

Multiword expressions and the lexicon



Jamie Y. Findlay
Linacre College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Hilary Term 2019

For my father.

Acknowledgements

It is a cliché to say that a project like this, in spite of only bearing a single name on the cover, was a collaborative effort, but that doesn't stop it being true all the same. I have been very fortunate to have been supported at every stage of my academic career not only intellectually, but also emotionally and financially, and I only hope that this thesis goes some way to vindicating the generosity that has been so graciously bestowed on me.

First and foremost, I must thank my supervisors, Ash Asudeh and Mary Dalrymple. Their intellectual influence can be felt throughout this thesis. They have made me a better linguist, a better writer, and a better thinker. Perhaps at least as importantly, though, I would like to thank them for all their other contributions to my well-being: for their encouragement, emotional support, career advice, and all the other things they have done for me that have made this work possible and have improved my life in innumerable ways. I really cannot imagine better supervisors, and I intend to spend the rest of my career striving to live up to the excellent example that they have set.

Heartily thanks should also go to my two examiners, Tom Wasow and Stephen Pulman, for a lively and stimulating *viva voce* exam, and for their detailed and helpful comments, which have resulted in a much improved thesis.

There are many others who have had an important hand in the intellectual development of this thesis: Doug Arnold, Sascha Bargmann, Alex Biswas, Kersti Börjars, Miriam Butt, Dag Haug, Stephen Jones, Ron Kaplan, Timm Lichte, Joey Lovstrand, John Lowe, Agnieszka Patejuk, Adam Przepiórkowski, Stephen Pulman, Manfred Sailer, Andy Spencer, Gert Webelhuth, and others here unnamed, to whom I can only apologise for the omission. Conversations, correspondences, and collaborations with these people have greatly improved the contents of this thesis, not to mention given me tremendous pleasure over the years.

My love for linguistics was kindled as an undergraduate at Oxford, and my teachers there, especially Mary Dalrymple, deserve all the credit for starting me on this path. A special mention should also go to Chiara Cappellaro: she was my first tutor and I was her first student. It seems like a very long time indeed since that first essay on Saussure!

Before I returned to Oxford to complete my MPhil, I spent a term at McGill, and there I had my eyes opened to a whole other world of linguistics. I would like to especially thank Alan Bale for giving me the best historical overview of 20th century syntactic theory I can imagine, and to Bernhard Schwarz for introducing me to the wonders of formal semantics.

For much of my (fairly long) stint as a graduate student, I have been lucky enough to be funded by both the Arts and Humanities Research Council, and later also by Linacre College. Without this support it would have been impossible for me to pursue further study, and so I cannot express my gratitude sufficiently to these organisations for enabling me to do what I love every day.

At Oxford, I have been fortunate enough to work with colleagues who I admire greatly, and who have been kind enough to share their guidance and expertise with me. Thanks go in particular to Kerstin Hoge, Matt Husband, Martin Maiden, and Louise Mycock. I have also had the privilege of making many fine friends along the way, and they have been invaluable in keeping me sane and making this whole endeavour more fun than it has any right to be: Alex Biswas, Stephen Jones, Yoolim Kim, Marjolein Poortvliet, Ana Werkmann Horvat, you have all been amazing. Of course, I also made the finest of friends, my partner, Alicia Cleary-Venables, whose blood, sweat, and tears are in this thesis almost as much as mine.

All of this faffing around with books and words started before I got to university, of course, and my family deserve the heartiest thanks of all, for everything that they have done and continue to do for me. My grannie, Joy, and my aunt, Tracey, have always been my staunchest supporters, and I know they will be breathing a sigh of relief now that this thesis is done. My mum always pushed me to do my best, and always showed me the love and belief I needed to succeed. Any intellectual worth I may have picked up over the years owes its origins to her. Finally, I want to thank my dad, who, although he may not have always understood what I was doing, was nonetheless unwavering in his support and in his pride. Crushingingly, he passed away during the final stages of this project, and so he never got to see it completed. This thesis is dedicated to him.

Abstract

The term ‘multiword expression’ (MWE) refers to a diverse group of linguistic phenomena connected by the fact that they do not fit neatly into the word-phrase dichotomy: like phrases, they appear to be made up of multiple words; but, like words, they have idiosyncrasies (of meaning, form, or both) which must be learned. Examples include idioms, like *spill the beans*, ‘reveal the secrets’; prepositional verbs, like *rely on*; or light verb constructions, like *have a break*. Such expressions are of interest to linguistic theory because they straddle the boundary between the productive and idiosyncratic components of language, i.e. between what is (or can be) computed and what must be memorised.

The thesis begins by investigating a number of properties of MWEs which any theory must give an account of: for example, they can be idiomatic in a variety of ways, such as by having unpredictable meanings, containing words which do not appear outside of the expression (*run amok*), or having a syntactic structure not attested elsewhere (*by and large*); they are also variably flexible, sometimes allowing modification (*they spilled the financial beans*), passivisation (*the beans were spilled*), or extraction (*the beans that they spilled*); and they have a specific psycholinguistic profile: idioms are processed faster than compositional phrases.

A theory of MWEs has to give an account of how the more idiomatic properties square with the more productive ones. This thesis aims to do just that. It is argued that there are essentially four families of theory one might entertain: one which treats MWEs as big words, one which treats them as phrases which are composed of special versions of the words they contain, one which obtains a literal parse of the phrase first and then translates the literal meaning into the idiomatic one, and one which treats them as stored phrases which possess internal structure. The first approach treats MWEs as wholly word-like (i.e. as atomic units), and the next two treat them as wholly phrase-like (i.e. as composed of smaller independent units). It is only the fourth which takes the tension inherent in MWEs seriously, recognising that they are units in their own right, but also have internal structure. I argue, therefore, that this is the kind of theory we should prefer.

The proposal I develop hinges formally on the use of a Tree-Adjoining Grammar (TAG), a tree-rewriting formalism which allows us to store (internally

structured) phrases as elementary objects. This is integrated into the architecture of Lexical Functional Grammar (LFG), replacing the usual context-free grammar used there. By treating TAG elementary trees as descriptions, we are able to reduce LFG lexical entries to nothing but a set of constraints (on all levels of linguistic structure). The resulting framework has a number of attractive properties: it inherits from LFG the ability to separate configurational from functional information in the syntax, and gains from TAG a natural capacity to represent MWEs, and the ability to 'lexicalise' the grammar (associate every elementary structure with a lexical item), which I argue is a desirable property for a lexicalist theory like LFG. Because of this last result, I call the theory Lexicalised LFG.

Contents

List of Figures	xiii
List of Tables	xvii
1 Introduction	1
1.1 What is a word?	2
1.2 Mismatches between the different conceptions of ‘word’	4
1.3 Multiword expressions	5
1.4 Outline of the thesis	9
2 Multiword expressions	11
2.1 MWEs or constructions?	12
2.1.1 Formal and substantive constructions	13
2.1.2 Lexical integrity	16
2.1.3 Summary	20
2.2 Properties of MWEs	21
2.2.1 Semantic idiomaticity	21
2.2.2 Syntactic idiomaticity	22
2.2.3 Lexical idiomaticity	23
2.2.4 The real words problem	24
2.2.5 Morphosyntactic flexibility	25
2.2.6 Syntactic flexibility	26
2.2.7 Decomposability	35
2.2.8 Psycholinguistic properties	41
2.2.9 Extended uses of idioms	43
2.3 Chapter summary	47
3 Approaches to MWEs	49
3.1 The monolexemic approach	50
3.2 The space of possibilities	55
3.3 Lexical ambiguity	58
3.3.1 Strengths of the lexical ambiguity approach	59
3.3.2 The collocational challenge	60

3.3.3	Non-decomposable MWEs	67
3.3.4	Other weaknesses	74
3.3.5	Summary	76
3.4	One-to-many lexical entries	77
3.5	Semantic mapping	78
3.5.1	Strengths of the semantic mapping approach	82
3.5.2	Weaknesses of the semantic mapping approach	88
3.5.3	Extended uses of idioms	92
3.5.4	Summary	96
3.6	Construction-based theories	96
3.6.1	Strengths of the construction-based approach	99
3.6.2	The real words problem	106
3.6.3	Summary	108
3.7	Comparing the approaches	108
3.7.1	Semantic idiomaticity	110
3.7.2	Syntactic idiomaticity	111
3.7.3	Lexical idiomaticity	112
3.7.4	The real words problem	113
3.7.5	Internal complexity	118
3.7.6	Psycholinguistic plausibility	121
3.7.7	Extended uses	122
3.7.8	Chapter summary	122
4	Theoretical background	125
4.1	Formal grammars	126
4.1.1	Formalisation	126
4.1.2	Theories of grammar	128
4.2	Tree-Adjoining Grammar	131
4.2.1	Introduction	131
4.2.2	Components of a grammar	133
4.2.3	Elementary structures	134
4.2.4	Combining operations	137
4.2.5	Summary	142
4.3	Lexical Functional Grammar	144
4.3.1	Introduction	144
4.3.2	Two levels of syntax	145
4.3.3	C-structure	147
4.3.4	F-structure	149
4.3.5	Connecting c-structure and f-structure	154

4.3.6	The correspondence architecture	165
4.3.7	Templates	166
4.3.8	Summary	167
4.4	Semantics	168
4.4.1	Meaning language	168
4.4.2	Glue Semantics	175
4.5	Chapter summary	193
5	Lexicalised LFG	197
5.1	Incorporating TAG into LFG	198
5.1.1	TAG-LFG	199
5.1.2	Structures and descriptions of structures	201
5.2	A language for tree descriptions	202
5.2.1	Phrase structure rules and lexical entries	203
5.2.2	A description-based TSG	205
5.2.3	Some linguistic constraints	207
5.3	Lexicalisation	209
5.3.1	Lexicalised vs. non-lexicalised	209
5.3.2	Lexicalising a CFG	211
5.3.3	Advantages of lexicalised grammars	214
5.3.4	<i>Lexical</i> Functional Grammar	216
5.4	Description-based TAG	217
5.4.1	Motivation	217
5.4.2	Partial descriptions of trees	219
5.4.3	Constraining the formalism	222
5.5	Lexicalised LFG	226
5.5.1	Noun phrases	229
5.5.2	Verbs	231
5.5.3	Modification	234
5.5.4	Long-distance dependencies	238
5.5.5	Lexical rules	243
5.6	Chapter summary	259
6	MWEs in Lexicalised LFG	261
6.1	Non-verbal MWEs	262
6.1.1	Syntax	263
6.1.2	Semantics	272
6.1.3	Non-verbal MWEs: a summary	295
6.2	Verbal MWEs	295
6.2.1	Decomposable expressions	296

6.2.2	Non-decomposable expressions	308
6.2.3	Verbal MWEs: a summary	319
6.3	Lexical flexibility	321
6.3.1	<i>Grist to the mill</i>	322
6.3.2	Other kinds of lexical flexibility	327
6.3.3	Summary	334
6.4	Chapter summary	335
7	Conclusion	339
7.1	Summary of the thesis	339
7.2	Assessing the proposal	345
7.3	Future research	347
Appendices		
A	Sources for attested examples	353
B	Semantic types	359
	Bibliography	361

List of Figures

3.1	Understanding “inferred” idioms	93
3.2	TAG derivation for <i>strings which Eddington claimed Benjamin pulled</i>	105
3.3	TAG derivation for <i>strings which Ross knew Eddington claimed Benjamin pulled</i>	106
4.1	The Chomsky hierarchy	131
4.2	C- and f-structures for <i>Worf has launched a torpedo</i>	146
4.3	Annotated c-structure for <i>Keiko loves the gift</i>	158
4.4	F-structure for <i>Keiko loves the gift</i>	160
4.5	Incorrect, non-minimal f-structure for <i>Keiko loves the gift</i>	160
4.6	Glue proof for <i>Janak was poked</i>	187
4.7	F-structure for <i>Odo apprehended the very dangerous criminal</i>	189
4.8	Glue proof for <i>the very dangerous criminal</i>	194
4.9	Glue proof for <i>Odo quickly apprehended the very dangerous criminal</i>	195
5.1	Lexical entry for <i>boy</i>	229
5.2	Possible lexical entry for <i>Benjamin</i>	230
5.3	Lexical entry for <i>some</i>	231
5.4	Lexical entry for <i>repaired</i>	232
5.5	Lexical entry for <i>said</i>	233
5.6	Lexical entry for perfective <i>has</i>	234
5.7	Lexical entry for <i>broken</i>	234
5.8	Preliminary lexical entry for <i>quickly</i>	236
5.9	Lexical entry for <i>quickly</i>	237
5.10	Lexical entry for <i>very</i>	238
5.11	Lexical entry for a relative clause headed by the object of <i>repaired</i>	239
5.12	Annotated c-structure for <i>the computer the processor of which Julian said Miles repaired</i>	244
5.13	F-structure for <i>the computer the processor of which Julian said Miles repaired</i>	245
5.14	Intermediate Glue proof 1 for <i>the computer the processor of which Julian said Miles repaired</i>	246

5.15	Final Glue proof for <i>the computer the processor of which Julian said Miles repaired</i>	247
5.16	Lexical entry for a relative clause headed by the object of <i>repaired</i> , without a relative pronoun	248
5.17	Lexical entry for a relative clause headed by the subject of <i>repaired</i>	249
5.18	Lexical rule for object relative clause (with relative pronoun) . . .	252
5.19	Lexical entry for <i>gave</i>	253
5.20	Lexical entry for a relative clause headed by the object of <i>gave</i> . .	254
5.21	Lexical rule for a subject relative clause	255
5.22	Lexical rule for the short passive	256
5.23	Lexical entry for <i>weighed</i>	258
6.1	Lexical entry for <i>by and large</i>	265
6.2	Lexical entry for <i>all of a sudden</i>	265
6.3	Lexical entry for <i>kith and kin</i>	269
6.4	Lexical entry for <i>fly in the ointment</i>	271
6.5	Schematic f-structure for <i>fly in the ointment</i>	271
6.6	Annotated c-structure for <i>some real kith and kin</i>	274
6.7	F-structure for <i>some real kith and kin</i>	274
6.8	Glue proof for <i>some real kith and kin</i>	274
6.9	F-structure for <i>the empirical fly in my theoretical ointment</i>	276
6.10	Glue proof for <i>the empirical fly in my theoretical ointment</i>	277
6.11	F-structure for <i>All of a sudden, Jadzia died</i>	278
6.12	Glue proof for <i>All of a sudden, Jadzia died</i>	279
6.13	F-structure for <i>All of a complete sudden, Jadzia died</i>	281
6.14	Glue proof for <i>All of a complete sudden, Jadzia died</i>	283
6.15	F-structure for <i>All of a random sudden, Jadzia died</i>	285
6.16	Glue proof for <i>All of a random sudden, Jadzia died</i>	287
6.17	F-structure for <i>All of a terrible sudden, Jadzia died</i>	291
6.18	Glue proof for <i>All of a terrible sudden, Jadzia died.</i>	293
6.19	Glue proof for the unattested reading of <i>all of a terrible sudden</i> . .	294
6.20	Lexical entry for <i>spilled...beans</i>	297
6.21	Annotated c-structure for <i>Julian spilled some important beans</i>	301
6.22	F-structure for <i>Julian spilled some important beans</i>	302
6.23	Glue proof for <i>Julian spilled some important beans</i>	303
6.24	Lexical rule for the short passive	304
6.25	Lexical rule for a subject relative clause	305
6.26	Lexical rule for an object relative clause (with a relative pronoun)	306
6.27	Lexical entries for a subject relative clause, object relative clause, and passive clause derived from <i>spill the beans</i>	307

6.28	Lexical entry for <i>kick the bucket</i>	308
6.29	F-structure for <i>Jadzia kicked the bucket</i>	309
6.30	F-structure for <i>Julian kicked the social bucket</i>	316
6.31	Glue proof for <i>Julian kicked the social bucket</i>	317
6.32	Lexical entry for <i>grist to the mill</i>	323
6.33	Lexical entry for verbal <i>grist to the mill</i>	325
6.34	Lexical entry for <i>gave</i>	325
6.35	Lexical entry for <i>kick up the backside</i> under the intensional approach	331

List of Tables

2.1	Fraser's (1970: 40–41) levels of analysis for idioms	30
3.1	The space of analytical possibilities for MWEs	56
3.2	1st person singular forms of MONEO 'advise' in Latin	78
3.3	Comparison of various approaches to MWEs	109
4.1	Grammatical functions in LFG	152

1

Introduction

Contents

1.1	What is a word?	2
1.2	Mismatches between the different conceptions of 'word' . . .	4
1.3	Multiword expressions	5
1.4	Outline of the thesis	9

Linguistic work in the tradition of formal language theory treats languages as sets of sentences, and assigns to the grammar the task of circumscribing this set. But for natural languages, this task is not one of simple enumeration, since, as Chomsky (1957) and others have long pointed out, the number of possible sentences in a natural languages is infinitely large, and if the grammar is to be represented in the mind/brain, which is a finite object, it cannot then be as a list of infinite length. Instead, we break down sentences into smaller units (words, morphemes, phonemes, . . .), and define procedures for generating the larger units out of the smaller ones.¹ Provided that these smaller units *are* finite in number, they *can* be listed. At the level of syntactic representation, these

¹This notion of generation is what gave rise to the moniker of 'generative' grammar, a label applied to many modern syntactic theories, emerging from work by Chomsky (1957, 1965). Even theories which have since diverged from the Chomskyan orthodoxy can still be viewed as generative in this foundational sense, e.g. Lexical Functional Grammar (LFG: Kaplan & Bresnan 1982) and Head-Driven Phrase Structure Grammar (HPSG: Pollard & Sag 1994).

smaller units are taken to be WORDS, and the relevant task is to generate all and only the grammatical sentences of a natural language from the set of words, which is otherwise known as the vocabulary or LEXICON.² Of crucial interest to a theory of syntax, then, is what exactly constitutes a word, since these are the building blocks it must work with.

1.1 What is a word?

What might seem like an elementary question in linguistics, therefore, is in fact a very difficult one to answer. As is so often the case, intuitive and pre-theoretical notions do not map neatly to rigorously defined theoretical ones. Di Sciullo & Williams (1987) point out that there are (at least) three distinct ways of viewing the notion of ‘word’, which lead us to different conclusions. We consider these now in turn.

One notorious problem in deciding what counts as a word is that of variation in morphology, e.g. due to inflection. Are the verbs in (1a) and (1b) two different words, or just different versions of the ‘same’ word, in some sense?

- (1) a. Benjamin **cuts** the wire.
 b. Jadzia and Kira **cut** the wire.

They seem to express the same meaning, and they have the same part of speech; the only difference is in the grammatical number of the subject (which they agree with). The general response from lexicographers, in one form or another, is to claim that they are two WORD FORMS which correspond to a single LEXEME. But of course this isn’t so much an answer to the question as an admission of defeat: they seem at once to be both different and the same, and so we invent

²Even if we do not accept this ‘generative-enumerative’ view of syntax (cf. Pullum & Scholz 2001; Pullum 2013), the fact remains that we must be able to account for how speakers recognise, understand, produce, and have grammaticality judgements about sentences they have never heard before, and the only plausible explanation is that this is achieved by recognising commonalities across sentences in terms of smaller units, both lexical and syntactic. Thus, I hope the reader will forgive the oversimplification in the text for the sake of avoiding cumbersome rhetorical equivocation throughout.

two different levels so that at one they can be different and at the other they can be the same. But now we no longer have a coherent notion of 'word': when we speak of words, we need to know whether we mean word forms or lexemes. This is precisely the point that Di Sciullo & Williams (1987) want to make: when we encounter the word 'word' we do not know which of several different interpretations is meant. To see words as word forms is, in their view, to see a word as a MORPHOLOGICAL OBJECT, which is different from the role of words in relating form to meaning, to which we now turn.

It is generally assumed that whatever procedure allows for the generation of sentences from words runs alongside or feeds into the procedure for generating the meaning of the sentence from the meaning of its parts. For, just as there is no upper bound on the number of sentences in a language, there is similarly no upper bound on the number of meanings, so we can list meanings no more readily than we can sentences. It is proposed, therefore, that there is a similar generative apparatus at work in the semantic component, composing the meanings of words into the meanings of sentences, mediated via the syntax (since syntactic structure can also carry meaning—*terrorist shoots President* means something different from *President shoots terrorist*, even though they both contain the same words; we consider the relationship between syntax and semantics in more detail in Chapter 4.4). Another way of viewing words, therefore, is as the input to the semantic component—as pairings of form and meaning. Such pairings are what Di Sciullo & Williams (1987) call LISTEMES.

There is a conceptual motivation at work here also. For, as has been remarked since Saussure (1916) and earlier, the pairing of form and meaning at the word level is in large part arbitrary: there is no essential reason why the word *dog* is any better suited to denote the set of domestic canines than any other sequence of sounds, as the existence of distinct words in other languages (French *chien*, German *Hund*, Estonian *kass*, etc.) exemplifies. Given this, such connections between form and meaning will have to be stored—that is, listed—since there is no way to infer one from the other. And such pairings are amenable to listing,

since the vocabulary of a language is finite.^{3,4} In this view, the lexicon is the repository of arbitrary form-meaning pairs, i.e. of listemes.

The final conception of 'word' which Di Sciullo & Williams (1987) discuss is the one which we started with: the building blocks of syntax. Words are the objects on which syntactic operations work, or over which syntactic generalisations are described. This is what Di Sciullo & Williams (1987) call the **SYNTACTIC ATOM** view of words.

1.2 Mismatches between the different conceptions of 'word'

We now have three different conceptions of 'word', summarised below:

- (2) a. **SYNTACTIC ATOM**: the units which syntax operates on; the input to syntax.
- b. **MORPHOLOGICAL OBJECT**: an inflected word form; the output of morphology.
- c. **LISTEME**: a pairing of form and meaning.

In the analysis of any given sentence, these three conceptions often coincide, so that a single listeme corresponds to a single morphological object, which corresponds to a single syntactic atom. Take sentence (3), for example:

- (3) Jadzia and Kira cut the wire.

³This is why dictionaries are not (wholly) futile projects, and why, on the other hand, no one has ever tried to write a dictionary of all the sentences of a language. Firstly, as we noted before, the number of potential sentences is infinite, and so any attempt at listing them is doomed from the start. But also, such a project would involve a vast amount of redundancy. The difference between the meanings of *John saw the book* and *John saw the table* is intimately related to the difference between the meanings of *book* and *table*, and to treat the two sentences as two completely unrelated form-meaning pairs would be absurd.

⁴This is true at least of the fundamental building blocks, whatever we take them to be (morphemes, roots, morphosyntactic features, etc.). Where morphology is highly productive, it may not be true that the stock of *words* in a language is finite: for example, if derivational rules can be applied recursively, or in polysynthetic languages, with their long, sentence-like words. However, the set of building blocks from which such derived forms are constructed is itself ultimately finite.

Each of the orthographic words (i.e. those separated by spaces) has its own, listed meaning, which contributes to the meaning of the sentence in systematic ways. Each one also corresponds to the terminal node of a syntactic tree, and each has a particular morphological form. That is, each ‘word’ is at the same time a listeme, syntactic atom, and morphological object.

However, things are not always so orderly. Consider an expression like *cut the mustard*, meaning ‘to suffice, to be good enough’. Even if a hearer knows the usual meaning of *cut*, as well as that of *mustard* and *the*, and the usual means of composing the meaning of VPs in English, they will still not know the meaning of the whole expression. If the meaning cannot be computed, it must be listed; thus, *cut the mustard* must be a listeme.⁵ However, the expression clearly contains multiple morphological objects: the verb inflects separately from the rest of the expression, so that we say (4a) and not (4b):

- (4) a. That cuts the mustard.
 b. *That cut the mustards.

Quite how many syntactic atoms we have here is unclear as well: depending on our theory, we might want to store the whole VP as a unit, along with its meaning, or alternatively combine the words as usual but have a convention that maps the literal meaning to an idiomatic one—i.e. it depends whether we want listemes and syntactic atoms to coincide. In either case, it is clear that we can have mismatches between the different concepts of ‘word’.

1.3 Multiword expressions

It is these so-called MULTIWORD EXPRESSIONS (MWEs) which this thesis is about—those expressions which must be listed, and so seem like words, but which consist of multiple morphological objects and potentially multiple syntactic

⁵Of course, it is possible to reverse engineer meanings for the three words so that when combined they provide precisely the correct meaning for the idiomatic expression. Indeed, as we will see in Chapter 3, some theories do just this. Still, I feel that this is a very strange view of compositionality, since it seems crucial to me that words should bear the same meaning across different contexts, and this is not the case here.

atoms, and so seem like phrases. A similar definition is given by Baldwin & Kim (2010: 269), based on Sag et al. (2002):

- (5) Multiword expressions (MWEs) are lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity.

By the term ‘lexical item’, the authors refer to the fact that MWEs must be stored in the mental lexicon, i.e. they are listemes. By ‘multiple lexemes’ they refer to the fact that the subparts of a MWE are recognisable as morphological objects and potentially syntactic atoms in their own right. This much matches our definition, but the second part expands it to require that the expression also exhibit some kind of idiomaticity.

In this thesis, I will primarily be interested in the first three kinds of idiomaticity identified by Baldwin & Kim: ‘lexical’, where one or more parts of the MWE are words which do not occur outside of the context of the expression itself (e.g. *by dint of, take umbrage*); syntactic, where the syntax of the MWE is not what would be expected based on the component words (e.g. *by and large, kingdom come, every which way*); and semantic, where the meaning of the MWE is not compositionally derived from the meanings its parts have elsewhere (e.g. *all of the above*, plus others such as *spill the beans, kick the bucket, shoot the breeze*). These kinds of idiomaticity relate directly to the representation of MWEs in the grammar, and thus raise questions bearing on competence rather than performance (in the sense of e.g. Chomsky 1965). They correspond to what Sag et al. (2002) call LEXICALISED PHRASES, which they describe as “hav[ing] at least partially idiosyncratic syntax or semantics, or contain[ing] ‘words’ which do not occur in isolation”. It is the question of exactly in what sense these phrases are ‘lexicalised’ which this thesis addresses.

The other two kinds of idiomaticity are much more to do with performance, and are thus arguably outside the scope of a theory of grammar. Pragmatic idiomaticity refers to the fact that certain MWEs are associated with particular

situations, and lose their idiomatic interpretation outside of it. For instance, *good morning* can be used as a greeting, but when it appears in other contexts (e.g. *Jake had a good morning*), it loses this force.

Statistical idiomaticity refers to the fact that certain combinations or orderings of words are conventionally preferred, even though there may be nothing semantically idiosyncratic about them. This corresponds to Sag et al.'s (2002) other type of MWE, the INSTITUTIONALISED PHRASE, which they describe as being “syntactically and semantically compositional, but occur[ing] with markedly high frequency (in a given context)”. For instance, when we are laying the table, we put out the *salt and pepper*, and not the *pepper and salt*. But this expression is perfectly compositional; the ordering preference is purely a statistical one. Such idiomaticity also includes the strong collocational tendencies of certain words, e.g. we frequently speak of *committing a crime*, but much more rarely of *doing a crime*. These are certainly facts about language use, and thus about language in the broadest sense, but they are not narrowly grammatical questions, and thus I will not return to them here. For more on these two types of idiomaticity, see Sag et al. (2002). On pragmatic idiomaticity specifically, see Kastovsky (1982) and Jackendoff (1997); and on statistical idiomaticity, see Fernando & Flavell (1981), Bauer (1983), and Cruse (1986).

Even with these limitations, our definition still covers a large number of different kinds of expression, which probably do not constitute a natural kind. Some examples of the phenomena which have been discussed under the MWE heading include the following:

1. **Idioms**: usually sub-sentential, and often VPs, but also including a variety of other different syntactic categories—e.g. *kick the bucket*, *pull strings*, *not have a leg to stand on*, *hot potato*, *ace in the hole*, *out of the loop*.
2. **Fixed expressions**: sometimes included under the heading of idioms; things that do not have an obvious literal interpretation, such as *by and large*, *every which way*, etc.

