} \rangle$
$e_3 = \text{SPL}(e_2)$	$= \langle \left[\begin{array}{c} \text{boil it} \\ \text{boiling\&int(it)} \end{array} \right] : -\text{V}, \{ \} \rangle$
$e_4 = \text{INS}(l_v, e_3)$	$= \langle \epsilon_v : +\text{V}+\text{d}-\text{v}, \{ \text{boil it} : -\text{V} \} \rangle$
$e_5 = \text{MRG}(e_4)$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \} \rangle$
$e_6 = \text{INS}(e_5, l_{\text{Brutus}})$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \epsilon_v : -\text{v}, \text{boil it} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e'_7 = \text{INS}(e_7, l_{\text{loudly}})$	$= \langle \epsilon_v : -\text{v}, \text{boil it} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k}, \text{loudly} : *V \} \rangle$
$e_8 = \text{SPL}(e'_7)$	$= \langle \left[\begin{array}{c} \epsilon_v \text{ boil it loudly} \\ \text{v\&int(boiling\&int(it)\&loud)\&ext(b)} \end{array} \right] : -\text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_9 = \text{INS}(l_{\text{T}}, e_8)$	$= \langle -\text{ed} : +\text{v}+\text{k}-\text{t}, \{ \epsilon_v \text{ boil it loudly} : -\text{v}, \text{Brutus} : -\text{k} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \epsilon_v \text{ boil it loudly} : \text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle -\text{ed} : -\text{t}, \epsilon_v \text{ boil it loudly} : \text{v}, \text{Brutus} : \text{k}, \{ \} \rangle$
$e_{12} = \text{SPL}(e_{11})$	$= \langle \left[\begin{array}{c} \text{Brutus boiled it loudly} \\ \text{F[v\&int(boiling\&int(it)\&loud)\&ext(b),\phi_{\text{T}}]} \end{array} \right] : -\text{t}, \{ \} \rangle$



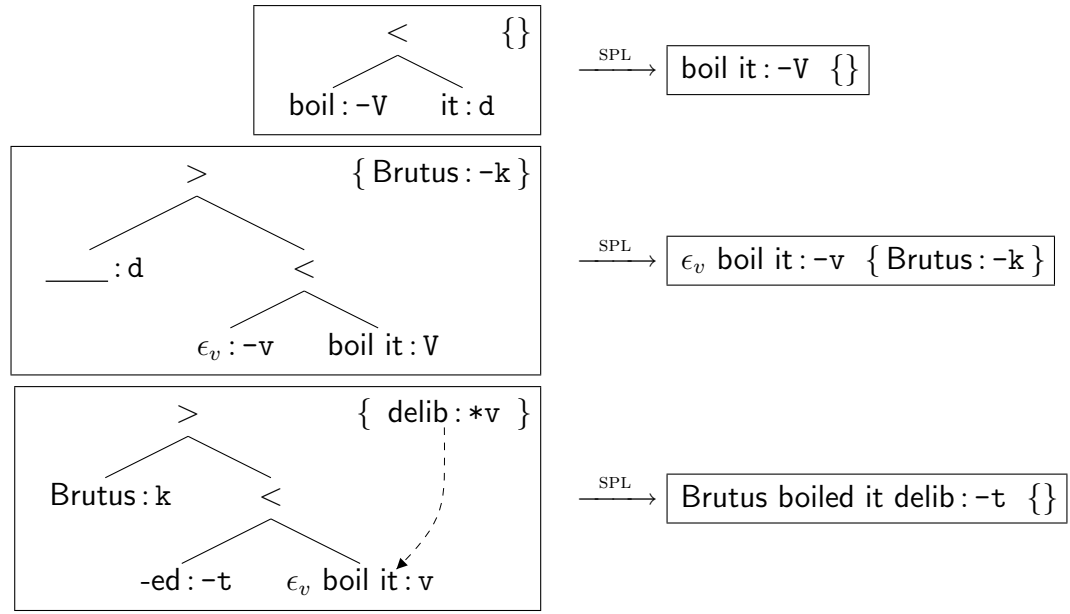
vP-modifying adjunct inserted in the *vP* phase

$e_1 = \text{INS}(l_{\text{boil}}, l_{\text{it}})$	$= \langle \text{boil} : +\text{d}-\text{V}, \{ \text{it} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{boil} : -\text{V}, \text{it} : \text{d}, \{ \} \rangle$
$e_3 = \text{SPL}(e_2)$	$= \langle \left[\begin{array}{c} \text{boil it} \\ \text{boiling\&int(it)} \end{array} \right] : -\text{V}, \{ \} \rangle$
$e_4 = \text{INS}(l_v, e_3)$	$= \langle \epsilon_v : +\text{V}+\text{d}-\text{v}, \{ \text{boil it} : -\text{V} \} \rangle$
$e_5 = \text{MRG}(e_4)$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \} \rangle$
$e_6 = \text{INS}(e_5, l_{\text{Brutus}})$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \epsilon_v : -\text{v}, \text{boil it} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e'_7 = \text{INS}(e_7, l_{\text{delib}})$	$= \langle \epsilon_v : -\text{v}, \text{boil it} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k}, \text{delib} : *v \} \rangle$
$e_8 = \text{SPL}(e'_7)$	$= \langle \left[\begin{array}{c} \epsilon_v \text{ boil it delib} \\ \text{v\&int(boiling\&int(it))\&ext(b)\&delib} \end{array} \right] : -\text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_9 = \text{INS}(l_{\text{T}}, e_8)$	$= \langle -\text{ed} : +\text{v}+\text{k}-\text{t}, \{ \epsilon_v \text{ boil it delib} : -\text{v}, \text{Brutus} : -\text{k} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \epsilon_v \text{ boil it delib} : \text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle -\text{ed} : -\text{t}, \epsilon_v \text{ boil it delib} : \text{v}, \text{Brutus} : \text{k}, \{ \} \rangle$
$e_{12} = \text{SPL}(e_{11})$	$= \langle \left[\begin{array}{c} \text{Brutus boiled it delib} \\ \text{F[v\&int(boiling\&int(it))\&ext(b)\&delib, \phi_{\text{T}}]} \end{array} \right] : -\text{t}, \{ \} \rangle$



*v*P-modifying adjunct inserted in the TP phase

$e_1 = \text{INS}(l_{\text{boil}}, l_{\text{it}})$	$= \langle \text{boil} : +\text{d}-\text{V}, \{ \text{it} : -\text{d} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{boil} : -\text{V}, \text{it} : \text{d}, \{ \} \rangle$
$e_3 = \text{SPL}(e_2)$	$= \langle \left[\begin{array}{c} \text{boil it} \\ \text{boiling\&int(it)} \end{array} \right] : -\text{V}, \{ \} \rangle$
$e_4 = \text{INS}(l_v, e_3)$	$= \langle \epsilon_v : +\text{V}+\text{d}-\text{v}, \{ \text{boil it} : -\text{V} \} \rangle$
$e_5 = \text{MRG}(e_4)$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \} \rangle$
$e_6 = \text{INS}(e_5, l_{\text{Brutus}})$	$= \langle \epsilon_v : +\text{d}-\text{v}, \text{boil it} : \text{V}, \{ \text{Brutus} : -\text{d}-\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \epsilon_v : -\text{v}, \text{boil it} : \text{V}, \text{---} : \text{d}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_8 = \text{SPL}(e_7)$	$= \langle \left[\begin{array}{c} \epsilon_v \text{ boil it} \\ v\&int(\text{boiling\&int(it)})\&ext(\text{b}) \end{array} \right] : -\text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_9 = \text{INS}(l_T, e_8)$	$= \langle -\text{ed} : +\text{v}+\text{k}-\text{t}, \{ \epsilon_v \text{ boil it} : -\text{v}, \text{Brutus} : -\text{k} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \epsilon_v \text{ boil it} : \text{v}, \{ \text{Brutus} : -\text{k} \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle -\text{ed} : -\text{t}, \epsilon_v \text{ boil it} : \text{v}, \text{Brutus} : \text{k}, \{ \} \rangle$
$e'_{11} = \text{INS}(e_{11}, l_{\text{delib}})$	$= \langle -\text{ed} : -\text{t}, \epsilon_v \text{ boil it} : \text{v}, \text{Brutus} : \text{k}, \{ \text{delib} : *v \} \rangle$
$e_{12} = \text{SPL}(e'_{11})$	$= \langle \left[\begin{array}{c} \text{Brutus boiled it delib} \\ F[v\&int(\text{boiling\&int(it)})\&ext(\text{b})\&delib, \phi_T] \end{array} \right] : -\text{t}, \{ \} \rangle$



Chapter 3

Adjunct Islands and Freezing Effects

3.1 Overview

The aim of this chapter is to show that the framework developed in chapter 2 can provide a unified account of two well-known conditions on extraction domains: the “adjunct island” effect and “freezing” effects. Descriptively speaking, extraction is generally problematic out of adjoined constituents and out of constituents that have moved. I note that it emerges from the framework independently motivated in chapter 2 that adjoined constituents are relevantly like constituents that have moved, unifying the two descriptive generalisations noted above.

Existing work in the MG formalism on conditions on extraction has looked at the impact on generative capacity of various combinations of island/locality constraints (Gärtner and Michaelis 2005, 2007). In such work, familiar descriptive constraints are simply stipulated individually (eg. the specifier island constraint, adjunct island constraint), the focus being not on the nature of the constraints but their effects; in this respect it differs from my goal here, which is to reduce two existing constraints to one.

The rest of the chapter is organised as follows. In §3.2 I outline previous accounts of the facts I aim to explain. In §3.3 I demonstrate how the system formulated in chapter 2 lets us introduce a single constraint that subsumes both adjunct island

effects and freezing effects. I explore an additional predicted constraint on remnant movement in §3.4, before concluding in §3.5.

3.2 Previous accounts of adjunct islands and freezing effects

In this section I review some of the influential accounts of adjunct island effects and freezing effects. As we will see, the histories of the two kinds of effects are somewhat intertwined, perhaps because they overlap to a reasonably large degree in the cases they cover: moved constituents from which extraction might *a priori* be possible have very frequently been moved to an adjoined position. But I have not found an instance of a complete reduction of the two to a single constraint in the existing literature.

Perspectives on adjunct islands in the literature can be broadly thought of as belonging to two groups: first, those based on the idea that adjuncts are somehow non-canonical in a sense that distinguishes them (possibly along with non-base structures) from both complements and specifiers; and second, those based on the idea that adjuncts are islands for roughly the same reasons that specifiers or left-branches are often islands. The proposal I present in this chapter is a strong version of the first idea.

3.2.1 Early work: non-canonical structures

Ross (1974) introduces the “Immediate Self-Domination Principle” (ISP), given in (3.1).

(3.1) The Immediate Self-Domination Principle (Ross 1974, p.102)

No element may be chopped out of a node which immediately dominates another node of the same type.

Ross takes this to account for the two kinds of violations exemplified by (3.2) and (3.3), which were previously (Ross 1969) accounted for separately by the Complex NP

Constraint⁶² (CNPC) and the Coordinate Structure Constraint (CSC), respectively.

- (3.2) a. She will execute [_{NP} [_{NP} anyone] [_S who mentions these points]]
b. * These points she will execute [_{NP} [_{NP} anyone] [_S who mentions ~~these points~~]]
- (3.3) a. We will take up [_{NP} either [_{NP} his proposals] or [_{NP} these points]] next
b. * These points we will take up [_{NP} either [_{NP} his proposals] or [_{NP} ~~these points~~]] next

The NPs ‘anyone who mentions these points’ and ‘either his proposals or these points’ immediately dominate one or more other NP nodes, so chopping ‘these points’ from either of these NPs violates the ISP.

Aside from the constructions in (3.2) and (3.3), the majority of the examples that Ross treats with the ISP are instances of movement to an adjoined position, where extraction from the moved (and thus, adjoined) constituent, is forbidden. For example, (3.4) shows the effect of attempting to extract from a heavy NP which has undergone “NP Shift”.

- (3.4) a. [_S [_S She will send to Inspector Smithers] [_{NP} a picture of the Waco Post Office]]
b. * The Waco Post Office [_S [_S she will send to Inspector Smithers] [_{NP} a picture of ~~the Waco Post Office~~]]

In general, the ISP will forbid extraction from any adjunct, so it captures the phenomenon that has come to be known as “adjunct island effects”.

In discussing (3.4) Ross foreshadows later trends by noting that “it is more difficult to chop constituents from the shifted constituent than it is to chop them from an unshifted one”, but for him the movement transformation itself is not what causes

⁶²Ross (1974, fn.36) notes that NPs with complement clauses may not have the structure shown in (3.2), in which case the ISP can only account for the relative clause cases of the CNPC.

problems for extraction; no movement contributes to the prohibition of the extractions in (3.2) or (3.3), for example. Wexler and Culicover (1981) introduce the “Generalised Freezing Principle” (GFP), given in (3.5), according to which the extraction in (3.4) is ruled out precisely because the shifted NP has been moved, and is therefore frozen. (Extraction is not possible out of a node which is “frozen”.)

(3.5) The Generalised Freezing Principle (Wexler and Culicover 1981, p.540)

A node is frozen if (i) its immediate structure is non-base, or (ii) it has been raised.

The ISP and the GFP cover, to a large extent, the same empirical ground. Both rule out cases such as (3.4) involving movement to an adjoined position.⁶³ What the ISP rules out that the GFP does not are cases like (3.2) and (3.3) where self-domination structures (a subcase being adjunction structures) are base-generated. What the GFP rules out that the ISP does not are cases of extraction from constituents that have been moved to non-adjoined positions, such as (3.6) where extraction of ‘who’ is degraded when it is contained in a larger phrase that has undergone *wh*-movement to a specifier of CP (Johnson 1986, p.75).⁶⁴

- (3.6) a. Who do you think I had told [a story about ~~who~~]?
 b. * Who do you wonder [which story about ~~who~~] I had told [~~which story about who~~]?

An intuition behind Wexler and Culicover’s proposal is revealed when they write that “we may think of the base grammar as providing *characteristic structures* of the language. Transformations sometimes distort these structures, but only these charac-

⁶³In fact, they also both rule out a number of derivations that are not typically considered under the heading of “freezing effects” anymore: in both cases, the node *dominating* (the new position of) the moved constituent is opaque for extraction.

⁶⁴Sentences of this form are usually attributed to Torrego (1985), who reports that extraction from *wh*-moved phrases is possible in Spanish. Particularly puzzling is that, according to Torrego, extraction from *wh*-moved subjects is no worse than extraction from *wh*-moved objects. See Corver (2005, fn.16), Lasnik and Saito (1992, p.102) and Chomsky (1986, pp.25–26) for some discussion.

teristic structures may be affected by transformations.” To the extent that adjunction structures in general have recently come to be thought of as non-base, it is tempting to think of freezing effects and adjunct island effects as two instances of a general prohibition on extraction from such structures. Chametzky (1996) explicitly follows this line of thought, referring to structures that either arise from transformations or involve adjunction configurations as instances of “non-canonical phrase structure”. We will also see below that Stepanov (2001, 2007) proposes that movement is forbidden out of (i) adjoined constituents, because of a particular view of adjuncts as, in a certain sense, “non-base” pieces of structure, and (ii) moved constituents; these two types of domains seem to echo the intuition that Wexler and Culicover (1981) and Chametzky (1996) appeal to, although Stepanov explicitly considers these to be two distinct kinds of effects.

3.2.2 The complement/non-complement distinction

Huang (1982), following ideas from Cattell (1976), proposes a unified “Condition on Extraction Domains” (CED), given in (3.7), to account for the pattern in (3.8).

(3.7) The Condition on Extraction Domains (Huang 1982, p.505)

A phrase *A* may be extracted out of a domain *B* only if *B* is properly governed.

- (3.8) a. What did you hear [a story about ~~what~~]?
 b. * What did [a story about ~~what~~] surprise you?
 c. * What did you cry [when you heard a story about ~~what~~]?

The details of proper government are not crucial to present purposes; suffice it to say that the CED aims to capture “a general asymmetry between complements on the one hand and non-complements (subjects and adjuncts) on the other” (Huang 1982, p.503). The complement ‘a story about (what)’ in (3.8a) is properly governed

(specifically, lexically governed by ‘hear’), so the extraction of ‘what’ is permitted. In contrast, the subject ‘a story about (what)’ in (3.8b), and the adjunct ‘when you heard a story about (what)’ in (3.8c), are not properly governed, so extraction from these domains is not allowed.⁶⁵ Thus Huang takes the notion of proper government, which receives some independent motivation from ECP effects, to unify “subject island effects”, previously accounted for by the Sentential Subject Constraint (Ross 1969) and then the more general Subject Condition (Chomsky 1973), and “adjunct island effects”, a subcase of the ISP discussed above.

The general idea of a distinction between complements and non-complements carries on through much of the GB literature (eg. Chomsky 1986, Lasnik and Saito 1992). Uriagereka (1999) aims to derive this distinction, which the CED comes close to stipulating, from minimalist assumptions about structure-building (see also Nunes and Uriagereka 2000). The proposal has two basic ingredients: first, it adopts the idea from Kayne (1994) that linear precedence is defined in terms of asymmetric c-command; and second, it rejects the idea from Chomsky (1995) that “spell-out” applies only once in the course of a derivation, and instead assumes that it can apply multiple times. From these premises, Uriagereka derives a distinction between complements and non-complements from which it plausibly follows that extraction is possible only out of the former.

Consider first the case of a simple specifier being selected, as indicated in Figure 39. In the resulting structure, the assumption that asymmetric c-command corresponds to precedence determines, for each terminal in the specifier (there happens to be only one), its linear ordering with respect to the terminals in the V' : ‘it’ asymmetrically

⁶⁵The contrast between (3.8a) and (3.8c) can not be attributed to the degree of embedding, because the extraction site in the following sentence is as embedded as that in (3.8c) and yet the result is as good as (3.8a). Therefore it must be the structural relationship between the matrix verb and the embedded clause in (3.8c) that is problematic.

- (i) What did you say that you heard a story about ~~what~~?

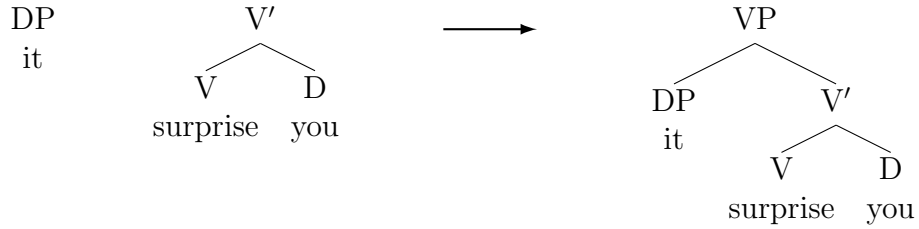


Figure 39

c-commands both ‘surprise’ and ‘you’.⁶⁶

Now consider the case of a complex specifier, as shown in Figure 40. In the resulting structure, there are a number of pairs of terminals, of which neither c-commands the other; ‘story’ and ‘surprise’, to take one example. By hypothesis, this structure is therefore unlinearisable.⁶⁷

The solution that Uriagereka proposes is to “spell-out” the complex specifier before it is merged. The ordering of the terminals *inside* the specifier is well-defined, just as the ordering of the terminals in the result of Figure 39 was, so we can linearise them and treat the result as a single terminal. Then the result of merging this “large” specifier, treated as a single terminal, as shown in Figure 41, is not formally different from the case in Figure 39; the problems raised by Figure 40 are avoided.

Consider now the consequences of this proposal for our understanding of extraction domains. We now know that any attempted derivation of (3.8b) must proceed as shown in Figure 41, rather than Figure 40. Since ‘who’ does not exist as an independent element in Figure 41, it is not possible to move or re-merge it into a specifier position of the matrix C; to do this would require a structure where ‘who’ remains an independent element, such as Figure 40, but this, as we have seen, is unlinearisable.

⁶⁶I put aside the question of how the linear ordering of two *mutually* c-commanding terminals, such as ‘surprise’ and ‘you’, is determined.

⁶⁷If we adopt the LCA in full from Kayne (1994), then ‘story’ and ‘surprise’ are linearised “indirectly” in virtue of the fact that the complex DP dominates ‘story’ and asymmetrically c-commands ‘surprise’. But Uriagereka aims to deduce this “induction step” of the LCA, taking only the “base step” as a premise.