3. **Prepositional verbs:** called INHERENTLY ADPOSITIONAL VERBS by Bhatia et al. (2018)—e.g. *rely on, consist of, deal with*.
4. **Particle or phrasal verbs:** called VERB-PARTICLE CONSTRUCTIONS by Bhatia et al. (2018)—e.g. *look up, take out, stand down*.
5. **Separable verbs:** e.g. German *anrufen*, Dutch *opbellen* ‘to call up’ (lit. ‘to up.call’) ~ *ich rufe dich an / ik bel je op* (‘I call you up’).
6. **Inherently reflexive verbs:** verbs which always include a reflexive particle even when it is impossible for them to be used non-reflexively—e.g. French *se suicider* ‘to commit suicide’ (lit. ‘to suicide oneself’; but one cannot ‘suicide’ anyone other than oneself, e.g. **je le suicide*, ‘I suicide him’).
7. **Light verb constructions:** semi-productive expressions which involve a semantically bleached verb like *take, have, make*—e.g. *take a bath, have a break, make a mistake*.
8. **Nominal compounds:** often unpredictable and non-systematic in their semantics, e.g. a *chocolate cake* is a cake made of chocolate, but a *birthday cake* is not a cake made of birthday (Jackendoff 2009).
9. **Complex proper nouns:** i.e. those consisting of more than one word—e.g. *New York, Jack the Ripper*.

For the sake of concreteness, the main focus of this work will be the first two kinds of expression, which are often grouped together under the heading of ‘idioms’. The reason for this is that these are in many respects the most complex of the categories listed above: they always exhibit semantic idiomaticity, and often include syntactic and lexical idiomaticity as well. In analysing such expressions, we will develop many of the tools needed to handle the other classes as well.

Such a move does mean taking semantic idiomaticity to be the MWE property *par excellence*, which might be thought of as questionable given the existence of

'collocations', i.e. MWEs which only exhibit statistical idiomaticity, and which therefore might be thought to have wholly predictable meanings. Even here, however, additional meanings may attach to the MWE which are absent from the same string viewed as a compositional expression. For example, *salt and pepper* can refer specifically to the tabletop condiments (and their containers) rather than merely to the two substances *qua* substances. It would be most uncooperative to respond to a request to put the salt and pepper on the table by pouring out a small pile of each. Regardless of what our conclusions are on this point, however, it is possible to use the same tools to describe both semantically idiomatic and compositional MWEs, as we will see in Chapter 6.2.1. Once again, by focussing on the more complex cases, we also develop the tools to handle the more straightforward ones.

1.4 Outline of the thesis

This thesis aims to make clear the analytical challenges posed by MWEs, and to provide my own analysis which meets as many of these as possible. The thesis can be divided in two along these lines: Chapters 1–3 outline the problem and assess how various theories fare when faced with them, while Chapters 4–6 focus on developing the analysis, motivating and presenting a framework I call Lexicalised LFG. Chapter 7 concludes. In this section, I give some details of what each chapter covers.

After this initial introductory chapter, we will look in Chapter 2 at the scope of the analytical problem, attempting to delimit the explananda. We will consider the distinction between MWEs and 'constructions' more generally, and conclude that MWEs are a specific subtype of construction, before going on to examine nine characteristics of MWEs which any theory needs to account for.

In Chapter 3, we will assess how well various proposals in the literature have achieved these aims. This chapter also outlines the space of possible theories, showing that there are only really three families of theories which are viable, and argues in favour of a construction-based account. Crucially, I argue against

what I perceive as the predominant approach in recent work, whereby idioms are made up of special versions of the words they (appear to) contain.

Chapter 4 describes the theoretical and formal tools which will be used in the analysis, introducing two syntactic theories, Lexical Functional Grammar (LFG) and Tree-Adjoining Grammar (TAG), and one theory of semantics and the syntax-semantics interface, Glue Semantics.

Chapter 5 shows how the descriptive advantages of TAG can be incorporated into LFG, and, in so doing, how we can lexicalise the grammar (a term to be defined). Chapter 6 then applies these tools to the analysis of MWEs, and shows how the combination of LFG and TAG affords us an elegant means of describing a variety of phenomena relating to both non-verbal and verbal MWEs.

Finally, Chapter 7 sums up, reviews the proposals made in the thesis, and suggests some interesting areas for future research.

2

Multiword expressions

Contents

2.1	MWEs or constructions?	12
2.1.1	Formal and substantive constructions	13
2.1.2	Lexical integrity	16
2.1.3	Summary	20
2.2	Properties of MWEs	21
2.2.1	Semantic idiomaticity	21
2.2.2	Syntactic idiomaticity	22
2.2.3	Lexical idiomaticity	23
2.2.4	The real words problem	24
2.2.5	Morphosyntactic flexibility	25
2.2.6	Syntactic flexibility	26
2.2.7	Decomposability	35
2.2.8	Psycholinguistic properties	41
2.2.9	Extended uses of idioms	43
2.3	Chapter summary	47

In this chapter, I go some way to defining what I mean by ‘multiword expression’ (MWE), starting by exploring how it fits into the larger category of ‘construction’. After this, I present a number of different properties which MWEs possess, and questions which they pose, in an effort to lay out the explananda for a theory of MWEs.

2.1 MWEs or constructions?

The idiosyncrasies of linguistic objects larger than the word have long been recognised, most particularly within the research tradition known broadly as CONSTRUCTION GRAMMAR (CxG: e.g. Lakoff 1984; Fillmore 1988; Fillmore et al. 1988; Goldberg 1995, 2006; Kay & Fillmore 1999; Jackendoff 2007; Boas & Sag 2012 among many others). Work in this area has tended to focus on detailed analysis of complex constructions which are often relegated to the periphery of other theoretical schools of thought.¹ Work has been done, for instance, on the *let alone* construction (e.g. *He doesn't get up for LUNCH, let alone BREAKFAST*—Fillmore et al. 1988), the *What's X doing Y?* construction (e.g. *What is this fly doing in my soup?*—Kay & Fillmore 1999), the *the X-er the Y-er* construction (e.g. *The more you tighten your grip, the more star systems will slip through your fingers*—Fillmore 1987), and many others. What unites these studies is that they understand the notion of 'construction' to constitute a meaningful unit of analysis above and beyond the word, and to be something which must be acquired and stored by the language learner: there is more to the grammar than words and rules for combining them; there are also larger objects which can be combined too.

But constructions do not only consist of the peripheral subregularities in language. In a thoroughgoing Construction Grammar, *everything* is a construction: from the more 'exceptional' cases mentioned above, to configurations which any formal theory of grammar would treat as core, such as the resultative construction (e.g. *She painted the barn red*—Goldberg & Jackendoff 2004), or even the 'transitive construction' (e.g. any kind of expression which includes a verb and its direct object—Fillmore & Kay 1995; Goldberg 2006), to individual words, all the way down to sub-lexical units such as morphemes. All of this is included in a CONSTRUCTICON, a broader version of what other theories call the lexicon. In this kind of theory, morphology and syntax are just one and the same kind of combinatory system; it is all about constructions. Goldberg (1995: 7) makes

¹For an excellent summary of the history of Construction Grammar and of its relation to other families of linguistic theorising, see Sag et al. (2012: 1–4).

this claim explicitly, when she says that “[i]n Construction Grammar, no strict division is assumed between the lexicon and syntax”. We will have more to say about this in Section 2.1.2, below.

Naturally, multiword expressions (MWEs) are also constructions in this broad sense. However, I want to draw a distinction between MWEs and constructions proper, rejecting the more totalising view of ‘construction’, whereby everything from morpheme to MWE falls under the same class. I will motivate this in two parts: firstly, the distinction between MWEs and other kinds of constructions is already present in early CxG work, in the contrast between formal and substantive idioms; secondly, the lexicalist hypothesis leads us to reject the radical CxG view.

2.1.1 Formal and substantive constructions

Fillmore et al. (1988: 504–506) discuss a variety of dimensions along which idiomatic expressions may vary. One of these is the SUBSTANTIVE vs. FORMAL contrast. Substantive idioms are also known as LEXICALLY FILLED idioms; these are the expressions whose lexical material is largely fixed. MWEs fall into this category. For the most part, regardless of whether their syntax is regular or idiosyncratic, their lexical form cannot be changed (except for inflection). For instance, *kick the bucket* is just a regular transitive verb phrase; it is its lexical content which makes it idiomatic, not anything about its structure—if we replace any of the three words by another, we simply have a new verb phrase; we don’t retain the same idiomatic meaning of ‘die’. And MWEs which have an idiosyncratic syntactic structure (on which see Section 2.2.2 below), such as *by and large* or *trip the light fantastic*, ‘dance’, still require the presence of specific lexical items to retain their idiomatic meaning—if we replace *by* with a different preposition, we do not retain the idiomatic reading; instead we simply have an ungrammatical string, e.g. **in/off/with and large*.

In contrast to the substantive idioms, we have the formal, or LEXICALLY OPEN idioms. They are the syntactic frames into which we can freely insert

any number of other expressions—both lexical items and larger phrases. Here the meaning is borne by the configuration itself, rather than by the words it contains, and this meaning is both independent of lexical selection, and interacts with it in predictable ways. For instance, in the formal idiom *the X-er the Y-er*, the value of X and Y can vary widely, while the construction still retains the same semantics (roughly involving a simultaneous conditional and comparative meaning; e.g. (1a) means ‘if you are DEGREE_x kind to him, he will impose on you to DEGREE_y’, where DEGREE_y increases whenever DEGREE_x does):

- (1) a. The kinder you are to him, the more he imposes on you.
 b. The more time I spend on this problem, the less I understand it.
 c. The more outrageous a politician’s promises are, the bigger his vote count is. (McCawley 1988: 176, 178)
- (2) a. The more carefully you do your work, the easier it will get.
 b. The bigger they come, the harder they fall. (Fillmore et al. 1988: 506)

Of course, these kinds of idiomatic expressions are not entirely devoid of substantive constraints: the two instances of *the* are obligatorily part of the construction, and it loses its idiomatic sense without them. Similarly, although *let alone* constructions involve complex syntactic and semantic structures which do not seem to be part of the regular grammar (Fillmore et al. 1988), they must, as the name suggests, contain the words *let alone*. So too with *What’s X doing Y?*, where the variables are very free, but the frame must contain the word *what*, an appropriate form of the verb *be*, and the word *doing* (Kay & Fillmore 1999). Truly fully formal idioms would be things like the resultative construction, where the frame is entirely schematic. For example, Goldberg & Jackendoff (2004) treat the English resultative as a family of constructions including the ‘causative property resultative’, exemplified in (3):

- (3) They wiped the table clean.

They give the properties of this construction as in (4) (Goldberg & Jackendoff 2004: 538):

- (4) **Syntax:** NP₁ V NP₂ AP₃
Semantics: X₁ CAUSE [Y₂ BECOME Z₃] (by V-ing)

This is totally general: no special words have to be present, merely this particular syntactic frame. As expected, therefore, many diverse examples are licensed:

- (5) a. Kim painted the house white.
 b. Robert scratched his face raw.
 c. Alex kissed it better.

But the label ‘formal idiom’ is perhaps now a little misleading: there is nothing idiomatic about this construction other than in the broadest sense of the word (that is, ‘natural’ or ‘native speaker-like’). Perhaps we might more generally speak of substantive and formal *constructions* (where constructions are at least word-sized; see the following section). The resultative construction, the transitive construction, or the *What’s X doing Y?* construction are all formal constructions, although only the latter is a formal idiom, but MWEs are by their very nature substantive. They require the presence of specific words, and it is by virtue of the presence of these words, and not (solely) by any other syntactic peculiarities, that they acquire their idiomatic meaning. Also substantive would be individual lexical entries, since these, too, get their meaning by virtue of having a particular content, rather than merely by being of a particular syntactic form.

For ease, then, let us reserve the word ‘construction’ for the formal construction, and instead of talking about substantive constructions, instead talk about MWEs and words. Given this typology, we can say that this thesis is about MWEs (substantive idioms), and how they differ from the other kind of substantive construction, the word.

2.1.2 Lexical integrity

So-called LEXICALIST grammatical theories, including Lexical Functional Grammar (LFG), the framework adopted here, adhere to some version of the LEXICALIST HYPOTHESIS (LH), *viz.* that there is a formal separation between the operations of the lexicon and the operations of the grammar, i.e. between morphology and syntax (first discussed by Chomsky 1970). The strong version of this hypothesis claims that all sub-parts of words are invisible to syntactic operations:

- (6) No constituent structure [i.e. phrase-structure] rule may order any element into or out of lexical categories such as N, A, V. That is, constituent structure rules are blind to the internal structure of lexical categories.

(Simpson 1983: 74)

This LEXICAL INTEGRITY PRINCIPLE implies a one-way transmission of information in the grammar: the syntax operates on the output of morphology (morphology feeds syntax), but has no access to its inner workings (syntax does not feed morphology, nor does it operate on sub-lexical parts). Bresnan & Mchombo (1995) discuss a number of tests which they claim vindicate lexical integrity, including for example the impossibility of gapping when applied to sub-lexical elements:

- (7) a. John liked the play and Mary, the movie.

- b. *John liked the play, and Mary *dis-* it.

(Simpson 1991: 51)

If this view of the architecture of the grammar is correct, then we have reason to reject the radical Construction Grammar position that everything from phrase to morpheme is the same kind of formal object, namely a construction, since clearly syntax and morphology operate on different kinds of entities.

However, the lexicalist hypothesis has recently been criticised as “both wrong and superfluous” by Bruening (2018). Bruening claims for instance that syntax can feed word-formation, and so the pipeline between morphology and syntax is not unidirectional, as the LH would have it. He argues this on the basis of examples such as (8):

- (8) a. What was your ‘**I don’t get paid enough for this shit**’ moment?
- b. Growing kids? The Yellow Pages is your **oh-boy-they-need-more-shoes-and-clothes-and-we-should-start-braces-for-their-teeth-now** directory.
- c. His general **ok-with-less-than-we-should-aim-for-ness** makes him an undesirable candidate.
- d. You can’t use the **termites-ate-the-walls** excuse because the home inspection didn’t find any! (Bruening 2018: 4, 4, 6, 7)

In these sentences, a whole phrase (even a whole sentence) is used as a prenominal modifier, in what looks like an adjectival position—that is, as a word. Spencer (1988, 1991: 414–417) argues that a phrase must have been lexicalised, i.e. conventionalised, in order to be used in this way, which would perhaps account for (8a). However, it is clear that this process is productive, as illustrated by (8b), and so not restricted to set expressions. What is more, it can provide input to further morphological processes, such as *-ness* suffixation, as in (8c). Bresnan & Mchombo (1995: 194) and Wiese (1996) suggest that such expressions are essentially quotative; this allows the modifying phrase to be ‘encapsulated’ and treated as a word. However, (8d) shows that the content of the ‘quoted’ phrase must be accessible to the syntax, since *termites* serves as the antecedent of an NP ellipsis (*... the home inspection didn’t find any (termites)*).² Thus, it seems like syntax must be able to feed morphology, *contra* the LH.

What is more, Bruening (2018: 8) points out that even if the quotation theory were right (that is, if it turns out that anything can be turned into a word by being quoted), then it is “not clear that the lexicalist hypothesis actually rules anything out”: the story the lexicalist theoretician is forced to tell is that syntax can’t feed morphology ... except when it can. Not a very useful constraint.

²If ellipsis is treated wholly semantically, as argued by Dalrymple et al. (1991), for example, then this objection carries less weight. Perhaps the contents of such quotative modifiers is invisible to the syntax but not to the semantics (although the determination of ‘parallelism’ necessary for the interpretation of ellipsis might still involve some syntactic machinery—Dalrymple et al. 1991: 9f.).

Based on this and a number of other arguments, Bruening argues that the lexicalist hypothesis does not really explain anything, and that, therefore, we should prefer a grammar with a single (syntactic) computational component which builds both words and phrases. That is, by Occam's Razor, having both a morphological and syntactic component is dispreferred over having just a syntactic component, all other things being equal—and Bruening believes he has shown the LH to be otiose, and thus all other things to indeed be equal. The radical Construction Grammarians reach the same conclusion, albeit for rather different reasons.

The ultimate force of Bruening's objections remains to be seen.³ Several of his arguments seem at best to merely muddy the waters—he does not offer many solutions to the problems he raises, and simply offering more challenging data is not a particularly compelling argument if the alternative he is selling does not have a ready-made solution either. For instance, although data like (8d) show that sometimes the content of phrasal compounds must be accessible to the grammar, examples like (9), discussed by Bresnan & Mchombo (1995: 194), show that sometimes it must not:

(9) His **I-told-you-so** attitude really grates.

In (9), the pronoun *I* explicitly does not refer to the speaker, as it would if it were interpreted normally by the grammar, but rather to the person who is described as having the attitude, exactly as it would in a quotation. At most, then, we might say that things are more complicated than either kind of theory can account for 'off the shelf'. But until further work is done to clarify matters, it seems premature to adjudicate one way or the other.

One important objection to Bruening's conclusion, and therefore to the radical CxG view as well, is that it fails to address the formal differences between syntax and morphology. As Asudeh et al. (2013: 4–5) point out, natural language morphology is computationally less powerful than natural language syntax: the

³See Müller (2018) for an early reply.

former can be described by a regular grammar, whereas the latter requires at least context-free power to describe.⁴ If words and phrases really are built by the same processes, this contrast is unexplained. For instance, why do normal words (i.e. not the phrasal compounds just discussed) not take advantage of the greater power available to them, and make use of nested and/or unbounded dependencies? Furthermore, morphological processes are not merely a subset of syntactic processes; there are also processes which are uniquely morphological. Why, if the two systems are identical, does syntax not make use of processes like root-and-pattern, for instance? These clear computational distinctions between morphology and syntax seem to plead for a formal separation of the two, *contra* the claims of Bruening (2018) and radical CxG, and along the lines suggested by the LH.

Since I do not believe the LH has yet been disproven, it remains a sensible guiding principle of grammatical theorising. We thus cannot countenance the radical CxG view which conflates morphological and syntactic processes, labelling both morphological and syntactic objects as ‘constructions’. In the rest of this thesis, I therefore assume that the processes which assemble words are distinct from those which assemble phrases; in other words, that morphology is separate from syntax. My attention will be mostly on syntax, and the theory presented in Chapter 5 is a theory of syntax, not of morphology. I will assume

⁴This claim requires some elaboration. Traditionally, two types of morphological processes have proved recalcitrant to a regular analysis: Semitic ‘root-and-pattern’ morphology, and reduplication. The former has been shown to be amenable to a finite-state analysis, albeit a comparatively complex one (see e.g. Beesley 1998). Reduplication is straightforward so long as it is *bounded*: if the morphology does not permit the copying of arbitrarily long strings of morphemes then it is simple (if cumbersome) to write a regular grammar that describes it. However, there are instances of unbounded copying, such as in Bambara (Culy 1985), which do fall outside the scope of regular grammars (although see Walther 2000 for a potential regular solution).

In contrast to morphology, the syntax of natural languages has been shown to be at least context free (Chomsky 1956, 1957; see also Partee et al. 1990: 480–482 for a different proof and an overview of the arguments). In fact, a point we return to in Chapter 4.2.1, some natural languages go beyond context freeness and require a degree of context sensitivity to describe their syntax (Shieber 1985).

Thus, we come to the claim in the text: natural language morphology is (basically) regular (although some processes in some languages may require more complexity to describe), whereas natural language syntax is (basically) context free (with some mildly context-sensitive extensions needed).

that there is a declarative, realisational theory of morphology operating in the background, along the lines proposed by Stump (2001), and that it interfaces with the syntax in much the manner described by Dalrymple (2015a).

However, above the level of the word things are different. There is nothing in the LH which prohibits the existence in the syntax of elementary objects which are larger than a single word. And the ample evidence from the many formal constructions discussed in the CxG literature and from the MWEs which are the focus of this thesis suggests that such objects must be taken seriously. These kinds of objects are more in line with the traditional notion of ‘construction’ which the early CxG research sought to rehabilitate. Thus, although I do not see the theory developed in this thesis as belonging to the framework of Construction Grammar, it is very clearly in the same spirit.

2.1.3 Summary

This thesis is about MWEs, which I have suggested we might think of in CxG terms as substantive idioms. This is in contrast both to single words (the other kind of substantive construction) and to formal constructions, the latter of which are not lexically anchored in the same way as substantive constructions are. The two kinds of construction are distinct and will be treated differently in the theory developed in this thesis. I will argue that, as substantive constructions, both MWEs and normal words have the same general formal properties. This is in keeping with Fillmore et al.’s (1988: 534) hope that “the structure-building principles of the so-called core and the machinery for building the phraseological units [...] may be of a uniform type, the former being a degenerate instance of the latter”. The theory advanced in Chapters 5 and 6 treats MWEs and words as the same kinds of formal objects, with words merely being simpler or more restricted, if not exactly “degenerate”, when compared with MWEs. Formal constructions, on the other hand, are not first-class objects in the lexicon (I do not assume a ‘constructicon’), but rather must be described by some other system. This is in keeping with the Lexicalist Hypothesis, whereby rules of

syntax are formally distinct from rules which operate within the lexicon—the latter providing input to the former.

Armed with a clearer idea of what this thesis is and is not about, let us now consider MWEs themselves in more detail. In the next section, we address the various properties of MWEs which any theory must contend with.

2.2 Properties of MWEs

There are a number of different properties which MWEs exhibit, axes along which they vary, and challenging questions which they pose, all of which ought to be accounted for by a complete theory of MWEs. No theory, including the one that I propose, succeeds in achieving all of this. Nonetheless, some fare better than others, as we will see in the next chapter, and I contend that the theory I advocate does best of all. Before we can compare approaches, however, we need to know on what grounds we are to compare them. In this section, we consider some of the major descriptive challenges that will recur throughout our discussion, and which our theory must deal with.

2.2.1 Semantic idiomaticity

Firstly, we must remember the three kinds of idiomaticity taken as definitional by Baldwin & Kim (2010): lexical, syntactic, and semantic.⁵ Foremost among these is semantic idiomaticity, where the meaning of the expression cannot be determined compositionally from the meanings that its parts have outside of the MWE. This is often taken as definitional for idioms, perhaps the most prominent class of MWE. For example, the meaning of *paint the town red*, ‘go out and party extravagantly’, cannot be arrived at from an understanding of the literal meanings of its constituent words, which would lead one to expect it to involve changing the colour of urban buildings. The idiomatic meaning cannot

⁵Recall that we are not considering the other two kinds of idiomaticity mentioned by Baldwin & Kim (2010), pragmatic and statistical, since these have to do with performance more than competence.

be computed, but must instead be stored. This is the motivation for thinking of these objects, larger than single words, as units, stored as a whole in the lexicon. Any analysis of MWEs must, before anything else, be able to account for how MWEs have the idiosyncratic semantics they do.

2.2.2 Syntactic idiomaticity

Another important level of idiosyncrasy is syntactic idiomaticity. This is where the MWE exhibits an external or internal syntax which is at odds with the rest of the grammar; on this basis, Fillmore et al. (1988: 505) call such expressions “extragrammatical idioms”.

An external syntax mismatch means that the category of the whole MWE is not what we would expect given the apparent categories of the parts which make it up. For example, *kingdom come*, a noun+verb string, does not behave distributionally like a clause; it can, for instance, appear as the object of a preposition, like a noun phrase:⁶

- (10) a. Even cars, trucks and buildings located 300 meters away were all blown to kingdom come and burned to a crisp. (WWW)
- b. I am [...] the only one—the god of kingdom come. (song lyrics)
- c. Your friendship means the world and I wouldn’t swap that for all the awards in kingdom come. (iWeb)

An internal syntax mismatch means exhibiting internal structure which is otherwise not attested in the grammar. For example, the expression *by and large* appears to involve the coordination of a preposition and an adjective, a structure which is otherwise impossible. In the idiom *trip the light fantastic*, meaning ‘dance’, it is not clear what the category of *light fantastic* should be:

⁶Where possible, and where appropriate, I use attested examples to illustrate the point under discussion. Many of these come from online corpora, and still others from the World Wide Web (indicated with ‘WWW’). Corpora are identified by abbreviations, explanations of which can be found in Appendix A. The Appendix also contains details of all other attested examples, including full URLs for Web sources. Example sentences drawn from the linguistics literature are identified by citation, while the sources for example sentences drawn from novels, song lyrics, etc. are listed in the Appendix.

(11) We [_{VP} tripped [_{NP} the [?? light fantastic]]] all night long.

Etymologically, the expression comes from the Milton poem *L'Allegro*, which includes the following lines:⁷

(12) Com, and trip it as ye go
On the light fantastick toe

Originally, therefore, *light* and *fantastic* were both adjectives, modifying a noun *toe*. But the MWE has dropped the noun, something which is not generally permitted in English. Should we now analyse *fantastic* as a noun modified by *light*, or *light* as a noun post-modified by *fantastic*? Different native speakers seem to have come to different conclusions. Neither is particularly natural, however.

Syntactic idiomaticity is related to semantic idiomaticity in that an idiosyncratic syntax can preclude the possibility of a literal, i.e. non-idiomatic, interpretation of a string. The synchronic grammar of English does not give a literal interpretation to *by and large*, for example, because there is no successful syntactic parse available for it outside of the MWE interpretation—it must, therefore, also be semantically idiomatic.

Clearly, this kind of syntactic unpredictability must also be encoded somewhere, and is another challenge faced by a theory of MWEs.

2.2.3 Lexical idiomaticity

Certain MWEs include ‘words’ which do not, in the synchronic grammar, have a life outside of the expression. Examples include *by dint of*, *take umbrage*, and *run amok*. Although each of the boldfaced terms did once exist independently in the language, they can no longer be used outside of these contexts. This is what we call lexical idiomaticity (such words are also referred to as ‘cranberry words’ by some authors, e.g. Richter & Sailer 2003, by analogy with the ‘cranberry morphemes’ discussed since the early structuralists). The other words in the

⁷The full text is available at the Poetry Foundation website: <https://www.poetryfoundation.org/poems/44731/lallegro>.

expressions (*by, of, take, run*) correspond to independently existing English words, which is why we can still confidently identify these as instances of MWEs. The semi-words, however, must be accounted for by a theory of MWEs—without also licensing their appearance elsewhere in the language.

2.2.4 The real words problem

This brings up one feature of MWEs that might be thought to pass without comment: although they can contain words which do not exist outside of the expression, MWEs nonetheless overwhelmingly contain words that do. And even in those expressions which contain lexical idiosyncrasies, there are always other independently existent words also present.⁸ But there is nothing in principle which rules out MWEs made up wholly of lexical isolates. Of course, if we encounter an otherwise unrecognisable sequence of sounds, we would in general analyse it as a new *word*, rather than a *multiword* expression; but if the sequence included recognisable inflection and other morphology, and always appeared together, it seems quite clear that we would think of it as a multiword string.