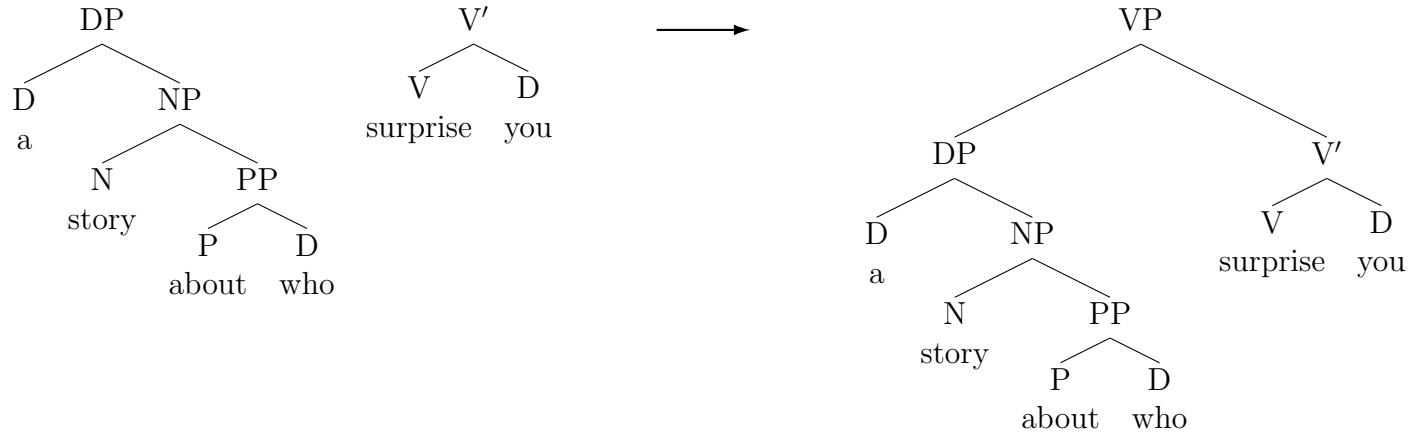


Figure 40

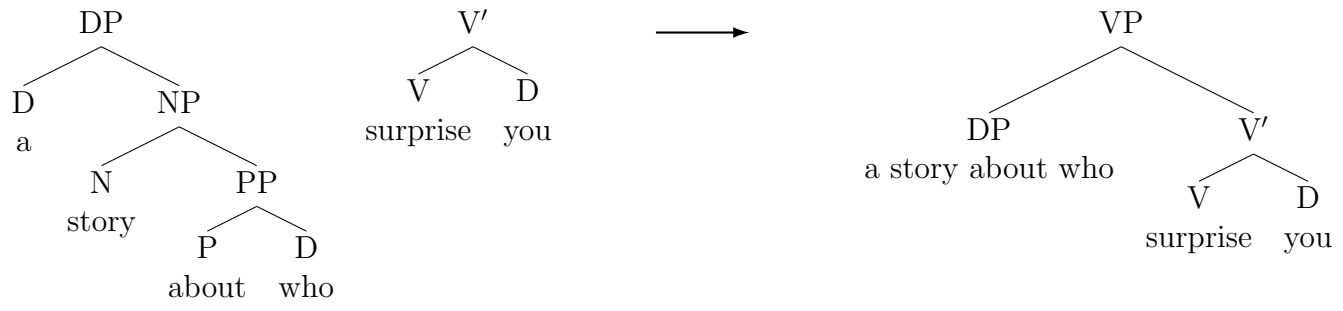


Figure 41

So there is no derivation where ‘who’ has been extracted from the complex specifier that does not cause problems for linearisation.⁶⁸

This result generalises to all specifiers and adjuncts by exactly the same logic.⁶⁹ (The fact that the result of combining DP with V’ was a VP, rather than another V’, played no role in the argument.) All specifiers and adjuncts must be “spelled out” before they are attached, making them islands. Complements need not be spelled out, however, because leaving their structure intact does not create any linearisation problems: notice that ‘who’ is embedded inside two complements in the complex specifier in Figure 41, but this does not hinder the ordering of the terminals as ‘a story about who’. Thus complements are predicted to permit extraction, while specifiers and adjuncts are not.

3.2.3 Subject islands as freezing effects

Stepanov (2001, 2007) argues against unifying subject and adjunct island effects on the basis of the complement/non-complement distinction. This argument rests mainly on the observation that the prohibition on extraction from adjuncts is much more robust, cross-linguistically, than that on extraction from subjects. Stepanov (2007, pp.89–91) provides data from a number of languages where extraction from subjects is acceptable but extraction from adjuncts is not. This is clearly unexpected under any approach that centres around the complement/non-complement distinction — whether derived, in the style of Uriagereka (1999) and Nunes and Uriagereka (2000),

⁶⁸This requires the additional assumption that constituents which have been “spelled out” cannot project — only “genuine” lexical items can — to rule out the alternative derivation where the V’ is linearised instead of the specifier DP, permitting extraction of ‘who’. (Of course, to the extent that asymmetric c-command corresponds to linear precedence, this alternative derivation would not produce a string where the DP precedes the V’; but see footnote 69)

⁶⁹Strictly speaking, it generalises to right adjuncts only if we adopt a looser interpretation of the significance of asymmetric c-command: specifically, we require that asymmetric c-command is a total order on the set of terminals, without stating exactly how this total order relates to linear precedence. This is in line with the fact that in more recent work, Uriagereka (2008) has referred more to the significance of creating *some* “one-dimensional” structure via the asymmetric c-command relation, and focussed less directly on linear order.

or stated more directly, in the style of Huang (1982).

Since the introduction of the “VP-internal” subject hypothesis (Koopman and Sportiche 1991), it has been possible to consider subject island effects as an instance of freezing effects in exactly the sense of Wexler and Culicover (1981) — at least, in cases where a subject has moved from its thematic positions to (say) specifier of TP.⁷⁰ Stepanov points out that many of the subjects from which extraction is possible are plausibly “in-situ”, in the sense that they have remained in their thematic position (say, specifier of *v*P). Thus extraction from the German subject in (3.9a) is possible for exactly the same reason that extraction from the English (would-be) subject in (3.10a) is possible, namely, both are in their base-generated positions, in contrast to the subjects in (3.9b) and (3.10b) (see Lasnik and Park 2003).

- (3.9) a. Was haben denn [_{vP} [was für Ameisen] einen Postbeamten
 what have PRT for ants a postman
 gebissen] ?
 bitten
 What kind of ants bit the postman?
- b. * Was haben [was für Ameisen] denn [_{vP} ~~was für Ameisen~~ einen
 what have for ants PRT a
 Postbeamten gebissen] ?
 postman bitten
 What kind of ants bit the postman?
- (3.10) a. Who is there [_{vP} [a picture of ~~who~~] on the wall] ?
- b. * Who is a [picture of ~~who~~] [_{vP} [~~a picture of who~~] on the wall] ?

Stepanov therefore proposes that subject island effects be totally reduced to freez-

⁷⁰I haven’t been able to determine exactly where this observation was first made. Many sources (Corver 2005, Müller 1998) cite Lasnik and Saito (1992), who relate the effect to the fact that canonical subjects in languages like English are not in a theta-marked position at s-structure, but as far as I can tell do not relate it to the *movement* of the subject. The matter is complicated somewhat by the fact that Stepanov (2001, 2007) follows Takahashi (1994) in expressing the constraint (that extraction from moved constituents is disallowed) in terms of chain uniformity rather than using the “freezing” terminology of Wexler and Culicover (1981), although Stepanov does note the connection. Perhaps the idea of relating the islandhood of canonical subjects to the fact that they have moved originated with Takahashi (1994), but he does not mention “freezing” or Wexler and Culicover either.

ing effects, with adjunct island effects being a completely different phenomenon. The crucial property of adjuncts that makes them islands for extraction, he argues, is that adjuncts are added to phrase markers “acyclically — that is, independently of the main cycle of the derivation” (Stepanov 2007, p.110). This follows on from the suggestion of Lebeaux (1988, 2000), discussed in §2.6: Stepanov strengthens Lebeaux’s suggestion, that adjuncts *may* be attached late, to a requirement that adjuncts *must* be attached late, after all substitution operations have been completed. This rules out (3.8c) since once the adjunct ‘when you heard a story about what’ is attached, no further substitution operations — of which the movement of ‘what’ to the specifier of CP would be one — are allowed.

(3.8c) *What did you cry [when you heard a story about ~~what~~]?

It is crucial for Stepanov that this prohibition of extraction from adjuncts, revolving around the late-attachment mechanism, is completely independent of his prohibition of extraction from (many) subjects, revolving around freezing (stated in terms of chain uniformity). But the property of adjuncts that he exploits is that they are “non-base” or “non-canonical” structures, just as moved (frozen) constituents are, making the theory reminiscent of the intuitions behind Wexler and Culicover (1981) and Chametzky (1996). This observation is not intended to imply that Stepanov has failed to separate the effects that he argues, on empirical grounds, should be separated. One can of course stipulate that two constraints are independent, no matter how easy it might be to unify them. But to the extent that the empirical grounds for separation might be called into question, it is interesting to note the potential for unification in the theoretical component of Stepanov’s work.

Jurka (2009) has argued against Stepanov’s claim that subject island effects are *completely* reducible to freezing effects, showing that the decrease in acceptability caused by extraction from subjects is greater than the decrease in acceptability caused by extraction from moved constituents in general. This suggests that extraction from

subjects is problematic for some reason beyond the fact that they have moved, as predicted by approaches like that of Uriagereka (1999). But more importantly for present purposes, Jurka did find that extraction from moved subjects is worse than extraction from in-situ subjects, suggesting that canonical subject island effects are actually the result of two independent constraints: one concerning extraction from left-branches, and one concerning extraction from moved constituents. In this chapter I address only the latter of these two constraints — “subject island effects” will not be important except insofar as they constitute examples of freezing effects. In particular, note that Stepanov’s typological observation, that subject islandhood and adjunct islandhood do not pattern together crosslinguistically, does not imply that freezing effects and adjunct islandhood do not pattern together crosslinguistically. Stepanov argues that adjuncts are uniformly islands, and that subjects are islands precisely when they move; this is completely consistent with the possibility that adjunct islandhood and freezing effects pattern together crosslinguistically.

In other cases, however, Jurka does not find any decrease in acceptability as a result of movement of a constituent (for example, extraction from a scrambled object). This is problematic for the assumption I make in this chapter that extraction from a moved constituent is always disallowed.⁷¹ Similarly, there have been claims that adjuncts are not, as Stepanov assumes, universally opaque for extraction (Cinque 1990, Yoshida 2006, Truswell 2007). I do not intend to advance these debates in this chapter.⁷² Rather, my aim is to investigate the formal relationship between the two constraints and explore a potential connection between them, so that the lay of the theoretical land on which these debates take place is better understood.

⁷¹As usual, saying that an operation is “always disallowed” should be understood to mean that it is “never part of a completely well-formed derivation”, thus producing *some* decrease in grammaticality whenever it applies.

⁷²I note briefly, however, two interesting possibilities for making contact with these empirical debates.

First, Truswell (2007) proposes a generalisation concerning which adjuncts do and do not permit extraction on the basis of the adjunct’s role in Davidsonian event-based logical forms, which also motivate the syntax of adjuncts that I adopt. Furthermore, the adjuncts that Truswell’s generalisa-

$$\begin{array}{c}
\langle \text{buy} : +d-V, \{\} \rangle \quad \langle \text{what} : -d-wh, \{\} \rangle \\
\hline
\langle \text{buy} : +d-V, \{ \text{what} : -d-wh \} \rangle \quad \text{INS} \\
\hline
\langle \text{buy} : -V, \text{what} , \{ \text{what} : -wh \} \rangle \quad \text{MRG} \quad \langle \text{yesterday} : *V, \{\} \rangle \\
\hline
\langle \text{buy} : -V, \text{what} , \{ \text{what} : -wh, \text{yesterday} : *V \} \rangle \quad \text{INS} \\
\hline
\langle \text{buy yesterday} : -V, \{ \text{what} : -wh \} \rangle \quad \text{SPL}
\end{array}$$

Figure 42

3.3 Constraining movement

In this section I will propose a single constraint that can be added to the formalism introduced in chapter 2 which will capture both adjunct island effects and freezing effects. This is best illustrated by first concentrating on the case of adjunct islands, and then showing that the constraint we are led to by the adjunct island facts alone also predicts freezing effects.

As mentioned earlier, the crucial property of the formalism we have now arrived at is that it establishes adjuncts and moving constituents as, in a certain sense, the same kind of thing. Roughly speaking, they are both things that end up in the set component of an expression to which SPL is applied. To sketch the basic idea, consider the derivation of ‘What did John buy yesterday’, which will begin as shown in Figure 42. Note that in the expression to which SPL applies here, both **what** and

tion predicts to permit extraction are in a sense more “closely” attached than those that do not, a distinction that is similar in spirit to the one I make between the “close” attachment of arguments (which permit extraction) and the more “loose” or “distant” attachment of adjuncts (which do not). I leave for future research the question of whether this similarity can be exploited in such a way that makes Truswell’s and my proposals genuinely compatible.

Second, reducing adjunct islands and freezing effects to the same constraint predicts that a language lacking one of these constraints will also lack the other. Japanese appears to be an interesting candidate: it has been claimed to permit extraction from adjuncts (Yoshida 2006) and there is some suggestion that it also lacks freezing effects (Yoshida 2006, p.177), or at least some other island constraints that are often thought to derive at least partially from freezing effects such as the subject island constraint (Lasnik and Saito 1992, pp.42–43). This possibility is not consistent, however, with the account I propose for the Japanese remnant movement data in §3.4.4.

yesterday are in the set component; the former because it is a “moving element” which is waiting there for a chance to check features and re-merge, the latter because it is an adjunct with a simple conjunctive interpretation. That these two kinds of constituents have been brought together in this way will let us, in the next section, introduce a single constraint that disallows movement out of either; in other words, a single constraint that subsumes adjunct island effects and freezing effects.

I first demonstrate that as things stand extraction from both arguments and adjuncts is unrestricted in §3.3.1. The distinction between arguments and adjuncts in my proposed formalism permits us to pinpoint the distinction upon which the possibility of extraction from a constituent should hinge. In §3.3.2 I introduce the required constraint and show how it rules out extraction from adjuncts; in §3.3.3 I show how this same constraint rules out extraction from moved constituents.

3.3.1 Extraction from adjuncts, as currently permitted

As things stand, the formalism makes no distinction between extraction from a complement and extraction from an adjunct. Both are derivable, although we would like to rule out the latter. The relevant part of the derivation of (3.11a) is shown in Figure 43.

- (3.11) a. Who do you think [that John saw ~~wh~~θ]?
- b. * Who do you sleep [because John saw ~~wh~~θ]?

The expression $\langle \text{that John saw} : -c, \{ \text{who} : -wh \} \rangle$ that the steps shown in Figure 43 begin with is the result of constructing the complement clause: roughly speaking, it is a CP with a *wh*-phrase waiting to re-merge at some higher point. When this is inserted into the next phase up, where the VP headed by ‘think’ is to be constructed, both *that John saw* and *who* end up in the set component of the resulting expression. By this stage these two units have exactly the same status: that one (*who*) is moving out of the other (*that John saw*) is not recorded. So in just the same way

$$\begin{array}{r}
\langle \text{think} : +c-V, \{\} \rangle \quad \langle \text{that John saw} : -c, \{ \text{who} : -wh \} \rangle \\
\hline
\langle \text{think} : +c-V, \{ \text{that John saw} : -c, \text{who} : -wh \} \rangle \quad \text{INS} \\
\hline
\langle \text{think} : -V, \text{that John saw}, \{ \text{who} : -wh \} \rangle \quad \text{MRG} \\
\hline
\langle \text{think that John saw} : -V, \{ \text{who} : -wh \} \rangle \quad \text{SPL}
\end{array}$$

Figure 43

$$\begin{array}{r}
\langle \text{sleep} : -V, \{\} \rangle \quad \langle \text{because John saw} : *V, \{ \text{who} : -wh \} \rangle \\
\hline
\langle \text{sleep} : -V, \{ \text{because John saw} : *V, \text{who} : -wh \} \rangle \quad \text{INS} \\
\hline
\langle \text{sleep because John saw} : -V, \{ \text{who} : -wh \} \rangle \quad \text{SPL}
\end{array}$$

Figure 44

that **that John saw** can merge into an argument position of **think** at the application of MRG in Figure 43, **who** will be able to merge into a higher specifier position to check its **-wh** feature when the rest of the structure is completed.

This is the right result in the case of a complement clause: it lets us derive (3.11a). But exactly the same mechanism will incorrectly allow movement out of an adjunct. The relevant part of the derivation of (3.11b) is shown in Figure 44; compare with Figure 43. The expression $\langle \text{because John saw} : *V, \{ \text{who} : -wh \} \rangle$ here represents the result of constructing the adjunct clause: roughly speaking, it is something that will adjoin to a VP, with a *wh*-phrase waiting to re-merge at some higher point. As in Figure 43, when this is inserted into the next phase up, both **because John saw** and **who** end up in the set component of the resulting expression, and no record of the relationship between these two units remains. Unlike **that John saw** in Figure 43, **because John saw** in Figure 44 does not merge into an argument position of the matrix verb before SPL applies — but this has *no bearing whatsoever* on the status of **who**, which will be free to merge into a higher specifier position as the derivation continues, exactly as in the case of the complement clause.

Therefore to capture the crucial distinction, we would like the ability of **who** to merge into higher specifier position to be contingent on the merging of **that John saw** into an argument position. This will mean that we need to maintain the structural relationship between **who** and the constituent it is moving out of (**that John saw** or **because John saw**) for longer than we currently do. In both the derivations in Figure 43 and Figure 44, the crucial information is lost as a result of the initial application of INS.

In the next subsection I will propose a modification to the formalism motivated solely by the need to prohibit extraction from adjuncts. In §3.3.3 I will then show that in doing so we also prohibit extraction from moved constituents with no further stipulation.

3.3.2 Prohibiting extraction from adjuncts

The key observation we have made is that in order to encode the prohibition on extraction from adjuncts we must “remember” what is moving out of what for longer than we currently do. The information that **who** is moving out of **that/because John saw** is present in the expressions with which we begin the derivations in Figure 43 and Figure 44 — encoded by the fact that **that/because John saw** is the head of the expression while **who** is in the set component — but is eliminated immediately by the INS operation. We must modify this operation in such a way that this information is maintained.

The basic idea of the modification I propose will be that the set component of an expression will no longer be a set of *units*, but rather a set of *other expressions*.⁷³ The entire expression $\langle \text{that John saw} : -c, \{ \text{who} : -wh \} \rangle$, for example, will be inserted “as is” into the set component of the expression headed by **think**, rather than having the asymmetric relationship between the two units **that John saw** : $-c$ and **who** : $-wh$ destroyed. The result of the application of INS that begins the derivation in Figure 43,

would therefore no longer be that repeated in (3.12a), but rather the more structured expression in (3.12b).

- (3.12) a. $\langle \text{think} : +c-V, \{ \text{that John saw} : -c, \text{who} : -wh \} \rangle$
 b. $\langle \text{think} : +c-V, \{ \langle \text{that John saw} : -c, \{ \text{who} : -wh \} \rangle \} \rangle$

Since we have now introduced a kind of recursion, an adjustment of notation will be useful: instead of writing $\langle x, a_1, a_2, \{y_1, \dots, y_n\} \rangle$, we draw a tree with a root node labelled $\langle x, a_1, a_2 \rangle$, which has n (unordered) children y_1, \dots, y_n . The effect of the INS function will therefore be to add a *child* to the root of a tree. The effect of the MRG function will be to check features between and establish an argument relation between the root of a tree and one of its *children*. And the effect of the SPL function will be to compose the interpretations of a head, its arguments, and any children that can be interpreted as adjuncts to it. Using *just this change in notation* — not yet incorporating recursion in the sense suggested in (3.12b) — the derivation in Figure 43 will be written as in Figure 45.

To create the recursive structure that will let us distinguish between extraction from complements and extraction from adjuncts, we now have the INS operation add a *tree* as a daughter of the root node. The analogue of the derivation in Figure 45 will now proceed as in Figure 46.

Let us say that a unit x is “ n -subordinate” to another unit y iff x can be reached from y by a downward path of length n ; and let us say more generally that x is subordinate to y iff x is n -subordinate to y for some n . Then what has crucially been maintained in the second expression in Figure 46, in contrast to Figure 45, is the information that $\text{who} : -wh$ is subordinate to $\text{that John saw} : -c$. More precisely, $\text{who} : -wh$ is 1-subordinate to $\text{that John saw} : -c$, and is 2-subordinate to $\text{think} : -V$. Next, MRG applies and draws **that John saw** into a distinguished argument

⁷³Unger (2010) adopts a system where expressions have a similar form, and also discusses the relationship between freezing effects and remnant movement in this context.

configuration. Let us understand this to have the consequence that **who** : **-wh** is no longer 2-subordinate to **think** : **-V** , but only 1-subordinate to it; intuitively, the idea is that “being n -subordinate” can be thought of as a kind of distance relation, and that this MRG step leaves the distance between **that John saw** and **who** unchanged while reducing the distance between **think** and **that John saw**, and therefore reduces the distance between **think** and **who**. Finally, SPL applies and composes the interpretations of **think** and **that John saw** as appropriate for their head-argument relation, and **who** : **-wh** remains, waiting to re-merge; this step is not significantly different from the corresponding final SPL step in Figure 45.

Now consider the (attempted) derivation of the adjunct island violation in (3.11b), previously derivable as shown in Figure 44. With our revisions in place, it will begin as shown in Figure 47. The expression produced by the initial application of INS is ready to be spelled out, since **because John saw** does not need to merged into an argument position; but **who** is still 2-subordinate to the head **sleep** : **-V** . To implement the adjunct island constraint, then, what we need to say is that applying SPL to an expression in which there exist units that are 2-subordinate to the head is prohibited. Units that are 1-subordinate because they are waiting to re-merge, such as **who** : **-wh** in Figure 46, are permitted; also permitted are units that are 1-subordinate because they are adjuncts to the head, ready to be interpreted conjunctively. But we prohibit these two “ways to be subordinate” (movement and adjunction) from compounding in such a way that there exist units that are 2-subordinate to the head.

Note that this does not rule out multiple adjuncts independently modifying a single head, as long as none of these adjuncts themselves have subordinate parts waiting to re-merge. The application of SPL in Figure 48, for example, is unproblematic, since no unit is 2-subordinate to **sleep** : **-V** .

I should emphasise that I have not said anything particularly insightful so far about the nature of adjunct island effects. This constraint on the expressions to

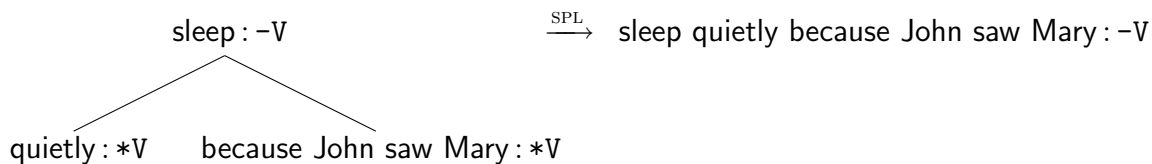


Figure 48

which SPL can apply is simply a restatement of the fact that extraction from adjuncts is prohibited. The attraction of it is that in combination with the implementation of movement that our formalism assumes, the very same constraint simultaneously prohibits extraction from moved constituents. This is what the next subsection will illustrate.

3.3.3 Freezing effects follow

An adjunct island constraint amounts to a prohibition on applying SPL to configurations in which a moving unit u is 1-subordinate to a unit u' , which is itself adjoined to, and thus 1-subordinate to, a head h ; in such configurations u is 2-subordinate to h . But the general constraint — that SPL may not apply to expressions where some unit is 2-subordinate to the head — also prohibits applying SPL to configurations in which a moving unit u is 1-subordinate to a unit u' , which is *moving out of*, and thus 1-subordinate to, a head h . This second kind of configuration is exactly the one that characterises freezing effects.

As a concrete example, I will suppose that (3.13) is ruled out because ‘who’ has moved out of a larger constituent which has itself moved: specifically, it has moved out of the subject, which has moved for Case/EPP reasons. I will show that this can be ruled out by the same constraint on SPL that we arrived at above to enforce adjunct islands.

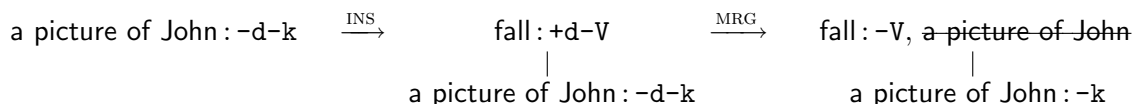


Figure 49

(3.13) * Who did [a picture of ~~who~~] fall (on the floor)?

First, though, let us consider the derivation of a sentence involving the relevant movement of the subject, but not yet with any additional *wh*-movement.

(3.14) A picture of John fell (on the floor)

The relevant part of the derivation of (3.14) is shown in Figure 49, beginning at the point where the subject **a picture of John** has been fully constructed, and it will need to merge into a theta position ($-\mathbf{d}$) and Case position ($-\mathbf{k}$). Note that since it will move out of its theta position (i.e. since the $-\mathbf{d}$ is not its last), a unit **a picture of John** : $-\mathbf{k}$ *remains as a daughter of the root*, just as it would previously have remained in the set component of the resulting expression. This causes no problem in the derivation of (3.14) where no movement out of the subject is required: SPL applies to the last expression shown in Figure 49, since we have completed construction of the VP by this point, and subject **a picture of John** re-merges into its Case position when the opportunity arises. (Note that this is just another instance of movement out of a complement, since the VP we have completed will be selected as the complement of, say, *v*.)

Now consider the attempted derivation of (3.13), where **who** must move out of the subject. By the end of the steps shown in Figure 50, the (as it happens, unpronounced) copy of **a picture of** has merged into an argument position of **fall**, so the phase is ready to be spelled out. However, as observed above, the unit **a picture of** : $-\mathbf{k}$ remains 1-subordinate to the head **fall** : $-\mathbf{V}$, and **who** : $-\mathbf{wh}$ remains 1-subordinate to

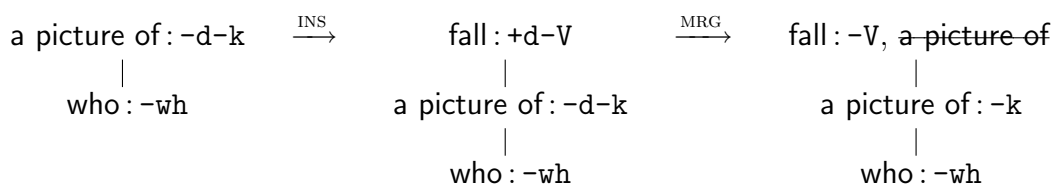


Figure 50

a picture of: -k ; therefore we have a unit, *who: -wh* , which is 2-subordinate to the head *fall: -V* and so our constraint from §3.3.2 prohibits application of SPL.

To see the similarity between adjunct island configurations and freezing configurations, note the similarity between the expressions we end up with in Figure 47 and in Figure 50. In each case the constituent out of which *who: -wh* “needs to move” remains 1-subordinate to the head — in the first case, because *because John saw: *V* is an adjunct, and in the second case, because *a picture of: -k* has not reached its final position. Contrast these in particular with the penultimate expression shown in Figure 46, to which SPL *is* able to apply, because the constituent *that John saw* out of which *who: -wh* needs to move is not subordinate to the expression’s root. The proposed constraint therefore makes a natural cut between (i) adjoined constituents and moving constituents, out of which movement is disallowed, and (ii) argument constituents, out of which movement is allowed.

Two further observations about the wider consequences of this constraint can also be made. First, to the extent that the unification I propose is valid, it makes the prediction that the cross-linguistic variation in the opacity of adjuncts should be matched by that of freezing effects. I leave for future research the question of whether the data can be (re-)analysed in a way that fits with this prediction.⁷⁴ Second, there are consequences for “remnant movement” configurations, to which I turn in the next

⁷⁴But see footnote 72.

section.

3.4 Remnant movement

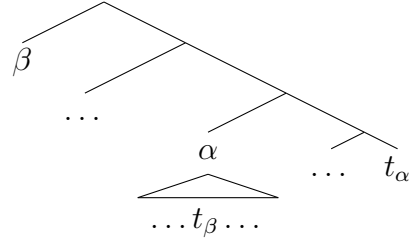
The constraint introduced in §3.3 actually bears on one further configuration besides extraction from adjuncts and extraction from moved constituents: the “remnant movement” configuration. Remnant movement is not predicted to be entirely disallowed, but a non-trivial restriction is imposed on it. This restriction is outlined in §3.4.1 and then illustrated in §3.4.2 with reference to VP-fronting in English, a form of remnant movement which is — roughly, with some exceptions that I discuss — correctly predicted to be possible. I then turn to some other instances of remnant movement that have been discussed in the literature, specifically some data from German in §3.4.3 and from Japanese in §3.4.4, which also appear to be consistent with the prediction.

3.4.1 Predictions for remnant movement

To a first approximation, the restriction we have imposed on the domain of SPL will also rule out remnant movement, since it is just another form of “moving out of a mover”. What distinguishes a “remnant movement” configuration from a “freezing” configuration is the relative height of the two movers’ final positions; see Figure 51. In each case some constituent α moves, and a subconstituent β moves out of α . In “freezing” configurations, the final position of α is below that of β ; by cyclicity/extension this means that movement of the entire constituent α occurs first. In “remnant movement” configurations, the final position of β is below that of α ; by cyclicity/extension this means that movement of the subconstituent β occurs first (followed by the movement of the “remnant” α with a trace of β inside it).

However, it turns out that the proposed restriction on SPL actually permits remnant movement to a limited degree. Specifically, remnant movement is allowed only

“Freezing”:



“Remnant movement”:

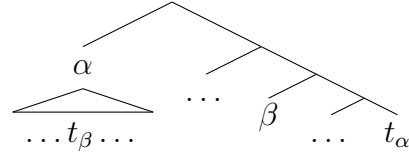


Figure 51

if the target position of the subconstituent’s movement is in the same maximal projection as the base position of the larger moving constituent. With respect to the diagrams in Figure 51, this means that t_α and β must be in the same maximal projection. This seems to be enough to let in a good portion of the cases where remnant movement is used in the literature.

3.4.2 English VP-fronting

The clearest example of remnant movement in English is the fronting of a VP out of which a subject has moved for Case.⁷⁵ Here and in what follows, the labels α and β indicate the corresponding constituents in Figure 51.

- (3.15) a. [Arrested t_β by the police] $_\alpha$, [John] $_\beta$ was t_α
 b. [Seem t_β to be tall] $_\alpha$, [John] $_\beta$ does t_α

These are consistent with the constraint stated above because the target position of the subconstituent’s movement (namely specifier of TP) is in the same maximal projection (namely TP) as the base position of the larger moving constituent (namely VP), as shown informally in Figure 52. This permits the relevant part of the derivation

⁷⁵If subjects are VP-internal, then *any* instance of VP-fronting will involve remnant movement, but for simplicity I use clearer cases of passivisation and raising.

to proceed as shown in Figure 53. This part of the derivation begins with the VP ‘seem to be tall’ (this has an extra $-f$ which will be checked in its fronted position), with a subconstituent ‘John’ that is waiting to check its $-k$ (Case) feature. The three steps shown are (i) inserting the VP into the phase of T, (ii) merging the VP into its (unpronounced) position as the complement of T, and (iii) merging ‘John’ into its (pronounced) position as the specifier of T. The next step will be to apply SPL to the last expression shown in Figure 53, which is permitted, since no unit is 2-subordinate to the root.

The crucial thing to observe about the derivation is that although the third expression shown in Figure 53 establishes the infamous “moving out of a mover” configuration, with $\text{John} : -k$ 2-subordinate to the root (cf. Figure 50), this configuration disappears before the point where SPL is next required to apply — because $\text{John} : -k$ merges into an argument position in this very phase. (More informally: as soon as the VP *seem to be tall* takes on its status as a mover, the movement out of it occurs, before SPL has a chance to “notice it”.⁷⁶) If the $-k$ feature of $\text{John} : -k$ were to be checked only in some higher phase, $\text{John} : -k$ would remain 2-subordinate to the root at the point of SPL, violating the constraint we have introduced.

Juan Uriagereka (p.c.) points out, however, a related situation where the constraint I have introduced appears to be too strong. On the assumption that modals and auxiliaries all head their own projections, I will incorrectly predict the VP-fronting in (3.16) to be disallowed.

(3.16) [Arrested t_β by the police] $_\alpha$, [John] $_\beta$ must have been t_α

⁷⁶Sneaking in some violations of freezing effects via the same route seems implausible. With respect to the diagrams in Figure 51, we would require a configuration where t_α and β are within the same maximal projection, which would mean that α itself would also need to be within that same maximal projection as well, violating anti-locality. This is actually just a special case of a more general corollary: if α were to move from one position to another within the same maximal projection, then movement out of it would not be restricted, because this doesn’t really count as movement in this formalism.

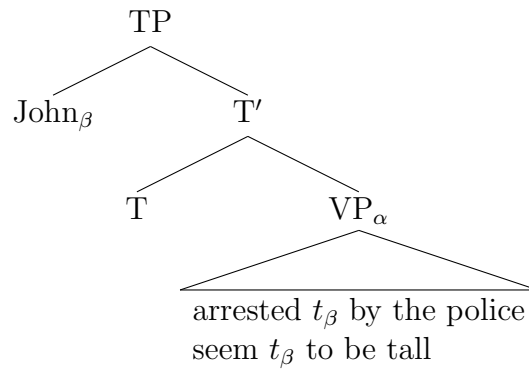


Figure 52

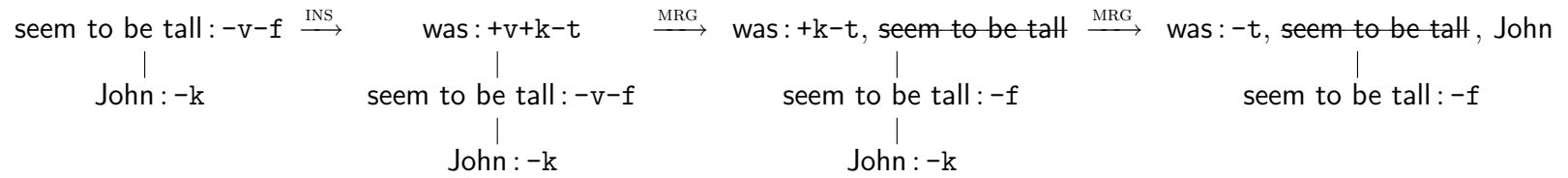


Figure 53

The reason for this is that the base position of the fronted VP will be the complement of the lowest auxiliary, ‘be(en)’. The position to which the subject has moved, however, will be outside the projection of ‘be(en)’; it will be in the projection of some higher auxiliary/tense head. This should not be allowed.

A solution to this problem that does not undermine the entire proposal will have to, it seems, deny that there is one phase for each auxiliary. This denial will have one of two forms. First, one might deny that each auxiliary heads its own maximal projection, and so in turn it is not the case that there is one phase for each auxiliary. Second, one might retract the strong hypothesis adopted so far that phases correspond perfectly with maximal projections. I favour the latter. Under standard assumptions about endocentricity and selection, maintaining the idea that every auxiliary projects a phrase lets us account for the ordering of auxiliaries and modals that we observe (‘have’ selects ‘be’, etc.), so this seems to be worth keeping. Furthermore, semantic composition in this part of the clause does seem to differ in at least some significant respects from semantic composition in the vP projection and below: the contribution of a tense head, for example, is plausibly “just another event predicate”, eg. $\lambda e.\mathbf{past}(e)$, in which case the result of interpreting the TP projection should not be of the form $\phi_T \& \mathbf{int}(\phi_{vP})$ as the current theory would dictate. In chapter 4 I introduce a distinct compositional principle that is used to interpret the TP projection (and others), according to which the tense head’s contribution (eg. $\lambda e.\mathbf{past}(e)$) is indeed directly conjoined with that of the vP . The discussion of the distinct compositional principle in chapter 4 does *not* take the additional step of dropping the assumption that each maximal projection is a phase, and therefore does not in and of itself avoid the problem posed by (3.16). Although I leave the problem unresolved here, it seems at least plausible that the underlying issue may be related to the observation that in roughly the area of auxiliaries we observe (i) apparent problems for the idea from chapter 2 that all complements are interpreted via the \mathbf{int} operator, and (ii) apparent

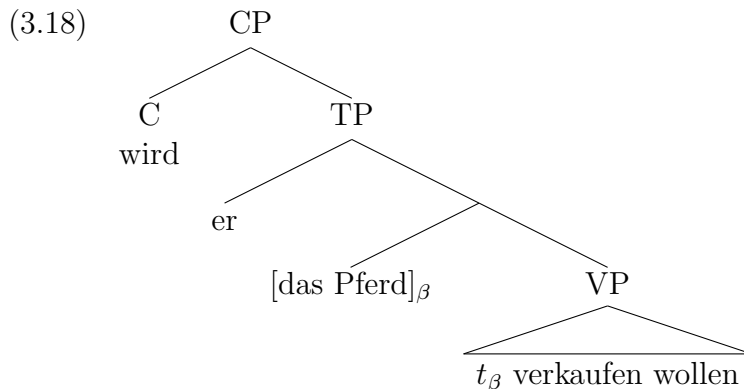
problems for the constraint on movement introduced in this chapter if the ideas from chapter 2 about phases are maintained as is.

3.4.3 German “incomplete category fronting”

A much-discussed case of remnant movement is that of “incomplete category fronting” in German, where a non-constituent appears to have been fronted, as in (3.17).

- (3.17) [t_β Verkaufen wollen] $_\alpha$ wird er [das Pferd] $_\beta$ t_α
 sell want will he the horse
 He will want to sell the horse (De Kuthy and Meurers 1998)

The object has moved out of the head-final VP ‘das Pferd verkaufen wollen’, before the VP is fronted. In order for our system to permit this movement, we require that the final position of the object ‘das Pferd’ is not outside the TP. This seems likely to be true, on the assumption that ‘wird’ is in the verb-second C head position. Specifically, the structure immediately before the fronting of the VP will be roughly as in (3.18).⁷⁷



It is similarly possible to leave behind a constituent more deeply embedded in the VP, such as a complement of the object as in (3.19).

- (3.19) [Ein Buch t_β schreiben] $_\alpha$ will niemand [darüber] $_\beta$ t_α
 a book write want no one about that

⁷⁷If ‘das Pferd’ is adjoined to VP, rather than in a specifier position of T, this also does not cause any problems: recall that the way such “stranded” adjuncts were treated in chapter 2 was essentially to consider them a part of the next phase up. I have omitted the category annotations on argument positions that make this possible in this chapter to reduce clutter.

No one wants to write a book about that (De Kuthy and Meurers 1998)

Again, for such movement to not violate the proposed constraint on SPL, we require that the moved PP ‘darüber’ not be outside the TP; and again, this seems reasonable since ‘will’ is in the verb-second C head position here.

The complement of an object can also be left behind, as in (3.19), in cases where only the object itself is fronted, as shown in (3.20). We can derive this as long as ‘über Syntax’ is not outside the VP, which seems safe to assume.

(3.20) [Ein Buch t_β] $_\alpha$ hat Hans sich [über Syntax] $_\beta$ t_α ausgeliehen
a book has Hans himself about syntax borrowed
Hans borrowed a book about syntax (De Kuthy and Meurers 1998)

And nothing significant changes when the movement of α is *wh*-movement:

(3.21) [Was für ein Buch t_β] $_\alpha$ hast du [über die Liebe] $_\beta$ t_α gelesen ?
what for a book have you about the love read
What sort of book about love did you read? (Müller 1998, p.231)

Note in general that if β were *not*, as we require, very close to t_α in these German examples, there would be no reason to regard them as interesting cases of apparent “non-constituent fronting”: it would be obvious that β had moved out of α . So it is tempting to tentatively conclude that nothing that has been labelled “incomplete category fronting” will turn out to be a violation of this constraint.

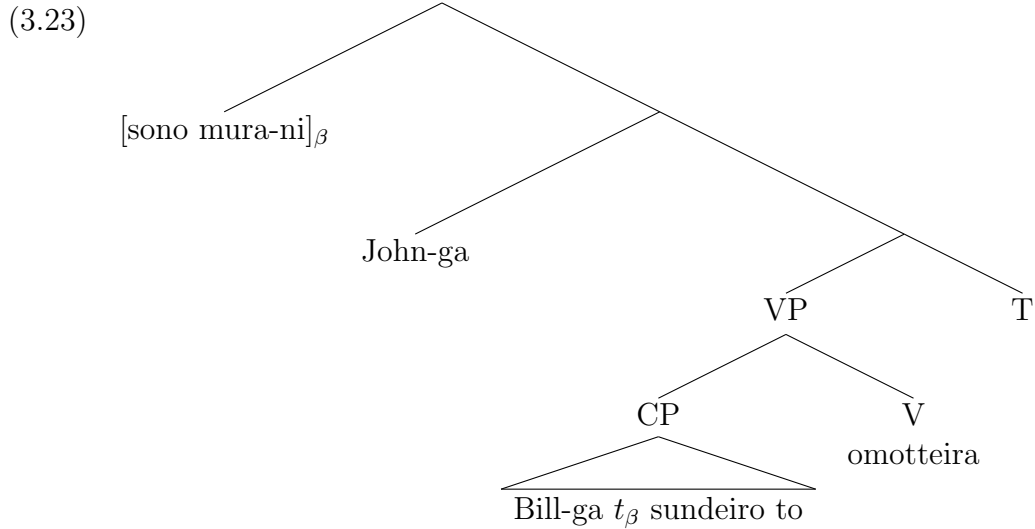
3.4.4 Japanese scrambling

Finally, I present an example of remnant movement which has been reported to be unacceptable, and which is correctly ruled out by the proposed constraint on SPL.

Takano (2000, p.143) gives the following unacceptable Japanese example. It is an instance of movement of β that seems to have gone outside the maximal projection containing the base position of α ; this is not permitted in the system presented above, so the prediction that this is underivable seems to be correct.

- (3.22) * [Bill-ga t_β sundeiro to] $_\alpha$ [sono mura-ni] $_\beta$ John-ga t_α omotteiru
 Bill live that that village in John think
 John thinks that Bill lives in that village

The structure of (3.22) immediately before the fronting of the embedded clause should be approximately as in (3.23); this is the structure of the acceptable sentence in (3.24) (Akira Omaki, p.c.) where only movement of β , and not the remnant movement of α , has taken place.



- (3.24) [sono mura-ni] $_\beta$ John-ga [Bill-ga t_β sundeiro to] $_\alpha$ omotteiru
 that village in John Bill live that think
 John thinks that Bill lives in that village

On the assumption that the subject ‘John-ga’ is in a specifier of T, the constituent β ‘sono mura-ni’ must have moved at least into the projection of T. This violates the requirement that its target position be in the same maximal projection (VP) as the base position of α — ruling out the remnant movement of α in (3.22).

3.5 Conclusion

In this chapter I have argued that two well-known generalisations concerning extraction domains can be reduced to a single constraint: first, the generalisation that

extraction from adjuncts is prohibited, and second, the generalisation that extraction from moved constituents is prohibited. Having integrated independently-motivated implementations of movement relations and adjunction, it emerges from the resulting system that adjoined constituents and moved constituents have a certain shared status. This allows us to add a single constraint to the theory to capture both the existing generalisations, and in turn provides further support for the syntactic framework that makes this result possible.