⁸Foreign expressions like *ad hoc* or *bon appétit* provide something of an exception to this generalisation, since none of these words exist in the language independently. A number of things can be said about this. Firstly, these may well not be true MWEs, instead being single words in the English speaker's lexicon, their multiword status being preserved by orthography alone. But even if they are, it usually turns out that at least one part of such expressions has been borrowed in other MWEs as well (*ad nauseam* and *bon mot*, for example), and so these words do have some degree of independent existence in the language. Beyond this, however, we must acknowledge that phrasal borrowings are a complex matter, with many attendant unanswered questions. For instance, there is always the complication of how much of the source language's grammar is known to the speaker; these expressions presumably started life as code-switches by speakers who knew some of the source language, and found that it had a means of expressing something that English lacked—and for contemporary speakers who have some knowledge of French or Latin grammar, say, there can still be an awareness of this foreignness. In this case, the expressions are not truly nativised, but retain some of their cross-lingual properties. Consider this playful example uttered by a Romance linguist of my acquaintance: "the expression '*ceteris paribus*' was bandied around a lot; and of course, in this instance, *ceteri* were indeed *pares*". The second clause here treats the MWE *ceteris paribus* as if it were a genuine instance of code-switching into Latin. This is very similar to the extended uses of idioms discussed below in Section 2.2.9—there, the original metaphor underlying the idiom has been 'reactivated'; here, it is the original code-switch which has been reactivated in some sense. This is clearly a complex area, where further study would be most fruitful, but I think it is clear that the generalisation in the text still holds for English-origin MWEs, where any cranberry words are the result of historical processes having caused them to fall out of general use, but having preserved them in the MWE.

Consider the potential MWE *flarg bliff*. On its own, we might take it to be the single word *flargbliff*. But if we observe inflection on the first element, and we never observe either word apart from the other, this appearance evaporates, and it becomes clear we are dealing with a MWE:

- (13) a. Edward flargs bliff with some aplomb.
 b. You shouldn't flarg bliff in front of the boss.
 c. The bliff she flarged was most impressive.

The fact, then, that there are apparently no MWEs that do not include at least one independently existent word is striking, and something we would like our theory to account for. In other words, it would be desirable if it were not left as a coincidence that MWEs are at least partly made up of pre-existing words from the language.

2.2.5 Morphosyntactic flexibility

Part of the interest of MWEs is that they have internal structure in a way that simplex words do not. One example of this is what I appealed to in the previous section, namely that the individual words in a MWE can independently be subject to morphological variation—most importantly, inflection:

- (14) a. Kim kicks/They kick the bucket. [Number]
 b. They kick/kicked/will kick the bucket. [Tense]

MWEs vary to the extent that they allow such flexibility: the adjective in *in short* does not inflect, for instance (**in shorter*), nor do expressions like *at last* permit internal modification (**at very last*—cf. *this is the very last one*).⁹

⁹The expression *at long last* would seem to be a counterexample to this claim, but it is perhaps better seen as a second listed expression, since no other adjective appears to be possible in the place of *long*:

(i) *at short/extended/... last

My thanks to Tom Wasow for reminding me of this putative counterexample.

One interesting interaction of the real words problem with morphosyntactic flexibility is that when parts of a MWE have an irregular paradigm in their literal usage, they will generally retain this irregular paradigm within the MWE as well:

- (15) a. Morn comes/came to the bar every day. [Literal]
 b. Morn comes/came a cropper. [Idiomatic]
- (16) a. Miles loses/lost every game. [Literal]
 b. Miles loses/lost his cool. [Idiomatic]

In each of the (a) examples, we have a literal use of the verb, where it exhibits an irregular past tense form. In the (b) sentences, the verb is used as part of a MWE, and retains its irregular form in the past tense.

This is not, however, as revealing as it may at first seem. Morphologically complex words also sometimes borrow part of their paradigm: for instance, *understand* inflects exactly as if it were made up of *under* + *stand*, using the irregular past tense form of *stand*.¹⁰

- (17) a. Julian stands/stood on the bridge.
 b. Julian understands/understood the implications.

Given that single words can have this property, it is perhaps hardly surprising that MWEs can access the morphological paradigms of their sub-parts as well. Nonetheless, MWEs remain problematic in that they allow *internal* inflection, which is not standard in English, even in morphologically complex words. That is, although the retention of forms may be a phenomenon shared with normal words, the fact that parts of the MWE which are not at the end of the expression can inflect remains a challenge which requires its own solution.

2.2.6 Syntactic flexibility

As well as allowing their parts to vary morphologically, some MWEs also allow them to be rearranged and/or separated by various syntactic processes. Firstly,

¹⁰I thank Andy Spencer for bringing the relevance of this fact to my attention.

certain modifiers may intervene between the parts of a MWE (we will say more about this in Section 2.2.7):

- (18) If your alternator has decided to **kick the proverbial bucket**, here's how to replace it yourself and save some money in the process. (iWeb)
- (18) To specifically design sermons to **pull emotional strings** so that "decisions" are made is ungodly, unbiblical, and unChristlike. (iWeb)

Voice alternations offer another example of such variability—some MWEs can passivise and retain their idiomatic meaning, for instance:

- (19) a. I pulled strings for you, my dear.
 b. Strings were pulled for you, my dear. Did you really think the Philharmonic would take on a beginner like you? (WWW)

Others can be questioned or separated by relativisation, even when their canonical form is a simple declarative verb phrase:

- (20) What strings did he pull to make me blind to the artificiality of his whole endeavour? (novel)
- (21) I will never forget all the strings you pulled for me. (WWW)

What is perhaps most striking about relativisation and question formation is that they are potentially unbounded in scope, which means that the parts of the MWE can end up arbitrarily far apart:

- (22) What strings did Jarvis say that Tony claimed that . . . he pulled?
- (23) I will never forget all the strings Benjamin told me that Kira said that . . . you pulled for me.

Although we have said that MWEs are units in some sense, such syntactic flexibility makes it clear that this unity needs to be explained carefully, since their subparts can be split up to a potentially extreme degree—this is not unity in the same way as words are units, for instance.

Just like morphosyntactic flexibility, syntactic flexibility varies from MWE to MWE. Sag et al. (2002) suggest there are three degrees of flexibility—they divide up their class of lexicalised phrases into three categories: fixed expressions, semi-fixed expressions, and syntactically flexible expressions. This is too limited, as we will see, but it captures some of the broad divisions between types.

Fixed expressions, as the name suggests, do not undergo any variation, morphological or syntactic. This category includes such expressions as *ad hoc*, *kingdom come*, and *in short*, as previously discussed.

Semi-fixed expressions allow for inflection, but cannot undergo specifically syntactic processes like passivisation or relativisation:

- (24) a. I shoot/He shoots the breeze (with Kim). [Person]
 b. We shoot/shot the breeze. [Tense]
 c. Part/Parts of speech [Number]
- (25) a. #The breeze was shot.
 b. #The speech of which it was part ...¹¹

This category is claimed to include non-decomposable idioms, such as *shoot the breeze* and *kick the bucket* (see Section 2.2.7 below for more on decomposability), and nominal MWEs, such as *part of speech* and *vicious circle*.

Syntactically flexible expressions, on the other hand, are supposed to permit the full gamut of morphosyntactic variability. This category includes decomposable idioms such as *pull strings*, particle verbs such as *look up*, and light verb constructions such as *take a bath*:

- (26) a. I pull/She pulls strings for him. [Person]
 b. I look/looked up the answer. [Tense]
 c. He takes/They take a bath every evening. [Number]

¹¹Here and throughout, asterisks are used to mark ungrammaticality (the grammar should not produce this string), and the hash is used to indicate semantic deviancy (the grammar should not produce this string with this meaning)—including that the intended idiomatic meaning is unavailable. For instance, *the speech of which it was part* is a valid string of English, with an interpretable meaning, but it no longer carries the idiomatic meaning relating specifically to grammatical categories.

- (27) a. What strings did you pull for him?
 b. I looked the answer up.
 c. The bath he took after dinner relaxed him.

Of course, a three-way classification is overly simplifying. Fraser (1970: 36–41), for example, suggests a six-way classification of idiomatic expressions, where each level corresponds to a certain family of syntactic constructions (conceived of as transformations by Fraser) which are permitted. Belonging to a higher level implies that you participate in all of the constructions associated with the lower levels, so that the ranking forms a hierarchical scale. Fraser (1970: 40–41) gives examples of idiomatic expressions (loosely, MWEs) belonging to each of the six levels, of which a selection are reproduced in Table 2.1 (p. 30). L5 idioms are the most flexible, although Fraser (1970: 33) claims that no idioms permit an idiomatic NP to be clefted, and so these expressions are still not unrestricted in what constructions they can enter into.¹² L4 idioms, on the other hand, do not permit what Fraser (1970: 38) calls ‘reconstitution’ transformations, the only one of which he discusses being the “action nominalization transformation”, e.g. *Benjamin laid down the law to Worf* → *Benjamin’s laying down of the law to Worf*. So, for instance, (28a), with an L5 idiom, is OK, but (28b), with an L4 one, is not acceptable with the idiomatic interpretation.

- (28) a. Jadzia’s spilling of the beans embarrassed Julian.
 b. #Miles’s breaking of the ice enabled us to get chatting.

L4 idioms, but not L3 idioms, allow for ‘extraction’ transformations, which for Fraser include passivisation, particle movement, and others (Fraser 1970: 38). Thus, (29a), L4, is OK, but (29b), L3, is not:¹³

- (29) a. The news was broken to him gently.
 b. #The house was brought down by his dramatic performance.

¹²Regular, fully productive words of the language, such as a verb like *love*, belong to a higher, “unrestricted” level, L6 (Fraser 1970: 36).

¹³These are Fraser’s judgements. The reader may, like me, disagree with him on a number of points of data; however, this in itself is not really important, as will be discussed shortly.

L5

keep one's word, kill the goose that lays the golden egg, lay down the law, let the cat out of the bag, make the best of a bad deal, make the punishment fit the crime, pop the question, pull some strings, read the riot act to, spill the beans, take liberties with

L4

add up to, ask for, auction off, bear down on, belong to, break the ice, break the news to, hit the nail on the head, lose sight of, make use of, pay attention to, poke fun at, qualify for, rely on

L3

bring down the house, give up someone for dead, keep up one's guard, put down one's foot, put down something to, put on some weight, teach new tricks to an old dog, the cat has someone's tongue

L2

bear witness to, do a good turn to, drop a line to, give chase to, give the benefit of the doubt to, give what for to, depend on, fish for, harp on

L1

kick the bucket, aspire to, insist on, stand for, encroach on, burn the candle at both ends, bank on, catch fire, clamor for, dance up a storm, give birth to, knock off work, pull up stakes, put pen to paper, stir up trouble, turn over a new leaf

L0

bite off one's tongue, blow one's cool, amount to, bear on, rail at, beware of, build castles in the air, face the music, get up one's energy, let off some steam, pluck up courage, stew in one's own juice, take up heart, turn a deaf ear to

Table 2.1: Fraser's (1970: 40–41) levels of analysis for idioms

This process continues down the scale; I will not discuss all of the levels here: see Fraser (1970: 36–41) for details. The important point is that we do not jump straight from MWEs which inflect but are not syntactically flexible to ones which are fully flexible—there is a much richer gradation.

Beyond this descriptive point, what makes Fraser’s proposal particularly interesting is that he claims the hierarchy is implicational. That is, each entry from a higher level inherits all of the transformations from lower levels. So, because *spill the beans* is L5, not only can it undergo the action nominalisation transformation, but it can also undergo passivisation, which it inherits from L4:

(30) Looks like the beans have been spilled on our dates for 2019! (WWW)

This claim makes Fraser’s typology scientifically much more interesting, since it now makes predictions. If an idiom is observed to undergo a transformation from L4, say, then we expect that it will also undergo the transformations from L1–L3 (L0 being the level at which *no* syntactic operations can apply). Of course, as Fraser (1970: 41) himself notes, we can disagree about which idiom occupies which level (I disagree with a number of the judgements he reports; (29b) seems acceptable to me, for instance); what is important for his proposal is that whatever level we assign an idiom to, the implicational scale holds. An idiom should inherit all of the transformations from the levels below it (wherever it may be placed), and should *not* take part in any of the transformations from higher levels. Predictably, of course, there are problems.

From Fraser’s description it seems that all idioms on a particular level should participate in all transformations associated with that level—but that doesn’t seem to be the case. For instance, Fraser’s (1970: 38) discussion of the ‘reconstitution’ action nominalisation transformation includes as definitional the fact that the verb phrase becomes nominal (with the subject of the verb becoming the specifier of the corresponding noun phrase). Presumably, therefore, relativisation also falls under the same level, L5, since this also involves nominalisation of the clause (in the sense that a sentence, e.g. *I had an idea* is transformed into a noun

phrase, e.g. *an idea I had*). But many of Fraser's L5 idioms, although they can definitely undergo action nominalisation, cannot appear as a relative clause:

- (31) a. Benjamin's reading of the riot act to Admiral Nechayev saved the day.
 b. ??The riot act that Benjamin read to Admiral Nechayev was a particularly long one.
- (32) a. Sisko's laying down of the law to Worf avoided further problems.
 b. ??The law that Sisko laid down to Worf was a particularly draconian one.

Even if we stick to transformations Fraser himself discusses, there are problems. L4, for instance, allows 'extraction' transformations, which for Fraser include both passivisation and preposing of prepositional phrases. But *hit the nail on the head* is acceptable in the passive, yet decidedly odd with the PP preposed:

- (33) a. I'd say the nail was hit on the head with this one! (WWW)
 b. #On the head you've hit the nail!

However, this kind of objection is not a knockdown argument. The fact that some idioms do not undergo all transformations on a given level does not mean that the existing hierarchy is untenable, merely that it needs to be made more detailed: that is, more levels need to be added. Perhaps there is an L5.5 which permits relativisation, while L5 itself does not, for instance.

A bigger problem would be gaps in the hierarchy: if an expression had properties of L5 but not of L2, say. This is apparently the case for some speakers—Horn (2003: 262), for instance, claims that the sentences in (34) can only have a literal interpretation:

- (34) a. The beans that Joe spilled caused us a lot of trouble.
 b. Which beans did Joe spill?

I presume, though, that he would find *Joe's spilling of the beans caused us a lot of trouble* perfectly acceptable with an idiomatic interpretation, which puts *spill the beans* in L5. Now, as discussed, relativisation is presumably also an L5 construction, so the unacceptability of the idiomatic reading for (34a) is just like in (31b) or (32b). However, I take *wh*-movement, as in (34b), to be an instance of 'extraction' in Fraser's terms (or possibly 'permutation'), which means that it is an L4 (respectively, L3) property. In this case, we have an example of an L5 idiom which does not possess an L4 (or L3) property, *contra* Fraser's hierarchy.

Personally, I find the sentences in (34) acceptable with the idiomatic interpretation, and it is not difficult to find attested examples either (see (35)–(36)), which means Fraser's hierarchy is preserved in this case.

- (35) a. The beans she spilled in that book included an account of how her alcoholic father beat her as a child. (WWW)
- b. I really want to see the beans he spills under oath. (WWW)
- c. Your friends will meet with you and laugh at you, wondering why you are so stupid to remain in that relationship in spite of all the beans that you 'spilled' regarding your spouse. (WWW)
- (36) a. What beans did she spill to the High Sparrow that he found interesting enough to let her leave? (WWW)
- b. Has anyone ever gotten someone to communicate with their horse? If so, what beans did your horse spill? (WWW)
- c. Of course, this only fueled fan theory that Marvel was muzzling the actor from spilling the beans. But which particular beans did he spill? (WWW)

If people with Horn's grammar exist, though,¹⁴ Fraser's hierarchy cannot be said to hold in general. While he may have detected a tendency, it is not necessarily

¹⁴And he is not alone in reporting these judgements: Schenk (1995: 262), for instance, marks (i) as unacceptable with the idiomatic reading:

(i) Which beans did Yvonne spill?

a strictly implicational hierarchy of the sort he imagined.

What, then, is the lesson to take away from this excursus? It seems to me that although there are undoubtedly tendencies in what syntactic operations cluster together—potentially relating to whether the operations in question are “meaningful” or “meaningless” (Schenk 1995), and how this interacts with the question of decomposability to be discussed below—MWEs are just too varied and too idiosyncratic (and thus disorderly) to be placed in any neat categories with respect to their syntactic flexibility. Note, for instance, that even where two different expressions both allow the same operations to apply, they may do so with varying degrees of ease. In my idiolect at least, *spill the beans* is much less happy as a relative clause than *pull strings*, for example:

- (37) a. The strings that Benjamin pulled got Ezri promoted.
 b. ?#The beans that Jadzia spilled caused a lot of drama.

And the contrast seems even stronger when it comes to topicalisation:

- (38) a. Some influential strings, Benjamin pulled.
 b. ?#Some explosive beans, Jadzia spilled.

Whether this is a question for the grammar proper to address, or whether it is a question of usage, for instance, I do not know. If anyone does want to advocate an implicational scale of the kind envisioned by Fraser, there certainly remain a number of unanswered questions. Before anything else, we need far richer empirical data: what kinds of grammar are attested and therefore need to be modelled? We would need to assess each possible permutation of grammaticality judgements: for example, are there speakers who accept *wh*-questions of *spill the beans* but not relative clauses? Or vice versa? And what exactly motivates the grouping of syntactic operations into levels? Are there any antecedent properties which should guide us, or do the categories emerge only from the observed behaviour of idioms? In the former case, the theory might count as explanatory; in the latter, it is purely descriptive. Must there be one hierarchy, or could

different speakers have different scales, Optimality Theory-style? (That is, is it a hard or a soft scale?) How do we handle the ‘extended’ uses to be discussed below, where many idioms can be reanalysed as productive metaphors, thus allowing for much more flexibility than is usually claimed?

Regardless of how such an endeavour might pan out, the general point remains: we do need to be able to model the variable syntactic flexibility of MWEs. Especially, we need to be able to represent the fact that MWEs which are ambiguous between a literal and an idiomatic parse do not exhibit the same degree of syntactic flexibility in both cases. That is, MWEs can possess their own restrictions on syntactic flexibility.

2.2.7 **Decomposability**

Many MWEs, and especially those that I am focussing on in this thesis, are non-compositional: their meaning is not the same as the literal interpretation of the string of words which makes them up, and sometimes there is no such literal interpretation. *Pulling strings*, for example, has nothing to do with literal pulling, or strings made of any kind of material, but rather with exploiting connections; *shooting the breeze* involves no firearms and no wind, but rather engaging in casual conversation; and there is no literal interpretation of *run amok*, or *kingdom come*, owing to lexical and syntactic idiomaticity, respectively.

At the same time, many (and perhaps most) idioms are motivated by some conceptual metaphor (Lakoff & Johnson 1980; Lakoff 1987), and so, while conventional, they are not entirely arbitrary. For example, according to the CONDUIT METAPHOR (Lakoff & Johnson 1980), THE MIND IS A CONTAINER and IDEAS ARE PHYSICAL ENTITIES, and communication involves moving ideas from one mind to another via language. In this context, *spill the beans* makes perfect sense as a metaphor for divulging information, since beans are individuated physical entities, just like ideas in this metaphor, and spilling them from a container is like releasing them accidentally from one’s mind (Lakoff 1987: 449–450). This might explain why we don’t say *slop the soup* instead, for instance,

since soup is not individuated in the same way beans are, but it doesn't explain why we don't use any other similar objects: why not *spill the peas*, for example? In other words, such underlying metaphors may constrain the form of idioms, but they do not determine them.

Importantly, though, where there is a one-to-one mapping between the participants in the literal and the metaphorical interpretations, the meaning of an idiom can be distributed across its parts, thus restoring a semblance of compositionality (Lakoff 1987: 451; Nunberg et al. 1994). For example, in *spill the beans* it is easy to assign an idiomatic meaning to each of the words, so that $[[\text{spill}]] \approx \text{divulge}$ and $[[\text{beans}]] \approx \text{secrets}$ (while *the* retains its literal meaning). Such idioms we call DECOMPOSABLE, since we can successfully decompose the meaning of the idiom and distribute its parts across the parts of the string.

NON-DECOMPOSABLE idioms, by contrast, do not permit such decomposition. The expression *kick the bucket*, for example, literally describes an event involving two participants, the kicker and the bucket, but its idiomatic interpretation describes an event with only one participant, the dier. This makes a mapping between domains impossible, and so there is no idiomatic interpretation we can give to the parts of the idiom individually: there is no idiomatic interpretation of *bucket*, nor of *kick*, but only of the whole VP.

Decomposability affects many other aspects of idiom behaviour, and Nunberg et al. (1994), for instance, have suggested that decomposability licenses greater syntactic flexibility. In the decomposable cases, the individual parts of the idiom have (figurative) referents, and it therefore makes sense to subject them to processes with semantic consequences which would not make sense if they were meaningless. For example, one can modify or quantify the subparts of an idiom:

- (39) a. We leave no **digital** stone unturned, we poke and prod every nook and cranny of the Interwebz. (WWW)
- b. Chad Michael Murray spilled the **baby** beans on Twitter. (WWW)
- c. Tom won't pull **family** strings to get himself out of debt. (WWW)

- (40) a. They pulled **some** last minute strings and managed to host us on the main stage before Goldie Looking Chain on the Saturday. (WWW)
- b. Emotiva has **lots of** bigger fish to fry at the current time which I fully understand. (WWW)
- c. Maybe by writing this book I'll offend a few people or touch **a few** nerves. (WWW)

It is important to notice that such modification really does target the subparts themselves: for instance, it is family connections which Tom won't exploit in (39c), and (40c) means 'I will upset a few people or annoy someone in a few ways', *not* 'I will cause the same irritation multiple times'—that is, the quantification has narrow scope, rather than wide scope over the whole proposition. This is in contrast to examples like (41), where the modification can be thought of as a domain delimiter, scoping over the whole sentence (examples taken from Findlay 2017a: 212):¹⁵

- (41) a. Musicians keep composing songs 'til they kick the **proverbial** bucket.
[= ... 'til, proverbially speaking, they kick the bucket.] (GloWbE)
- b. Britney Spears came apart at the **mental** seams. (WWW)
[= Mentally, Britney Spears came apart at the seams.]
- c. Let's say you want to return the **oral sex** favour he happily gives to you. (WWW)
[= In the domain of oral sex, you want to return the favour.]

¹⁵Quite how these apparently nominal modifiers come to have the right semantic type to modify a whole clause is far from obvious. Such examples are highly reminiscent of the 'occasional-type' adjectives first discussed by Bolinger (1967), exemplified in (i):

- (i) An occasional sailor comes into the bar.
[= Occasionally, a sailor comes into the bar.]

This might suggest one fruitful direction to pursue. For a contemporary overview of the issues relating to these kinds of adjectives, see Gehrke (To appear). My analysis of sentences like (41) in Chapter 6.2.2 explains how the adjectives obtain the correct scope compositionally, but does not address the question of how they come to exhibit such polymorphic typing.

Ernst (1981) describes this distinction as INTERNAL vs. EXTERNAL modification: the examples in (39)–(40) are of internal modification, since they modify parts inside the idiom; the examples in (41), on the other hand, contain external modification, since the modification scopes over the whole expression, and so is in some sense outside of it. Because their parts do not carry meaning on their own, non-decomposable idioms are compatible with external modification (as above), but not internal:

- (42) a. #They kicked the rusty/tragic bucket.
 b. #We shot all the breezes last night.

Similar arguments apply to other kinds of syntactic manipulations. Kuno & Takami (2004: 127) argue that the subject of a passive construction is associated with a particular information structural role, namely topic. If an idiom is passivised, then, whatever appears as the subject must be able to bear this role. And if the putative subject is in fact semantically empty, this will be problematic. Hence, non-decomposable idioms are unable to passivise, whereas decomposable ones can:

- (43) a. The beans haven't been spilled yet on who is to play SRK's leading lady in the film. (WWW)
 b. #Logs were sawn all night long.

Similarly, if the head of a relative clause does not have a meaning on its own, it cannot stand in the appropriate relation to whatever predicate it is an argument of:

- (44) a. Most are not aware of the sinister strings that were pulled to orchestrate this war. (WWW)
 b. #The fat that we chewed last night was very interesting/challenging/pleasant.

It makes sense to talk about not being aware of the connections in (44a), but (44b) is uninterpretable, because *fat* has no interpretation on its own.

However, decomposability cannot be the whole story when it comes to determining syntactic flexibility. As we have already seen, there is more nuance to the contrast than a binary ‘flexible’ vs. ‘non-flexible’, which is what the decomposable vs. non-decomposable divide would suggest. There are certainly decomposable idioms which are less flexible—think of the speakers, like Horn (2003), for whom *wh*-movement is unacceptable with idiomatic *spill the beans*, for example. What is more, there are also non-decomposable idioms which show flexibility. In German, for instance, Bargmann & Sailer (2018: 4) claim that “there seem to be no syntactic restrictions [...] that correlate with semantic non-decomposability”; this means that non-decomposable idioms like *den Löffel abgeben* ‘die’ (lit. ‘pass on the spoon’) can passivise, for example:

- (45) Hier wurde der Löffel ab-gegeben.
here was the spoon on-passed
 ‘Someone died here.’ (Bargmann & Sailer 2018: 4)

Bargmann & Sailer (2018: 20–21) also provide attested examples of English non-decomposable idioms occurring in the passive:

- (46) When you are dead, you don’t have to worry about death anymore. [...]
 The bucket will be kicked. (Bargmann & Sailer 2018: 5)

They claim that this is possible because the constraint on English passives is not quite as restrictive as it is sometimes presented as being. As Bargmann & Sailer (2018: 11–12) point out, it is not always necessary for English passive subjects to be semantically contentful, as the following sentences show (data from Kay et al. 2015):

- (47) a. **There** was believed to be another worker at the site besides the neighbors who witnessed the incident.

- b. It was rumored that Great Britain, in apparent violation of the terms of the Clayton-Bulwer treaty, had taken possession of certain islands in the Bay of Honduras.¹⁶

In Bargmann & Sailer's (2018) analysis, the meaning of an idiom is redundantly marked on all of its parts, and so the NP component of a non-decomposable V+NP idiom contributes the meaning of the whole expression (as does the verb). This means that it *can* be felicitous as the subject of a passive, provided the meaning of the idiom as a whole is sufficiently discourse-old, and can therefore serve as a topic. Examples like (48) illustrate this point most clearly:

- (48) There was really no need for the police to have a cow, but **a cow was had**, resulting in kettling, CS gas and 182 arrests.

(attested example from Bargmann & Sailer 2018: 21)

Here the idiom is obviously discourse-old, since it is used in its canonical form immediately beforehand.

In fact, even *kick the bucket*, the prototypical non-decomposable and inflexible idiom, has been shown to appear in a number of different non-canonical syntactic configurations with its idiomatic interpretation intact (attested examples from Bargmann 2017):

- (49) a. For those of you who don't know, my computer's motherboard finally kicked its last bucket.
 b. They say famous people die in threes, and I've believed them since that summer in 1997 when Nusrat Fateh Ali Khan, Mother Teresa and Lady Di all kicked their respective buckets in unison and the world ran out of flowers.
 c. It's clear that Stein's bucket is going to be kicked right from the start. Despite this, the death still carries emotional weight [...].

¹⁶This might be a less compelling example than Kay et al. (2015) intend, since *rumored* does not have a transitive active form (**We rumored it that Great Britain ...*). Thus, despite having the surface form of a passive, one might be inclined to argue that (47b) is not in fact a true passive.

However, several of these examples have a distinctively ‘playful’ character, and this may suggest that something else is going on here. I discuss what this might be in more detail in Section 2.2.9. For now, it suffices to observe that, although decomposability and flexibility are certainly connected, the correlation between them is nevertheless far from as clearcut as earlier researchers might have hoped.

2.2.8 Psycholinguistic properties

Idioms have received, and continue to receive, a relatively large amount of attention in the psycholinguistic literature. The findings of these studies have been far from conclusive, and over the years different relationships between the literal and figurative meanings of idioms have been proposed: that the literal meaning is processed first (Bobrow & Bell 1973), that the figurative meaning is processed first (Gibbs 1980), or that both are processed in parallel (Swinney & Cutler 1979), among others. However, despite this theoretical indecision, one fairly robust empirical finding has emerged: the *IDIOM ADVANTAGE*. In general, we recognise the figurative meaning of an idiom faster than its literal meaning (assuming it has one), and faster than we recognise the meaning of non-idiomatic expressions (Swinney & Cutler 1979; Estill & Kemper 1982; Cronk & Schweigert 1992; Carrol & Conklin 2017).¹⁷ We also *prefer* idiomatic readings to literal ones, and sometimes do not even compute a literal meaning for an expression at all (Gibbs 1986). That is, in general, idioms are quicker and easier to process than literal expressions of similar syntactic complexity. This has led some authors to suggest that idioms are lexicalised, i.e. stored as large units (Swinney & Cutler 1979; Cronk & Schweigert 1992; Cronk et al. 1993). In this thesis I will argue for just such a position, although I believe that it is also motivated on independent grounds, regardless of the psycholinguistic facts (see Chapter 3).