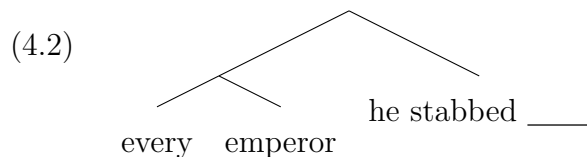
Chapter 4

Quantification via Conjunctivist Interpretation

The aim of this chapter is to show how the Conjunctivist approach to semantic composition can be extended to deal with simple quantificational sentences. As mentioned in chapter 1, one might worry that such “variable-binding” phenomena will pose serious problems for the system developed in chapter 2, because of the pervasive existential closure of variables and the unusual treatment of syntactic movement. The additions to the existing picture of the syntax/semantics interface that will be required are less drastic than one might expect, and in the course of formulating them we encounter interesting questions about the syntax of quantifier raising and its relation to other approaches to quantification.

As is relatively standard, I will associate sentences such as (4.1) with structures of the sort informally sketched in (4.2), where the quantificational DP has raised to a position c-commanding the rest of the sentence (May 1977, 1985).

(4.1) He stabbed every emperor



The main idea will be that the sentence in (4.1) is true if and only if the conjunction of three particular conditions can be met: one condition contributed by ‘every’, one by ‘emperor’ understood as an internal argument, and one by ‘he stabbed ____’

understood as a sentential expression. The effect of conjoining these three conditions will be to emulate the Tarskian satisfaction conditions for quantified logical formulas (Tarski 1983), by asking, for each of a range (dictated by ‘emperor’) of assignments of values to variables, whether the assignment satisfies a formula corresponding to ‘he stabbed ____’. This is a version of the approach already presented by Pietroski (2005, 2006), modified to fit into the syntactic framework presented in previous chapters.

I begin by presenting more formally the meanings that were used in previous chapters (**stabbing** & $\text{int}(\mathbf{c})$, etc.) in §4.1. In §4.2 I then rehearse the Tarskian method of recursively specifying satisfaction conditions by quantifying over assignment variants, in the setting of the language specified in §4.1 — without yet asking how to reconcile the general ideas with the Conjunctivist hypothesis. This question is taken up in §4.3 and §4.4, where I present a particular Conjunctivist division of the Tarskian labour, along the lines mentioned above. The account is shown to extend straightforwardly to multiply-quantified sentences in §4.5; I conclude by discussing some consequences in §4.6.

4.1 Basic semantic values

Syntactic units in previous chapters paired a string with a formula (which I will call “the meaning of the unit”) in a particular language that has thus far been only informally introduced.⁷⁸ A first definition, which will be extended and revised over the course of this chapter, is given in Figure 54.

As this language was used in previous chapters, the set C of constants would contain **stabbing**, **b**, **c**, and so on. I will assume that each lexical item has one of these constants specified as its translation. Formulas can have as their values elements

⁷⁸Or, to the extent that it was formally defined, formulas in this language were treated as a sort of shorthand for certain lambda terms over first-order logic: see (1.30) and (1.36). That was convenient in earlier chapters to facilitate comparisons with other approaches to semantic composition, and the intuitions thereby gained remain useful, but in this chapter I will define the “real” semantics for these formulas.

Given a set C of constants, the set of formulas is the smallest set \mathcal{F} such that:

- if $c \in C$ then $c \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\text{int}(\phi) \in \mathcal{F}$ and $\text{ext}(\phi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ and $\psi \in \mathcal{F}$ then $(\phi \& \psi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\langle \phi \rangle \in \mathcal{F}$

We write $\text{Val}(x, \phi)$ to mean that x is a value of the formula ϕ . Given specifications of the values of all constants $c \in C$, the values of the other formulas are specified as follows:

- $\text{Val}(x, \text{int}(\phi))$ iff $\exists y[\text{Internal}(x, y) \wedge \text{Val}(y, \phi)]$
- $\text{Val}(x, \text{ext}(\phi))$ iff $\exists y[\text{External}(x, y) \wedge \text{Val}(y, \phi)]$
- $\text{Val}(x, (\phi \& \psi))$ iff $\text{Val}(x, \phi) \wedge \text{Val}(x, \psi)$
- $\text{Val}(\top, \langle \phi \rangle)$ iff $\exists x[\text{Val}(x, \phi)]$ (and otherwise, $\text{Val}(\perp, \langle \phi \rangle)$)

Figure 54: A first definition of the formal language (to be revised)

of an assumed domain \mathcal{D} . For example, **stabbing** is the translation of ‘stab’, and e is a value of **stabbing** if and only if (iff) e is (an element of \mathcal{D} which is) an event of stabbing; **b** is the translation of ‘Brutus’ and the only value of **b** is Brutus (who is also an element of \mathcal{D}). We can put this differently by saying that **stabbing** has x as a value, or that $\text{Val}(x, \text{stabbing})$, iff x is an event of stabbing (taking the terminology and notation from Larson and Segal (1995)). Note that the conditions for value-hood of a formula ϕ are specified in terms of the conditions for value-hood of the subformulas of ϕ (i.e. those formulas that “license” ϕ as a formula according to the syntactic rules for the construction of formulas). This form of compositionality will be maintained in revised versions these definitions.

The values of the formula in (4.3a) are specified as in (4.3b).⁷⁹

(4.3) a. **stabbing** & $\text{int}(\mathbf{c})$ & $\text{ext}(\mathbf{b})$

- b. $\text{Val}(e, \mathbf{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b}))$
iff $\text{Val}(e, \mathbf{stabbing}) \wedge \text{Val}(e, \text{int}(\mathbf{c})) \wedge \text{Val}(e, \text{ext}(\mathbf{b}))$
iff $\text{Val}(e, \mathbf{stabbing}) \wedge \exists x[\text{Internal}(e, x) \wedge \text{Val}(x, \mathbf{c})] \wedge$
 $\exists y[\text{External}(e, y) \wedge \text{Val}(y, \mathbf{b})]$

As desired, this says that e is a value of the formula in (4.3a) iff: e is a stabbing event, e has Caesar (since Caesar is the only value of \mathbf{c}) as its internal participant, and e has Brutus as its external participant.

I have not yet discussed, however, formulas of the form $\langle\phi\rangle$. Intuitively, $\langle\phi\rangle$ is the result of applying existential closure to the predicate corresponding to ϕ (eg. in the case of (4.3a), a certain predicate of events). We assume that the domain \mathcal{D} contains two elements \top and \perp to represent truth and falsity, respectively. The formula in (4.4a) will have \top as a value iff something is a value of the formula in (4.3a), and thus, iff there was a stabbing of Caesar by Brutus; if a formula of this form does not have \top as a value, then it has \perp as a value.

- (4.4) a. $\langle\mathbf{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b})\rangle$
b. $\text{Val}(\top, \langle\mathbf{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b})\rangle)$
iff $\exists e[\text{Val}(e, \mathbf{stabbing} \ \& \ \text{int}(\mathbf{c}) \ \& \ \text{ext}(\mathbf{b}))]$
iff $\exists e[\text{Val}(e, \mathbf{stabbing}) \wedge \exists x[\text{Internal}(e, x) \wedge \text{Val}(x, \mathbf{c})] \wedge$
 $\exists y[\text{External}(e, y) \wedge \text{Val}(y, \mathbf{b})]]]$

I leave the details for §4.3, but one semantic effect of the TP projection will be to apply this “existential closure” to the formula associated with the VP.

If we were aiming only to account for simple sentences of the sort presented so far (such as ‘Brutus stabbed Caesar’), there would be no particular need for this $\langle\dots\rangle$ operator. Although there would not be any sense in which truth or falsity of

⁷⁹For ease of exposition, throughout this chapter I will ignore the additional complexity that comes from adopting the more articulated vP structures from §2.A. Adopting these more complex formulas instead of the one in (4.3a) will have no significant effect on the account of quantification that is the main subject of this chapter.

formulas, and therefore of natural language sentences, had been defined, it would suffice to assume that the pragmatic effect of uttering a declarative sentence with meaning ϕ is to claim that some value of ϕ exists. It will be important for the account of quantification, however, to have some encoding of whether or not such a claim is true (relative to a certain encoding of relevant contextual factors).⁸⁰

4.2 Assignments and assignment variants

4.2.1 Pronouns and demonstratives

Before turning to quantificational DPs we must first expand our formal language in a way that will let us deal with pronouns and demonstratives; a revised definition is given in Figure 55. The existing sense of value-hood is replaced by value-hood *relative to an assignment function*. Formally, an assignment function (or just “an assignment”) is a total function from \mathbb{N} (the set of natural numbers) to \mathcal{D} . We also add variables \mathbf{x}_i to the set of formulas. For any $i \in \mathbb{N}$, x is a value of the formula \mathbf{x}_i relative to the assignment function σ iff $\sigma(i) = x$.

With these modifications in place, we can deal with a sentence like (4.5) in the familiar way, by assuming indices on pronouns and demonstratives (Heim and Kratzer 1998, Larson and Segal 1995). As usual the translation of a pronoun or demonstrative with index i is the variable \mathbf{x}_i , and so the sentence in (4.5) will be associated with the formula in (4.6a), the values of which are specified in (4.6b).

(4.5) He₃ stabbed that₇

(4.6) a. $\langle \mathbf{stabbing} \ \& \ \mathbf{int}(\mathbf{x}_7) \ \& \ \mathbf{ext}(\mathbf{x}_3) \rangle$

b. $\text{Val}(\top, \langle \mathbf{stabbing} \ \& \ \mathbf{int}(\mathbf{x}_7) \ \& \ \mathbf{ext}(\mathbf{x}_3) \rangle, \sigma)$

iff $\exists e[\text{Val}(e, \mathbf{stabbing} \ \& \ \mathbf{int}(\mathbf{x}_7) \ \& \ \mathbf{ext}(\mathbf{x}_3), \sigma)]$

⁸⁰It also makes possible a straightforward Conjunctivist account of sentential coordination with ‘and’ and ‘or’ (Pietroski 2005, 2006).

Given a set C of constants, the set of formulas is the smallest set \mathcal{F} such that:

- if $c \in C$ then $c \in \mathcal{F}$
- if $i \in \mathbb{N}$ then $\mathbf{x}_i \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\text{int}(\phi) \in \mathcal{F}$ and $\text{ext}(\phi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ and $\psi \in \mathcal{F}$ then $(\phi \& \psi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\langle \phi \rangle \in \mathcal{F}$

We write $\text{Val}(x, \phi, \sigma)$ to mean that x is a value of the formula ϕ relative to an assignment function σ . Given specifications of the values of all constants $c \in C$, the values of the other formulas are defined as follows:

- $\text{Val}(x, \mathbf{x}_i, \sigma)$ iff $\sigma(i) = x$
- $\text{Val}(x, \text{int}(\phi), \sigma)$ iff $\exists y[\text{Internal}(x, y) \wedge \text{Val}(y, \phi, \sigma)]$
- $\text{Val}(x, \text{ext}(\phi), \sigma)$ iff $\exists y[\text{External}(x, y) \wedge \text{Val}(y, \phi, \sigma)]$
- $\text{Val}(x, (\phi \& \psi), \sigma)$ iff $\text{Val}(x, \phi, \sigma) \wedge \text{Val}(x, \psi, \sigma)$
- $\text{Val}(\top, \langle \phi \rangle, \sigma)$ iff $\exists x[\text{Val}(x, \phi, \sigma)]$ (and otherwise, $\text{Val}(\perp, \langle \phi \rangle, \sigma)$)

Figure 55: A revised definition of the formal language with values relative to assignments (to be further revised)

$$\begin{aligned}
& \text{iff } \exists e[\text{Val}(e, \mathbf{stabbing}, \sigma) \wedge \exists x[\text{Internal}(e, x) \wedge \text{Val}(x, \mathbf{x}_7, \sigma)] \wedge \\
& \quad \exists y[\text{External}(e, y) \wedge \text{Val}(y, \mathbf{x}_3, \sigma)]] \\
& \text{iff } \exists e[\text{Val}(e, \mathbf{stabbing}, \sigma) \wedge \exists x[\text{Internal}(e, x) \wedge x = \sigma(7)] \wedge \\
& \quad \exists y[\text{External}(e, y) \wedge y = \sigma(3)]] \\
& \text{iff } \exists e[\text{Val}(e, \mathbf{stabbing}, \sigma) \wedge \text{Internal}(e, \sigma(7)) \wedge \text{External}(e, \sigma(3))]
\end{aligned}$$

Therefore \top is a value of the formula in (4.6a), relative to any assignment function σ , iff there exists a stabbing event with $\sigma(7)$ as its internal participant and $\sigma(3)$ as its external participant.

4.2.2 Tarskian assignment variants

Having relativised the notion of value-hood to assignments, it is now possible to state the conditions under which the formula we eventually associate with the sentence in (4.7) should have \top as a value. For now I leave aside the questions of what this formula is — this will require further revisions of the formal language — and how it is constructed via the syntactic derivation of (4.7). The aim in this subsection is just to carve out relatively explicitly the work that will need to be done by the changes, to the formal language and to the syntax/semantics interface, that I will present in the rest of the chapter.

(4.7) Brutus stabbed every emperor

The approach here is due to Tarski (1983); see also Larson and Segal (1995), Partee (1975) and Hodges (2008) for helpful discussion. It is useful to define the notion of an *i*-variant of an assignment: given an assignment σ , another assignment σ' is an *i*-variant of σ (written: $\sigma' \approx_i \sigma$) iff $\sigma(j) = \sigma'(j)$ for all natural numbers $j \neq i$. Put differently, an *i*-variant of σ is an assignment that differs from σ at most in the *i* position. Given a treatment of pronouns and demonstratives of the sort just introduced, the crucial idea is that of *satisfaction of a formula by an assignment*; the intuition behind this idea is that an assignment σ satisfies a formula ϕ iff ϕ is true under the assignment of values to variables encoded by σ . Then the conditions under which an assignment σ satisfies (4.7) can be stated as in (4.8) (where the choice of index 7 is arbitrary, and we temporarily ignore the distinction between sentences and formulas).

(4.8) σ satisfies ‘Brutus stabbed every emperor’
iff for every σ' such that $\sigma' \approx_7 \sigma$ and $\sigma'(7)$ is an emperor,
 σ' satisfies ‘Brutus stabbed him₇’

This idea generalises straightforwardly to sentences containing more than one quantifier: the conditions under which an assignment satisfies (4.9) (restricting our attention to its “surface scope” reading) can be stated as in (4.10).

(4.9) Some senator stabbed every emperor

(4.10) σ satisfies ‘Some senator stabbed every emperor’
iff for some σ' such that $\sigma' \approx_3 \sigma$ and $\sigma'(3)$ is a senator,
 σ' satisfies ‘he₃ stabbed every emperor’
iff for some σ' such that $\sigma' \approx_3 \sigma$ and $\sigma'(3)$ is a senator,
for every σ'' such that $\sigma'' \approx_7 \sigma'$ and $\sigma''(7)$ is an emperor,
 σ'' satisfies ‘he₃ stabbed him₇’

Corresponding to Tarski’s notion of an assignment satisfying a formula we have, in the system being developed here, the notion of a formula ϕ having \top as a value relative to σ , written $\text{Val}(\top, \phi, \sigma)$.⁸¹ Therefore the conditions under which we would like \top to be a value, relative to σ , of the formula to be eventually associated with (4.7) — call it ψ for now — can be stated as in (4.11). I include now the constant **emperor**, of which x is a value iff x is an emperor, as the translation of ‘emperor’.

(4.11) $\text{Val}(\top, \psi, \sigma)$
iff for every σ' such that $\sigma' \approx_7 \sigma$ and $\text{Val}(\sigma'(7), \mathbf{emperor}, \sigma)$,
 $\text{Val}(\top, \langle \mathbf{stabbing} \ \& \ \text{int}(\mathbf{x}_7) \ \& \ \text{ext}(\mathbf{b}) \rangle, \sigma')$
iff for every σ' such that $\sigma' \approx_7 \sigma$ and $\sigma'(7)$ is an emperor,
 $\exists e[\text{Val}(e, \mathbf{stabbing}, \sigma') \wedge \mathbf{Internal}(e, \sigma'(7)) \wedge$
 $\quad \exists y[\mathbf{External}(e, y) \wedge \text{Val}(y, \mathbf{b}, \sigma')]]$
iff for every x such that x is an emperor,
 $\exists e[\text{Val}(e, \mathbf{stabbing}, \sigma) \wedge \mathbf{Internal}(e, x) \wedge \exists y[\mathbf{External}(e, y) \wedge$
 $\quad \text{Val}(y, \mathbf{b}, \sigma)]]$

This is the condition that for every emperor, there was a stabbing with that emperor as its internal participant and with Brutus as its external participant (i.e. for every emperor, there was a stabbing of that emperor by Brutus).

In order to use this treatment of quantification as a basis for adding to the theory developed in previous chapters, two problems must be solved. The first, in one form or another, is encountered by almost any theory of natural language quantification: the satisfaction condition for ‘Brutus stabbed every emperor’ (i.e. the condition under which this sentence has \top as a value) is stated in terms of the satisfaction condition for (basically) ‘Brutus stabbed him’, despite the fact that the latter is not obviously a constituent, syntactically speaking, of the former. The second is specific to the Conjectivist approach to semantic composition that I have adopted: even given a solution to the constituency problem, how can the satisfaction condition of ‘Brutus stabbed every emperor’ be derived, using only the spare machinery Conjectivism provides (or minimal additions to it), from that of ‘Brutus stabbed him’ and the semantic contributions of ‘every’ and ‘emperor’? I discuss the first of these problems briefly here, and then turn to the second in the rest of the chapter.

To sharpen the first problem, concerning syntactic constituency: recall that the value-hood conditions of complex formulas are specified in terms of the value-hood conditions of their subformulas. On the face of things, it would seem that the formula derived for (4.7) (repeated below) will be (4.12a). This does not have the formula in (4.12b) as a subformula; but as we have just noted, the value-hood conditions for whatever formula *is* eventually associated with (4.7) can only be stated in terms of those of (4.12b).

(4.7) Brutus stabbed every emperor

(4.12) a. $\langle \mathbf{stabbing} \ \& \ \mathbf{int}(\mathbf{every} \ \& \ \mathbf{int}(\mathbf{emperor})) \ \& \ \mathbf{ext}(\mathbf{b}) \rangle$

⁸¹Tarski had no truth *values* in his definitions; instead a sentence is defined to be true if and only if all assignments satisfy it.

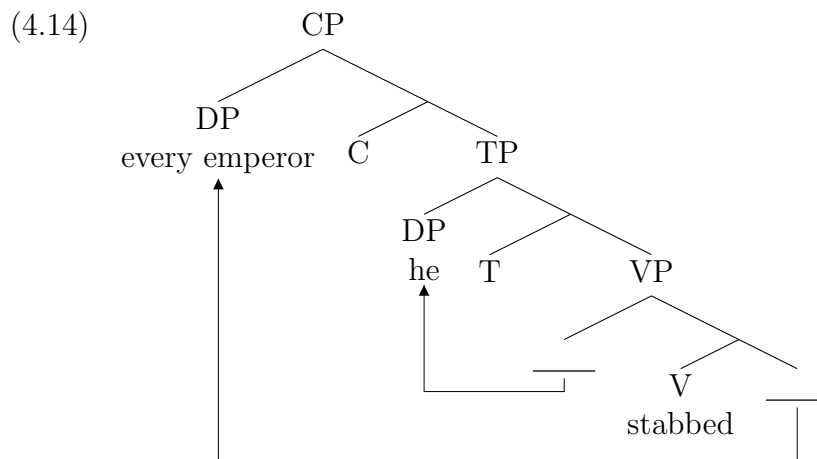
b. $\langle \text{stabbing} \ \& \ \text{int}(x_7) \ \& \ \text{ext}(\mathbf{b}) \rangle$

I will therefore follow many others (Montague 1974, May 1977, Cooper 1983) in assuming that something with the semantics of ‘Brutus stabbed him’ does indeed figure, in one way or another, in the derivation of (4.7).⁸² In particular I adopt the quantifier-raising approach from May.

4.3 Syntactic details

I now begin to address the question of how a formula with a satisfaction condition of the sort discussed in §4.2 can be derived, using Conjunctivist tools, from the syntactic derivation of the sentence in (4.13). First, this requires a clear picture of exactly what the syntactic derivation looks like. As mentioned above I will assume a quantifier-raising structure; more specifically, the structure informally indicated in (4.14).⁸³

(4.13) He stabbed every emperor



⁸²This can be contrasted with other approaches, often found in the categorial grammar literature, where it is not assumed that there is, in any sense, a sentence-like subconstituent of (4.7). Instead ‘Brutus stabbed’ is treated as a constituent denoting a function that maps individual denotations to sentential denotations (i.e. it denotes a “property”, just as a verb phrase often does). This is particularly clear in Combinatory Categorial Grammar (Steedman 1996, 2000), but equally true in the “deductive” versions of categorial grammar (Morrill 1994, Moortgat 1997) where a full sentence is derived with the help of a hypothesised gap-filler which is then retracted.

⁸³Also, recall that I assume a simplified VP structure for ease of exposition in this chapter, but nothing significant hinges on any differences between the *v*P shell structure discussed in §2.A and

This can be represented in the syntactic formalism of previous chapters as follows. I will assume that quantifier-raising checks features of a new type, q . Therefore the C head and the determiner ‘every’ will bear $+q$ and $-q$ features, respectively. The relevant unit for ‘every’ will therefore be $\text{every:}+n-d-q$, since it selects an NP complement, merges into a theta/DP position and then into a quantificational position. Representing the empty phonological content of the C head by ϵ_C , the C head will be the lexical unit $\epsilon_C:+t+q-c$, since it selects a TP complement and then a quantifying specifier. Although I will assume it is semantically vacuous, I also include movement of the subject to the specifier of TP for Case, encoded by k features; therefore the T head is $-ed:+V+k-t$, and the subject pronoun here is $\text{he:} -d-k$. This setup clearly abstracts away from many details — I leave open the question of how the tense affix is attached to the verb, and in general we will need an account of both Case movement and quantifier-raising of both subjects and objects⁸⁴ — but it is enough to illustrate the relevant aspects of the syntax. The full derivation will proceed as shown in Figure 56. (For ease of exposition I ignore, throughout this chapter, the additional hierarchical structure within expressions that was introduced in chapter 3 to constrain movement.)