¹⁷Even this apparently robust result has its dissenters, however: Cacciari & Tabossi (1988), for instance, believe that the findings are due to poor experimental designs which do not adequately control for familiarity/frequency effects, and Kyriacou et al. (2018) report that the idiom advantage disappears when ‘contextual congruity’ and predictability (as measured by a cloze test) are controlled for.

Not all MWEs are alike, however: light verb constructions (LVCs), for instance, seem to be *harder* to process than literal expressions, in contrast to idioms (Wittenberg & Piñango 2011; Wittenberg et al. 2014). This is not unexpected, however, since LVCs are semi-productive. Languages like English have a very limited stock of light verbs, which can be used with a variety of complements with more or less the same bleached meaning:

- (50) a. have a bath/walk/break/smoke/conversation/...
 b. take a bath/walk/break/chance/breather/...
 c. make a(n) mistake/effort/speech/...

The role of the light verb is to combine with its complement to produce a new verbal meaning based on the complement's nominal meaning: to *take a bath* is essentially to *bathe*, for instance; to *have a conversation* is to *converse*; etc. But this then requires a degree of compositionality. It does not seem sensible to list every such expression separately, since the meaning is often totally predictable from the meaning of the noun,¹⁸ which makes these kinds of expressions different from idioms, where no such compositional analysis is possible. If LVCs involve complex type-shifting/meaning-altering operations on their nominal complements, it is hardly surprising that they may be more costly in terms of processing than straightforward transitive VPs.

Our theory of MWEs should be compatible with (and ideally explain) the idiom advantage. But we must bear in mind that the exact relationship between formal models and psycholinguistic processing is not straightforward or well understood, and ultimately depends on a number of assumptions made about the parsing procedure, so evidence from one may not be wholly probative for the other.

¹⁸Although something must be said about which nouns and which light verbs can co-occur: #*make a breather* does not have the idiomatic meaning of *take a breather*, for instance. My thanks to Mary Dalrymple for reminding me of this issue.

2.2.9 Extended uses of idioms

As we have seen, idioms are fast to recognise and easy to comprehend. They are used frequently in everyday speech, being very often related to “a recurrent situation of particular social interest”, and indicative of “relatively informal or colloquial registers” and of “popular speech and oral culture” (Nunberg et al. 1994: 493). That is, they are commonplace and unaffected.

However, consider the following, much more poetic, uses of idioms:

- (51) a. The strings we’ve been pulling to keep you out of prison are fraying badly.
- b. That horse you’re flogging isn’t quite dead yet, but it’s definitely not well.
- c. I know the bathwater was really dirty, but you still shouldn’t have thrown the baby out with it.

(Egan 2008: 391)

- (52) a. The cat is out of the bag, claws fully exposed. (iWeb)
- b. I had butterflies in my stomach and they had big wings on them. (WWW)
- c. This month, Meriden City Council’s PLA chickens came home to roost, and they laid a big stinky egg on Meriden taxpayers. (WWW)
- d. Sometimes [...] the person that you’d take a bullet for is behind the trigger. (song lyrics)

Understanding these examples cannot be achieved by ‘direct access’ of the idiom meaning alone; it also requires accessing the metaphorical underpinning of the idioms involved, and then extending it. For instance, a cat with its claws exposed is dangerous and ready to strike; if the cat stands metaphorically for a secret, then we understand that the secret itself, or its being revealed, is dangerous, and perhaps imminently so.

Compare (53a) with (53b):

- (53) a. If you let this cat out of the bag, a lot of people are going to get scratched. (Egan 2008)
- b. If you let this cat out of the bag, a lot of people are going to get hurt.

In (53b), we can interpret the idiom directly, in isolation as it were, obtaining the meaning ‘if you reveal this secret’ for the first part of the sentence, and then interpret the continuation based on this: if the secret gets out, a lot of people are going to get hurt (by the secret). (53a) is different, however. If we follow the same procedure, we are left with the meaning that a lot of people are going to get literally scratched as a result of the secret being revealed. While of course this is one possible reading, it is not the only one (and surely not the preferred one in most contexts). In order to obtain the correct one, which is essentially the same as for (53b), we have to interpret the apodosis with respect to the *literal* meaning of the protasis, even though the sentence as a whole ultimately receives an idiomatic interpretation.

The same pattern obtains with modification of the subparts of idioms. Consider the following examples:

- (54) a. She skated through college and pulled family strings to secure her short-lived, high-priced jobs in management consulting and media. (iWeb)
- b. In spite of its conservatism, many people were eager to jump on the horse-drawn Reagan bandwagon. (Ernst 1981: 52)
- c. With the recession, oil companies are having to tighten their Gucci belts. (Ernst 1981: 60)

In (54a), we can give a directly idiomatic interpretation of *strings* as ‘connections’, and then modify this with *family*: it is family connections which she exploited. In (54b), on the other hand, we cannot simply interpret *bandwagon* as meaning ‘(political) movement’, since it is not true that the movement is horse-drawn.

Rather, we are to infer from the fact that the metaphorical bandwagon is horse-drawn that the movement it stands for is old-fashioned. The situation is even clearer in (54c), where there isn't even any appropriate idiom-internal meaning to assign to *belts* and then modify with *Gucci*; rather, we must understand the modification as applying to the literal meaning of *belts*, and then make some further inference that this means that the oil companies are rich (because wearing a Gucci belt suggests wealth).

Example (54c) is actually a special version of what Ernst (1981) calls CONJUNCTION MODIFICATION. In these kinds of examples, two propositions are expressed, one idiomatic, and one literal:

- (55) a. (from an article on the making of the movie *Jaws*)
 Bruce, a shark, found it a part he could really sink his three rows of teeth into. (Ernst 1981: 52)
- b. (from a review of a production of *Twelfth Night*)
 Malvolio deserves *almost* everything he gets, but ... there is that little stab of shame we feel at the end for having had such fun pulling his cross-gartered leg for so long. (Ernst 1981: 52)
- c. In spite of the treatment the other refugees received from the rescue party in the desert, he bit his thirst-swollen tongue and kept to himself. (Ernst 1981: 59)
- d. The \$6,000,000 man came over and lent us a helping electronic hand. (Ernst 1981: 59)

For instance, (55a) expresses the two propositions 'Bruce found it a part he could really sink his teeth into (idiomatically speaking)' and 'Bruce has three rows of teeth (literally speaking)'. Similarly, (55d) says that the Six Million Dollar Man lent us a helping hand (idiomatically) and that he has an electronic hand (literally).¹⁹ Our 'Gucci belts' example, (54c), is like these cases except that the

¹⁹The *Six Million Dollar Man* was a US TV show about an astronaut who is given bionic implants after an accident, in order to "rebuild" him. This process costs \$6,000,000, hence the name. See https://en.wikipedia.org/wiki/The_Six_Million_Dollar_Man for more.

'literal' conjunct is then interpreted figuratively: that is, we understand that the oil companies literally have Gucci belts, but then reject this as a possible interpretation, and instead understand it metaphorically as meaning that they are rich. For more on the wide range of ways in which idiom modification can be interpreted, see Bargmann et al. (2018).

The playful language of these examples is very often dismissed from linguistic theorising as mere 'word play', being 'metalinguistic', i.e. beyond the scope of the grammar. There is no doubt that such examples have a different character from the canonical examples we have seen so far, both phenomenologically and experimentally—for instance, McGlone et al. (1994) and Glucksberg (2001) show that figurative modifications of idioms reverse the usual tendency for idiomatic expressions to be understood faster than literal ones—but this is not to say they are beyond the reach of linguistic theory.²⁰ They are also intimately linked to the ambiguous status of idiomatic expressions, as being both lexically encoded and at the same time vividly metaphorical.

In joint work, colleagues and I propose that there are in fact two modes of idiom interpretation, which we call 'core' and 'extended' (Findlay 2018; Findlay et al. 2018, 2019). The core uses involve treating idioms as lexically encoded, while the extended uses involve treating them as productive metaphors.²¹ The 'playful' uses discussed in this section are examples of the extended use of idioms. One consequence of 'reactivating' the metaphors underlying idioms in this way is that the idioms become, for instance, much more syntactically flexible, and permit freer substitution of synonyms, because their literal meaning increases in importance and their precise form becomes less significant. This is why apparent counterexamples to the non-passivisability of certain idioms, such as (56) and (57), repeated from (46) and (48), have a decidedly affected feel:

- (56) When you are dead, you don't have to worry about death anymore. [...] The bucket will be kicked.

²⁰ "Word play" should not be a sign on a linguistic trash can", as Bargmann et al. (2018) put it.

²¹ A similar distinction is drawn in *Deliberate Metaphor Theory* (Steen 2008, 2017).

- (57) There was really no need for the police to have a cow, but a cow was had, resulting in kettling, CS gas and 182 arrests.

In their core use, such idioms cannot be passivised; it is only by treating them as live metaphors that they can be. And this move has cognitive and stylistic implications—it is a marked way of using an idiom.

Although these ‘extended’ cases are a fascinating area of fertile research, I will have little more to say about them in this thesis. Instead, we will focus on the ‘core’ cases, arguing for a particular version of the “direct access” family of theories mentioned by Findlay et al. (2018); that is, the view that idioms are lexically encoded, rather than inferred or generated by some mapping procedure from literal expressions. However, in Chapter 3.5.3, I briefly discuss how non-direct access theories can offer some insight into the phenomena introduced in this section, and sketch a means of combining the two families of theory (to be presented fully in Findlay et al. in preparation).

2.3 Chapter summary

In this chapter, I have clarified where I draw the line between MWEs and constructions, and then outlined several different properties which MWEs possess and that demand an explanation from a theory of MWEs. Some of these relate to the tension between their fixedness on the one hand and their flexibility on the other, or, to put it another way, between their single-word and multi-word statuses. We will have much more to say about this in the following chapter. Others have to do with their relationship to their literal meanings, or to how they are processed. I will have less to say about these, but they will nonetheless be important to bear in mind when comparing theories. This is the major focus of the next chapter: now that we understand what a theory of MWEs has to account for, we turn to the question of what such a theory might look like, and to the task of assessing how well different kinds of theory fare when faced with the data.

3

Approaches to MWEs

Contents

3.1	The monolexemic approach	50
3.2	The space of possibilities	55
3.3	Lexical ambiguity	58
3.3.1	Strengths of the lexical ambiguity approach	59
3.3.2	The collocational challenge	60
3.3.3	Non-decomposable MWEs	67
3.3.4	Other weaknesses	74
3.3.5	Summary	76
3.4	One-to-many lexical entries	77
3.5	Semantic mapping	78
3.5.1	Strengths of the semantic mapping approach	82
3.5.2	Weaknesses of the semantic mapping approach	88
3.5.3	Extended uses of idioms	92
3.5.4	Summary	96
3.6	Construction-based theories	96
3.6.1	Strengths of the construction-based approach	99
3.6.2	The real words problem	106
3.6.3	Summary	108
3.7	Comparing the approaches	108
3.7.1	Semantic idiomaticity	110
3.7.2	Syntactic idiomaticity	111
3.7.3	Lexical idiomaticity	112
3.7.4	The real words problem	113
3.7.5	Internal complexity	118
3.7.6	Psycholinguistic plausibility	121
3.7.7	Extended uses	122
3.7.8	Chapter summary	122

We now have an idea of what facts there are to explain in the domain of multiword expressions (MWEs). In this chapter, we examine the different kinds of possible analysis that have, and that could be, put forward to capture them. In Section 3.2, I consider the logical space of possibilities for such a theory, based on the tripartite division of ‘word’ introduced in Chapter 1, arguing that we can rule out most on empirical or conceptual grounds, leaving us with three real families of contenders. Subsequent sections introduce these three in detail, and report on relevant proposals in the literature. Finally, Section 3.7 compares these approaches, and ultimately concludes that a construction-based theory is preferable. Before this, however, I begin by considering the ‘simplest’ theory, that MWEs are really just big words.

3.1 The monolexemic approach

The simplest approach to MWEs is, essentially, to deny their multiword status. Since they need to be listed in the lexicon, we treat them like the usual kind of lexical entry; that is, we treat them as words. This is referred to as the WORDS-WITH-SPACES approach by Sag et al. (2002), and the LEXICAL ITEM (LI) view by Egan (2008). Whatever we call it, the key component is a denial of the multiword status of MWEs, so that they in fact behave like single words. For this reason, I call it the monolexemic approach.

Note that this kind of analysis actually goes against Baldwin & Kim’s (2010: 269) definition, given in Chapter 1 and repeated below:

- (1) Multiword expressions (MWEs) are lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity.

The monolexemic analysis claims that such expressions explicitly *cannot* be decomposed into multiple lexemes, at least synchronically. In other words, this kind of analysis contends that such appearances of multiplicity are really

misleading. If we wish to pursue this line of thought, therefore, we should modify the definition in (1) by adding 'superficially' at the start of clause (a). We wish to describe the same set of expressions, but claim not that they *are* made up of multiple lexemes, but merely that they appear to be.

Assuming this definitional concession can be permitted, we may now ask how well the the monolexemic approach fares when it comes to explaining the properties of MWEs outlined in the previous chapter. The answer is that despite its apparent naïveté, it actually performs better than we might think.

Specifically, the question of idiomaticity, of any of the three types, poses no problem, since words are the prototypical instances of idiosyncratic pairings of form and meaning which must be listed. That is, since each idiom is its own (single) listeme, it is perfectly straightforward to record any idiosyncratic properties the idiom might have.

What is more, such approaches predict a more rapid processing of MWEs than of standardly compositional strings of multiple words, since the latter involves multiple instances of lexical retrieval followed by syntactic combination, while the former will only involve a single instance of retrieval. This is in keeping with the general consensus in the psycholinguistic literature that idioms are processed faster than compositional expressions.

All of these advantages obtain, however, because the monolexemic approach favours one side of the MWE tension over the other. Any approach to MWEs has to confront the tension between MWEs' single-word and multi-word statuses, between their fixedness and their flexibility, but the monolexemic approach deals with this tension by dissolving it in favour of one side. By relegating MWEs to the status of single words, it favours idiosyncrasy over regularity, and so those properties which are related to idiomaticity and fixedness are well covered, but those related to flexibility and productivity are completely ignored.

When it comes to morphosyntactic or syntactic flexibility, then, the monolexemic approach fares much worse. Inflection is a problem, since words in English don't inflect internally, but MWEs clearly do: for example, the past tense of *kick*

the bucket is kicked the bucket, not **kick the bucketed* as we would expect if this was a simplex word. If one insists on the monolexemic approach, one is forced to claim that all MWEs which inflect are irregular words, and do not follow the usual paradigms. This is unsatisfactory, however, given that the inflection in MWEs always appears on the appropriate subpart of the expression—verbal morphology appears on verbs, nominal on nouns, etc.—and that it follows the morphological pattern of that subpart, regular or otherwise—the past tense of *come a cropper*, for example, is *came a cropper*, with the usual irregular past tense form of *come*. Without any internal structure in the MWE, one cannot make any principled statements about this, and such patterns are nothing but a large-scale coincidence, meaning that a huge generalisation has been lost.

For much the same reason—the proposed lack of internal structure in MWEs—the monolexemic approach totally fails to deal with syntactic flexibility. For example, on the assumption that *spill the beans* is a single morphological word, a verb, there is no reason that the passive of *spilled the beans* should be *the beans were spilled*, since no other verb in English passivises by moving its second half to the subject position and adding the perfect participle morphology to the remaining first half. Similar remarks can be made about syntactic flexibility across clausal boundaries: there should be no more reason for (2a) to be grammatical than (2b), for example, since, by hypothesis, *pull strings* is as much a word as *contribute*:

- (2) a. The strings that Shepard pulled got the thing made. (GloWbE)
 b. *The bute that Shepard contried got the thing made.

In a theory which assumes that syntactic atoms do not have internal (syntactic) structure (as is claimed in lexicalist theories, for instance), the subparts of MWEs should be as inaccessible as the subparts of words.

If one believes the Nunberg et al. (1994) hypothesis that syntactic flexibility is a reflex of semantic decomposability, then one might respond that the contrast in (2) is a false equivalence: *strings* in *pull strings* is meaningful, whereas *bute* in *contribute* is not, hence the former can head a relative clause while the

latter cannot. But this requires that the monolexemic approach has some way of modelling semantic decomposability, and this, too, is lacking. Although formally there may be nothing stopping us from associating multiple meaning contributions with the same lexical item, without having some internal structure in the MWE we have no way of associating a particular meaning with a particular part of the expression. Much the same goes for internal modification—there is no way of modifying a subpart of the word if there is no internal structure. And although there are productive processes of apparent infixation in English, as in (3), this is (a) usually assumed to be restricted to a very narrow set of emotive/expletive modifiers, not the wide range we see in some MWEs,¹ and (b) governed by prosodic, not syntactic/semantic constraints (McCarthy 1982).

- (3) abso-fucking-lutely, inde-goddamn-pendent, halle-bloody-lujah
(McMillan 1980)

These problems are in part related to the fact that the monolexemic approach totally fails to address the real words problem. In this approach, MWEs are explicitly claimed *not* to be made up of multiple words, so the fact that *kick* appears in *kick the bucket* becomes no more significant than that the string *ran* appears in the word *cranberry*.² This is clearly unsatisfactory, and gets to the heart of the problems with the monolexemic approach: *multiword* expressions really are made up of multiple words, i.e. multiple, separate morphological objects, and if we do not model this then we fail to capture an important component of the puzzle.

The monolexemic approach is therefore wholly inappropriate as a general theory of MWEs, since it does not really resolve the tension inherent in their analysis at all, instead simply ignoring their regularities in favour of their

¹McMillan (1980) does point out that more neutral terms also appear, such as *guaran-absolutely-tee*. However, these seem to me much less natural, and more clearly in the domain of word play, which suggests they require a rather different analysis.

²Nunberg et al. (1994: 492–3, fn. 2) make a similar point when discussing figuration: even when people have no idea *why* a particular metaphor is used, as with *bite the bullet* or *kick the bucket*, they nonetheless perceive that there *is* a metaphor present, and it is not simply that *bullet* or *kick* are accidental homophones of their non-idiomatic counterparts.

idiosyncrasies. But just because it is unfit for purpose generally, this does not necessarily mean that it has no role to play in the analysis. There are, it would appear, certain MWEs which really do behave like single words. These are what Sag et al. (2002) describe as fixed expressions, and include for instance *by and large*, *in short*, *kingdom come*, and *every which way*, as well as foreign phrases like *ad hoc*, *terra firma*, or *bon appétit*. These expressions do not inflect, and cannot be split up by syntactic processes like passivisation or internal modification, and so some of the key objections to the monolexemic approach are of no consequence.

However, I do not believe the approach is very satisfying even in this narrow domain. These expressions are still clearly made up of things which are recognisably separate words (excepting perhaps the foreign expressions, which it is conceivable that a non-literate speaker who does not also speak the foreign language in question may not view as made up of more than one word), and the monolexemic approach really has nothing to say about the real words problem. This certainly makes the approach conceptually less appealing, even if it achieved adequate coverage of the facts. But it is not at all clear that it even does that. For although these fixed expressions are not susceptible to the full gamut of syntactic manipulations, there are nonetheless ample examples of speakers manipulating their internal structure productively:³

- (4) a. I've worked for both good companies and bad. But Bhandal was a nightmare in every psychotic which way. (WWW)
- b. The hat had fallen off his head at some point, his ears twitching in every little which way. (WWW)
- c. I've been out of balance every which ways. (WWW)
- (5) a. Bon bloody appétit! (WWW)
[literal: said of some Halloween snacks decorated with fake blood]
- b. Bon appétit, Marik! Bon bloody appétit! (WWW)
[expressive: said in anger/frustration]

³The difference in accenting in examples (5a) and (5b) is present in the originals.

- c. Mine, it appeared, had suffered from the full thud of my five-eleven frame hitting terra extremely firma.

(Francis 1978: 179; quoted in McMillan 1980: 179)

- (6) a. A: By and large, the economy seems to be doing well.
 B: By but not so large: have you seen the latest unemployment figures?
 (Glucksberg 1993: 7)
- b. But by and indeed large, the Space-Sim has long since perished. (iWeb)

Some of these may count as ‘extended’ uses in the sense described above (although it is not immediately apparent how expressions without a literal interpretation can be understood as metaphors), but they surely do not all have the feel of word play. Under either analysis, however, it remains important that speakers have (access to) knowledge of the internal structure of these expressions. It is possible, of course, that in indulging in such word play, speakers merely recognise the superficial resemblance of these expressions to multiple words, and perform some kind of reanalysis. The most straightforward explanation, however, is that such expressions are represented with internal structure even in the basic uses.

Fundamentally, the intuition behind the monolexemic approach, namely that MWEs are atomic at the level of listeme, syntactic atom, *and* morphological object, is simply wrong. Even if the approach could be retained for certain kinds of expressions (given a willingness to accept implausible analyses) its lack of faithfulness to our intuitions should speak against it. We *are* aware of the multiword status of MWEs, and this means that, at least at the level of morphological object, such multiplicity must be represented.

3.2 The space of possibilities

As we might expect, then, multiword expressions should be analysed as containing multiple ‘words’, at least in the sense of morphological objects. But what of the other two of Di Sciullo & Williams’s (1987) levels? Do MWEs correspond to a

	A	B	C	D	E	F	G	H
Morphological Object	–	–	–	–	+	+	+	+
Listeme	–	–	+	+	–	–	+	+
Syntactic Atom	–	+	–	+	–	+	–	+

Table 3.1: The space of analytical possibilities for MWEs

single listeme or a single syntactic atom, or can their ‘multi-ness’ be represented at these levels as well? Let us consider the full space of analytical possibilities, given these two dimensions of variation: conception of word/lexical item, and multiplicity/singularity. The cross-tabulation of the two is given in Table 3.1, where + represents multiplicity at that level, and – represents singularity. Each of the columns represents a different theoretical approach to MWEs, a different balance between their unitary and divided natures, and a different location of whatever complexity is introduced. In principle, there are eight such families of theories: call them theories A–H.

In reality, we can reduce the space of theories quite dramatically. Some can be ruled out empirically: as we saw above, it is important that *multiword* expressions contain multiple morphological objects. Otherwise, our theory will suffer from the “flexibility problem” described by Sag et al. (2002) and outlined above, as well as the other issues discussed in the previous section. This eliminates theories A–D from consideration.

In addition, we can rule out another possibility on logical grounds. Theory G has MWEs being represented by multiple listemes, but still entering the syntax only once, as a single syntactic atom. It is not clear what this would entail, formally speaking, or, in fact, how it would differ from having only one listing in the lexicon—multiple listemes which can only be introduced into a parse simultaneously, and as one unit, may as well be a single listeme. In other words, [Listeme +] entails [Syntactic Atom +].

This leaves us with three theories, namely E, F, and H, which, broadly speaking, correspond to three kinds of accounts which have been proposed in the literature. E represents the CONSTRUCTION-BASED approach, where MWEs

are treated as pre-composed, singular units (e.g. syntactic trees) of more than one morphological object (e.g. terminal nodes). Thus, they are listed once, and as a single syntactic atom, but this object contains complex internal structure. Examples of this kind of approach include Abeillé (1995) and Jackendoff (1997). H represents the totally lexical approach, what I call the `LEXICAL AMBIGUITY` approach, whereby MWEs contain special versions of the words that make them up, constrained so that they only appear in the requisite contexts. Examples of this approach include Kay et al. (2015) and Lichte & Kallmeyer (2016). F represents something in the middle, where a single listeme corresponds to or contributes more than one syntactic atom. Examples of this include Pulman (1993), where idioms are listed as semantic rules (one per idiom) which operate on normally derived syntactic objects, i.e. ones containing multiple syntactic atoms; or Biswas (2017), which treats periphrasis as an instance of a single lexeme contributing more than one syntactic object in the process of its morphological realisation.

Since we have decided that MWEs should contain multiple morphological objects, the choice between these theories ultimately comes down to whether we want to have multiple listemes and/or multiple syntactic atoms: the construction-based approach, E, says neither; F says multiple syntactic atoms; and the lexical ambiguity approach, H, says both.

In the following three sections, I will examine each of these families of theories in turn, outlining to what extent they address the analytical challenges described in the previous chapter. We will conclude in Section 3.7 by taking a different perspective, and considering each of the properties discussed in Chapter 2 in turn, and looking at what makes different kinds of theories well suited or not so well suited to explaining them. We will find that the construction-based approach offers the best overall coverage, but that there are still aspects which remain unsatisfactory.

3.3 Lexical ambiguity

Approach H, the lexical ambiguity approach, involves taking quite the opposite stance to the monolexemic approach: in the most radical version of that theory, every row in Table 3.1 contained a –; here, every row contains a +. This emphasises the phrase-like rather than word-like properties of MWEs. The fundamental intuition is this: since MWEs often exhibit flexibility and modification possibilities which suggest that each of their parts is independent and has its own meaning (at least in the decomposable cases), they are modelled as being made up of separate words which combine to give the appropriate meaning for the whole expression. In this kind of theory, words like *pull* and *strings* become ambiguous, meaning either ‘pull’ and ‘strings’, as in the literal expression, or ‘exploit’ and ‘connections’, as in the idiom.⁴ This is what Egan (2008) calls the CHUNKS THEORY, and what I will call the lexical ambiguity approach. Such a move is on the face of it appealing, since it retains the usual one-to-one mapping between morphological objects, syntactic atoms, and listemes, while reducing the peculiarities of MWEs to lexical idiosyncrasies, the kinds of things any theory of syntax needs to assume.

There are two distinct ways that we might think *pull* could be ambiguous between ‘pull’ and ‘exploit’: either there are two different lexical entries (*pull*₁ and *pull*₂), i.e. a situation of homonymy, or there are two meanings associated with the same entry, i.e. a situation of polysemy. Both kinds of theory are attested in the literature: for an example of the first kind of theory, see Sailer (2000) and Waszczuk & Savary (2015); and for the second, see Gazdar et al. (1985)

⁴When I say that *pull* means ‘exploit’, this does not necessarily mean that this version of *pull* has precisely the same meaning as the word *exploit*, although it may do. To make the distinction clear, in our formal meaning language we should use separate predicates, so that the denotation of idiomatic *pull* is **pull-id** or something similar, and not **exploit**, which is the denotation of *exploit*. This is common practice in the frame semantics and Minimal Recursion Semantics literature (e.g. Riehemann 2001; Ruppenhofer et al. 2010; Kay & Sag 2012). For the purposes of the present discussion, and when it does not introduce confusion, I will keep to the more transparent approach taken here, where it is implied that the logical constant corresponding to the meaning of the idiom word is the same as that corresponding to the best ‘translation’. In my own formal treatment, given in Chapter 6, however, I will once again make use of idiom-specific predicate names, to avoid any imprecision.

and Lichte & Kallmeyer (2016). While these two approaches are theoretically distinct, it happens that they are inter-translatable in any formal theory which allows disjunction in its description language—instead of having two lexical entries, one can simply have a single entry with a massive disjunction.⁵ Because of this formal isomorphism, I will treat them as instances of the same kind of general theory in this section, but will attempt to draw attention to areas where the two accounts may diverge.⁶

3.3.1 Strengths of the lexical ambiguity approach

The lexical ambiguity approach deals naturally with the syntactic flexibility demonstrated by decomposable idioms, since the idiom parts are just normal words, and so all the usual syntactic processes can apply: whatever one's story is about relative clauses or voice alternations, it carries over without the need for special rules. Morphosyntactic flexibility also follows without further comment, since the internal structure of MWEs is affirmed in this kind of theory, and so idiom verbs will naturally inflect in the same way as any other verb. Lexical idiomaticity, too, becomes a non-problem, since the idiom-only words can be assigned appropriate meanings in the lexicon (in this case, of course, the lexical ambiguity approach does not actually involve *ambiguity*, since the words will have no idiom-external meanings).

Such an approach is a vast improvement over the monolexic approach, therefore, since it maintains many of the former's advantages while also solving

⁵Theories which treat category information as distinct from other kinds of lexical information (e.g. Lexical Functional Grammar) impose at least some constraints on this isomorphism, since *bank*, N, 'financial institution' must be a separate lexical entry from *bank*, V, 'to tilt sideways' (but not necessarily from *bank*, N, 'land along the side of a river').

⁶We might wonder what the theoretical implications of this formal isomorphism are for such frameworks. After all, this makes them essentially agnostic with respect to the distinction between polysemy and homonymy. So, we may ask, is such a difference important? Certainly there is an important distinction to be drawn when it comes to diachrony, where the histories of homonyms will be distinct, while that of polysemes will be the same; but is the distinction also important synchronically? Pylkkänen et al. (2006) have shown using MEG experiments that there is a difference in the way homonymous and polysemous words are processed on-line, which suggests that it is, at least insofar as formal theories of grammar are expected to align with processing facts.

the flexibility problem: idioms are flexible in precisely the same way normal phrases are, because, quite simply, they are themselves normal phrases. However, this once again dissolves the MWE tension by coming down entirely on one side of it, albeit the opposite side this time. As we might expect, this has its own attendant shortcomings, and we now turn our attention to these. Some of them can be addressed to a greater or lesser extent, but many inhere in the philosophy behind the approach itself, and so are insurmountable.

3.3.2 The collocational challenge

One technical issue for this approach is the need to restrict the distribution of idiom words, so as to prevent over-generation. For example, if one version of *pull* means ‘exploit’ and one version of *strings* means ‘connections’, we want to prevent the generation of sentences like (7), which lack the idiomatic interpretation:

- (7) a. #You shouldn’t pull his good nature.
 b. #Peter was impressed by Claudia’s many strings.

This is what Bargmann & Sailer (2018: 12) refer to as the COLLOCATIONAL CHALLENGE.

An early proposal for solving this problem was put forward by Gazdar et al. (1985: sec. 10.7). They make use of partial functions in their meaning language, so that idiomatic meanings are only defined in the correct context. Thus, *pull* corresponds to one of two constants in the meaning language, **pull** or **pull-id**; and *strings*, to **strings** or **strings-id** (ignoring the question of plurality). We then define them in such a way that all combinations which do not involve both idiomatic parts are undefined:

- (8) a. **pull-id(strings-id)** : defined
 b. **pull(strings-id)** : undefined
 c. **pull-id(strings)** : undefined

Thus, unless idiomatic *strings* is the argument of idiomatic *pull*, neither can have their idiomatic interpretation.

This is not without its challenges, however. Firstly, as Lichte & Kallmeyer (2016: 120) point out, there is “no genuinely semantic motivation” for having a separate **pull-id** predicate, when it “conceptually coincides” with the existing predicate **exploit**. In the partial function theory, it is vital that idiom meanings are *not* identical to their apparent translations, even though, at least superficially, they seem to be. These different meaning constants “essentially reflect morphological properties” (Lichte & Kallmeyer 2016: 120): they are different because the form of words used is different, not because the meaning is. Such a confusion of form and meaning in the meaning language itself could be seen as a deficiency.

However, if idiom meanings truly are different from their putative translations, then this is no real objection. McGinnis (2002), for instance, argues that idiom meanings have aspectual restrictions which are based on the normal interpretation of their syntactic forms, rather than on the idiomatic meaning. Thus, for instance, *kick the bucket* has the same aspectual properties as other VPs with definite direct objects, rather than of the semantically intransitive *die* (Nunberg 1977; Marantz 1997: 212):

- (9) a. Ezri was dying for weeks.
 b. #Ezri was kicking the bucket for weeks.

The sentence in (9b) cannot be interpreted idiomatically—or, rather, the idiomatic interpretation would retain the punctual aspect of literal *kick the bucket*, meaning that Ezri would have to have been dying over and over again. This of course runs counter to our world knowledge about how many times one can die, making the idiomatic interpretation highly anomalous. If such mismatches of aspect are widespread, then this offers another motivation for positing unique constants for idiom meanings which are different in at least this respect from their ‘translations’ by other, pre-existing terms in the meaning language.

A more troubling problem for the partial function approach is that it necessarily either under- or over-generates when one tries to make sense of relative clauses like (10) (Pulman 1993: 50f.):

(10) He tried to break the ice which inhibited our conversation.

This is perfectly acceptable with the idiomatic interpretation ('he tried to relieve the tension which inhibited our conversation'). However, if the relative pronoun *which* is assigned the idiomatic meaning of *ice*, then the predicate in the relative clause, *inhibited*, would be incompatible with it, since only one expression in **inhibit(ice-id)** is idiomatic, and idiomatic **ice-id** will only be defined when it is the argument of idiomatic **break-id**. So the partial function account under-generates, since it predicts that (10) cannot have an idiomatic interpretation. But if we amend the analysis to make **ice-id** compatible with literal **inhibit**, then we expect that a sentence like (11) should allow for an idiomatic interpretation, contrary to fact:

(11) #The ice inhibited our conversation.

Now the theory over-generates. All in all, the partial function approach captures some intuition about the mutual selectional restrictions idiom parts place on one another, but it is ultimately too restrictive. A different strategy is required if the lexical ambiguity approach is to be maintained.

Many modern researchers have suggested such alternative strategies; usually, they involve some kind of extended selectional restriction (see, for instance, Krenn & Erbach 1994, Sag 2007, and Soehn & Sailer 2008 [2003] for a selection of work in Head-Driven Phrase Structure Grammar (HPSG) on this theme). In the remainder of this section, we will consider this family of approaches in more detail. But when all is said and done, I do not believe that they are successful in their aims either. For while they can be made compatible with the facts, it is not at all apparent that doing so leaves us with any kind of satisfying

generalisations. Before I demonstrate this, however, we begin with the simple cases, and show how the approach is intended to work.

Modern lexicalist theories all have some means of identifying particular lexemes, e.g. the *lex-id*/LID feature in HPSG (Pollard & Sag 1994) and Sign-Based Construction Grammar (SBCG) (Boas & Sag 2012), or the PRED feature in LFG.⁷ All that is required to achieve the desired effect is for idiomatic *break* to say that idiomatic *ice* must be its object, and, conversely, for idiomatic *ice* to say that it must be selected as the object of idiomatic *break*. For the LFG-versed reader, the partial lexical entries in (12) and (13) give an example of how this could be achieved in that framework; Lichte & Kallmeyer (2016) show how the same can be achieved in LTAG with a paired frame semantics.

$$(12) \quad \textit{break} \quad V \quad (\uparrow \text{PRED}) = \textit{'break}_{id}' \\ (\uparrow \text{OBJ PRED FN}) =_c \textit{ice}_{id}$$

$$(13) \quad \textit{ice} \quad N \quad (\uparrow \text{PRED}) = \textit{'ice}_{id}' \\ ((\text{OBJ } \uparrow) \text{PRED FN}) =_c \textit{break}_{id}$$

The main advantage of this kind of approach is that it makes the requirements syntactic rather than semantic. This means that the problem of under-generation for the partial function approach (and the over-generation problem which results from trying to solve it) does not arise: we do not need partial functions in the semantics, because the restrictions are no longer represented at this level, and so non-idiomatic predicates like **inhibit** can remain agnostic about whether their arguments are idiomatic or not, because there is no formal difference in the semantics. Examples like (10) are, correctly, still licensed because they are in the proper *syntactic* configuration.

Just as in the partial function account, the constraints are defined in both directions (not only does *break* require *ice*, but *ice* also requires *break*); this means that there is no over-generation either: neither part of the idiom can appear without the other.

⁷In fact, this might be one of the few remaining functions of PRED, assuming its roles in the semantics and in subcategorisation have been subsumed by other mechanisms. See Chapter 4.3.4 for further discussion.

However, such constraints, as they stand, are too restrictive. Recall that there are idioms which can passivise: since we can say both (14a) and (14b), it would be insufficient for idiomatic *pull* to require that idiomatic *strings* be its object—in (14b), it is its subject instead.

- (14) a. Kira pulled some strings for me.
 b. Some strings were pulled for me.

What we need to say instead is that *strings_{id}* must be the internal argument of *pull_{id}*, however that argument is realised in the syntax: as the object in the active voice, or the subject in the passive. Note, however, that this already requires reintroducing semantics into our description, albeit only at the level of argument structure, and this will cause problems down the line.

Interestingly, the two options just discussed—selection in terms of grammatical functions or in terms of semantic arguments—might be thought to give us a useful way of characterising the distinction between those VP idioms where the object can be passivised and those where it cannot: the former state their selectional restrictions in terms of argument structure, allowing for syntactic flexibility, whereas the latter state them more strictly, in terms of grammatical functions. However, there is far more to the collocational challenge than just this contrast.

Firstly, there is the question of syntactic idiomaticity. In order to be able to state the selectional restrictions between two words, we must be able to ascertain what (syntactic) relationship they stand in. But how do we do this in cases where the syntactic structure is esoteric? What grammatical relationship holds between the words in *every which way*, for instance? Or in *trip the light fantastic*? To assign them an arbitrary structure in order to then describe the selectional restrictions in terms of that structure is to beg the question. (For more on syntactic idiomaticity, see Section 3.3.4.1, below.)

Secondly, many VP idioms of the *break the ice* kind described above can also be split across a relative clause boundary:

(15) Who's at the centre of the strings [that were quietly pulled]? (TV script)

But it is generally assumed that there is no direct syntactic or semantic relationship between the head noun of a relative clause and the predicate inside it. Instead, a relative pronoun, overt or covert, is taken to be the argument of the within-clause predicate, and this is related anaphorically to the head noun. So the head noun and the predicate are related only indirectly, via the pronoun; Falk (2010) describes this state of affairs as the ANAPHORICALLY MEDIATED theory of relative clauses. In such a theory, even the argument structure-based description of the dependency between the two parts of the idiom is insufficient: *strings_{id}* is not the internal argument of *pull_{id}* in (15); the (covert) relative pronoun is instead. *Strings_{id}* is merely coreferential with that pronoun.

Perhaps, therefore, we should return to a semantic constraint, and stipulate that the internal argument of idiomatic *pull* must be *coreferential* with idiomatic *strings*, as it is in (15). Unfortunately, this is now too permissive; (16) cannot receive an idiomatic interpretation, for instance, even with coreference:⁸

(16) #Those are some impressive strings_{*i*}—you should pull them_{*i*} for me!

The correct description seems instead to be a mixed one: *strings_{id}* is either the internal argument of *pull_{id}*, or it is modified by an (arbitrarily deeply embedded) relative clause headed by *pull_{id}*, which has a pro-form as its internal argument which is coreferential with *strings_{id}*. This, at least, seems to describe the facts. Although the second disjunct may be cumbersome, and does not look like a very natural generalisation, it is, nonetheless, statable.⁹

⁸This is essentially just a recapitulation of the point made by Pulman (1993) regarding over-generation discussed above. The only difference is that here we at least require explicit coreference between a pronoun and the idiom chunk—although this ultimately makes no difference to the acceptability.

⁹For the LFG-literate reader, (i) provides the requisite description to be associated with idiomatic *strings*:

(i) $\{((\text{ARG}_2 \uparrow_\sigma)_{\sigma^{-1}} \text{ PRED FN}) =_c \text{pull}_{id} \mid$
 $(\uparrow \text{ ADJ} \in \text{COMP}^*) = \% \text{PULL-CLAUSE}$
 $(\% \text{PULL-CLAUSE PRED FN}) =_c \text{pull}_{id}$
 $((\% \text{PULL-CLAUSE}_\sigma \text{ ARG}_2)_{\sigma^{-1}} \text{ PRED FN}) =_c \text{PRO}$
 $(\% \text{PULL-CLAUSE}_\sigma \text{ ARG}_2 \text{ INDEX}) =_c (\uparrow_\sigma \text{ INDEX})\}$

Some researchers might, at this point, be satisfied. We have observed a generalisation, and can give a characterisation of it which can be represented explicitly in our theory. My protest that the generalisation is ‘unnatural’, or any objection as to the ‘inelegance’ of the theoretical account, would be thought of as merely a question of aesthetics, and thus not a proper subject for scientific debate. I wish to disagree. It is well known that many modern syntactic frameworks are extremely computationally powerful;¹⁰ this means that they can describe almost any arbitrary relationship between two words. It is therefore hardly surprising, and also far from revealing, that they can describe this one too. What we should value in theories is not only adequate description of the facts (although this is obviously essential), but also some notion of explanation.¹¹ By the same token, we should prefer general solutions over *ad hoc* ones. Formally powerful theories are always in danger of providing inadequate answers to the questions we pose, since a mere description of the facts is always possible, but also rarely enlightening.

When assessing theories, therefore, theoretical ‘ugliness’ is not merely an aesthetic concern, but can be indicative of missed generalisations.¹² Disjunctive rules are not, in general, good rules. A theory of the orbit of the planets which just gives eight different orbital speeds for each of the planets is not as good as one which states a single rule for calculating said speeds in terms of mass, distance from the sun, etc., even though it would be a perfectly accurate

The first disjunct just says that idiomatic *strings* must be the internal argument of idiomatic *pull*. The second disjunct breaks down as follows: first, we identify some *f*-structure inside the ADJ set of *strings*, which can be embedded in any number of COMP clauses, including zero, and give it the local name %PULL-CLAUSE (see Crouch et al. 2012 on local names). Next, we state that this *f*-structure is headed by idiomatic *pull*, that its internal argument is pronominal (at *f*-structure; it need not be realised by an overt pronoun at *c*-structure), and that it shares an index with idiomatic *strings*. This final line enforces the requirement of coreference, using the INDEX feature introduced by Dalrymple et al. (2018).

¹⁰For example, HPSG is NP complete (Blackburn & Spaan 1993), and LFG (as a formalism) is context sensitive—Berwick (1982) shows that in, the worst case, LFG grammars are intractable, although a recent proposal by Wedekind & Kaplan (2018) constrains the formalism in linguistically-motivated ways in order to avoid this outcome.

¹¹This is of course far from a new idea, going back to Chomsky (1964) in linguistics, and arguably in some sense to William of Ockham and the law of parsimony which bears his name.

¹²My thanks to Agnieszka Patejuk and Gert Webelhuth for discussing this point with me, and pushing me to clarify my thoughts on it.

statement of the facts. So too here. I contend that the collocational challenge is not one easily met by lexical ambiguity approaches, and that it is no real defence of such theories to say that it can be met by *ad hoc* stipulations and disjunctive rules. These may address the letter of the challenge, but surely not its spirit.

3.3.3 Non-decomposable MWEs

The lexical ambiguity approach is particularly well suited to decomposable idioms, since in this case we really do want to assign individuated meanings to the parts of the MWE, and so it makes sense to do this in the lexicon. For other kinds of MWE, however, the approach is much less appealing; it means, for instance, that in non-decomposable idioms, we are forced to make a decision about where to locate the meaning—a decision which seems ultimately to be an arbitrary one. Which of the three words in *kick the bucket* ought to host the meaning ‘die’, for example? Perhaps the head, *kick*, which is at least also a verb? But there is nothing in the approach itself which makes this any more appropriate a host than *bucket*, or even *the*. And then there is the question of what to do with the other words: must they be semantically empty, or contain some vacuous identity function? The same holds for even more syntactically fixed expressions, like *by and large*, which, ultimately, means something like ‘largely’. Again, we want to associate a simplex meaning with a complex expression, but that is not something this approach is designed to do.

One solution would be to have different accounts of the different types of MWE. This is what Sailer (2000) does, for instance, treating idioms like *kick the bucket* as fixed phrasal expressions, and idioms like *spill the beans* as being formed compositionally from special versions of *spill* and *beans*. But many recent works (e.g. Kay et al. 2015; Lichte & Kallmeyer 2016; Bargmann & Sailer 2018) have sought to extend the lexical ambiguity approach even to the non-decomposable cases where it has less *prima facie* appeal. In this section, I examine how successful these attempts have been. The theories divide into three camps, depending on how they achieve the goal of deriving a simplex meaning from

complex parts, but I conclude that none of the possibilities are very satisfying. Once again, by insisting on treating MWEs as wholly the sum of their parts, and thereby denying them any independent existence, such theories are forced to tie themselves in theoretical knots solving problems that otherwise would not arise.

3.3.3.1 One meaning, one place

The first possibility when treating non-decomposable MWEs in the lexical ambiguity approach is to assign the meaning of the expression to one of the words and to have the others either be empty (have no meaning) or contain a vacuous identity function. For instance, perhaps *kick* means ‘die’, while *the* and *bucket* make no semantic contribution. This is the approach taken by Kay et al. (2015), for instance. What is more, if we assume the Resource Sensitivity Hypothesis of Asudeh (2004, 2012), which says that all meaning contributions by words must be used exactly once, then this is the most natural position to adopt.¹³

There are a number of issues with such an approach. Firstly, the decision about which word should host the meaning is apparently arbitrary. Since the other words make no contribution (or a vacuous one), it makes no difference whether *kick* means ‘die’ or whether *bucket* does instead, or *the*, for that matter. This degree of arbitrariness is undesirable in a theory.¹⁴

¹³Arnold (2015) proposes a Glue-based approach (which is therefore resource-sensitive; see Chapter 4.4.2) where *kick* carries the meaning of *kick the bucket*, but where the other words in the expression, *the* and *bucket*, are *not* semantically empty. He does so by using MANAGER RESOURCES (Asudeh 2004), which can ‘throw away’ the meaning associated with ‘the bucket’. However, since such tools can, if used unconstrainedly, allow one to ignore the requirements of resource sensitivity altogether, they need to be invoked only where there is strong motivation. Asudeh (2004, 2012) motivates their use in a resource-sensitive setting by associating them with marked morphological material, e.g. particular complementisers in Irish; but there is no such marking in the case of idioms. Arnold’s approach also suffers from major empirical failings, as outlined by Arnold himself, and discussed in more detail in Findlay (2017a: Appendix).

¹⁴The point already observed about idioms retaining the aspect of their literal verbs may be one argument for hosting the idiom meaning on the verb, *kick*. In this case, we could advocate a polysemy-based theory whereby one lexical item contains two different ‘lexical’ meanings which share the same aspectual meaning—corresponding to the IDIOSYNCRATIC and STRUCTURAL components of meaning discussed by Levin & Rappaport Hovav (1998) and McGinnis (2002). But this won’t help in cases where the aspect is not so obviously carried over: for instance, it seems to me that *shoot the NP_[SG]* shares with *kick the bucket* the property of being punctual and being an achievement, so that literal *We were shooting the barrel for hours last night* can only have

Secondly, this approach necessitates adding a large amount of ambiguity to the lexicon, either in terms of additional lexical entries or additional meanings associated with existing entries. Potentially, in fact, as many new meanings as there are words in idioms. For, as discussed above, we need to restrict idiom meanings so that they don't appear outside of the idioms they belong to. Now we have empty idiom meanings, we need to do the same for them: we don't want semantically vacuous *the* to be available elsewhere, for instance, otherwise (17) should be ambiguous between a definite and a generic reading (i.e. the reading corresponding to (18), where the definite article is absent):

(17) The hippos are mad.

(18) Hippos are mad.

However, this requirement means that the *the* in *kick the bucket* cannot be the same *the* as in *shoot the breeze*, or *cut the mustard*, etc., because the selectional restrictions are specific to each idiom (one of them needs to be the specifier of idiomatic *bucket*, one of them the specifier of idiomatic *breeze*, etc.). But this seems to be an intolerable explosion of ambiguity. Either the lexicon expands in size hugely, or else the lexical entries for simple words like *the* become exceedingly complicated, incorporating a huge disjunction, one for each idiomatic context. On anyone's theory of idioms, we need at least one listed item per idiom, so some expansion of the lexicon (or 'the list') is necessary. But here we face the worst case scenario: the lexicon is expanded by one entry per word in an idiom. This surely demands some justification. And yet in the non-decomposable case such justification seems decidedly absent. For decomposable idioms, there is independent motivation for the ambiguity: there is a sense in which *beans* means 'secrets' in the context *spill the beans*, since it can be individually modified, quantified, etc. But there is no such independent motivation available in the case

an iterative meaning; and yet idiomatic *We were shooting the breeze for hours last night* is perfectly acceptable with a durative meaning. What is more, in non-verbal MWEs we do not have aspect to guide us. Where should the meaning of *by and large* be hosted, for instance?

of non-decomposable idioms like *shoot the breeze*; the empty meanings are there purely to serve the theory. Such *ad hoc* theorising is surely to be avoided.

3.3.3.2 One meaning, everywhere

An alternative is presented by Lichte & Kallmeyer (2016), who argue for what they call ‘idiomatic mirroring’, whereby the entire meaning of a non-decomposable idiom is found in all of the components. This is possible because of the unification-based semantics they are working with (Kallmeyer & Romero 2008): multiple words can make identical semantic contributions, which are then unified together, ultimately corresponding to only a single meaning (making this theory incompatible with the Resource Sensitivity Hypothesis). For example, both *kick* and *bucket* contribute the meaning ‘die’. Presumably, *the* must also mean ‘die’, although the authors ultimately avoid answering this question one way or the other (Lichte & Kallmeyer 2016: 125).

This approach has the advantage of not requiring a large number of semantically empty words, but it still introduces widespread ambiguity into the lexicon. Is this any more motivated than in the previous version of the theory? Lichte & Kallmeyer (2016: 14–15) argue that, at least in the case of *kick the bucket*, there is some independent motivation for having the meaning ‘die’ associated with *bucket* as well as with *kick*, since it can carry this meaning even outside the idiom, notably in the expression *bucket list* (a list of things one wants to do before one dies). However, a limited analogical creation like this doesn’t show that the word *bucket* meaning ‘die’ has an independent existence outside of the idiom.¹⁵ Firstly, in Lichte & Kallmeyer’s (2016) account it is required to have a verbal meaning identical to that of idiomatic *kick*, i.e. ‘die’, not a nominal meaning, i.e. ‘death’. But *bucket* is not a verb, and cannot be used verbally; in the *bucket list*

¹⁵For what it is worth, this also seems to be a relatively recent coinage: the *Oxford English Dictionary* dates it to the 2007 film *The Bucket List*, with the earliest citation being from a 2006 press release about the film—see the ‘draft additions’ section of the OED Online entry for ‘bucket, n.2’ (<http://www.oed.com/view/Entry/24163>; accessed 13 March 2019).

example, it is part of a nominal compound—*death list* would be an acceptable literal alternative, but *die list* certainly would not.¹⁶

Secondly, although the ‘die’/‘death’ sense of *bucket* may marginally persist in *bucket list*, it is certainly not generally available to the productive processes of the grammar—there is no *#bucket book* in which to write one’s bucket list, for instance. And this phenomenon is not widespread across other non-decomposable idioms either: I am not aware of any similar analogical extensions based on *shoot the breeze* or *bite the dust*, for instance—there are no *#breeze rooms* where we can have a chat, or *#dust yards* where we store the dead.

We have to conclude, then, that associating the idiom meaning with each of the parts of a non-decomposable idiom is just as unmotivated as associating empty meanings with them—it serves the theory, but has no independent justification.

A final damning point for this approach is that we still blow up the lexicon. For, in addition to separate selectional restrictions, the two *thes* in *kick the bucket* and *shoot the breeze* now also have different meanings: one means ‘die’ and one means ‘chat’. So once again we face a huge proliferation of ambiguity with no motivation other than to make the theory work.

3.3.3.3 Multiple meanings after all

The final alternative is that advocated by Bargmann & Sailer (2018): the apparently simplex meanings of non-decomposable idioms are treated as being in fact complex. Using Lexical Resource Semantics (Richter & Sailer 2004), Bargmann & Sailer (2018: 12–15) outline what they call a REDUNDANCY-BASED analysis,

¹⁶In the frame semantics setting which Lichte & Kallmeyer (2016) use, this distinction is obscured, and so it might be thought that the difference between *die* and *death* is a morpho-syntactic one, rather than a semantic one. Proponents of this position might point out that predicates which can apply to the one can apply to the other too, with the same kinds of morpho-syntactic modifications: if there is a *sudden death*, then someone *died suddenly*; if someone *died peacefully* then they had a *peaceful death*. However, this is not true for all predicates: if we describe something as an *ominous death*, this does not imply that someone *died ominously*, for instance: the adverb here can only have a manner reading, but this is not the natural interpretation of the adjective, which says that it is the death itself, rather than the manner of death, which augurs badly. Even more striking are intensional adjectives like *fake*: the existence of a *fake death* certainly does not entail that someone *died fakely*.

whereby different parts of the idiom contribute different components of the overall meaning. For instance, they give the following semantic analysis of the words in *kick the bucket*:

- (19) a. $kick_{id}: \langle s, \mathbf{die}'_{id}, \mathbf{die}'_{id}(s, \alpha), \exists s(\beta) \rangle$
 b. $the_{id}: \langle s, \exists s(\beta) \rangle$
 c. $bucket_{id} \langle s, \mathbf{die}'_{id}, \mathbf{die}'_{id}(s, \alpha) \rangle$

Each word is associated with a list of its semantic contributions: here idiomatic *kick* contributes a situation variable, a predicate \mathbf{die}'_{id} , a formula for composing this predicate with its arguments, and an existential closure of the situation variable; the determiner contributes a situation variable and the existential closure of it; and the noun *bucket* contributes a situation variable and a predicate, plus the formula for combining it with its arguments. (α and β are meta-variables which must be instantiated to one of the expressions introduced in the contribution lists of whatever words make up the sentence in which they appear; e.g. α will be assigned the semantic contribution of the subject of the sentence.)

All of the required meanings are contained in $kick_{id}$, but are redundantly marked in the other parts of the idiom as well—the parts which are identical must be unified in the final analysis. This reliance on unification is similar to the approach taken by Lichte & Kallmeyer (2016), but the crucial difference is that the meanings of all the parts of the idiom are, formally at least, just like their meanings outside the idiom. For instance, verbs normally contribute a situation, a predicate, and an existential closure of that situation, just as *kick* does here. Determiners normally contribute variables and quantifications over those variables, and nouns normally contribute variables and predicates on those variables, and this is what *the* and *bucket* do here. Thus, some of the *ad hoc* nature of other implementations is alleviated.

However, any such gain is only at the expense of further ad hocery elsewhere. For while it is now true that these words are formally like their non-idiom variants, they are substantively very different. The definite article does not

usually quantify over situations, and certainly does not usually have (merely) existential force. And nouns usually introduce variables over individuals, not situations (except perhaps in the case of event-related nouns like *destruction*). Once again, this approach makes the data fit the theory, rather than the other way around. The fact that it is slightly more sophisticated in this instance, forcing the data to look formally appropriate, does not offset this fact.

What is more, the approach still suffers from the lexical explosion problem. Although it might be thought that by having idiomatic *the* introduce a situation variable we can avoid it combining spuriously with nouns, thus avoiding the definite/generic ambiguity highlighted above, this is not in fact the case. And this is for the very simple reason that the *kick the bucket* example which the analysis is set up to allow exemplifies precisely the problematic configuration. The variable contributed by the idiomatic determiner in *kick the bucket* is unified with the variable contributed by the verb, *kick*. But if this is the case, then there is nothing stopping the same thing happening in (20), using the idiomatic *the*, and thus giving the generic reading even when the determiner is present:

(20) John loves the hippos.