I assume, as is standard, that the movement of the quantifier to check q features is phonologically vacuous (or “covert”). I have nothing original to say about why certain movement steps are reflected in both phonological and semantic interpretation (eg. most English *wh*-movement), others are reflected only in phonological interpretation (eg. subjects’ movement for Case) and others only in semantic interpretation (eg. quantifier raising). But whatever the underlying reason — perhaps these facts

this simpler alternative.

⁸⁴The Shortest Move Condition in the MG formalism makes this relatively difficult to achieve. If both subjects and objects are to move through theta, Case and quantificational positions, we are forced to say that (i) the object’s Case position must be lower than the subject’s theta position (to avoid competition for $+k$ features), and (ii) the object’s quantificational position must be lower than the subject’s Case position (to avoid competition for $+q$ features). See Kobele (2006) for details of how this pans out.

$e_1 = \text{INS}(\ell_{\text{every}}, \ell_{\text{emp}})$	$= \langle \text{every} : +\text{n}-\text{d}-\text{q}, \{ \text{emperor} : -\text{n} \} \rangle$
$e_2 = \text{MRG}(e_1)$	$= \langle \text{every} : -\text{d}-\text{q}, \text{emperor} : \text{n}, \{ \} \rangle$
$e_3 = \text{SPL}(e_2)$	$= \langle \text{every emperor} : -\text{d}-\text{q}, \{ \} \rangle$
$e_4 = \text{INS}(\ell_{\text{stab}}, e_3)$	$= \langle \text{stab} : +\text{d}+\text{d}-\text{V}, \{ \text{every emperor} : -\text{d}-\text{q} \} \rangle$
$e_5 = \text{MRG}(e_2)$	$= \langle \text{stab} : +\text{d}-\text{V}, \text{every emperor} : \text{d}, \{ ____ : -\text{q} \} \rangle$
$e_6 = \text{INS}(e_5, \ell_{\text{he}})$	$= \langle \text{stab} : +\text{d}-\text{V}, \text{every emperor} : \text{d}, \{ ____ : -\text{q}, \text{he} : -\text{d}-\text{k} \} \rangle$
$e_7 = \text{MRG}(e_6)$	$= \langle \text{stab} : -\text{V}, \text{every emperor} : \text{d}, ____ : \text{d}, \{ ____ : -\text{q}, \text{he} : -\text{k} \} \rangle$
$e_8 = \text{SPL}(e_7)$	$= \langle \text{stab every emperor} : -\text{V}, \{ ____ : -\text{q}, \text{he} : -\text{k} \} \rangle$
$e_9 = \text{INS}(\ell_{\text{T}}, e_8)$	$= \langle -\text{ed} : +\text{V}+\text{k}-\text{t}, \{ \text{stab every emperor} : -\text{V}, ____ : -\text{q}, \text{he} : -\text{k} \} \rangle$
$e_{10} = \text{MRG}(e_9)$	$= \langle -\text{ed} : +\text{k}-\text{t}, \text{stab every emperor} : \text{V}, \{ ____ : -\text{q}, \text{he} : -\text{k} \} \rangle$
$e_{11} = \text{MRG}(e_{10})$	$= \langle -\text{ed} : -\text{t}, \text{stab every emperor} : \text{V}, \text{he} : \text{k}, \{ ____ : -\text{q} \} \rangle$
$e_{12} = \text{SPL}(e_{11})$	$= \langle \text{he stabbed every emperor} : -\text{t}, \{ ____ : -\text{q} \} \rangle$
$e_{13} = \text{INS}(\ell_{\text{C}}, e_{12})$	$= \langle \epsilon_{\text{C}} : +\text{t}+\text{q}-\text{c}, \{ \text{he stabbed every emperor} : -\text{t}, ____ : -\text{q} \} \rangle$
$e_{14} = \text{MRG}(e_{13})$	$= \langle \epsilon_{\text{C}} : +\text{q}-\text{c}, \text{he stabbed every emperor} : \text{t}, \{ ____ : -\text{q} \} \rangle$
$e_{15} = \text{MRG}(e_{14})$	$= \langle \epsilon_{\text{C}} : -\text{c}, \text{he stabbed every emperor} : \text{t}, ____ : \text{q}, \{ \} \rangle$
$e_{16} = \text{SPL}(e_{15})$	$= \langle ____ \epsilon_{\text{C}} \text{ he stabbed every emperor} : -\text{c}, \{ \} \rangle$

Figure 56: A complete derivation showing quantifier-raising.

can be derived from some completely independent principles of linearisation, perhaps we will need a stipulation for each feature category (eg. q -driven movement is covert, k -driven movement is overt), or perhaps we will need a stipulation for each feature instance on each lexical item — it is worth noting that when the derivation is construed as yielding a sound-meaning pair, phonologically vacuous (or “covert”) movement is a natural counterpart of semantically vacuous movement, which I have assumed since chapter 2.⁸⁵ In other words, to the extent that we accept movement that is reflected phonologically but not semantically (eg. subjects’ movement for Case), we should not be surprised to find movement that is reflected semantically but not phonologically (eg. quantifier raising).

⁸⁵I use the term “phonologically vacuous” movement to bring out the parallel with semantically vacuous movement, but this should not be confused with *string vacuous* movement. The movement illustrated in (i) is (generally taken to be) phonologically *non-vacuous*, in the sense that the

Note also that while I have arrived at this derivation by asking how the idea of quantifier-raising, as illustrated in (4.14), can be represented in the MG formalism, we would have arrived at something very similar if we had instead set out to implement “Cooper storage” (Cooper 1983). This interesting convergence arises from (i) the fact that the implementation of syntactic movement that I adopt can, as noted in §1.5.2, be thought of as roughly applying Cooper storage to strings, and (ii) the assumption that every unit has a meaning component as well as a string component (although only the string components are shown in Figure 56).⁸⁶ I will continue to use the terminology of quantifier raising for concreteness, but for the purposes of this chapter I suspect that the differences between the various existing approaches to the syntax of quantification are irrelevant.

Semantic composition inside the VP proceeds as in previous chapters. Assuming a version, explained in more detail shortly, of the normal treatment of “traces” as semantic variables, the formula associated with the VP will be that in (4.15). More formally: when SPL to is applied to e_7 in Figure 56, the formula in (4.15) is constructed as the meaning component of the unit **stab every emperor** : -V that is the head of the resulting expression, e_8 .

$$(4.15) \quad \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3)$$

Also, the earlier application of SPL to e_2 in Figure 56 will produce **every** & **int(emperor)** (about which more will be said below) as the meaning component of the **every emperor** : -d-q unit, by exactly the same principles.

phonological contribution of ‘who’ is indeed realised in the specifier of CP position rather than in the surface subject position; it happens to be string vacuous because no other phonological output intervenes between these two positions.

(i) Who ____ left?

⁸⁶See Kobele (2006) for a similar implementation of quantification in MGs that assumes a slightly looser connection between syntactic movement and the quantifier store. Kobele’s proposal covers far more empirical ground than I aim to here, including facts that may turn out to be inconsistent with the stronger connection I adopt.

I have not so far discussed the TP or CP projections. Let us assume for now — so that we can focus on the CP projection where the actual quantification must be fleshed out — that the semantic effect of the TP projection is to add an event predicate such as **past** or **present** to specify tense, and to existentially close the result. More specifically, the formula produced at the end of the TP projection, when SPL is applied to e_{11} , will be that in (4.16).

$$(4.16) \quad \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3) \rangle$$

I will return to the details of how this is produced later, but under this assumption the challenge presented for the CP projection is how to combine the formula in (4.16) with semantic contributions of the DP ‘every emperor’ and the C head — perhaps taking into account the structural relationships among these constituents — to produce a formula ψ with the satisfaction condition given in (4.17).

$$(4.17) \quad \text{Val}(\top, \psi, \sigma) \\ \text{iff for every } \sigma' \text{ such that } \sigma' \approx_7 \sigma \text{ and } \text{Val}(\sigma'(7), \mathbf{emperor}, \sigma), \\ \text{Val}(\top, \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3) \rangle, \sigma')$$

We can see that this value-hood condition is stated in terms of (i) the value-hood condition of the formula associated with the TP, (ii) that of the formula **emperor** which is associated with the word ‘emperor’, and (iii) a certain quantificational force which it will presumably make sense to attribute to the semantic contribution of the word ‘every’. But it is unclear, to say the least, how the Conjunctivist principles of semantic composition will combine them to yield a formula with the properties of the hypothetical ψ in (4.17). The structure of the CP phase, to which SPL will be applied in order to somehow produce something with the properties of ψ in (4.17), is illustrated in (4.18). This raises two problems to be addressed.

when the relevant determiner takes arguments, and (ii) indexes traces left by phrases headed by this determiner. The re-merging of the quantificational phrase is unrelated to either of these indexing steps.

The second problem raised by the structure in (4.18) concerns the strong relationship I have so far assumed between syntactic projection (i.e. headedness) and semantic composition. Semantic composition in the VP domain proceeded under the assumption that each phrase’s syntactic head contributes its meaning “un-shifted”, whereas complements and specifiers contributed “shifted” meanings. Following this pattern, we would expect the formula in (4.19) to be the result of applying SPL to the structure in (4.18). I include now the index on the *int* operator applied to the restrictor as just discussed, and for the moment write ϕ_C for whatever constant is associated with the C head.

$$(4.19) \quad \phi_C \& \text{ext}(\mathbf{every} \& \text{int}_7(\mathbf{emperor})) \& \text{int}(\langle \mathbf{past} \& \mathbf{stabbing} \& \text{int}(x_7) \& \text{ext}(x_3) \rangle)$$

At least intuitively speaking, there appears to be too much “embedding” here. It seems difficult to maintain the conventional intuition that the determiner ‘every’ is the semantic backbone of the sentence, since **every** is within the scope of **ext**. I will assume that the strict correlation between syntactic headedness and semantic composition can not be maintained here, and will suggest shortly that some new compositional principle is required.

Note that the possibility of a consistent role in semantic composition for all syntactic heads is one that might in principle be accepted or rejected independently of any choice between (say) Conjunctivism and function-oriented alternatives; and more standard approaches based on function application are typically also forced to reject it. In an approach based on function application one would probably be led to suspect, on the basis of semantics inside the VP domain, that a syntactic head denotes a function which is applied to the denotations of the syntactic arguments of that head. But the standard account to quantification in such a framework forces

us to reject this correlation. Since ‘every emperor’ is not the head of the structure in (4.18), we would expect **every(emperor)** (or whatever its denotation is) to be passed as an *argument* to some function; but instead it takes the denotation of its sister as an argument to produce something similar to (4.20). The semantic reflection of the head-argument relation in syntax here is argument-function, rather than function-argument as it usually is in the VP domain.⁸⁷

$$(4.20) \quad \mathbf{every}(\mathbf{emperor})(\lambda x_7.\mathbf{stabbed}(x_7)(x_3))$$

The new compositional principle that I adopt for the CP phase does not “reverse” the expected semantic relationship between syntactic heads and their arguments, as (4.20) does, but rather just “ignores” the syntactic head-argument configurations and composes the phase’s pieces, un-shifted — via conjunction, of course — and then applies existential closure. The result, when applied to (4.18), is the formula in (4.21); assuming, as is relatively standard, that the C head is semantically vacuous,⁸⁸ this is equivalent to (4.22).

$$(4.21) \quad \langle \phi_C \ \& \ \mathbf{every} \ \& \ \mathbf{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3) \rangle \rangle$$

$$(4.22) \quad \langle \mathbf{every} \ \& \ \mathbf{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3) \rangle \rangle$$

Here, **every** is in a position to act as the semantic head or backbone of the sentence, **emperor** is appropriately placed to act as one of its arguments, and while the TP’s meaning is not marked as an argument, it is bundled and “separated” from **every** by the existential closure.

Since (4.22) is almost bare conjunction, one may well ask if it would not be simpler to assume that quantifiers raise to an *adjoined* position, so that they can compose conjunctively as would already be expected from previous assumptions. I address

⁸⁷One might wonder whether the assumption is justified that when a quantifier re-merges, it is not the syntactic head of the newly-formed constituent. At least to the extent that quantifier-raising is thought to be normal syntactic movement, this seems reasonable since in seemingly all other instances of movement it is the attractor, not the mover, that projects. In particular, this assumption is built in to the MG formalism. But see also the suggestion of “reprojection” from Hornstein and Uriagereka (2002).

this concern in §4.6.1, after working through the details of the interpretation of these formulas.

The new compositional principle (that is to say, simply, the new subcase of SPL) is illustrated in (4.23b), alongside the existing one in (4.23a). Note that while (4.23b) is fresh stipulation about how a piece of syntactic structure can be interpreted, there is a sense in which it makes use of the same basic semantic operations as (4.23a): conjunction, which is clearly central to both and remains the only binary connective, and existential closure, which (although obscured somewhat by the current notation) is invoked by both the $\langle \dots \rangle$ in (4.23b) and by the int and ext operators in (4.23a).⁸⁹

$$(4.23) \quad \text{a.} \quad \begin{array}{c} > \\ \swarrow \quad \searrow \\ \psi_2 \quad < \\ \quad \swarrow \quad \searrow \\ \quad \phi \quad \psi_1 \end{array} \quad \xrightarrow{\text{SPL}} \quad \phi \ \& \ \text{int}(\psi_1) \ \& \ \text{ext}(\psi_2)$$

$$\text{b.} \quad \begin{array}{c} > \\ \swarrow \quad \searrow \\ \psi_2 \quad < \\ \quad \swarrow \quad \searrow \\ \quad \phi \quad \psi_1 \end{array} \quad \xrightarrow{\text{SPL}} \quad \langle \phi \ \& \ \psi_1 \ \& \ \psi_2 \rangle$$

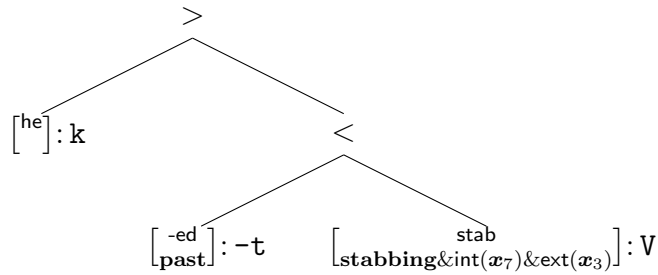
Note that this new compositional principle gives exactly the result we would like in the TP phase, namely (4.16) (repeated below): the meaning of the VP is conjoined with the constant **past** that is associated with the T head, and the result is existentially closed. I assume the occurrence of the subject in the specifier position of TP for Case (checking k features) is semantically vacuous.

$$(4.16) \quad \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \text{int}(\mathbf{x}_7) \ \& \ \text{ext}(\mathbf{x}_3) \rangle$$

⁸⁸To say that the C head is semantically vacuous is just to say that $\text{Val}(x, \phi_C, \sigma)$ for any $x \in \mathcal{D}$ for any assignment σ .

⁸⁹I leave for future research a proper investigation of whether these commonalities might permit a less stipulative formulation of the two possibilities in (4.23a) and (4.23b).

(4.24)



Of course, we now need to specify what determines, for each phase, which of these two rules applies. I have little to offer on this issue at the moment, so I simply stipulate which rule is appropriate for phases headed by each feature type.⁹⁰ Phases headed by (units bearing) $-V$ or $-v$ are interpreted via (4.23a), and phases headed by $-c$ or $-t$ are interpreted via (4.23b), for example.⁹¹

To take stock: having noted problems with maintaining the compositional principle from chapter 2 in the domain of quantification, I have introduced a single new compositional principle which (i) produces the desired result for the TP phase, namely the formula in (4.16) for the sentence under discussion, and (ii) produces the formula in (4.22) for the CP phase, which I have claimed can be given an appropriate interpretation. In the next section I flesh out the details of this claim.

4.4 Semantic details

Without further ado, the final specification of the formal language in which meanings will be expressed is given in Figure 57.

The only change to the set of formulas is the addition of the int_i operators, as

⁹⁰Attributing distinctive properties to various feature types in this manner is a departure from the way features are usually treated in the MG formalism: usually the only significance of a feature is that it may match or fail to match another feature. It seems likely that this purely formal conception of feature types will need to be given up at some point however, since otherwise it will be impossible to define notions such as a clause.

⁹¹Possible sources of this distinction might include (i) a split between lexical and functional heads, the former using (4.23a) and the latter (4.23b), but $-v$ and $-d$ appear to be counter-examples to this pattern; and (ii) sort mismatches between the relevant meaning components forcing shifting via (4.23a), and (4.23b) applying elsewhere. This second possibility would go some way toward justifying the departure from (4.23a) in roughly the same way that (4.20) is sometimes justified by a type mismatch.

Given a set C of constants, the set of formulas is the smallest set \mathcal{F} such that:

- if $c \in C$ then $c \in \mathcal{F}$
- if $i \in \mathbb{N}$ then $\mathbf{x}_i \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\text{int}(\phi) \in \mathcal{F}$ and $\text{ext}(\phi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ and $i \in \mathbb{N}$ then $\text{int}_i(\phi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ and $\psi \in \mathcal{F}$ then $(\phi \ \& \ \psi) \in \mathcal{F}$
- if $\phi \in \mathcal{F}$ then $\langle \phi \rangle \in \mathcal{F}$

For some things X , we write $\text{Val}(X, \phi, \Sigma)$ to mean that they *are values of* the formula ϕ *relative to* an assignment relation Σ . Given specifications of the values of all constants $c \in C$, the values of the other formulas are defined as follows:

- $\text{Val}(X, \mathbf{x}_i, \Sigma)$ iff $\Sigma(i) = X$
- $\text{Val}(X, \text{int}(\phi), \Sigma)$ iff $\exists Y[\text{Internal}(X, Y) \wedge \text{Val}(Y, \phi, \Sigma)]$
- $\text{Val}(X, \text{ext}(\phi), \Sigma)$ iff $\exists Y[\text{External}(X, Y) \wedge \text{Val}(Y, \phi, \Sigma)]$
- $\text{Val}(X, \text{int}_i(\phi), \Sigma)$ iff $\exists Y[\text{Internal}_i(X, Y) \wedge \forall y[Yy \leftrightarrow \text{Val}(y, \phi, \Sigma)]]$
- $\text{Val}(X, (\phi \ \& \ \psi), \Sigma)$ iff $\text{Val}(X, \phi, \Sigma) \wedge \text{Val}(X, \psi, \Sigma)$
- $\text{Val}(X, \langle \phi \rangle, \Sigma)$ iff $\forall v_r[Xv_r \rightarrow (v = \top \leftrightarrow \exists Y[\text{Val}(Y, \phi, \Sigma[r])])]$

where the relations **Internal**, **External** and **Internal_i** are “lifted” to apply to pluralities according to the following definitions in terms of their first-order restrictions:

- **Internal**(X, Y) iff
 $\forall x[Xx \rightarrow \exists y[Yy \wedge \text{Internal}(x, y)]] \wedge \forall y[Yy \rightarrow \exists x[Xx \wedge \text{Internal}(x, y)]]$
- **External**(X, Y) iff
 $\forall x[Xx \rightarrow \exists y[Yy \wedge \text{External}(x, y)]] \wedge \forall y[Yy \rightarrow \exists x[Xx \wedge \text{External}(x, y)]]$
- **Internal_i**(X, Y) iff
 $\forall x[Xx \rightarrow \exists y[Yy \wedge \text{Internal}_i(x, y)]] \wedge \forall y[Yy \rightarrow \exists x[Xx \wedge \text{Internal}_i(x, y)]]$

Figure 57: The final definition of the formal language

mentioned above. Formulas of the form $\text{int}_i(\phi)$ will have an interpretation similar, but not identical, to those of the form $\text{int}(\phi)$.

The most significant change to the interpretation of the language is that the conditions for value-hood are now stated using second-order variables, interpreted in the style of Boolos (1984, 1998). Understood this way, a second-order variable like ‘ X ’ in the metalanguage does *not* range over sets, as in the standard semantics for second-order logic (Enderton 2001, 2009); nor over plural entities or sums of any sort (Link 1983, 1998, Landman 1996). Such a variable ranges over the very same things that first-order variables range over, but it does so in such a way that it can, relative to a single (metalanguage) assignment, have more than one of these things as values.⁹² Rather than have *a set* of things as its *value*, the variable ‘ X ’ might have *the things* as its *values*. Since the variable ‘ X ’ may have multiple values (in the metalanguage sense), if $\text{Val}(X, \phi, \Sigma)$ then the formula ϕ may have multiple values (in the sense defined in Figure 57). I will not provide a thorough exposition of this view of second-order logic, or repeat from elsewhere arguments that its use is independently motivated by considerations of the semantics of plurality (see Boolos 1984, 1998, Schein 1993, Pietroski 2003a, 2005), but one brief point is worth noting in order to highlight its difference from the more standard interpretation based on sets. Assume a domain of all and only (Zermelo-Fraenkel) sets, and consider the formula ‘ $\exists X \forall x [Xx \leftrightarrow x \notin x]$ ’. This is false if we understand ‘ X ’ to range over sets of things in the domain (i.e. sets of sets) and understand ‘ Xx ’ to mean that $x \in X$, since there is no set of all non-self-elemental sets. But if we understand ‘ X ’ to range, plurally, over the things that ‘ x ’ ranges over, and understand ‘ Xx ’ to mean that the thing x is one of the things X , then the formula is true, because there are indeed some non-self-elemental sets. The two interpretations of second-order variables therefore cannot be mere terminological variants.