Once again, we need special selectional restrictions to prevent this happening, and so we still need as many versions of *the* as there are MWEs which contain it.

3.3.3.4 Interim summary

The lexical ambiguity approach unavoidably introduces mass ambiguity throughout the lexicon, but this should be seen as a last resort in linguistic theorising—after all, “[i]t is very much the lazy man’s approach in philosophy to posit ambiguities when in trouble” (Kripke 1977: 268), and I should say the same holds for linguistics.¹⁷ We should therefore only appeal to this kind of theory when there is good empirical motivation to do so. While the modification and quantification facts might be thought to provide such a motivation in the case of

¹⁷Cf. what has come to be called ‘Grice’s Razor’: “senses are not to be multiplied beyond necessity” (Grice 1978: 118–119).

decomposable idioms, there is no such justification in the non-decomposable cases. Without such motivation, this approach is guilty of forcing the data to fit the theory.

3.3.4 Other weaknesses

The lexical ambiguity approach suffers from a similar problem to the monolexic approach: by situating itself entirely on one side of the MWE tension between phrase-like and word-like properties, it solves some problems elegantly, but singularly fails to give a satisfactory solution to others, instead provoking awkward theoretical wrangling. In this section, I highlight a few other problems which this perspective gives rise to.

3.3.4.1 Syntactic idiomaticity

As alluded to above, the lexical ambiguity approach does not naturally explain syntactic idiomaticity. Since idiom words are represented as words just like any other, the expectation would be that they should inhabit exactly the same kinds of syntactic structures as any other (especially under the polysemy version of the theory), and so the existence of this kind of idiomaticity is *prima facie* surprising.

Nevertheless, idiosyncratic syntactic configurations can be easily accommodated in this kind of theory by, for example, making use of bespoke phrase structure rules (or whatever the equivalent is in one's favourite formalism), which only allow the appropriate idiomatic realisation of their terminals. But such a move is a major weakening of the strength of such a theory. The lexical ambiguity approach is predicated on the fact that words are the natural locus of idiosyncratic pairings of form and meaning; idiom words are treated as nothing unusual in that respect—they are normal words in all respects except for their extreme selectional restrictions. But if we are forced to introduce special phrase structure rules to account for expressions like *trip the light fantastic*, then idioms are no longer restricted to the lexicon, instead extending to the syntax as well.

A theory which treats words as syntactic atoms will not be able to represent idiosyncratic syntactic structure straightforwardly in the lexicon, because words are not (directly) associated with any syntactic structure beyond their category. A theory which treats larger structures as the building blocks of syntax will have an easier time accounting for such anomalies, and this is what I will argue for in Section 3.6.

3.3.4.2 The real words problem

When it comes to the real words problem, the homophony-based version of the lexical ambiguity approach has no real answer. While here MWEs are at least claimed to be made up of multiple words, giving the theory an advantage over the monolexic approach, the fact that they are almost invariably homonyms of other, pre-existing, words is left as a total coincidence, at least in the synchronic grammar. But what about the polysemy-based version of the theory? This at least has the virtue of predicting the correct inflectional patterns of idiom words: if *come* in *come a cropper* is the same lexical item as literal *come*, then they will of course share the same past tense form. However, lexically idiomatic words pose a problem: if *cropper* has no synchronic existence outside of the MWE, then clearly idiom words need not be polysemes of existing words. And if arbitrary non-words are allowed in MWEs, then it is hard to see how they can be ruled out in general. The real words problem relates closely to the historical development of MWEs, which will be discussed further in Section 3.7.4.

3.3.4.3 Psycholinguistic properties

The psycholinguistic finding that idioms are easier to process than literal expressions is difficult to accommodate in a lexical ambiguity approach: since this treats idiom composition as the same process as regular, literal composition, we would not expect there to be any difference in processing times between the two, contrary to the experimental findings.

Lichte & Kallmeyer (2016: 116) argue that a polysemy-based lexical ambiguity approach actually makes the correct predictions, since it delays ambiguity resolution when compared to a construction-based approach (see below, Section 3.6), the former locating it in the semantics, the latter in the syntax. But of the psycholinguistic papers Lichte & Kallmeyer cite to back up this claim, two out of the three, Wittenberg & Piñango (2011) and Wittenberg et al. (2014), are studies of a very specific kind of MWE, the light verb construction (LVC). As discussed in Chapter 2.2.8, LVCs are different from idioms in that they are partially productive. Given this fact, it seems likely that (at least some) LVCs involve a special meaning associated with the light verb that ‘verbalises’ its nominal complement.¹⁸ And in this case, it is hardly surprising that they may be more psycholinguistically costly than a normal transitive VP. But this has no bearing on idioms being *less* costly, as has been shown to be the case, and so Lichte & Kallmeyer’s (2016) contention that the lexical ambiguity approach makes the right predictions for idioms in general can hardly be sustained.

3.3.5 Summary

Although it appears to offer a natural analysis of decomposable idioms, the lexical ambiguity approach is rife with problems. The collocational challenge is not easily met, and so even the description of decomposable idioms is problematic. And the approach does not readily extend to non-decomposable expressions, *pace* recent work which purports to do just this. Many of its strengths arise from treating MWEs as made up of multiple morphological objects (thus accounting for their morphosyntactic behaviour), but this property is shared by the other approaches outlined in the following sections. Its weaknesses, though, stem from a meta-theoretical misstep: MWEs exhibit a tension between a unitary and a divided nature; just as the monolexemic approach fails by emphasising their unitary nature to the exclusion of their divided one, so too the lexical

¹⁸See Lowe (2015a) for a proposal along these lines involving light verbs in a slightly different domain, that of complex predicates.

ambiguity approach fails by doing the opposite. What one wants to say about *pull strings* is that there is something idiosyncratic about the whole phrase, not independently about *pull* and about *strings* as individual words. In the lexical ambiguity approach, the MWE itself is a consequence of the existence of these two lexical entries, and not the other way around. This is putting the cart before the horse, and ultimately ensures that the approach fails to give a satisfying account of the phenomenon.

3.4 One-to-many lexical entries

Lexical ambiguity approaches claim that MWEs are made up of multiple syntactic atoms, each their own listeme. An alternative family of theories maintains that MWEs contain multiple syntactic atoms, but instead proposes that they are listed only once. There are two major versions of this theory: that the listing happens in the lexicon, before the MWE enters the syntax, or that the listing happens in the semantics, and so the MWE is built normally but subsequently recognised as special. In this section we consider the former; the next section addresses the latter.

Periphrasis involves the realisation of particular collections of morphosyntactic features as multiple words, when other combinations of those same features are realised as single words. For instance, in Latin the combination of features PERFECT and PASSIVE is realised periphrastically, even though the perfect active and the imperfective passive are both realised synthetically; Table 3.2 shows the first person singular forms for the various cells (Bonami 2015: 64):

These parts are both separate morphological objects (*monitus* agrees with the subject, and the form of the verb ‘to be’ which accompanies it inflects for person, number, tense, etc.) and separate syntactic atoms (they can be separated by intervening material), but they are the realisation of a single cell of the lexeme’s paradigm, i.e. they correspond to a single listeme. We need a theory which can allow for the morphology to produce multiple syntactic atoms which correspond to a single lexeme. Ackerman et al. (2011), Dalrymple (2015b), and Biswas (2017)

		MONEŌ active 'advise'	MONEŌ passive 'be advised'
[−PERFECT]	PRESENT	moneō	moneor
	PAST	monēbam	monēbar
	FUTURE	monēbo	monēbor
[+PERFECT]	PRESENT	monuī	monitus sum
	PAST	monueram	monitus eram
	FUTURE	monuerō	monitus erō

Table 3.2: 1st person singular forms of MONEO 'advise' in Latin

propose (variants of) such a theory, whereby single lexemes can, when their paradigms contain periphrastic realisations, generate multiple syntactic atoms. I will not go into detail on the implementation of these theories of the morphology-to-syntax mapping, since periphrasis is not the focus of this thesis, but such theories do raise an interesting possibility: perhaps this technique could be applied not just to periphrasis but to MWEs more generally.

Unfortunately I do not believe this is the case. The reason is that this approach recapitulates the collocational challenge discussed in Section 3.3.2. Because there are multiple syntactic atoms, we need to be able to describe the relationship between them so as to constrain them to appear in the correct configurations. For periphrasis, this is relatively straightforward, since the relationships between the parts are generally fairly constrained, and usually quite local. As we saw above, this is not the case for idioms. Without a satisfactory response to the collocational challenge, it doesn't matter whether the syntactic atoms come from one listeme or several: they still cannot be appropriately constrained, and thus the theory fails to account for the facts.

3.5 Semantic mapping

There is a quite different kind of theory which also maintains that MWEs are listed once, but contain multiple syntactic atoms and morphological objects. This

is what I will call the SEMANTIC MAPPING approach.¹⁹ In this view, MWEs are simply conventions about meaning: in English, for example, there is a convention that when we speak about kicking (contextually salient) buckets, we are (or can be) actually speaking about dying. Because these are conventions about semantics, we can leave other parts of the grammar, e.g. the lexicon and syntax, untouched. This has a number of advantages, as we will see below, but such total isolation from the grammar also brings with it its own problems.

One classic representative of such an approach is the QUASI-INFERENCE theory of Pulman (1993). In this theory, knowledge of an idiom is possession of a quasi-inference rule taking us from literal to idiomatic meanings. This is a post-semantic rule of the form $X \approx Y$ which says ‘if you have a semantic expression X , you may infer instead a semantic expression Y ’. We might understand this as a kind of rewriting rule on semantic expressions (see discussion in Lichte & Kallmeyer 2016: 134–137). Unlike a regular inference rule, $X \Rightarrow Y$, the left-hand side of the rule does not have to be true in order to license the right-hand side (indeed, in the normal state of affairs, it will not be—generally no bucket need have been struck by a foot for *Charlie kicked the bucket* to be true)—this is why Pulman calls them *quasi*-inference rules.

As a concrete example (taken from Pulman 1993: 262f.), to know the meaning of *the cat is out of the bag* is to have the following quasi-inference rule in one’s grammar:²⁰

¹⁹Sailer (To appear) uses the term “mapping approach”, *tout court*, but since what we are mapping is meanings (to other meanings), I think it is useful to make this explicit.

²⁰Pulman gives the rules with universal quantification on the left-hand side and existential on the right, which emphasises their inference-like properties: the universal quantification presumably scopes over the whole rule, not merely the left-hand side, illustrated schematically in (i):

$$(i) \quad \forall x[\dots \approx \exists y[\dots]]$$

One might replace the universal quantifier with an existential one to emphasise the ‘rewriting rule’ interpretation, as in (ii):

$$(ii) \quad [\exists x[\dots]] \approx [\exists y[\dots]]$$

However, we will see in Section 3.5.1 one reason to prefer the interpretation in (i) to that in (ii), since it helps to keep track of internal modification.

$$(21) \quad \forall x, y. \mathbf{cat}(x) \wedge \mathbf{bag}(y) \wedge \mathbf{out-of}(x, y) \approx \exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z)$$

That is, if your semantics gives you some meaning where a cat is out of a bag, then you can instead infer that a secret has been revealed.

To determine the meaning of *John let the cat out of the bag*, the usual rules of semantic composition apply, giving the literal meaning of the sentence (involving releasing felines from sacks). Then, the rule in (21) is triggered, and, along with other general meaning postulates, we can follow the line of reasoning in (22) to arrive at the idiomatic interpretation (Pulman 1993: 262–263):

(22) From:

John let the cat out of the bag

let(john, ($\exists c, b. \mathbf{cat}(c) \wedge \mathbf{bag}(b) \wedge \mathbf{out-of}(c, b)$))

via:

$\exists c, b. \mathbf{cat}(c) \wedge \mathbf{bag}(b) \wedge \mathbf{out-of}(c, b)$

and:

$\forall x, y. \mathbf{cat}(x) \wedge \mathbf{bag}(y) \wedge \mathbf{out-of}(x, y) \approx \exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z)$

to:

let(john, ($\exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z)$))

Egan (2008) presents a very similar theory, albeit couched in less formal and more philosophical terms. In his account, idioms are said to be interpreted via a *PRETENCE*, essentially a conventional fiction that speakers of a language share. For instance, the pretence behind *kick the bucket* is the following (Egan 2008: 387):

(23) If somebody dies, pretend that there's some salient bucket that they kicked.

Now, when we hear that someone kicked the bucket and we realise that the pretence is in force, we can infer what would have to be the case in the real world for it to be true in the pretence that someone kicked a bucket; namely, that someone died.

For the simple cases, Egan's (2008) approach is largely equivalent to Pulman's (1993): a pretence is like a kind of quasi-inference rule.²¹ Once again, it is based on meaning, not form (although Egan 2008: 398ff. claims that the form of words used is important for recognising that the pretence is in force). Where the pretence-based account differs, and where it comes into its own, is with the extended uses of idioms, as discussed in Chapter 2.2.9. Recall that these are non-canonical uses of idioms such as the following (Egan 2008: 395):

- (24) Livia didn't quite kick the bucket, but she took a good strong swing at it.

Here the second clause 'extends' the idiom, by picking up on the metaphor, or pretence, which underlies it. We will discuss such extended cases in more detail below.

Although neither theory spells out how the underlying mechanisms are supposed to work in any great formal detail, Pulman's certainly has the advantage of being couched in terms of first-order predicate logic, which is formally explicit. For the next two sections, therefore, I take the quasi-inference theory as the representative of this family of approaches, returning to the pretence theory in Section 3.5.3. In these two sections, I discuss the strengths and weaknesses of the purely semantic approach for handling what I am calling the 'core' uses of idioms. In Section 3.5.3, I briefly discuss the unique role the semantic mapping theory has in accounting for the 'extended' uses.

²¹Although, as Sascha Bargmann (pers. comm.) has pointed out to me, the two approaches do differ in the ordering they assume between the literal and the figurative meanings of idioms: the quasi-inference theory is presented as mapping the literal meaning to the figurative one ('If you understand that someone kicked the bucket, infer that they died'), while the pretence theory is presented in the other direction ('If someone dies, pretend they kick the bucket'). Of course, both must ultimately work in either direction, to account for production as well as comprehension, but it is nonetheless interesting to note that the theories at least differ in what they take as their point of departure (the literal or the figurative meaning).

3.5.1 Strengths of the semantic mapping approach

3.5.1.1 Semantic idiomaticity

The semantic mapping approach obviously handles semantic idiomaticity very well, since this is precisely what it is designed to deal with. The mapping procedures—whether quasi-inference rules, pretences, or something else—are arbitrarily powerful: any kind of literal meaning can be inspected and replaced by any kind of idiomatic meaning, and so no kind of semantic idiomaticity poses a problem.

3.5.1.2 Flexibility

Much in the same way as the lexical ambiguity approach, the semantic approach allows morphosyntactic and syntactic flexibility readily: this is because both approaches insist that MWEs are made up of multiple syntactic atoms and therefore constructed by the usual kinds of syntactic rules. Since the words involved are normal words, and the syntactic structures normal structures, they will behave in exactly the same way as they do outside of MWEs, and so we expect the full gamut of morphosyntactic variation and syntactic manipulability. However, by relegating the MWE component to the semantics, the current approach has nothing to say about any *restrictions* on syntactic flexibility, a weakness which we will discuss shortly.

3.5.1.3 The real words problem

The semantic mapping approach, unlike the other theories we have considered so far, offers a particularly neat solution to the real words problem, since the words in MWEs are precisely the normal words found outside the expressions. In this approach, MWEs are assembled entirely normally from the regular vocabulary of a language; it is only afterwards that the idiomatic semantics emerges. For this reason, MWEs *must* be made up of words which have a life outside the MWE, otherwise there will be no literal meaning to act as input to the quasi-inference rules. This limits the scope of such a theory, since, as we

have noted, there *are* MWEs which include words which do not exist outside of them. We will have more to say about this in Section 3.5.2.

3.5.1.4 Internal modification

The modification of subparts of composable MWEs is correctly predicted to be possible by the purely semantic approach, since these are represented as separate parts of the overall meaning. Once they have been translated into their idiomatic meaning, either via a pretence or a quasi-inference rule, they are available for modification just like anything else. Above and beyond this, the pretence theory purports to offer an explanation for more complex kinds of modification, which we will have more to say about in Section 3.5.3.

Unfortunately, as it is presented in Pulman (1993), the quasi-inference approach to internal modification is deficient. It is true, as Pulman points out, that quasi-inference rules will apply equally well with additional premises; i.e., if $X \approx Y$, then $X \wedge A \approx Y \wedge A$. However, because there is no connection between literal discourse referents and their idiomatic equivalents articulated in the quasi-inference rules, there is no way to get the meanings of internal modifications right. For example, as it stands, in an inference rule like (21) (repeated below as (25)), there is no connection between the variables x and z , and thus no way to match up a modifier of *cat* with its idiomatic meaning *secret*, as in *let the financial cat out of the bag*.

$$(25) \quad \forall x, y. \mathbf{cat}(x) \wedge \mathbf{bag}(y) \wedge \mathbf{out-of}(x, y) \approx \exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z)$$

This is easily remedied, however: either part of the definition of the quasi-inference connective \approx is that it replaces *all* instances of the variables on the left-hand side with the appropriate new variable (thus capturing arguments of predicates which are outside of the quasi-inference rule), or we simply add the requisite equalities to the right-hand side, e.g.:²²

²²For this second option to work, it is of course crucial that the universal quantifier scopes over the whole rule, not merely the left-hand side.

$$(26) \quad \forall x, y. \mathbf{cat}(x) \wedge \mathbf{bag}(y) \wedge \mathbf{out-of}(x, y) \approx \\ \exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z) \wedge x = z$$

Now whatever is predicated of the cat, x , e.g. $\mathbf{financial}(x)$, will also be predicated of the secret, z , since it has been made explicit that they refer to one and the same thing.

This does not necessarily solve all of the translational problems, however. Notice that in example (22), the final representation is (27):

$$(27) \quad \mathbf{let}(\mathbf{john}, (\exists a, z. \mathbf{secret}(z) \wedge \mathbf{revealed}(a, z)))$$

But this isn't quite right: (27) says that John let someone reveal a secret, but in *John let the cat out of the bag*, it must be John who revealed the secret, not someone else (i.e. $a = \mathbf{john}$). The existentially closed agent argument is appropriate for instances like (28), where no agent is specified, but not where the agent is made explicit in the sentence.

$$(28) \quad \text{Now the cat is out of the bag.}$$

This points to one of the problems with the semantic approach: there is no connection to the syntax, and so any relation between the literal meaning and idiomatic meaning that cannot be expressed semantically cannot be captured. I talk more about the problems raised by this total isolation of modules in Section 3.5.2, below.

3.5.1.5 Lexical flexibility

Some idioms allow for lexical variation in one or more of their parts, even in their canonical uses:

- (29) a. Put/lay/spread your cards on the table.
b. Give someone a kick up the backside/arise/bum/behind.

These idioms allow almost any semantically appropriate lexeme to occupy certain slots—any verb that involves transfer onto a surface, any noun that

means *derrière*, etc. This looks like semantic, rather than lexical, selection, and so the semantic approach offers a much neater solution than other theories. In any theory where the form of an idiom is itself stored in the lexicon, such lexical flexibility must simply be stipulated, and all the options listed. But this is inelegant, and also potentially empirically inadequate. For instance, if the idiom selects for a particular meaning, we would expect new coinages which fit that meaning to be acceptable, whereas if the lexical alternations are merely listed, this would not necessarily follow—some synonyms might be acceptable and others not. In fact, for the idiom in (29b), there seems to be no synonym of *backside* which isn't used, including relatively recent coinages like *booty* and *tushy*:²³

- (30) a. Even instructors sometimes need a kick up the booty to reassess, re-motivate and make things fresh again. (WWW)
- b. That provided the kick up the bootie I needed and I found an extra gear. (WWW)
- (31) a. Introducing the emo-tinged Aussies giving indie rock a feminist kick up the tushie. (WWW)
- b. Thank you so much for being the kick up the tushy that I needed. (WWW)

If this is the case, then the semantic approach offers a neat solution: idioms are recognised by the meanings they express, not by the form they take. So as long as the sentence expresses a meaning that matches a quasi-inference rule or describes a situation encoded in a pretence, it doesn't matter which words are used to express this. Indeed, sometimes idioms can be extremely lexically deformed and still succeed in coming off:

- (32) a. But the feline escaped from the sack, or some one anticipated the intentions of the company and its friends and the result has been evident on the market during several weeks past. (WWW)

²³Earliest OED citations 1959 and 1962 respectively.

- b. Then the manure really entered the ventilation system.

(Pulman 1993: 266)

However, although examples like (32) are possible, they seem to require some additional cognitive effort to recognise, and have a distinctly affected feel to them (especially euphemistic examples like (32b)). They are therefore perhaps best seen as outside of the ‘core’ use of idioms. Unfortunately, however, the semantic approach predicts that they should be the norm: after all, if meaning is all that matters, the choice of words used to encode it should be quite free. Something needs to be said about why such examples are exceptional, therefore.

Even more problematic is that some idioms are not at all lexically flexible, with even a slight deviation from the canonical form resulting in a loss of idiomaticity:

- (33) a. Antos kicked the bucket.
 b. #Antos booted the bucket.
 c. #Antos kicked the pail.

It seems, therefore, that the semantic mapping approach radically overgenerates, and needs to be constrained in some way—this is what Pulman identifies as the “problem of false positives” (Pulman 1993: 265ff.).

Pulman (1993: 267–268) proposes the following mechanism to achieve this: we associate idiom mapping rules with sets of lexemes which must be present before the rule can be triggered; only the fixed words are included, and the more flexible ones are omitted. For the lexically flexible idioms in (29), therefore, these sets might be the following:

- (34) a. {cards, on, table}
 b. {kick, up}

Each set is associated with the relevant rule, and then the presence of the words in the set causes the rule to be triggered. Of course, the semantics of the left-hand side of the rule still need to be matched (the appropriate meaning

needs to be present), and so even though the *kick up the backside* rule will be triggered by sentences like (35), which contain all the right lexemes, it will fail to be implemented, since its left-hand side will still make reference to backsides and kicks in the appropriate relation.

(35) Max gave the football a mighty kick up onto the wall.

This solution allows a neat characterisation of the degree of lexical flexibility which each idiom permits. For the most highly distortable idioms, for example, the lexical index might be the empty set: their rules would then be triggered in all contexts, and would fire provided the right meaning can be inferred, regardless of which words are used to express it. More fixed idioms like *kick the bucket* would, presumably, have all of their lexemes represented in the sets associated with their rules, thus ensuring that they appear in the correct form.²⁴

3.5.1.6 Summary

The semantic mapping approach can account for the semantic idiomaticity of MWEs, and, like the lexical ambiguity approach, predicts the existence of morphological and syntactic flexibility, since the words involved are just normal words. For the first time, we also have an explanation for the real words problem: since idiomatic meanings are generated from literal ones, we must be able to generate the literal ones in the first place, and this requires that real words are used. Equally, by locating the encoding of MWEs in the semantics rather than in the lexicon *per se*, we are able to give a neater account of the lexical flexibility of some expressions: since what matters is the meaning, it is inconsequential

²⁴Egan (2008: 398ff.) suggests that deviation from the canonical form of an idiom, by which he means both lexical and syntactic variation, is infelicitous insofar as it is “gratuitous”: deviation for some communicative purpose is acceptable, but otherwise there is pressure not to interpret a deformed utterance idiomatically (or else why would a cooperative speaker not have used the canonical form?). However, this doesn’t really explain the fact that different idioms apparently have different degrees of tolerance for lexical variation—this seems to be a conventional or encoded property, not one that follows from more general principles. After all, Egan concedes (2008: 400) that speakers don’t just deviate to be communicative, but also to be funny, original, witty, etc.; but why should it be funnier/more original/wittier/... to replace *backside* with *rear*, as in (29b), than to replace *bucket* with *pail*, as in (33c)?

what words are used to produce this meaning. Of course, MWEs differ with respect to how lexically permissive they are, and this can be captured by adding *some* lexical specification to the mapping rules.

The semantic mapping approach appears therefore to fare very well. However, this is only true for a certain subclass of MWEs: namely, those which are ambiguous between a literal and an idiomatic interpretation. Once again, we focus on the regularity of MWEs over and above their idiosyncrasies. Any kind of idiosyncratic constraints on even these regular, ambiguous idioms is impossible to capture. And, more generally, the fact that we cannot state constraints on other aspects of MWEs besides meaning makes describing lexical or syntactic idiomaticity highly problematic.

3.5.2 Weaknesses of the semantic mapping approach: separation from syntax and the lexicon

The total dissociation of the grammatical and semantic aspects of MWEs is a hallmark of the semantic mapping approach, and allows it to readily explain lexical flexibility and give an account of the real words problem. But this separation is also a weakness, since there are times when we *do* seem to need to make statements about other aspects of the grammar when describing MWEs.

3.5.2.1 Lexical and syntactic idiomaticity

Firstly, as mentioned above, other kinds of idiomaticity are a challenge. Since the semantic mapping approach operates purely on meanings, it requires that the grammar produce a semantic interpretation for every sentence containing a MWE, which of course also requires there to be a valid syntactic parse. But, as we have seen, there are lexically or syntactically idiomatic MWEs which will not, under normal circumstances, receive any semantic interpretation. The grammar cannot give a literal interpretation to *by and large* or *come a cropper*, for instance, since it has no way of interpreting the conjunction of a preposition and an adjective, and it lacks a lexical entry for *cropper*. If there is no semantic

interpretation, there is nothing for the semantic mapping approach to operate on, and it has nothing to say about the expressions.

Of course, we can explicitly force the grammar to give an interpretation to such expressions, expanding the meaning of *and* so that it can combine these specific words, or inventing an arbitrary meaning to give to *cropper* purely so it can trigger the left-hand side of a quasi-inference rule, but these are not very satisfying solutions, and undermine the advantage of the semantic approach, which is precisely that it does *not* require *ad hoc* extensions to the syntactic or semantic combinatorial systems. In other words, while the semantic mapping approach might be able to account for the behaviour of a certain class of idioms, it cannot serve as a general theory of MWEs.

3.5.2.2 Constraints on syntactic flexibility

Secondly, even idioms which are syntactically regular can nonetheless be idiosyncratically restricted in what syntactic arrangements they can be found in. *Kick the bucket*, for example, does not passivise in its idiomatic interpretation, but this is not because the verb *kick* is defective in any way—the passive is perfectly acceptable with a literal interpretation:

- (36) a. Old Man Mose kicked the bucket.
 b. #The bucket was kicked (by Old Man Mose).

- (37) a. Dukat kicked the ball.
 b. The ball was kicked (by Dukat).

Similarly, some idioms can be split up by relativisation or topicalisation more felicitously than others:

- (38) a. To the best of our knowledge [...], the daughter was not aware of the strings that were being pulled for her [by her father]. (iWeb)
 b. #The daughter was not aware of the beans that were being spilled about her by her father.

- (39) a. Those strings, he wouldn't pull for you. (Nunberg et al. 1994: 501)
 b. #Those beans, he wouldn't spill. (Maher 2013: 10)

As it stands, the semantic approach has no way of enforcing any syntactic constraints, since it makes no mention of syntax, and so it predicts that all verbs should be as (in)flexible in their idiomatic usage as in their literal usage. Since this does not appear to be the case, this is a problem for the semantic approach, unless some other account of the differences can be found.

Pulman (1993: 268) suggests one such explanation: that the unacceptability of idiomatic #*The bucket was kicked* is due to the fact that *the bucket* does not refer to any discourse entity, and so it is incompatible with the information-structural effect of being made a passive subject (*viz.* being a topic). But we have already seen that this cannot be the whole story, since passive sentences with non-thematic subjects are perfectly acceptable:²⁵

- (40) It is considered polite to remove your shoes when inside. (Arnold 2015)
- (41) a. There was believed to be another worker at the site besides the neighbors who witnessed the incident.
 b. It was rumored that Great Britain, in apparent violation of the terms of the Clayton-Bulwer treaty, had taken possession of certain islands in the Bay of Honduras. (Kay et al. 2015)

What is more, this cannot be the explanation for the contrast in (38) or (39), where both idioms are decomposable, and so *beans* and *strings* both refer to actual discourse entities. It really does seem like we need to be able to state

²⁵One might object that some of these examples are not 'real' passives, since they are not related to active forms (e.g. **We rumoured it that...* is ungrammatical). Two things can be said in response to this. Firstly, some *do* correspond to active forms, even if the alternating pronoun remains pleonastic in the active: *We considered it polite...*, *We believed there to be...*, etc. Secondly, even if none of the examples did correspond in this way, the point made in the text would still stand: *formally* speaking, the passive construction is not incompatible with a non-thematic subject. If the argument is that passive *The bucket was kicked* is blocked by a structure against non-thematic passive subjects, and that the putative counter-examples given to this claim are not 'real' passives, one could rightly ask why *The bucket was kicked* couldn't just be one of these 'fake' passives instead.

(arbitrary?) constraints on the syntactic flexibility of MWEs, and this is beyond the scope of the semantic mapping approach.²⁶

3.5.2.3 Psycholinguistic evidence

Finally, any theory which posits some kind of post-syntactic step to MWE interpretation, as the semantic mapping approach does, runs counter to the psycholinguistic claims that idiom processing is faster than literal interpretation. If idiom comprehension involves literal composition and then an *additional* step on top of this, we would expect the process to be slower/more costly. The fact that we recognise idioms as well formed faster than we do free combinations is evidence for some kind of direct (i.e. non-mediated) access, even if this is just due to familiarity effects (Tabossi et al. 2009).

3.5.2.4 Summary

It is clear that the scope of the semantic mapping approach must be narrower than we might like. It cannot explain syntactically idiosyncratic expressions like *trip the light fantastic*, and it offers no insights into lexically idiosyncratic expressions like *leave X in the lurch*, in both cases because it requires that the expressions it operates on have literal interpretations. Unfortunately, though, even for those expressions which do, like *spill the beans*, *pull strings*, or *kick the bucket*, it cannot explain the fact that they differ with respect to what kinds of syntactic construction they can appear in (regardless of whether this is a categorical phenomenon or merely a preference). Finally, it makes entirely the wrong psycholinguistic predictions, since it claims that idiomatic interpretation is *more* complicated than literal interpretation, when the experimental findings imply the reverse.

²⁶I have no doubt that there are patterns to be found here, and quite probably a historical explanation for the different degrees to which different MWEs adhere to their 'canonical' forms, i.e. their degree of frozenness. But this historical development is not part of our synchronic knowledge; what we need is some way of representing its consequences, which may take on any number of different forms. It is this which is lacking in the semantic approach.

However, all is not doom and gloom for the mapping approach. Although it cannot be the whole answer when it comes to core uses of idioms, and MWEs more generally, it nevertheless has a crucial role to play in the analysis of so-called extended uses, to which we now turn.

3.5.3 Extended uses of idioms

In Chapter 2.2.9, I introduced the distinction between core and extended uses of idioms. As mentioned there, this thesis is about the core uses, and so this has been the focus of the review in this chapter. On these terms, the semantic mapping approach is deficient. However, it is worth mentioning that it has great potential when it comes to explaining the extended uses, since these seem to rely precisely on some mapping between literal and idiomatic meaning that is the hallmark of the semantic mapping approach.

In recent work, colleagues and I have been developing an approach to figurative extensions which makes use of the intuitions behind Egan's (2008) pretence-based theory (Findlay 2018; Findlay et al. 2018, 2019, in preparation). Since the formal details are not finalised, I will not present them here. I will, however, give a brief account of the underlying ideas, so as to indicate the potential of the pretence-based theory in accounting for extended uses of idioms.

3.5.3.1 Inferred idioms

Pulman (1993: 266) introduces the term *INFERRED IDIOMS* for those instances where idioms are not mentioned explicitly, but merely alluded to. Sometimes this can be via extreme lexical distortion, as in (32b) (repeated below as (42a)), but sometimes it goes further than this, as in (42b)–(42d):

(42) a. Then the manure really entered the ventilation system.

(Pulman 1993: 266)

b. Shit finally completes 29-month journey towards fan. (WWW)

c. Good gawd its another porcine flyer. (WWW)

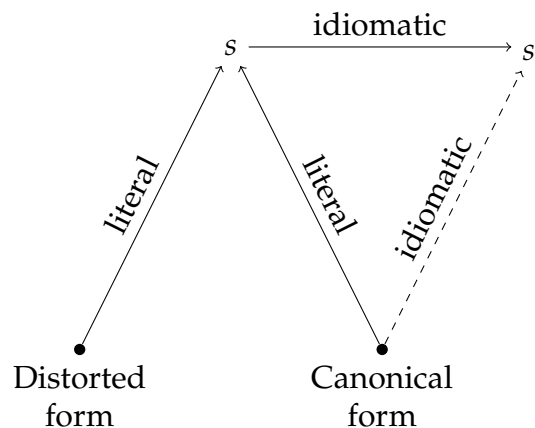


Figure 3.1: Understanding “inferred” idioms

- d. Bandwagons are interesting vehicles to be on and just as interesting to watch drive past if only to see who is on board. (WWW)

These seem to be instances where the literal meaning of the idiom is more important than its form. The canonical form of certain idioms can be taken to describe a literal situation (someone really striking a bucket with the foot; shit actually hitting a fan). By referring to the same situation with different words, it is possible to indirectly access the idiom. For instance, *porcine flyers* are flying pigs; in the idiom *when pigs fly/pigs might fly*, the literal situation described includes pigs which are flying. Thus, there are situations described by both the canonical idiom and example (42c); this link is apparently enough to get us to the idiomatic interpretation.

But this requires reference to meanings, and, ultimately, to situations, not just forms of words. The semantic mapping approach is thus clearly implicated here. Figure 3.1 shows schematically how this might work. Both the canonical and distorted forms of the idiom describe the same literal situation, s . The idiom mapping $s \rightarrow s'$ is what is given by the semantic mapping approach; it is a mapping from situations to situations, not between linguistic expressions. In the semantic mapping approach, the link from the canonical form to the idiomatic situation s' is not present; instead, all idiomatic interpretation goes via the literal situation, s . We have seen reasons to think this is not right, and

that idioms should directly encode their idiomatic meanings. But assuming that the mapping $s \rightarrow s'$ must still be present—to account for the inferred idioms—this leaves us with two paths for the interpretation of (non-inferred) idioms to follow: directly from the idiom straight to s' , or indirectly, via s . We contend that this is exactly right, and that the two paths correspond to the difference between core and extended uses of idioms.

3.5.3.2 Idiom extensions

Idiom extensions are where we have the canonical form of the idiom (potentially modified) immediately followed by a continuation which is based on the metaphor underlying it. Recall some examples (repeated from Chapter 2.2.9):

- (43) a. The cat is out of the bag, claws fully exposed. (iWeb)
 b. I had butterflies in my stomach and they had big wings on them. (WWW)
 c. This month, Meriden City Council's PLA chickens came home to roost, and they laid a big stinky egg on Meriden taxpayers. (WWW)
 d. Sometimes [...] the person that you'd take a bullet for is behind the trigger. (song lyrics)

It is unclear how any syntactically- or lexically-based theory of MWEs could account for these data, since metaphor is fundamentally a relation between meanings, not between forms. What is more, these kinds of grammatical theories assume that idiomatic expressions are simply ambiguous between a literal and idiomatic interpretation, but here, both seem to be relevant simultaneously: the extension relies on the literal meaning, but the overall interpretation is idiomatic. Under a purely semantic theory, however, both kinds of meaning are potentially available: the sentence is parsed literally, and this is then interpreted via the mapping mechanism.

It seems, then, that such figurative extensions of idioms require something like the semantic mapping approach, even if it remains inadequate for the core

uses. In fact, the psycholinguistic findings on figurative distortions of idioms (e.g. *shatter the ice* as an intensified version of *break the ice*) also support this: these kinds of expressions lose the processing advantage of regular idiom usage (McGlone et al. 1994). It is plausible that the idiom extensions exemplified in (43) share this processing cost, although I am not aware of any experimental work which confirms this.

Findlay (2018) and Findlay et al. (2018) offer the beginnings of a formalisation of this idea, grounded in situation semantics. We propose a hybrid approach to idiom interpretation: on the one hand, idiom meanings are directly encoded, accounting for the psycholinguistic findings and the possibility of explaining lexical or syntactic idiosyncrasies; on the other hand, however, idioms also encode the mapping between the literal and idiomatic situations they describe. This is achieved by associating the idiom with a PRETENCE STATEMENT, which is just a formula that describes the metaphorical mapping between the literal and idiomatic situations. For instance, the pretence statement for *kick the bucket* is as follows:

$$(44) \quad \lambda s \lambda s'. \forall x [\mathbf{kick}(s', x(s'), \iota y [\mathbf{bucket}(s', y(s'))]) \leftrightarrow \mathbf{die}(s, x(s))]$$

The denotation of this pretence statement is a set of pairs of situations, in one of which someone kicks a bucket, and in the other of which they die. (All predicates and individual variables take an additional situation argument, which relativises them to that situation.)

By recasting some ideas from the literature on analogical reasoning (Gentner 1983; Gentner & Maravilla 2018) in terms of model-theoretic semantics, it is possible to capture the inferences that the figurative extensions call for with some degree of explicitness. We can see analogical extensions as relations holding between pairs of situations, and can constrain these situations so that the members of each pair must stand in the appropriate pretence relation, as given in the idiom's pretence statement.

This is not the place for a full exposition of the theory, and our discussion has already taken us quite far from the main focus of this chapter. But this section has hopefully achieved two things: firstly, to illustrate how, even if the semantic mapping approach is not appropriate as a general theory of the core uses of idioms, it may have an important role in explaining the extended uses; and secondly, to suggest that such extended uses are amenable to formal linguistic analysis, and need not be relegated to some metalinguistic rubbish heap. Of course, I have not provided a full account of this second point, but I have indicated the direction such an account could take, and hope that future work will illuminate this point more explicitly (for instance, Findlay et al. in preparation).

3.5.4 **Summary**

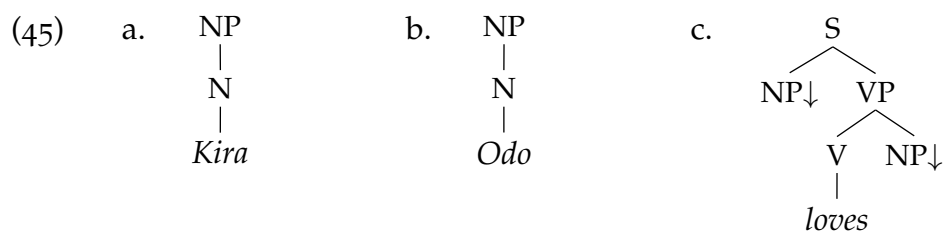
The semantic mapping approach is unparalleled when it comes to explaining the playful and poetic manipulations of idioms which speakers use. It also offers an explanation of the real words problem, and has a potentially neat story to tell about lexical flexibility. However, its scope is limited to just those expressions which are ambiguous between a literal and an idiomatic interpretation—in this respect its response to the real words problem is too extreme. And even within that class, it is unable to account for differing degrees of syntactic flexibility. It has nothing to say about other kinds of MWEs except idioms, of course, since these will very often not have a literal interpretation, or have one which is not in any sense metaphorical. It may well be that the semantic approach represents one way of analysing a subclass of idioms, but one which is dispreferred, since the idiomatic meaning is also directly encoded. The figurative extensions exploit this underused path, and that is what lends them their affected, playful feel.

3.6 **Construction-based theories**

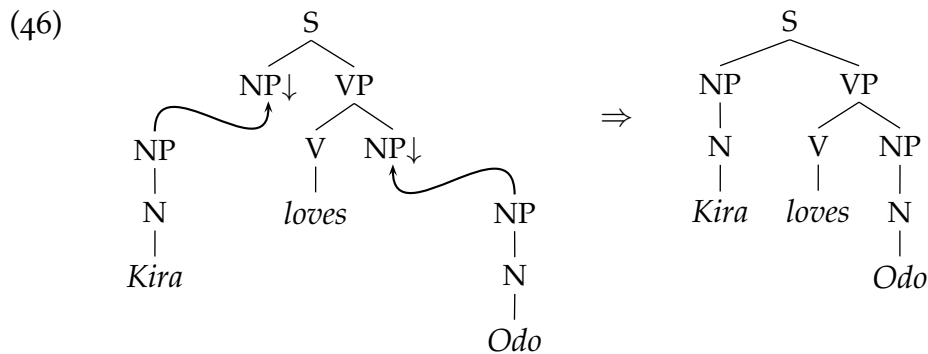
It seems clear to me that it is inadequate to come down entirely on one side of the tension between a MWE's lexical and phrasal statuses: treating them as 'words

with spaces' as in the monolexemic approach is doomed to failure because of the relative flexibility of MWEs, while treating them as normal phrases fails to explain their cohesion, the lack of one-to-one mapping between syntactic parts and semantic parts in the case of non-decomposable idioms, and the various idiosyncrasies of form which suggest MWEs cannot (always) be constructed using the normal rules of the grammar. Ideally, our analysis should allow us to model both aspects at the same time, thus representing the tension rather than attempting to sidestep it. The CONSTRUCTION-BASED approach offers just such a possibility: like in the monolexemic approach, the MWE is stored as a unit; unlike in that approach, it also brings with it some internal structure. Abeillé (1988, 1995) and Jackendoff (1997) are recent proponents of such a theory, although it has its roots in the earlier generative work of Fraser (1970).

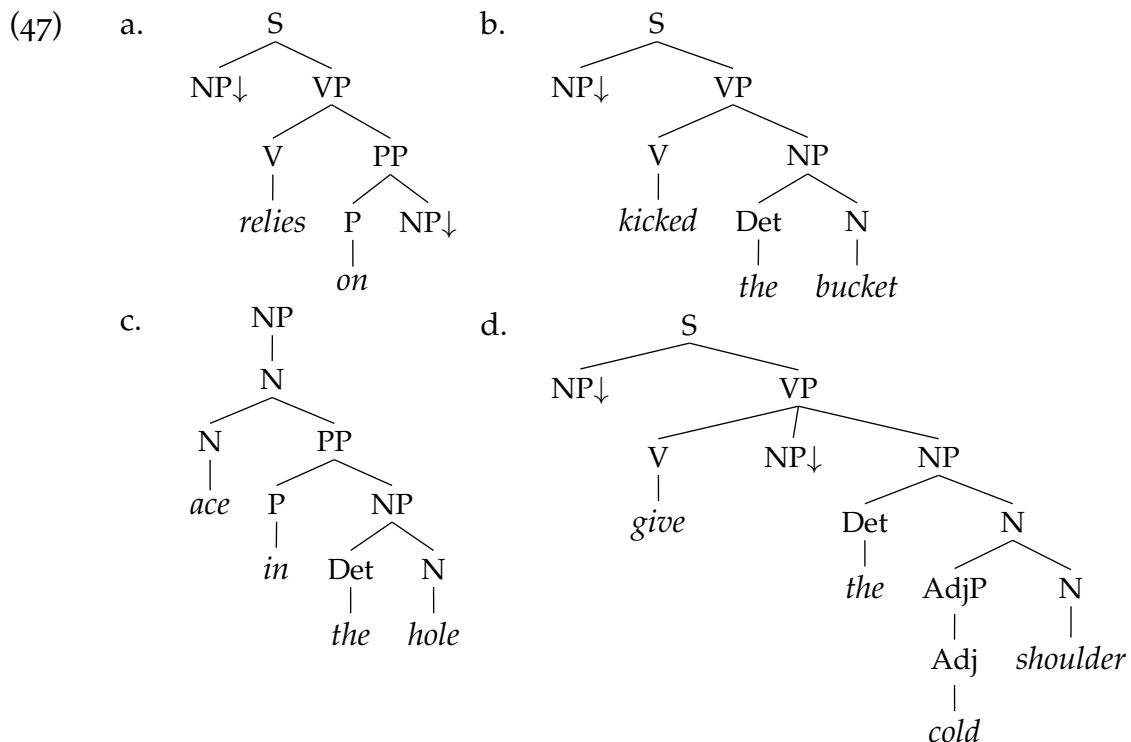
To illustrate the general approach in this section, we will make use of a pared-down version of Tree-Adjoining Grammar (TAG: Joshi et al. 1975; Abeillé & Rambow 2000b); Chapter 4.2 gives a fuller exposition of the formalism. In TAG, all lexical items are stored as trees (called ELEMENTARY TREES). Some examples are given in (45):



Verbal entries, like (45c), contain slots for their arguments: in this case, *kicked* is represented by the tree corresponding to the whole clause, with an open NP slot for its subject and for its object. A simple sentence is produced by inserting, or SUBSTITUTING, trees into these argument slots, or SUBSTITUTION SITES, which are indicated by a down arrow (↓). (46) shows such a derivation:



The terminal node labelled *Kira*, *Odo*, or *kicked* in each of (45a)–(45c), i.e. the morphological form which the lexical entry corresponds to, is called the **ANCHOR** of that tree. Crucially, there is nothing stopping us having *multiply*-anchored trees, and so having multiple morphological objects associated with a single lexical entry. This gives us a particularly neat way of representing MWEs directly. (47) shows some potential lexical entries for MWEs:



Each of these objects represents a single listed entry (which can therefore be associated with an idiosyncratic semantics), and a single, albeit complex, syntactic unit; but each also contains multiple terminal nodes, i.e. multiple morphological words. In this way, the two sides of the MWE tension are

addressed simultaneously: MWEs are phrasal, in that they correspond to complex syntactic objects and contain multiple morphological words, but they are also word-like, in that they are listed, and single words and MWEs are both stored as the same kind of formal object, namely as syntactic trees. It would seem that this approach has much to recommend it, therefore.

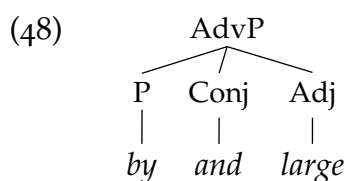
3.6.1 Strengths of the construction-based approach

In fact, this approach fares very well when compared against our criteria. In common with all the other approaches, it succeeds in handling semantic idiomaticity. Since the expressions are each stored as a whole, this means we can associate with them whatever semantics we need to, directly in the lexicon. Whatever means we choose for associating meanings with elementary trees, that same approach will readily transfer over to the multiply-anchored trees of MWEs.²⁷

The difference between decomposable and non-decomposable idioms can also be straightforwardly represented by differences in the mapping between syntax and semantics. An entire tree can correspond to a single meaning, representing the non-decomposable case, or the parts of the tree can correspond to the parts of a more complex meaning (in synchronous TAG this would be represented by mapping the meaning either from the root node of the tree or from sub-trees inside it). However, this has no bearing on the syntactic flexibility of the idiom, at least not without further stipulation, and so does not in itself explain the generalisation proposed by Nunberg et al. (1994) that decomposability and syntactic flexibility go hand in hand. As mentioned in Chapter 2.2.7, however, there may be reasons to question how strict this connection really is.

²⁷There are numerous proposals for how to model semantics in TAG, e.g. synchronous TAG (Shieber & Schabes 1990), unification-based semantics (Kallmeyer & Romero 2008), or GlueTAG (Frank & van Genabith 2001). See Abeillé (1995) for an account of the semantics of French idioms in terms of synchronous TAG.

Syntactic idiomaticity is no challenge for the construction-based approach, either, since we encode syntactic information directly in the lexical entries. Given that a lexical entry is a syntactic tree, it is trivial to make use of bespoke or idiosyncratic trees to account for syntactic peculiarities. For example, we might represent *by and large* as (48):

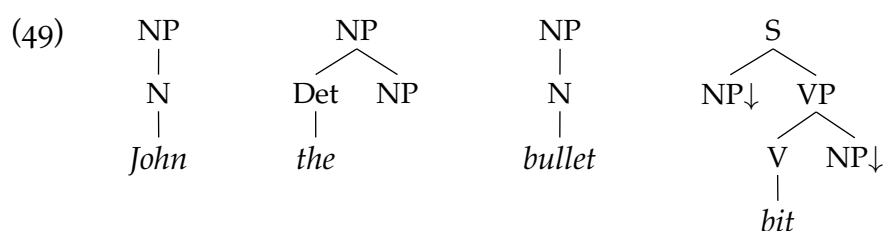


This encodes both its external and internal syntactic idiosyncrasy (its root node is labelled AdvP even though it contains no adverbs, and it has a structure not found in other lexical entries). Note that this is similar to the approach discussed above in relation to lexical ambiguity, where idiosyncratic phrase structure rules were used to account for syntactic idiomaticity. However, TAG lexical entries are already intrinsically syntactic, and so adding additional/bespoke phrasal structure to them is quite natural. By contrast, in the lexical ambiguity approach such a change cannot be stored lexically and instead necessitates a complexification of a separate, syntactic component of the grammar.

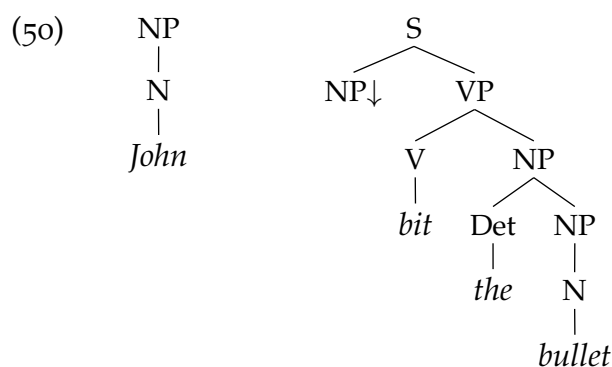
Lexical idiomaticity poses no problem, because there is no restriction forcing us to pick pre-existing words to anchor the elementary trees. However, this obviously means this approach falters when it comes to the real words problem, which I will discuss further in Section 3.6.2.

Since the construction-based approach maintains that MWEs contain multiple morphological objects, morphosyntactic variation can readily be explained, too. However one chooses to implement a theory of morphological realisation, MWEs are no different from any other lexical item. Where there is lexical idiomaticity, we must suppose that the idiosyncratic words also provide their own paradigm, but in other cases the expression can be allowed to access the pre-existing morphological lexicon, thus accounting for the regular inflectional paradigms we see (*come~came a cropper*, etc.).

The construction-based approach also has a natural explanation for the psycholinguistic findings. Since MWEs are treated as single units, this approach shares with the monolexic one the prediction that constructing a sentence with a MWE interpretation will be less computationally taxing than constructing it without, because there are simply fewer units involved. For instance, composition of *John bit the bullet* with a literal interpretation involves the following elementary trees:



To combine them all therefore requires three syntactic operations. For the idiomatic reading of *John bit the bullet*, however, we need only the following elementary trees, and then require just one operation of substitution:



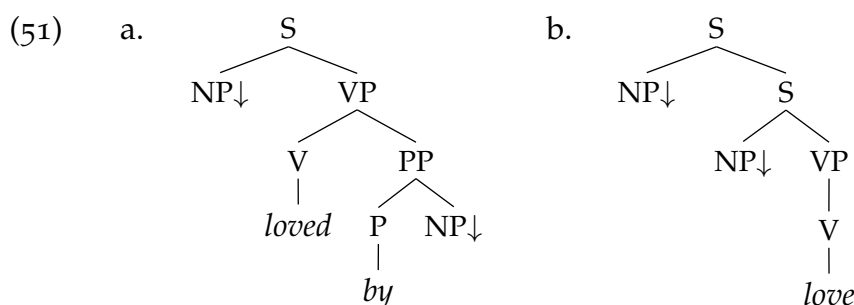
Since the idiom is stored *en bloc*, we in effect pre-combine certain words in the lexicon, and thus reduce the amount of computation that must be done online by the syntax. This predicts, correctly, it would seem, that processing the same string will be faster if it has an idiomatic interpretation than if it does not.

3.6.1.1 Syntactic flexibility

Syntactic flexibility can be handled in a number of ways by construction-based approaches, but again is in principle no challenge. In this section, I will explain how TAG can handle variable degrees of syntactic flexibility and

how it can account for the possibility of unbounded dependencies between the parts of idioms.

In TAG, lexical entries are associated with multiple different syntactic realisations; for example, *love* is not just associated with the tree in (45c), but also with trees for the use of the verb in the passive voice, or in a *wh*-question on its object, etc.:

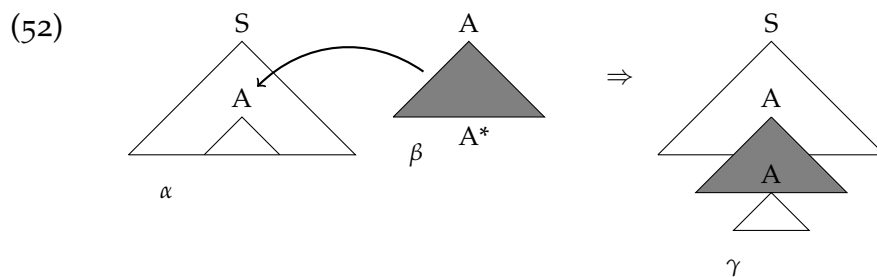


Groupings of such trees are known as TREE FAMILIES, and represent generalisations of syntactic patterns. For instance, most transitive verbs will be associated with the same tree family, since they share the same syntactic realisations, e.g. *wh*-questions and relative clauses on their two argument positions, long and short passives, cleft constructions, etc.²⁸ However, some lexical items have idiosyncracies which mean that the tree families associated with them must be restricted in some way; for example, unpassivisable verbs like *weigh* or *cost* (in the relevant senses), will not include any passive trees in their tree families. There are various ways of operationalising this, but it is often achieved by annotating the abstracted tree schemata (trees absent their lexical heads) with feature structures, and marking the relevant lexical items with features that are incompatible with features on the trees that need to be blocked (Abeillé 1988). Equipped with these tools, we face no special challenges from the variable behaviour of MWEs in this regard. MWEs can, just like any lexical items, be associated with tree families representing their syntactic flexibility, and these

²⁸These tree families are often given mnemonic names: for instance, the transitive tree family is called 'noVn1' by Abeillé (1990), or 'TnxoVnx1' in the wide-coverage LTAG grammar developed for English as part of the XTAG project (Bleam et al. 2001).

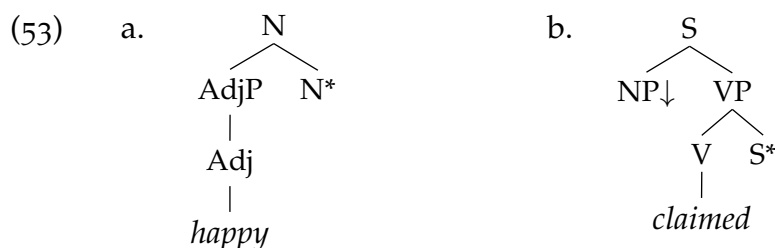
can be more or less restricted by idiosyncratic constraints. Although MWEs will in general be more degenerate than single-word entries (having more reduced tree families), this is a difference of amount, not of kind. Again, MWEs and ‘normal’ words are not qualitatively different.