⁹²Boolos (1998, p.72) neatly summarises the idea when he writes: “neither the use of plurals nor the employment of second-order logic commits us to the existence of extra items [such as plural

This prompts a corresponding adjustment in what constitutes an assignment of values to variables in our (object) language. In place of an assignment function σ from \mathbb{N} to \mathcal{D} , we now have an assignment *relation* $\Sigma \subseteq (\mathbb{N} \times \mathcal{D})$. The domain of this relation must be all of \mathbb{N} , but it need not be a function; so while, for every $i \in \mathbb{N}$, it must pair at least one thing with i , it may pair more than one thing with i . I will write $\Sigma(i)$ for the *things*, in Boolos’s sense, that Σ pairs with i .⁹³

The “lifted” versions of **Internal** and **External** are relatively intuitive extensions of their existing versions. Adopting the notational convention from Boolos (1984, 1998) for discussing pluralities informally in prose: these_X bear the **Internal** relation to those_Y iff (i) for every one_x of these_X , there is one_y of those_Y , such that it_x bears the **Internal** relation to it_y , and (ii) for every one_y of those_Y , there is one_x of these_X , such that it_x bears the **Internal** relation to it_y . Note that this does not require that there is a one-to-one correspondence between these_X and those_Y : if one_x of these_X bears the **Internal** relation (individually) to many a different one_y of those_Y , then there may be many more of those_Y than there are of these_X . It only requires that there are no “leftovers” among these_X that bear the **Internal** relation to none of those_Y , and similarly that there are no “leftovers” among those_Y .

The interpretation of the new formulas of the form $\text{int}_i(\phi)$ relies on a new sort of object that I need to assume is present in our domain \mathcal{D} .⁹⁴ Intuitively, these new objects are truth values supplemented with zero or more restrictions on assignments (put differently, conditions that an assignment might meet). Formally, the set of

entities] beyond those to which we are already committed. We need not construe second-order quantifiers as ranging over anything other than the objects over which our first-order quantifiers range . . . a second-order quantifier needn’t be taken to be a kind of first-order quantifier in disguise, having items of a special kind, collections, in its range. It is not as though there were two sorts of things in the world, individuals and collections of them, which our first- and second-order variables, respectively, range over and which our singular and plural forms, respectively, denote. There are, rather, two (at least) different ways of referring to the same things . . .”.

⁹³For some things X and some things Y , we define: $X = Y$ iff $\forall x[Xx \leftrightarrow Yx]$.

restricted truth values (RTVs) can be modelled as

$$\{\top, \perp\} \times [\mathbb{N} \multimap \mathcal{D}]$$

where $[A \multimap B]$ is the set of partial functions from the set A to the set B . If $\alpha \in \mathcal{D}$ and $\beta \in \mathcal{D}$, then one example of a RTV is

$$\langle \top, \{\langle 7, \alpha \rangle, \langle 9, \beta \rangle\} \rangle$$

which we can understand as the truth value \top paired with two conditions that might be imposed on an assignment: that it assign (only) α to the 7 position, and that it assign (only) β to the 9 position. We can write $\top_{[\alpha/7, \beta/9]}$ for this RTV, so that in general: given an RTV v_r and an assignment Σ , we can denote by $\Sigma[r]$ the assignment just like Σ but “updated” to conform to the restrictions encoded by $r \in [\mathbb{N} \multimap \mathcal{D}]$. For example, if $v_r = \top_{[\alpha/7, \beta/9]}$ then $\Sigma[r] = \Sigma[\alpha/7, \beta/9]$ is the assignment that assigns (only) α to the 7 position, assigns (only) β to the 9 position, and is otherwise just like Σ . For RTVs of the form $\langle v, \emptyset \rangle$ we can write just v instead of v_\emptyset ; as this notation suggests, we can think of “plain” truth values as just special cases of RTVs.

RTVs, rather than only the plain truth values \top and \perp , will be the potential values of formulas of the form $\langle \phi \rangle$. The idea, stated formally in the relevant clause of Figure 57, is that some RTVs V are values of $\langle \phi \rangle$ relative to Σ iff for each v_r among V : v is \top if there exists some things $_Y$ such that they $_Y$ are values of ϕ relative to $\Sigma[r]$, and v is \perp otherwise.

⁹⁴This treatment where (parts of) assignments of values to variables belong to the domain — which will permit a treatment of quantification that is technically categoric, since variations on assignments can be encoded by model-theoretic objects — is similar to that of Kobele (2006, in press).

Let us consider then, the formula

$$\psi = \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3) \rangle$$

which is a subformula of (4.22). As before, $\text{Val}(\top, \psi, \Sigma)$ iff there was a stabbing of $\Sigma(7)$ by $\Sigma(3)$ (and otherwise $\text{Val}(\perp, \psi, \Sigma)$). To illustrate the effect of adding a restriction — specifically, the restriction that α be assigned to the 7 position — observe that

$$\begin{aligned} & \text{Val}(\top_{[\alpha/7]}, \psi, \Sigma) \\ & \text{iff } \exists Y [\text{Val}(Y, \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathbf{int}(x_7) \ \& \ \mathbf{ext}(x_3), \Sigma[\alpha/7])] \\ & \text{iff there was a stabbing of } (\Sigma[\alpha/7])(7) \text{ by } (\Sigma[\alpha/7])(3) \\ & \text{iff there was a stabbing of } \alpha \text{ by } \Sigma(3) \end{aligned}$$

and $\text{Val}(\perp_{[\alpha/7]}, \psi, \Sigma)$ iff there was no stabbing of α by $\Sigma(3)$. Similarly, $\text{Val}(\top_{[\alpha/7, \beta/3]}, \psi, \Sigma)$ iff there was a stabbing of α by β .

The crucial point to note is this: asking whether a formula has \top_r (where $r \neq \emptyset$) as a value relative to Σ is, in effect, asking whether the formula has \top as a value relative to a particular *variant* of Σ , namely $\Sigma[r]$. So by taking RTVs as potential values of certain formulas we have the tools to implement part of the Tarskian “assignment-variant” approach to quantification. Of course, to finish the job we need to consider whether ψ above has \top as a value relative to (in general) not just one assignment but many; to foreshadow, the relevant intuitions will be that $\mathbf{int}_7(\mathbf{emperor})$ will specify the assignments to consider, and **every** will impose constraints on the patterns of \top -value-hood across these assignments.

We have been considering questions of the (assignment-relative) value-hood of a single RTV, effectively ignoring the universal quantification component of the definition of the values of a formula of the form $\langle \phi \rangle$. But these formulas, like all others, may have multiple values. For some things $_X$ to be values of such a formula, each one $_x$ of them $_X$ must be a value in the way just discussed with reference to $\top_{[\alpha/7]}$, $\top_{[\alpha/7, \beta/3]}$

and so on.

As well as being potential values for formulas of the form $\langle \phi \rangle$, RTVs are also the potential values of formulas of the form $\text{int}_i(\phi)$. (Since two such formulas are conjoined in (4.22), this is to be expected; and in turn, RTVs will also be the potential values of the third conjunct in that formula, namely **every**.) The requirement that some things must meet in order to be values of $\text{int}_i(\phi)$, however, is not distributive in the way that the requirement for value-hood of $\langle \phi \rangle$ is. Some things_X can be values of $\text{int}_i(\phi)$ even though no one_x of them_X is, just as some things_Y can form a circle, cluster in the forest, or rain down, even though no one_y of them_Y forms anything, clusters anywhere or rains down (Gillon 1987, Schein 1993, 2006).

Specifically, some things_X are values of $\text{int}_i(\phi)$, relative to Σ , iff there are some things_Y such that: they_Y bear the **Internal**_i relation to them_X, and they_Y are all and only the values of ϕ relative to Σ . Less formally, some things are values of $\text{int}_i(\phi)$ iff their *i*-internal participants are all and only the values of ϕ .⁹⁵ So some things are values of $\text{int}_7(\mathbf{emperor})$ iff their *i*-internal participants are all and only the values of **emperor**, i.e. are all and only the emperors.

Now we must ask: what is it to be the *i*-internal participant of something? Given a RTV v_r and some $\alpha \in \mathcal{D}$, let us define **Internal**_i such that **Internal**_i(v_r, α) iff $r(i)$ is defined and $r(i) = \alpha$. Put differently, α is the *i*-internal participant of v_r iff r includes a requirement that α be assigned to the *i* position. Recall now the “two-way exhaustive” definition of when some things bear the **Internal**_i relation to some other things: these are the *i*-internal participants of those iff every one of those has one of these as its *i*-internal participant, and every one of these is the *i*-internal participant of one of those. Unpacking this in terms of RTVs’ restrictions on assignments: these_X are the *i*-internal participants of those_Y RTVs iff every one_{v_r} of those_Y requires that one_x of these_X be assigned to the *i* position, and every one_x of these_X is required by

⁹⁵This commits us to distributive conditions for value-hood of formulas to which int_i is applied.

one_{*v_r*} of those_{*V*} to be assigned to the *i* position.

Recall now that some things are values of $\text{int}_7(\mathbf{emperor})$ iff their *i*-internal participants are (all and only) the emperors. Therefore, some things_{*V*} are values of $\text{int}_7(\mathbf{emperor})$ iff they require that (all and only) the emperors be assigned to the 7 position. Suppose that α and β are the (only) emperors. Consider some things_{*X*} such that: they_{*X*} are only two, one of them_{*X*} is $\top_{[\alpha/7]}$, and one of them_{*X*} is $\perp_{[\beta/7]}$. Then they_{*X*} are values of $\text{int}_7(\mathbf{emperor})$, since they require that all and only α and β be assigned to the 7 position. So are the things_{*Y*} among which are only $\perp_{[\alpha/7, \gamma/8]}$ and $\top_{[\beta/7, \delta/3]}$. But consider some things_{*Z*} such that: they_{*Z*} are three, one of them is $\top_{[\alpha/7]}$, one of them is $\perp_{[\beta/7]}$, and one of them is $\top_{[\gamma/7]}$; these_{*Z*} are not values of $\text{int}_7(\mathbf{emperor})$ because one_{*z*} of them_{*Z*} requires that a non-emperor be assigned to the 7 position. Neither are the things_{*W*} such that they_{*W*} are one, and the only one of them_{*W*} is $\top_{[\alpha/7]}$; even though these_{*W*} do not require that a non-emperor be assigned to the 7 position, they_{*W*} fail to require that all emperors be assigned to the 7 position.

Having considered the values of $\text{int}_7(\mathbf{emperor})$ and the values of ψ above, it remains to discuss the values of **every**. Some things_{*X*} are values of **every** (relative to any assignment) iff every one_{*x*} of them is an RTV \top_r for some *r*. Similarly, they_{*X*} are values of **some** iff at least one_{*x*} of them is an RTV \top_r for some *r*, and they_{*X*} are values of **most** iff those_{*Y*} of them_{*X*} that are RTVs \top_r for some *r* outnumber those_{*Z*} of them_{*X*} that are not.

Now, finally, we can consider the overall conditions under which \top is a value (relative to an assignment) of the formula in (4.22). The first steps of “unpacking” are straightforward:

$$\begin{aligned} & \text{Val}(\top, \langle \mathbf{every} \& \text{int}_7(\mathbf{emperor}) \& \langle \mathbf{past} \& \mathbf{stabbing} \& \text{int}(\mathbf{x}_7) \& \text{ext}(\mathbf{x}_3) \rangle \rangle, \Sigma) \\ & \text{iff } \exists X [\text{Val}(X, \mathbf{every} \& \text{int}_7(\mathbf{emperor}) \& \langle \mathbf{past} \& \mathbf{stabbing} \& \text{int}(\mathbf{x}_7) \& \\ & \quad \text{ext}(\mathbf{x}_3) \rangle, \Sigma)] \\ & \text{iff } \exists X [\text{Val}(X, \mathbf{every}, \Sigma) \wedge \text{Val}(X, \text{int}_7(\mathbf{emperor}), \Sigma) \wedge \text{Val}(X, \langle \mathbf{past} \& \end{aligned}$$

stabbing & int(\mathbf{x}_7) & ext(\mathbf{x}_3), Σ]

Then, treating each of the three conjuncts here one at a time:

Val(X , **every**, Σ)

iff for every v_r among X , $v = \top$

Val(X , int₇(**emperor**), Σ)

iff $\exists Y$ [Internal₇(X , Y) \wedge $\forall y$ [$Yy \leftrightarrow$ Val(y , **emperor**, Σ)]]

iff these _{X} assign all and only the emperors to the 7 position

iff for every v_r among X , $r(7)$ is an emperor, and for every emperor e

there is a v_r among X such that $r(7) = e$

Val(X , **<past & stabbing & int(\mathbf{x}_7) & ext(\mathbf{x}_3)>**, Σ)

iff $\forall v_r$ [$Xv_r \rightarrow$

($v = \top \leftrightarrow \exists Y$ [Val(Y , **<past & stabbing & int(\mathbf{x}_7) & ext(\mathbf{x}_3)>**, $\Sigma[r]$)])]

iff for every v_r among X ,

$v = \top$ iff there was a stabbing of $\Sigma[r](7)$ by $\Sigma[r](3)$

If there exist some things _{X} that meet these three conditions, then \top is a value of the formula in (4.22) (relative to Σ). That these are appropriate conditions to associate with the original sentence ‘He stabbed every emperor’, is best illustrated via some examples.

Suppose that α and β are the only two emperors, both of which were indeed stabbed by $\Sigma(3)$. Consider some things _{X} such that: they _{X} are only two, $\top_{[\alpha/7]}$ is one of them _{X} , and $\top_{[\beta/7]}$ is one of them _{X} . Then

Val(X , **every** & int₇(**emperor**) & **<past & stabbing & int(\mathbf{x}_7) & ext(\mathbf{x}_3)>**, Σ)

and therefore

$$\text{Val}(\top, \langle \mathbf{every} \ \& \ \text{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \text{int}(\mathbf{x}_7) \ \& \ \text{ext}(\mathbf{x}_3) \rangle \rangle, \Sigma)$$

since things_X meet all three of the conditions above.

If some things_X are only one, and it_x is $\top_{[\alpha/\gamma]}$, things_X do not make the sentence true because — while things_X are values of the other two relevant formulas — things_X are not values of $\text{int}_7(\mathbf{emperor})$, because things_X “leave out” the other emperor, β . The same goes for the things_X among which is only the one thing $x \top$; and for any things_X among which is $\top_{[\gamma/\gamma]}$ where γ is not an emperor.

Return now to the things_X which are two, namely $\top_{[\alpha/\gamma]}$ and $\top_{[\beta/\gamma]}$. If there were an additional emperor δ , then things_X would still be a value of **every**, and would still be a value of $\langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \text{int}(\mathbf{x}_7) \ \& \ \text{ext}(\mathbf{x}_3) \rangle$ — since things_X still answer the “stabbed by $\Sigma(3)$ or not” question correctly in the cases where things_X do answer it, namely for α and β — but things_X would not be values of $\text{int}_7(\mathbf{emperor})$, because they miss the other emperor, δ .

If $\Sigma(3)$ did not in fact stab β , then $\top_{[\beta/\gamma]}$ would not be a value of (or among any things that are values of) $\langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \text{int}(\mathbf{x}_7) \ \& \ \text{ext}(\mathbf{x}_3) \rangle$. The disjunct RTV $\perp_{[\beta/\gamma]}$ would be, but this (unlike $\top_{[\beta/\gamma]}$) is not a value of (or among any things that are values of) **every**.

4.5 Multiple quantifiers

Semantic formulas of the sort introduced so far extend without further adjustments to sentences with multiple quantifiers. Some assumptions underlying the MG formalism, however, cause problems when the syntax I have proposed for quantifier-raising must be extended to deal with multiple quantifiers. In this section I first show that there is a simple way to use the formal language developed so far to at least express the

desired meanings for multiply-quantified sentences, and then discuss the syntactic difficulties.

4.5.1 Semantics

Recall the observation that the Tarskian method of defining satisfaction relative to sequence variants deals with multiple quantifiers easily, as shown in (4.10), repeated here.

(4.9) Some senator stabbed every emperor

(4.10) σ satisfies ‘Some senator stabbed every emperor’
iff for some σ' such that $\sigma' \approx_3 \sigma$ and $\sigma'(3)$ is a senator,
 σ' satisfies ‘he₃ stabbed every emperor’
iff for some σ' such that $\sigma' \approx_3 \sigma$ and $\sigma'(3)$ is a senator,
for every σ'' such that $\sigma'' \approx_7 \sigma'$ and $\sigma''(7)$ is an emperor,
 σ'' satisfies ‘he₃ stabbed him₇’

Note that the higher-scoping quantifier ‘some senator’ induces 3-variants of the assignment σ in just the same way that it would if it were the only quantifier in the sentence, and the 3-variants σ' are checked for satisfaction of the sentence ‘he₃ stabbed every emperor’, which is exactly the sentence treated in the previous section. Thus the satisfaction condition for (4.9) can be specified in terms of that of ‘he₃ stabbed every emperor’, in just the same way that the satisfaction condition of the latter is specified in terms of that of ‘he₃ stabbed him₇’.

If the formula to be eventually associated with (4.9) is χ , then we can write the desired conditions under which $\text{Val}(\top, \chi, \Sigma)$ as follows; compare with (4.17).

(4.25) $\text{Val}(\top, \chi, \Sigma)$
iff for some Σ' such that $\Sigma' \approx_3 \Sigma$ and $\text{Val}(\Sigma'(3), \mathbf{senator}, \Sigma),$
 $\text{Val}(\top, \psi, \Sigma')$

where $\psi = \langle \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(\mathbf{x}_7) \ \& \ \mathit{ext}(\mathbf{x}_3) \rangle \rangle$

Here ψ is the formula discussed in detail in the previous section, which we have seen has \top is a value relative to Σ' if and only if $\Sigma'(3)$ stabbed every emperor. So the conditions under which ψ has \top as a value will contribute to the specification of the conditions under which χ has \top as a value, in just the same way that $\langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(\mathbf{x}_7) \ \& \ \mathit{ext}(\mathbf{x}_3) \rangle$ contributed to the specification of the conditions under which ψ has \top as a value.

The formula χ that should be associated with (4.9) is therefore the one given in (4.26).

$$(4.26) \quad \chi = \langle \mathbf{some} \ \& \ \mathit{int}_3(\mathbf{senator}) \ \& \ \langle \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(\mathbf{x}_7) \ \& \ \mathit{ext}(\mathbf{x}_3) \rangle \rangle \rangle$$

Now, for any assignment Σ , $\text{Val}(\top, \chi, \Sigma)$ if and only if there exist some things $_X$ meeting the following three conditions:

$\text{Val}(X, \mathbf{some}, \Sigma)$

iff for some v_r among X , $v = \top$

$\text{Val}(X, \mathit{int}_3(\mathbf{senator}), \Sigma)$

iff $\exists Y[\text{Internal}_3(X, Y) \wedge \forall y[Yy \leftrightarrow \text{Val}(y, \mathbf{senator}, \Sigma)]]$

iff these $_X$ assign all and only the senators to the 3 position

iff for every v_r among X , $r(3)$ is a senator, and for every senator s there

is a v_r among X such that $r(3) = s$

$\text{Val}(X, \langle \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(\mathbf{x}_7) \ \& \ \mathit{ext}(\mathbf{x}_3) \rangle \rangle, \Sigma)$

iff $\forall v_r[Xv_r \rightarrow$

$(v = \top \leftrightarrow \exists Y[\text{Val}(Y, \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(\mathbf{x}_7) \ \& \ \mathit{ext}(\mathbf{x}_3) \rangle, \Sigma[r])])$]

iff for every v_r among X ,

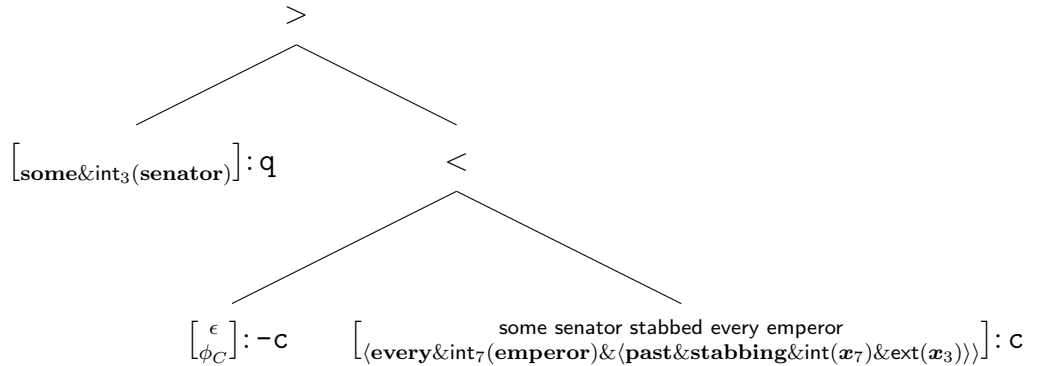
$$v = \top \text{ iff } \Sigma[r](3) \text{ stabbed every emperor}$$

Some things_X which satisfy these conditions will be some RTVs. Each one of these RTVs must assign a senator to the index 3, and no senator can be left out; and each one of these RTVs must be such that it is of the form \top_r if the senator it assigns stabbed every emperor and is of the form \perp_r otherwise. Finally, at least one of these things_X RTVs must be of the form \top_r . If and only if some such things_X exist, then $\text{Val}(\top, \chi, \Sigma)$ for any assignment Σ .

4.5.2 Syntax

Given the form of the formula in (4.26), and bearing in mind the way in which the formula for a single-quantifier sentence was derived in the previous section, the obvious way to suppose that (4.26) is constructed from a syntactic derivation is via an application of SPL to a phase with the structure illustrated in (4.27).

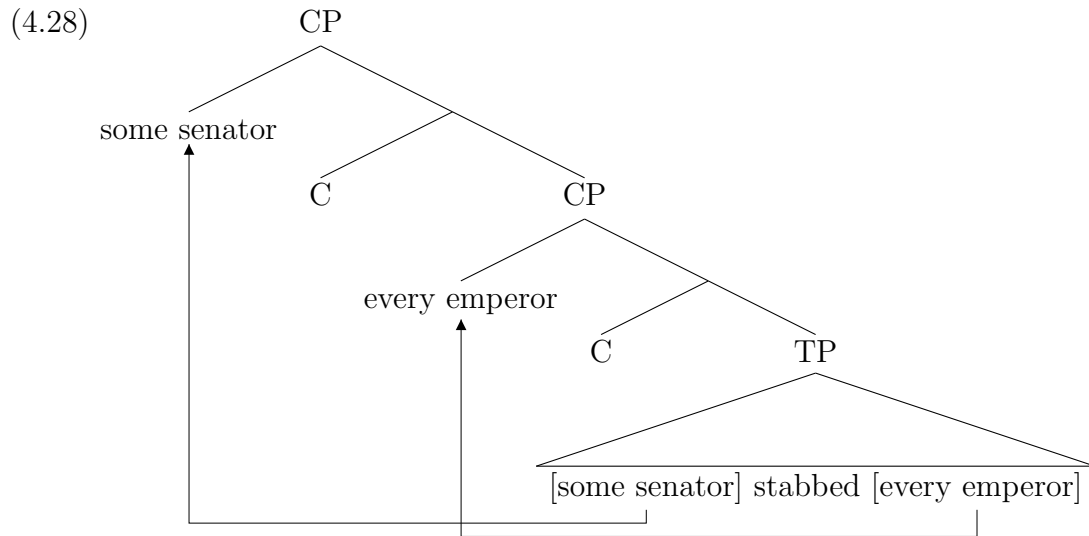
(4.27)



The complement of this phase is the unit produced by the final application of SPL in the derivation of the single-quantifier sentence considered above, modulo the difference between \mathbf{x}_3 being contributed as an “unbound” variable previously (and pronounced as ‘he’) but as the residue of quantifier raising (and pronounced as ‘some

senator’) in (4.27).⁹⁶

The structure in (4.27) certainly yields the desired formula, but it forces us to certain syntactic assumptions. The C head in (4.27) has the CP derived for the single-quantifier sentence as its complement, and the higher-scoping quantifier as its specifier; this means we have a structure where each quantifier moves to the specifier position of its own maximal projection, as illustrated informally in (4.28).



The uppermost two heads here are labelled C for concreteness; all that matters is that these heads attract quantifiers to their specifier positions and are semantically vacuous. This does however force us to assume two lexical entries for C heads (or whatever they are), one which takes a TP as a complement (+t+q-c) to accommodate the lowest quantifier in any sentence, and one which takes a CP as a complement (+c+q-c) to accommodate subsequent quantifiers.

⁹⁶More explicitly: whereas the final application of SPL in the derivation of ‘He stabbed every emperor’ produces

$$\left\langle \left[\text{he stabbed every emperor} \right]_{\langle \text{past} \& \text{stabbing} \& \text{int}(x_7) \& \text{ext}(x_3) \rangle} : -c, \{ \} \right\rangle$$

the penultimate application of SPL in the multiple-quantifier sentence produces

$$\left\langle \left[\text{he stabbed every emperor} \right]_{\langle \text{past} \& \text{stabbing} \& \text{int}(x_7) \& \text{ext}(x_3) \rangle} : -c, \left\{ \left[\text{some senator} \right]_{\langle \text{int}_3(\text{senator}) \rangle} : -q \right\} \right\rangle$$

A more serious problem is that it is not possible to produce this sort of structure if we assume that all quantifiers move to check the same type of feature (i.e. $-q$ features). Under this assumption, the expression produced at the end of the TP phase, when all that remains is for the quantifiers to re-merge, for the doubly-quantified sentence under discussion will be (4.29). (For clarity I write only the phonological components of units, and abstract away from distinctions between pronounced and unpronounced positions.)

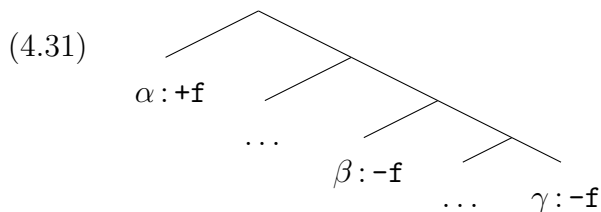
(4.29) $\langle \text{some senator stabbed every emperor} : -t, \{ \text{some senator} : -q, \text{every emperor} : -q \} \rangle$

The problem here is that since the two quantifiers are “competing” to re-merge and check features when the next $+q$ feature becomes available, the Shortest Move Condition (SMC) will prevent this derivation from continuing. More precisely, the MRG function will be unable to apply to the following expression, which will be produced immediately after the (first) C head takes the TP as its complement.

(4.30) $\langle \epsilon_C : +q-c, \text{some senator stabbed} \dots : t, \{ \text{some senator} : -q, \text{every emperor} : -q \} \rangle$

The SMC states that in order for the MRG function to apply to an expression with a $+f$ feature as the first feature on its head, there must be a *unique* unit in the set component of the expression looking to check a $-f$ feature. There is no such unique unit in (4.30).

The MG formalism’s SMC is a particularly strong version of a minimality constraint (Rizzi 1990, Chomsky 1995), but even with a more standard, weaker version, the configuration we are considering will be problematic. To illustrate the relationship between the SMC and more common minimality constraints, consider the abstract structure in (4.31).



According to standard minimality restrictions, when two elements are competing for a single feature, as β and γ are for the single $+f$ feature here, only the closer of the two, here β , is allowed to move. Moving γ here would violate minimality, but moving β is permissible. According to the SMC, however, *neither* of these two potential movements is allowed: if such competition exists, the derivation is stuck, since no rule can apply to this expression.

The SMC is therefore a problematic constraint for the proposal sketched in (4.28), since a situation analogous to (4.31) will arise when the first $+q$ feature becomes available. But consider also the consequence of a more standard conception of minimality when the derivation reaches this situation: it would predict that only the quantifier coming from the higher base position should be able to re-merge into this lower quantificational position (and that the quantifier with the lower base position will have to wait to re-merge later), predicting obligatory “inverse scope”. So, while it is true that the MG formalism’s strong SMC causes a problem here, the additional strength of the SMC relative to other conceptions of minimality is not the primary cause of the problem: some workaround is required even for more standard versions of minimality.

One workaround that will at least permit us to express a derivation within the MG framework is simply to stipulate that the features that the two quantifiers are looking to check upon re-merging are distinct — not both $-q$, but rather one $-q$ and one $-r$ say. This requires us to say that there exist distinct versions of (what I have been calling) the C head that attract bearers of each of these distinct quantifying features. This will make it possible to represent any particular single derivation we wish to, but since a sentence can contain unboundedly many quantifiers, this is not an attractive option.⁹⁷

⁹⁷Also, to the extent that we might want to attribute certain properties to particular feature types — for example, to say that positions where k features are checked are semantically vacuous, or that positions where q features are checked are phonologically vacuous — this would force us to duplicate such statements.

A second workaround would be to remove the problematic SMC prohibiting application of MRG to (4.30), and leave it undetermined which of the two potential re-mergers checks features first. The result would be that MRG is no longer a *function* mapping expressions to expressions, but merely a relation. Specifically, “applying” MRG to the expression in (4.30) would yield either of the two expressions in (4.32).

- (4.32) a. $\langle \epsilon_C : -c, \text{some senator stabbed} \dots : t, \text{every emperor} : q, \{ \text{some senator} : -q \} \rangle$
 b. $\langle \epsilon_C : -c, \text{some senator stabbed} \dots : t, \text{some senator} : q, \{ \text{every emperor} : -q \} \rangle$

Whichever of the two quantifiers merges here (‘every emperor’ in the first case, ‘some senator’ in the second) will have narrower scope; the other will re-merge later and take wider scope. Notice that by removing the SMC in this way we have not only weakened the formalism’s minimality restrictions to the point where it permits movement of β in structures like (4.31), as common minimality constraints do; we have removed any minimality constraint altogether, allowing either β or γ to move in a configuration like (4.31). In the case of quantifier-raising, this is a reasonable result, since both scope possibilities, at least in a simple sentence like this, are available. In other situations where minimality constraints are typically invoked however (eg. superiority effects), this is almost certainly too weak.

Note, however, that these problems with the syntax of quantifier-raising are independent of the details of the Conjunctivist approach to quantification. The root of the problem is the need for multiple quantifiers to raise outside the TP domain. This is certainly not a unique requirement of Conjunctivism.

4.6 Discussion

The central goal of this chapter has been to show that quantification can be handled within the broad outlines of the Conjunctivist approach (i.e. based around monadic predicate conjunction and existential closure, with occasional appeal to binary rela-

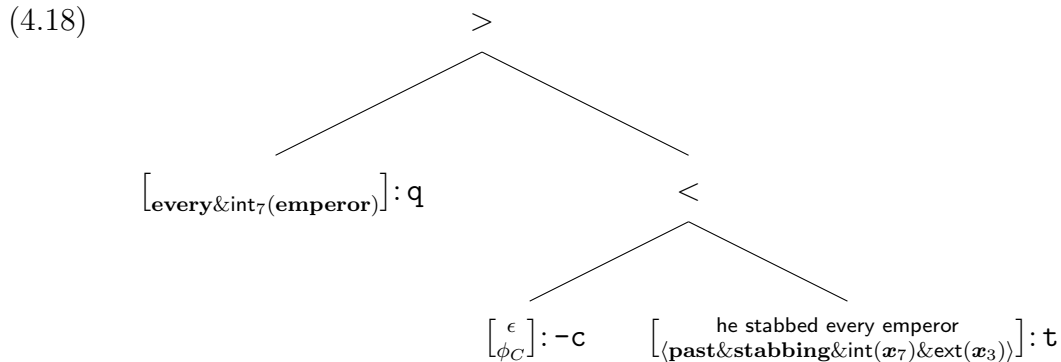
tions introduced by syntactic configurations).

While many of the details, particular syntactic details, of the account presented here are largely orthogonal to this central goal, the process of developing this account reveals some interesting consequences for our understanding of quantifier-raising, which I discuss in the remainder of this section.

4.6.1 Quantifiers as adjuncts

Recall that the formula combining a quantifier with a sentence simply conjoins the quantifier’s formula with the (existentially closed) formula produced at the TP phase, as in (4.22) — a departure from the assumption that complements and specifiers have their meanings adjusted via *int* and *ext*, given the structure in (4.18).

$$(4.22) \quad \langle \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor}) \ \& \ \langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(x_7) \ \& \ \mathit{ext}(x_3) \rangle \rangle$$



Since I have already assumed in previous chapters that adjuncts are interpreted as bare conjuncts, one may wonder whether it would be simpler to suppose that quantifiers are adjoined instead of selected as specifiers, without introducing the new “flat” compositional principle.

This alternative appears at first to be workable. We would be led to the formula in (4.33) for the CP phase.

$$(4.33) \quad \phi_C \ \& \ \mathit{int}(\langle \mathbf{past} \ \& \ \mathbf{stabbing} \ \& \ \mathit{int}(x_7) \ \& \ \mathit{ext}(x_3) \rangle) \ \& \ \mathbf{every} \ \& \ \mathit{int}_7(\mathbf{emperor})$$

Maintaining the assumption that ϕ_C is vacuous, this amounts to a formula “headed” by **every**, as desired, with the restrictor **emperor** and the “open sentence” marked,

quantifier sentence was ready to slot in to the formula for a two-quantifier sentence; recall (4.25) and (4.26).

Note that something very similar to this second, flatter compositional principle seems unavoidable for the TP phase. It appears very unlikely that we would want to derive the formula in (4.35) when interpreting the relationship between the T head and its VP complement.

(4.35) **past** & int(**stabbing** & int(x_7) & ext(x_3))

Given that we are apparently forced to introduce a compositional principle of flat conjunction and existential closure for the TP phase, and having observed that the existing argument-shifting compositional principle is inappropriate for the higher CP phases as noted in discussion of (4.33) and (4.34), the theoretical cost of introducing the second TP-tailored principle would be reduced if it could be used for the higher quantificational phases; §4.4 showed exactly this.

4.6.2 The target position of quantifier-raising

Although early versions of quantifier-raising (May 1977, 1985) postulated movement to a high position outside the TP roughly analogous to *wh*-movement, more recent proposals have suggested that quantifiers move to (and take scope from) a lower position inside the TP, either via scrambling (Johnson 2000) or via regular A-movement (Hornstein 1995). I have adopted the more traditional version in this chapter; there are both advantages and disadvantages to adopting the shorter-movement alternative instead.

The disadvantages concern the feasibility of quantifying from such low positions in an event-based semantics. As we have seen, quantifiers must operate on something roughly “truth-evaluable” — more correctly, something for which Tarskian conditions for satisfaction by assignments are defined. In an event-based framework, this means an existentially-closed (or equivalent) formula, rather than, say, an event predicate.

But if the contribution of the T head is an event predicate specifying tense (via whatever compositional principle), then the event predicate will have to remain unclosed until at least the T head, which is *higher* than many of the candidate scope-taking positions under short-QR proposals: in order to derive inverse scope readings of transitive sentences in the system of Hornstein (1995), for example, both the subject and object quantifiers must take scope *inside* the *vP*. If quantifiers raise to positions in the TP projection as suggested by Johnson (2000), the problem is not necessarily insoluble but at least awkward: for a subject to take scope in its specifier position of TP, for example, we would need the event predicate to be existentially closed at the T' level, immediately after the T head contributes a tense predicate. This is presumably manageable, but clashes with the idea that maximal projections are the relevant phases of interpretation.

The advantage of a short-QR alternative is that it has the potential to provide a way out of the SMC problems encountered with multiple quantifiers, as discussed in §4.5.2. An approach where quantificational positions are interspersed with thematic and Case positions can avoid the situation illustrated in (4.29), repeated here, where multiple quantifiers have, as their only remaining feature, a $-q$ feature.

(4.29) $\langle \text{some senator stabbed every emperor} : -t, \{ \text{some senator} : -q, \text{every emperor} : -q \} \rangle$

Kobele (2006) presents exactly this sort of account, along the lines of Hornstein (1995), within the MG formalism (using an event-free semantics). Kobele's system avoids the "competition" for $+q$ features that arises in (4.29) by having the object's quantificational position lower than the subject's Case position, i.e. lower than the point at which the subject begins to look for a $+q$ features. To avoid the same sort of competition for Case features, it is likewise necessary to assume that the object's Case position is lower than the subject's thematic position, i.e. lower than the point at which the subject begins to look for a $+k$ feature.

Chapter 5

Conclusion

In this thesis I have proposed a novel account of the syntactic phenomena of movement and adjunction, and the relationship between them. Two independently-appealing ideas that provided points of departure were the distinctive Conjunctivist hypothesis concerning the semantic composition of neo-Davidsonian logical forms, and a formal syntactic framework in which merging and moving are unified into a single operation.

In chapter 2 I presented a modification to the syntactic framework taking advantage of the fact that the unification of merge and move steps leaves room for the possibility of a constituent being only loosely associated with a phrase: more specifically, since bringing together merge and move steps was achieved via the introduction of an insertion operation that does not check features, it is possible to hypothesise that to be adjoined is to be (merely) inserted. In the resulting system adjuncts have a degree of syntactic freedom that arguments do not, from which the otherwise puzzling ability of adjuncts to be “either inside or outside” a maximal projection emerges naturally. This picture gains plausibility from the role that adjuncts play in Conjunctivist semantic composition, in particular the idea that the interpretation of adjuncts is, in a well-defined sense, simpler than that of complements of specifiers.

A consequence of the proposal in chapter 2 is that adjuncts and moving constituents have, in a certain sense, a common status; the proposal is therefore strengthened by any evidence that these two classes of constituents behave alike in certain

contexts. In chapter 3 I argued that evidence of exactly this sort can be found if we consider conditions on the domains out of which extraction can occur. Since extraction out of both moved constituents and adjoined constituents is generally prohibited, the framework developed in chapter 2 permits two otherwise unrelated constraints to be unified as one.

Since it adopts non-standard conceptions of movement and of semantic composition, a burden of proof arguably lies with this novel framework to show that quantificational expressions can be adequately dealt with. In chapter 4 I showed that a non-standard but descriptively adequate account of quantification can be constructed within the broad outlines of the Conjunctivist approach. This empirical ground is therefore not given up by adopting the framework that permitted the positive results of previous chapters.

More generally, this thesis constitutes a case study in (i) deriving explanations for syntactic patterns from a restrictive and independently motivated theory of the compositional semantics that syntactic structures must support, and (ii) using a computationally explicit framework to rigourously investigate the primitives and consequences of our theories, taking a “methodologically minimalist” approach to motivate revisions. The emerging picture that is suggested is one where some well-known central facts about the syntax and semantics of natural language hang together in a way that they otherwise would not.

Bibliography

- Abeillé, A. and Rambow, O. (2000). *Tree Adjoining Grammars*. CSLI Publications, Stanford, CA.
- Ades, A. E. and Steedman, M. (1982). On the order of words. *Linguistics and Philosophy*, 4:517–588.
- Ajdukiewicz, K. (1935). Die syntaktische Konnexität. *Studia Philosophica*, 1:1–27. English translation in McCall (1967).
- Bach, E. (1976). An extension of classical transformational grammar. In *Problems of Linguistic MetaTheory* (Proceedings of the 1976 Conference), pages 183–224. Michigan State University.
- Bach, E. (1979). Control in Montague Grammar. *Linguistic Inquiry*, 10(4):515–531.
- Bach, E. and Cooper, R. (1978). The NP-S analysis of relative clauses and compositional semantics. *Linguistics and Philosophy*, 2(1):145–150.
- Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.
- Baker, M. (1997). Thematic roles and syntactic structure. In Haegemann, L., editor, *Elements of Grammar*, pages 73–137. Kluwer, Dordrecht.
- Baker, M. (2003). *Lexical Categories: Nouns, verbs and adjectives*. Cambridge University Press, Cambridge.
- Bar-Hillel, Y. (1953). A Quasi-Arithmetical Notation for Syntactic Description. *Language*, 29(1):47–58.
- Bar-Hillel, Y., Gaifman, C., and Shamir, E. (1960). On categorial and phrase structure grammar. *Bulletin of the Research Council of Israel*, 9F(1):1–16.
- Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219.
- Belletti, A. (1988). The case of unaccusatives. *Linguistic Inquiry*, 19:1–34.
- Boeckx, C. (2008). *Bare Syntax*. Oxford University Press, Oxford.

- Boeckx, C. and Grohmann, K. K. (2007). Putting phases into perspective. *Syntax*, 10:204–222.
- Boolos, G. (1984). To be is to be the value of a variable (or some values of some variables). *The Journal of Philosophy*, 81:430–449. Reprinted in Boolos (1998).
- Boolos, G. (1998). *Logic, Logic and Logic*. Harvard University Press, Cambridge, MA.
- Bouchard, D. (1998). The distribution and interpretation of adjectives in French. *Probus*, 10:139–183.
- Bošković, Z. and Lasnik, H., editors (2007). *Minimalist Syntax: The Essential Readings*. Blackwell, Malden, MA.
- Burzio, L. (1986). *Italian Syntax*. Reidel, Dordrecht.
- Carlson, G. (1984). Thematic roles and their role in semantic interpretation. *Linguistics*, 22:259–279.
- Carnie, A. (2007). *Syntax: A Generative Introduction*. Blackwell, Malden, MA, second edition.
- Carpenter, B. (1997). *Type-Logical Semantics*. MIT Press, Cambridge, MA.
- Castañeda, H. N. (1967). Comments. In Rescher (1967).
- Cattell, R. (1976). Constraints on movement rules. *Language*, 52(1):18–50.
- Chametzky, R. A. (1996). *A Theory of Phrase Markers and the Extended Base*. State University of New York Press, Albany, NY.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton, The Hague.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
- Chomsky, N. (1973). Conditions on transformations. In Anderson, S. R. and Kiparsky, P., editors, *A Festschrift for Morris Halle*, pages 232–286. Holt, Rinehart and Winston, New York.
- Chomsky, N. (1975). Questions of form and interpretation. *Linguistic Analysis*, 1(1):75–109.
- Chomsky, N. (1981). *Lectures of Government and Binding*. Foris, Dordrecht.
- Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge, MA.
- Chomsky, N. (1994). Bare phrase structure. *MIT Occasional Papers in Linguistics*, 5.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge, MA.

- Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M. J., editor, *Ken Hale: A Life in language*. MIT Press, Cambridge, MA.
- Chomsky, N. (2002). *On Nature and Language*. Cambridge University Press, Cambridge.
- Chomsky, N. (2004). Beyond explanatory adequacy. In Belletti, A., editor, *Structures and Beyond*. Oxford University Press, Oxford.
- Chung, S. and Ladusaw, W. A. (2004). *Restriction and Saturation*. MIT Press, Cambridge, MA.
- Cinque, G. (1990). *Types of A'-dependencies*. MIT Press, Cambridge, MA.
- Cinque, G. (1994). Evidence for partial N-movement in the Romance DP. In Cinque, G., Koster, J., Pollock, J.-Y., and Zanuttini, R., editors, *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*. Georgetown University Press, Washington D.C.
- Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.
- Cooper, R. (1983). *Quantification and Syntactic Theory*. Reidel, Dordrecht.
- Cornell, T. L. (2004). Lambek calculus for transformational grammar. *Research on Language and Computation*, 2(1):105–126.
- Corver, N. (2005). Freezing effects. In Everaert and van Riemsdijk (2005), pages 383–406.
- Culicover, P. W. and Jackendoff, R. (2005). *Simpler Syntax*. Oxford University Press, New York.
- Davidson, D. (1967). The logical form of action sentences. In Rescher (1967), pages 81–95.
- Davidson, D. (1985). Adverbs of action. In Vermazen, B. and Hintikka, M., editors, *Essays on Davidson: Actions and Events*. Clarendon Press, Oxford.
- De Kuthy, K. and Meurers, W. D. (1998). Incomplete category fronting in German without remnant movement. In Schröder, B., Lenders, W., Hess, W., and Portele, T., editors, *Computers, Linguistics, and Phonetics between Language and Speech (Proceedings of the 4th Conference on Natural Language Processing, KONVENS 98)*, volume 1 of *Computer Studies in Language and Speech*, pages 57–68. Peter Lang, Frankfurt a.M.
- Dowty, D. (1982). Grammatical relations in Montague Grammar. In Jacobson, P. and Pullum, G. K., editors, *The Nature of Syntactic Representations*, pages 79–130. Reidel, Dordrecht.

- Dowty, D. (1988). Type-raising, functional composition and non-constituent coordination. In Oehrle et al. (1988), pages 153–198.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, 67:547–619.
- Enderton, H. B. (2001). *A Mathematical Introduction to Logic*. Academic Press, San Diego, second edition.
- Enderton, H. B. (2009). Second-order and higher-order logic. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University, fall 2009 edition.
- Epstein, S. D., Groat, E., Kawashima, R., and Kitahara, H. (1998). *A Derivational Approach to Syntactic Relations*. Oxford University Press, Oxford.
- Epstein, S. D. and Hornstein, N., editors (1999). *Working Minimalism*. MIT Press, Cambridge, MA.
- Epstein, S. D. and Seely, T. D., editors (2002). *Derivation and Explanation in the Minimalist Program*. Blackwell, Oxford.
- Everaert, M. and van Riemsdijk, H., editors (2005). *The Blackwell Companion to Syntax*. Wiley-Blackwell, Malden, MA.
- Fisher, C., Hall, D. G., Rakowitz, S., and Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92:333–375.
- Fox, D. and Pesetsky, D. (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics*, 31(1–2):1–45.
- Frank, R. (2002). *Phrase Structure Composition and Syntactic Dependencies*. MIT Press, Cambridge, MA.
- Freidin, R. (1986). Fundamental issues in the theory of binding. In Lust, B., editor, *Studies in the acquisition of anaphora*, volume 1, pages 151–188. Reidel, Dordrecht.
- Frey, W. and Gärtner, H.-M. (2002). On the treatment of scrambling and adjunction in minimalist grammars. In Jäger, G., Monachesi, P., Penn, G., and Wintner, S., editors, *Proceedings of Formal Grammar 2002*, pages 41–52.
- Fukui, N. (1986). *A Theory of Category Projection and Its Applications*. PhD thesis, MIT.
- Gärtner, H.-M. and Michaelis, J. (2003). A note on countercyclicity and Minimalist Grammars. In Jäger, G., Monachesi, P., Penn, G., and Wintner, S., editors, *Proceedings of FGVienna: The 8th Conference on Formal Grammar*, pages 103–114.

- Gärtner, H.-M. and Michaelis, J. (2005). A note on the complexity of constraint interaction: Locality conditions and Minimalist Grammars. In Blache, P., Stabler, E. P., Busquets, J., and Moot, R., editors, *Logical Aspects of Computational Linguistics*, volume 3492 of *Lecture Notes in Computer Science*, pages 114–130. Springer.
- Gärtner, H.-M. and Michaelis, J. (2007). Locality conditions and the complexity of Minimalist Grammars: A preliminary survey. In *Model-Theoretic Syntax at 10*, Proceedings of the ESSLLI Workshop (Dublin), pages 87–98.
- Geach, P. (1972). A program for syntax. In Davidson, D. and Harman, G., editors, *Semantics of Natural Language*, pages 483–497. Reidel, Dordrecht.
- Gillon, B. S. (1987). The readings of plural noun phrases in english. *Linguistics and Philosophy*, 10(2):199–219.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1:3–55.
- Gropen, J., Pinker, S., Hollander, M., and Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, 41:153–195.
- Haegeman, L. (1994). *Introduction to Government and Binding Theory*. Blackwell, Malden, MA, second edition.
- Hale, K. and Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. In Hale, K. and Keyser, S. J., editors, *The View from Building 20*, pages 53–109. MIT Press, Cambridge, MA.
- Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell, Oxford.
- Herburger, E. (2000). *What Counts: Focus and Quantification*. MIT Press, Cambridge, MA.
- Higginbotham, J. (1985). On semantics. *Linguistic Inquiry*, 16(4):547–593.
- Higginbotham, J. and May, R. (1981). Questions, quantifiers and crossing. *The Linguistic Review*, 1(1):41–80.
- Hinzen, W. (2009). Hierarchy, merge and truth. In Piattelli-Palmarini, M., Uriagereka, J., and Salaburu, P., editors, *Of Minds and Language: A dialogue with Noam Chomsky in the Basque Country*, pages 123–141. Oxford University Press, New York.
- Hodges, W. (2008). Tarski’s truth definitions. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University, fall 2008 edition.

- Hornstein, N. (1995). *Logical Form*. Blackwell, Oxford.
- Hornstein, N. (2001). *Move! A minimalist theory of construal*. Blackwell, Oxford.
- Hornstein, N. (2009). *A Theory of Syntax: Minimal Operations and Universal Grammar*. Cambridge University Press, New York.
- Hornstein, N. and Nunes, J. (2008). Adjunction, labeling, and bare phrase structure. *Biolinguistics*, 2(1):57–86.
- Hornstein, N. and Pietroski, P. (2009). Basic operations: Minimal syntax-semantics. *Catalan Journal of Linguistics*, 8:113–139.
- Hornstein, N. and Uriagereka, J. (2002). Reprojections. In Epstein and Seely (2002), pages 106–132.
- Huang, C. T. J. (1982). *Logical relations in Chinese and the theory of grammar*. PhD thesis, MIT.
- Hunter, T. and Conroy, A. (2009). Children’s restrictions on the meanings of novel determiners: An investigation of conservativity. In Chandler, J., Franchini, M., Lord, S., and Rheiner, G.-M., editors, *BUCLD 33 Proceedings*, pages 245–255, Somerville, MA. Cascadilla Press.
- Jäger, G. (2005). *Anaphora and Type Logical Grammar*. Springer, Dordrecht.
- Johnson, K. (1986). *A case for movement*. PhD thesis, MIT.
- Johnson, K. (2000). How far will quantifiers go? In Martin, R., Michaels, D., and Uriagereka, J., editors, *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, pages 187–210. MIT Press, Cambridge, MA.
- Joshi, A. K. (1987). An Introduction to Tree Adjoining Grammars. In Manaster-Ramer (1987), pages 87–115.
- Joshi, A. K. and Schabes, Y. (1997). Tree-adjoining grammars. In Rozenberg, G. and Salomaa, A., editors, *Handbook of Formal Languages*, volume 3, pages 69–124. Springer, New York, NY.
- Jurka, J. (2009). Gradient acceptability and subject islands in German. Ms., University of Maryland.
- Kako, E. and Wagner, L. (2001). The semantics of syntactic structures. *Trends in Cognitive Sciences*, 5:102–108.
- Kamp, H. (1975). Two theories of adjectives. In Keenan, E., editor, *Formal Semantics of Natural Language*, pages 23–55. Cambridge University Press, Cambridge.
- Kamp, H. and Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57:129–191.

- Kayne, R. (1984). *Connectedness and Binary Branching*. Foris, Dordrecht.
- Kayne, R. (1994). *The antisymmetry of syntax*. MIT Press, Cambridge, MA.
- Keenan, E. and Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9:253–326.
- Kitahara, H. (1997). *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, MA.
- Klein, E. and Sag, I. A. (1985). Type-driven translation. *Linguistics and Philosophy*, 8(2):163–201.
- Kobele, G. M. (2006). *Generating Copies: An investigation into Structural Identity in Language and Grammar*. PhD thesis, UCLA.
- Kobele, G. M. (2009). Syntactic identity in survive minimalism: Ellipsis and the derivational identity hypothesis. In Putnam, M., editor, *Towards a Derivational Syntax: Survive Minimalism*, pages 195–229. John Benjamins, Philadelphia.
- Kobele, G. M. (in press). Inverse linking via function composition. *Natural Language Semantics*.
- Koopman, H. and Sportiche, D. (1991). The position of subjects. *Lingua*, 85:211–258.
- Kratzer, A. (1996). Severing the external argument from its verb. In Rooryck, J. and Zaring, L., editors, *Phrase Structure and the Lexicon*. Kluwer, Dordrecht.
- Krifka, M. (1992). Thematic relations as links between nominal reference and temporal constitution. In Sag, I. A. and Szabolcsi, A., editors, *Lexical Matters*. CSLI Publications, Stanford, CA.
- Kroch, A. (1987). Unbounded dependencies and subjacency in tree adjoining grammar. In Manaster-Ramer (1987), pages 143–172.
- Kroch, A. (1989). Asymmetries in long distance extraction in tree adjoining grammar. In Baltin, M. and Kroch, A., editors, *Alternative conceptions of phrase structure*, pages 66–98. University of Chicago Press, Chicago.
- Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170.
- Lambek, J. (1961). On the calculus of syntactic types. In Jakobson, R., editor, *Structure of Language and its Mathematical Aspects*, pages 166–178. American Mathematical Society, Providence, RI.
- Landman, F. (1996). Plurality. In Lappin, S., editor, *The Handbook of Contemporary Semantic Theory*, pages 425–457. Blackwell, Oxford.

- Landman, F. (2000). *Events and Plurality: The Jerusalem Lectures*. Kluwer, Dordrecht.
- Larson, R. and Segal, G. (1995). *Knowledge of Meaning*. MIT Press, Cambridge, MA.
- Larson, R. K. (1998). Events and modification in nominals. In *Proceedings of SALT VIII*, pages 145–168, Ithaca, NY. CLC Publications.
- Lasnik, H. (1998). Some reconstruction riddles. In *U. Penn. Working Papers in Linguistics*, volume 5.1, pages 83–98.
- Lasnik, H. and Park, M.-K. (2003). The EPP and the subject condition under sluicing. *Linguistic Inquiry*, 34(4):649–660.
- Lasnik, H. and Saito, M. (1992). *Move α : Conditions on its application and output*. MIT Press, Cambridge, MA.
- Lasnik, H. and Uriagereka, J. (1988). *A Course in GB Syntax*. MIT Press, Cambridge, MA.
- Lasnik, H. and Uriagereka, J. (2005). *A Course in Minimalist Syntax*. Blackwell, Malden, MA.
- Lebeaux, D. (1988). *Language acquisition and the form of the grammar*. PhD thesis, University of Massachusetts, Amherst.
- Lebeaux, D. (2000). *Language acquisition and the form of the grammar*. John Benjamins, Philadelphia.
- Lecomte, A. (2004). Rebuilding MP on a logical ground. *Research on Language and Computation*, 2(1):27–55.
- Lecomte, A. and Retoré, C. (2001). Extending lambek grammars: a logical account of minimalist grammars. In *ACL '01: Proceedings of the 39th Annual Meeting of Association for Computational Linguistics*, pages 362–369, Morristown, NJ. Association for Computational Linguistics.
- Levin, B. and Rappaport Hovav, M. (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press, Cambridge, MA.
- Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, Use and Interpretation of Language*, pages 302–323. de Gruyter, Berlin.
- Link, G. (1998). *Algebraic Semantics in Language and Philosophy*. CSLI Publications, Stanford, CA.
- Manaster-Ramer, A., editor (1987). *Mathematics of Language*. John Benjamins, Amsterdam.

- Martin, R., Michaels, D., and Uriagereka, J., editors (2000). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. MIT Press, Cambridge, MA.
- May, R. (1977). *The Grammar of Quantification*. PhD thesis, MIT.
- May, R. (1985). *Logical Form*. MIT Press, Cambridge, MA.
- McCall, S., editor (1967). *Polish Logic, 1920–1939*. Oxford University Press, New York.
- McNally, L. and Boleda, G. (2004). Relational adjectives as properties of kinds. *Empirical Issues in Formal Syntax and Semantics*, 5:179–196.
- Michaelis, J. (2001a). Derivational minimalism is mildly context-sensitive. In Moortgat, M., editor, *Logical Aspects of Computational Linguistics*, volume 2014 of *Lecture Notes in Computer Science*, pages 179–198. Springer.
- Michaelis, J. (2001b). *On Formal Properties of Minimalist Grammars*. PhD thesis, Universität Potsdam.
- Montague, R. (1974). *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven, CT. Edited and with an introduction by Richmond H. Thomason.
- Moortgat, M. (1988). *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*. Foris, Dordrecht.
- Moortgat, M. (1996). Multimodal linguistic inference. *Journal of Language, Logic and Information*, 5:349–385.
- Moortgat, M. (1997). Categorial type logics. In van Benthem, J. and ter Meulen, A., editors, *Handbook of Logic and Language*, pages 93–177. Elsevier Science, New York.
- Morrill, G. (1994). *Type Logical Grammar: Categorial Logic of Signs*. Kluwer, Dordrecht.
- Müller, G. (1998). *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Kluwer, Dordrecht.
- Muysken, P. (1982). Parametrizing the notion ‘head’. *Journal of Linguistic Research*, 2:57–75.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17:357–374.
- Nunes, J. and Uriagereka, J. (2000). Cyclicity and extraction domains. *Syntax*, 3(1):20–43.

- Oehrle, R. T., Bach, E., and Wheeler, D., editors (1988). *Categorial Grammars and Natural Language*. Reidel, Dordrecht.
- Parsons, T. (1990). *Events in the semantics of English*. MIT Press, Cambridge, MA.
- Partee, B. (1975). Montague grammar and transformational grammar. *Linguistic Inquiry*, 6(2):203–300.
- Partee, B. (1986). Noun phrase interpretation and type-shifting principles. In Groenendijk, J., de Jongh, D., and Stokhof, M., editors, *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, pages 115–143. Foris, Dordrecht.
- Partee, B. and Rooth, M. (1983). Generalized conjunction and type ambiguity. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, Use and Interpretation of Language*, pages 361–383. de Gruyter, Berlin.
- Pietroski, P. M. (2003a). Quantification and second-order monadicity. *Philosophical Perspectives*, 17(1):259–298.
- Pietroski, P. M. (2003b). Small verbs, complex events. In Antony, L. and Hornstein, N., editors, *Chomsky and His Critics*, pages 179–214. Blackwell, Oxford.
- Pietroski, P. M. (2005). *Events and Semantic Architecture*. Oxford University Press, Oxford.
- Pietroski, P. M. (2006). Interpreting concatenation and concatenates. *Philosophical Issues*, 16(1):221–245.
- Pietroski, P. M. (forthcoming). Minimal semantic instructions. In Boeckx, C., editor, *The Oxford Handbook of Linguistic Minimalism*. Oxford University Press, Oxford.
- Pollard, C. (1984). *Generalized phrase structure grammars, head grammars and natural language*. PhD thesis, Stanford University.
- Rescher, N., editor (1967). *The Logic of Decision and Action*. University of Pittsburgh Press, Pittsburgh.
- Retoré, C. and Stabler, E. (1999). Resource logics and minimalist grammars. Technical Report 3780, INRIA.
- Retoré, C. and Stabler, E. (2004). Generative grammars in resource logics. *Research on Language and Computation*, 2(1):3–25.
- Rizzi, L. (1990). *Relativized Minimality*. MIT Press, Cambridge, MA.
- Roberts, I. (2000). Caricaturing dissent. *Natural Language and Linguistic Theory*, 18(4):849–857.
- Ross, J. R. (1969). *Constraints on Variables in Syntax*. PhD thesis, MIT.

- Ross, J. R. (1974). Three batons for cognitive psychology. In Weimer, W. B. and Palermo, D., editors, *Cognition and the Symbolic Processes*, pages 63–124. Lawrence Erlbaum, Hillsdale, NJ.
- Schein, B. (1993). *Plurals and Events*. MIT Press, Cambridge, MA.
- Schein, B. (2002). Events and the semantic content of thematic relations. In Preyer, G. and Peter, G., editors, *Logical Form and Language*, pages 263–344. Oxford University Press, Oxford.
- Schein, B. (2006). Plurals. In Lepore, E. and Smith, B. C., editors, *The Oxford Handbook of Philosophy of Language*, pages 716–767. Oxford University Press, Oxford.
- Speas, M. (1986). *Adjunctions and projections in syntax*. PhD thesis, MIT.
- Stabler, E. P. (1997). Derivational minimalism. In Retoré, C., editor, *Logical Aspects of Computational Linguistics*, volume 1328 of *Lecture Notes in Computer Science*, pages 68–95. Springer.
- Stabler, E. P. (2001). Recognizing head movement. In de Groote, P., Morrill, G., and Retoré, C., editors, *Logical Aspects of Computational Linguistics*, volume 2099 of *Lecture Notes in Computer Science*, pages 254–260. Springer.
- Stabler, E. P. (2006). Sideways without copying. In Monachesi, P., Penn, G., Satta, G., and Wintner, S., editors, *Proceedings of the 11th Conference on Formal Grammar*.
- Stabler, E. P. and Keenan, E. L. (2003). Structural similarity within and among languages. *Theoretical Computer Science*, 293:345–363.
- Steedman, M. (1988). Combinators and grammars. In Oehrle et al. (1988), pages 417–442.
- Steedman, M. (1996). *Surface Structure and Interpretation*. MIT Press, Cambridge, MA.
- Steedman, M. (2000). *The Syntactic Process*. MIT Press, Cambridge, MA.
- Stepanov, A. (2001). *Cyclic Domains in Syntactic Theory*. PhD thesis, University of Connecticut.
- Stepanov, A. (2007). The End of the CED? Minimalism and extraction domains. *Syntax*, 10(1):80–126.
- Stroik, T. (1999). The Survive Principle. *Linguistic Analysis*, 29:278–303.
- Stroik, T. (2009). *Locality in Minimalist Syntax*. MIT Press, Cambridge, MA.
- Takahashi, D. (1994). *Minimality of Movement*. PhD thesis, University of Connecticut.

- Takano, Y. (2000). Illicit remnant movement: An argument for feature-driven movement. *Linguistic Inquiry*, 31(1):141–156.
- Tarski, A. (1983). The concept of truth in formal languages. In Corcoran, J., editor, *Logic, Semantics and Metamathematics*, pages 152–278. Hackett Publishing Company, Indianapolis. Translated by J.H. Woodger.
- Torrego, E. (1985). On empty categories in nominals. Ms., University of Massachusetts.
- Truswell, R. (2007). Extraction from adjuncts and the structure of events. *Lingua*, 117:1355–1377.
- Unger, C. (2010). *A computational approach to the syntax of displacement and the semantics of scope*. PhD thesis, Utrecht Institute of Linguistics.
- Uriagereka, J. (1998). *Rhyme and Reason: An Introduction to Minimalist Syntax*. MIT Press, Cambridge, MA.
- Uriagereka, J. (1999). Multiple spell-out. In Epstein and Hornstein (1999), pages 251–282.
- Uriagereka, J. (2008). *Syntactic Anchors: On Semantic Structuring*. Cambridge University Press, New York.
- Uriagereka, J. and Pietroski, P. (2002). Dimensions of natural language. In Uriagereka, J., editor, *Derivations: Exploring the Dynamics of Syntax*, pages 266–287. Routledge, London.
- Valois, D. (2005). Adjectives: Order within DP and attributive APs. In Everaert and van Riemsdijk (2005), pages 61–82.
- Vergnaud, J.-R. (1974). *French Relative Clauses*. PhD thesis, MIT.
- Wexler, K. and Culicover, P. (1981). *Formal Principles of Language Acquisition*. MIT Press, Cambridge, MA.
- Williams, A. (2005). *Complex Causatives and Verbal Valence*. PhD thesis, University of Pennsylvania.
- Williams, A. (2008). Patients in Igbo and Mandarin. In Doelling, J., Heyde-Zybatow, T., and Schaefer, M., editors, *Event structures in linguistic form and interpretation*, pages 3–30. Mouton de Gruyter, Berlin.
- Yoshida, M. (2006). *Constraints and mechanisms in long-distance dependency formation*. PhD thesis, University of Maryland.