The TAG approach to long-distance dependencies naturally explains the behaviour of MWEs in this respect as well. To see this, we must introduce the second combining operation which TAG employs: ADJUNCTION. As well as substituting a tree into a (non-terminal) frontier node in another tree, we can also adjoin certain kinds of trees, called AUXILIARY TREES, into *non*-frontier nodes. This allows us to expand trees from the inside out. Adjunction is shown schematically in (52) (after Abeillé & Rambow 2000a: 9):

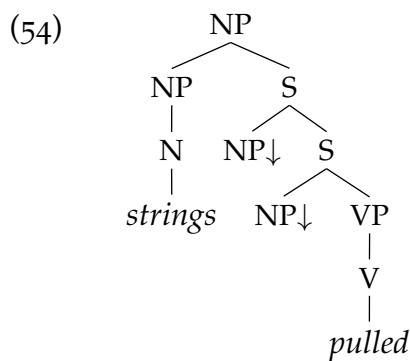


To be an auxiliary tree, a tree has to have a root node and frontier node which share a label; the frontier node is called the FOOT NODE, and is marked with an asterisk. In this case, β is an auxiliary tree, with root and foot nodes labelled A. To adjoin β into α , we remove the subtree rooted in A from α , replace it with β , and then attach the subtree which we removed to the foot node of β . This produces a larger tree, γ . In effect, the auxiliary tree is inserted at the adjunction site and ‘expands’ the node around itself.

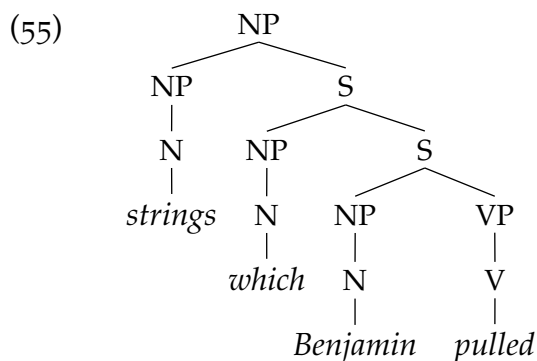
Elementary trees for modifiers like adjectives are auxiliary trees. So too are the elementary trees for sentential embedding verbs like *know*, *think*, or *claim*:



This latter fact explains how long-distance dependencies are constructed in TAG—since adjunction allows trees to grow from the inside out, *wh*-dependencies and the like can be represented locally, in the same elementary tree, and then the parts can be separated through adjunction, so that they can ultimately end up arbitrarily far apart. We discuss this in more detail in Chapter 4.2.4, but for now we can give a brief example of how this works in the MWE domain. Take *pull strings*: its tree family will contain a relative clause tree headed by *strings* and with *pull* as the verb in the dependent clause:



This tree can be used to form simple relative clauses like (55) by substitution of the appropriate NPs:



However, we can also adjoin sentence-embedding verbs at the lower S node, which expands the tree in the middle, and thereby causes the parts of the idiom to be separated; this is shown in Figure 3.2. And this can continue indefinitely (Figure 3.3 shows a further repetition), meaning the parts of the idiom can be separated by an arbitrarily large distance. Unlike in the lexical ambiguity approach, we do not have to characterise the relationship between

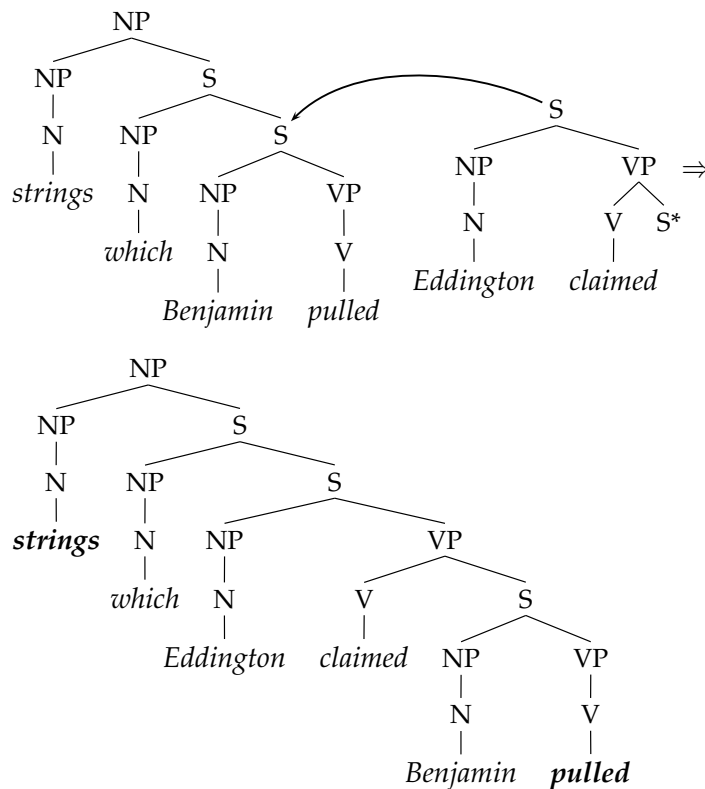


Figure 3.2: TAG derivation for *strings which Eddington claimed Benjamin pulled*

strings and *pulled* except in the local cases. Once we have specified the small set of elementary trees associated with a verbal MWE, adjunction will do the rest when it comes to explaining the possibility of long-distance dependencies existing between its parts. Again, we maintain the balance between the two sides of MWEs: because the parts must appear in the same, single elementary tree, they cannot appear separately; but because these elementary trees contain complex structure, they can be manipulated (by adjunction), and so the parts of the MWE can end up distant from one another.

This reliance on general principles makes a prediction about one kind of variability we should not expect to find, namely MWEs which appear in relative clauses or *wh*-questions, but only of some fixed depth (e.g. depth one). The important factor in determining whether MWEs participate in long-distance dependencies or not is whether they have the appropriate elementary tree in their tree family. But once they do, then adjunction can apply freely, and allow the parts to be separated to an unlimited extent. It is all or nothing: either the parts of

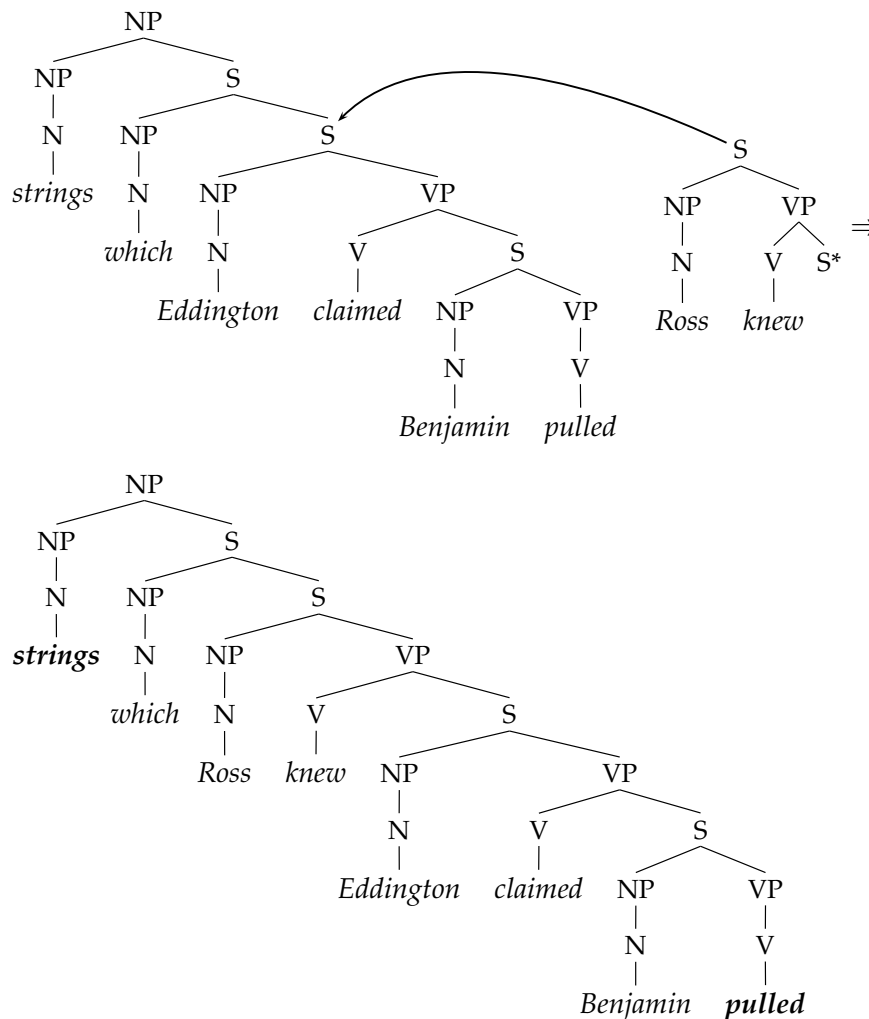


Figure 3.3: TAG derivation for *strings which Ross knew Eddington claimed Benjamin pulled*

the MWE cannot be separated, or they can be separated at an arbitrary distance. As far as I know, this prediction is correct: there are MWEs which cannot be so separated, and there are MWEs which can be separated at any distance, but there are none which set restrictions on the depth of embedding permitted.

3.6.2 The real words problem

Although the construction-based approach offers natural solutions to many of the challenges posed by MWEs, it shares with the monolexemic and lexical ambiguity approaches a failure to tackle the real words problem. In discussing lexical idiomaticity above, I claimed that elementary trees for MWEs did not need to be anchored by lexical forms that had any independent existence outside

of the MWE. As long as they provide the necessary morphological information, any ‘words’ can anchor an elementary tree. But if we allow this, then we are of course denying that there is any principled solution available to the real words problem. If we allow in MWEs words that appear nowhere else, then there is no formal reason why they must also be accompanied by co-anchors which do. Like the monolexemic or lexical ambiguity approaches, the construction-based approach is too unconstrained in what it permits in the representation of MWEs. By insisting on the lexical status of these expressions, the approach implies that MWEs are just as arbitrary as any other lexical items, and thus can be just as idiosyncratic in their form.

Another way of putting this is that the construction-based approach once again treats it as accidental that the form of words inside MWEs matches those of words outside. There is no connection between regular *kick* and the *kick* in *kick the bucket* (other than that they inflect in the same way); just like in the lexical ambiguity approach, their identical shape is merely a coincidence, at least in the synchronic grammar. However, one advantage of the construction-based approach is that, because it treats MWEs as whole units, rather than conspiracies of smaller words, it does not explode the size of the lexicon in the same way as the lexical ambiguity approach does. The determiner in *kick the bucket*, for example, does not exist as an independent object in the lexicon, and so we don’t need to constrain its distribution. There is only the whole expression, *kick the bucket*; its parts have no status.

The construction-based approach does, also, give us the opportunity to formally state the constraint implied by the real words problem, albeit purely stipulatively. Specifically, in a TAG setting, we can require that elementary trees must be such that at least one of the anchors appears in a lexical entry in which it is the sole anchor (i.e. has some independent existence outside of the MWE). Singly-anchored trees trivially satisfy this constraint, while multiply-anchored MWE trees do so provided they feature at least one word which

exists independently in the language.²⁹

3.6.3 Summary

By expressing both the lexical and phrasal aspects of MWEs simultaneously, the construction-based approach allows us to account for both the idiosyncrasies of MWEs, namely their various kinds of idiomaticity, and also their more regular properties such as their syntactic flexibility. Its main failing as an approach is in accounting for the real words problem, but this is a failing it shares with all but the semantic approach, and it fares better on the question of psycholinguistic plausibility than the latter does.

* * *

Now that we have a better understanding of what the different approaches are, and how they tackle the challenges posed by MWEs, we turn in the following section to a direct comparison between them, and ask which kind of theory we should prefer.

3.7 Comparing the approaches

In Table 3.3, I summarise the above discussion, assessing the various approaches against the criteria identified earlier. I have collapsed morphosyntactic flexibility, syntactic flexibility, and decomposability into a single column, ‘internal complexity’, since approaches (broadly) covary in how well they handle these issues—in large part it amounts to whether they assign a complex structure to the MWE or whether they treat it as indivisible. In the table, a tick indicates that the approach handles the problem satisfactorily, a cross that it does not. Parentheses indicate that the symbol enclosed within them is only accurate given certain caveats to be discussed in the text. In the rest of this chapter, we take a look at each of the headings in Table 3.3, and review how the different approaches succeed or fail in addressing them.

²⁹This will, however, pose a problem for foreign phrases like *ad hoc* or *bon appétit*, where neither of the words has an independent existence in the borrowing language.

	Semantic idiomaticity	Syntactic idiomaticity	Lexical idiomaticity	Real words problem	Internal complexity	Psycholinguistic plausibility	Extended uses
Monolexemic	✓	✓	✓	×	×	✓	×
Lexical ambiguity	✓	(×)	✓	×	✓	×	×
Semantic mapping	✓	(×)	×	(✓)	(✓)	×	✓
Construction-based	✓	✓	✓	(×)	✓	✓	×

Table 3.3: Comparison of various approaches to MWEs

3.7.1 Semantic idiomaticity

All approaches handle semantic idiomaticity, this being the most canonical feature of MWEs, and thus of central importance to any theory which attempts to account for them. This is not, therefore, in and of itself a point of contrast among them. However, exactly how they achieve coverage of this kind of idiomaticity does differ.

The monolexemic, semantic mapping, and construction-based approaches have in common that they associate the meaning of a MWE with the whole expression, in one form or another, and thus store the meaning in one place. As I noted above, this seems to be a good thing, since it is the expression as a whole which has a particular meaning; it is not that certain lexical items happen to exist independently in such a way as to conspire to create larger meanings only when put together with one another, as the lexical ambiguity approach implies.

Another concern when it comes to semantic idiomaticity is how to handle internal modification. The appearance of MWE-internal modifiers is a question of syntactic flexibility, but the effect that such modification has, when it is truly internal, surely relates to the semantics and to the question of what the appropriate representation of MWE semantics actually is. This question divides the pack another way: now, the monolexemic approach stands alone as the only one which is not in principle capable of distributing the meaning of a MWE across its parts. The others all have some means of associating meanings with the individual components of the expression, because they all affirm the existence of some subparts over which to distribute the meaning. In the lexical ambiguity approach these are individual words, in the semantic approach they are conjuncts in the semantic representation, and in the TAG analysis they are separate terminal nodes. All of these provide some means of subdividing the meaning if needed, and thus offer some way of handling internal modification.

Taking these two concerns jointly, then, the semantic and construction-based approaches emerge as preferable, since they both deal with both concerns adequately.

3.7.2 Syntactic idiomaticity

The monolexemic and construction-based approaches unquestionably deal with syntactic idiomaticity, since they are able to spell out any arbitrary ordering of elements within the MWE. The latter fares slightly better, however, since it also allows us to represent syntactic structure and not just word-order variation.

The semantic mapping approach simply does not have anything to say about syntactic idiomaticity. It requires a valid parse of the sentence to be available before it even begins its work, and so anything that is not independently accounted for by the syntax of the language will not have an interpretation, idiomatic or otherwise.

The lexical ambiguity approach suffers from a similar problem, in that if MWE parts are just regular words, we expect them to be subject to the normal rules of syntax. In both cases, we can fix the problem by adding extra rules to the syntax which describe precisely the structures which the syntactic idiosyncrasy requires, but this is *ad hoc*. Of course, the monolexemic and construction-based approaches allow arbitrary syntactic variation to be encoded, so are in some sense equally *ad hoc*; the crucial difference, though, is that for these approaches syntactic idiosyncrasies are expected, since the structures are encoded lexically, and the lexicon is exactly the place we expect to find idiosyncratic behaviour. The logic of the lexical ambiguity or semantic mapping approaches, though, would lead us to expect that syntactic idiomaticity should not exist, since the part of the grammar claimed to be idiomatic is not syntactic: in the first it is the lexicon, narrowly construed, and in the second it is the repository of quasi-inference rules or pretences. To augment either theory with idiomatic syntactic rules would detract from both their underlying logic and their formal parsimony.

In sum, then, on the grounds of syntactic idiomaticity, we should prefer the construction-based approach, since all of the others have shortcomings of one form or another.

3.7.3 Lexical idiomaticity

Lexical idiomaticity groups together the monolexemic, lexical ambiguity, and construction-based approaches, as opposed to the semantic mapping approach. The former all allow us to insert any arbitrary form into our MWE, either as the node of a tree in construction-based theories, or as a lexical entry in its own right in the lexical ambiguity approach, or simply as part of a larger word in the monolexemic account. The semantic mapping approach, by contrast, once again expects the grammar to produce a sensible parse for the sentence on which it can then operate. If a word *only* appears in a MWE, it won't get any interpretation except via whatever mechanism we are using to account for MWEs, and thus the semantic mapping approach cannot get started.

Between those theories which succeed here, we can discern a sort of scale, depending on how tightly connected the words in a MWE are, and thus how embedded in the MWE the lexically idiomatic item is. The monolexemic approach treats the whole MWE as a single word, and so lexical idiomaticity is no more surprising than the fact that the string *ter* appears in *caterpillar* but not independently outside of it. The TAG approach treats the MWE as made up of more than one word (= terminal node), but those words are necessarily connected in a particular syntactic configuration. In this way, the lexical isolate cannot appear outside of the MWE, but it is still a word in its own right. In the lexical ambiguity approach, though, the unusual word is a fully fledged word in just the same way as any other, and thus special arrangements must be made so as to prevent it appearing outside of the context of the MWE. However, the lexical ambiguity approach also needs such arrangements in the general case to prevent any MWE component word appearing with its idiomatic meaning outside of the expression, and so lexically idiomatic items are not particularly unusual in this respect.

Thus, although the lexical ambiguity approach has an air of the stipulative about it here, this concern is not unique to the question of lexical

idiomaticity, and so all three approaches fare about the same when this issue is viewed in isolation.

3.7.4 The real words problem

In many ways, the real words problem and the question of lexical idiomaticity are complementary, and this is reflected in the patterning observed in Table 3.3. The problem is this: to account for lexical idiomaticity, arbitrary non-words must be allowed inside MWEs. But to account for the real words problem, we need to restrict the occurrence of such non-words. These demands are clearly pulling in opposite directions, and unless a theory takes the tension seriously, it will tend to simply come down on one side or the other, as, I believe, is the case here. The monolexic, lexical ambiguity, and construction-based approaches impose no restriction on what can appear as a ‘word’ in a multiword expression. But this means they have no response to the real words problem. On the other hand, the semantic mapping approach explicitly requires that *all* words in a MWE be real words, since there is no such thing as an ‘idiom word’ or the equivalent. This in turn means that the real words restriction is imposed rather too stringently, and lexical idiomaticity is ruled out.

In order to delve further into the nature of the real words problem, we might ask *why* MWEs overwhelmingly contain independently-existing words. A plausible explanation is a historical one. Suppose that all MWEs with idiosyncratic semantics were originally just productive metaphors, i.e. literal expressions that were interpreted figuratively. This clearly makes sense for idioms where a conceptual metaphor is apparent even now, but it also seems to be the origin of such apparently non-metaphorical MWEs as prepositional verbs. For example, *rely* originally had the etymologically appropriate meaning of bringing together (from the Anglo-Norman *relier*, ‘to re-link’), and thus, by extension, of bringing together so that one thing supports the other. We therefore have examples of

quite literal uses of *rely (up)on* in the sense of ‘rest upon’, as in the following, from John Davies’s 1609 *The holy roode*:³⁰

(56) Ah see how his most holy Hand **relies Vpon** his knees, to vnder-prop his Charge.

It is not a great leap from this literal sense of support to the extended one which is the basis of the modern meaning of *rely on*. Even here, then, when the origins of the metaphor are not immediately apparent synchronically, such figuration is nonetheless present in the history of the expression.

We might, then, make the following diachronic claim: every MWE begins life as a figurative expression, but becomes more and more conventionalised the more it is repeated, eventually reaching the point where speakers might not even recognise its metaphorical origins. For example, we would say that *kick the bucket* was originally interpreted as a literal phrase involving pails being struck by feet, which was sometimes interpreted metaphorically as meaning ‘die’ (possibly by way of the image of dangling by the feet from a gibbet or a butcher’s beam, otherwise known as a bucket³¹), but in the same way as any productive, innovative metaphor is interpreted, namely online, as it occurs. Over time, certain expressions gain traction, become popular, and are used more and more frequently—not for any principled linguistic reason, necessarily, but because of all manner of other psychosocial factors to do with cultural trends and transmission. Let us say that *kick the bucket* is one such expression. In that case, its meaning becomes more and more conventionalised, so that it is no longer processed online as a spontaneous figuration, and instead is recognised as a pre-existing part of the language.³² At this point, given their unpredictable

³⁰See the OED Online entry for ‘rely, v.1’ (<http://www.oed.com/view/Entry/162029>; accessed 5 February 2019).

³¹See the OED Online entry for ‘bucket, n.2’ (<http://www.oed.com/view/Entry/24163>; accessed 13 June 2018).

³²Others have made the same suggestion about the path fixed expressions take from poetic to conventionalised. See Perrin (2003) for one such proposal regarding idioms, and Vande Wiele (2016) for an overview of how this process might be looked at from a Relevance Theoretic standpoint. Compare also the ‘career of metaphor’ hypothesis of Bowdle & Gentner (2005), which makes related claims about the processing of conventionalised metaphors.

semantics, the learner might start to see such expressions as single lexical items (or as made up of different versions of the words they seem to contain). But now the meaning of the expression is stored rather than computed, the link to the original metaphor may start to be lost, and the expression becomes a MWE.

This story would also explain why most instances of lexical idiomaticity in MWEs involve words which *used* to exist independently in the language, but no longer do, rather than completely random words with no independent history. Given the path just outlined for MWE formation, we have a route for such 'fossilised' words to be preserved: namely, they are used productively, with their own original meaning, in a piece of figurative language which becomes conventionalised. But once it is conventionalised, the original or literal meanings of its parts are largely irrelevant. In this way, it doesn't matter if the individual words themselves cease to be used independently, since understanding the MWE is not dependent on understanding each of its component parts. But there is no such route for purely nonsense words to enter a MWE.

It is true that not all instances of lexical idiomaticity are down to such fossilisation processes. One other large source of such idiomaticity is foreign phrases borrowed from other languages, where the words quite possibly didn't have a widespread life in the borrowing language outside of the MWE. However, in their original usage they were presumably intended to be understood—perhaps by bilinguals, or perhaps by an educated élite who were expected to know another language (as with French and Latin among the English upper classes, for example)—and thus the same historical path could be posited.³³ But the same route is not open to words which only exist and *have only ever existed* in such an expression. It is the overwhelming lack of such nonsense words which the real words problem is about.

³³In cases where pre-existing MWEs are borrowed, of course, this will not be the case, the expression having already been conventionalised in the source language, and thus borrowed in the same way as individual words are.

If the historical narrative suggested above is true, then it is natural to ask to what extent our different families of theory are compatible with it. A TAG-based explanation is easy to imagine. MWEs start life as multiple, singly-anchored trees, representing their literal use as productive metaphors. But at the point of conventionalisation, speakers may instead hypothesise single, multiply-anchored trees, since the expressions acquire some degree of unity and have a conventionalised meaning, just like any other lexical item.

To some extent the same story could be told with the monolexemic approach, where instead of fusing into a multiply-anchored tree, the individual words fuse into a single word. But this has a number of shortcomings: to merge the component parts of a MWE over time like this would leave unexplained why inflection survives in the right place, why syntactic alternations are possible, and all the other issues which the monolexemic approach fails to account for. Although it is true that MWEs sometimes have restrictions on their behaviour which their literal counterparts do not, they are never so rigid as would be expected based on the monolexemic account.

A historical account based on the lexical ambiguity approach needs to explain how such ambiguity arises. One such story might go as follows. Assume that at the point of conventionalisation, instead of the expression being reanalysed as a single elementary tree, as in the TAG approach, the component words are reanalysed as having their idiomatic meanings, thus generating the new senses of these words and creating the ambiguity. This explains the fossilisation facts, since once a word has split between its idiomatic and non-idiomatic versions, there is nothing forcing their histories to continue down the same path, and so it is not unexpected that sometimes one form will fall out of use while the other endures.

This seems plausible, but there are a number of issues which make it less compelling than the TAG version. Firstly, the problems regarding non-decomposable idioms remain: in *kick the bucket*, what meaning would we assign the new version of *bucket*, or *the*, for that matter, that would follow naturally from the idiomatic meaning? Secondly, the distributional facts need explaining.

The fact that *beans* can't appear with the meaning **secrets** outside of *spill the beans*, or the fact that *amok* can't appear outside of *run amok* full stop, needs to be accounted for. In the TAG-based approach, this follows naturally, since once the expression becomes a single elementary tree, it must appear together if it appears at all, by definition. But in the lexical ambiguity approach, additional mechanisms are called upon, essentially imposing extreme selectional restrictions on the various components of a MWE. But where do these come from, and why do they follow from the conventionalisation of the figurative expression which is the source of the MWE? There may be satisfactory answers to these questions, but they don't seem to fall out naturally in the same way as in the TAG approach.

The semantic mapping approach assumes that although the figurative meaning of MWEs may become conventionalised, it does not really become lexicalised. What has to be learned are conventions about meanings, and this, as we have mentioned, means that idiosyncrasies of form remain unexplained. This is what it means to say that the semantic mapping approach's solution to the real words problem is too strict; it insists that idiomatic expressions are essentially still metaphorical, albeit conventionalised. This explains the presence of real words in MWEs, but would not predict the fossilisation that we observe, whereby historically obsolete words endure in MWEs.

Taken on its own, we are forced to say that none of the theories discussed really has a satisfactory solution to the real words problem: the monolexemic, lexical ambiguity, and construction-based approaches fail to account for it, and the semantic mapping approach does so at the cost of excluding the possibility of lexical or syntactic idiomaticity. However, if we assume that the phenomenon is actually the result of historical processes, and thus does not need a synchronic explanation, we might look more favourably on the construction-based theories, since here we can tell the most plausible story about how MWEs come into existence, and why this should mean they contain independently-existing words.

3.7.5 Internal complexity

Our theories differ in whether they assign internal structure to MWEs only in that the monolexemic approach does not, while the others do. Provided that such expressions *do* have some internal complexity, it is generally possible to account for why individual parts vary morphologically, are able to bear their own meanings, or can be manipulated by syntactic processes. Nevertheless, as with the question of semantic idiomaticity, it can still be informative to look at exactly *how* the theories explain the various kinds of flexibility, and to see whether the predictions they make naturally align with the data.

The monolexemic approach is pretty hopeless here, since it does not envisage MWEs as having any internal structure. This means that there is no reason for the sub-parts to inflect in the appropriate places, and no way to make them subject to any regular syntactic manipulation. It would be possible to design *ad hoc* rules which reorder elements, but even then it would be very difficult to interleave other material from outside the MWE. And it certainly makes no sense to assign individuated meanings to the internal parts of a MWE, since there are no such parts. In short, the monolexemic approach fails to account for the possibility of morphosyntactic or syntactic flexibility, or the facts of decomposability.

The semantic mapping approach is the most neutral of the theories when it comes to internal complexity, since it essentially says nothing about the morphosyntax of the expressions it operates on, and thus whatever inflectional patterns or syntactic operations would normally apply continue to do so. It also correctly expects that individual parts could have specific idiom-internal meanings, dependent on the nature of the idiomatic mapping. This neutrality is entirely sensible for MWEs with no formal idiomaticity, and correctly predicts that irregular lexemes continue to inflect irregularly even in idiom contexts, e.g. *come~came a cropper*, *lose~lost one's cool*. However, any lexical idiomaticity remains unaccounted for, and thus so does the morphosyntax of any MWE-only words. What is more, the semantic mapping approach predicts that MWEs should be as syntactically flexible as their compositional counterparts, e.g. verbs

which can passivise in their literal use should be passivisable when used inside an idiomatic expression, contrary to fact. It seems that our theory of MWEs should be able to say something about their form, not just about their meaning, and the semantic mapping approach fails in this regard.

The lexical ambiguity approach does not struggle with lexically idiomatic material, since any idiom-only words will be given their own full lexical entry, this being the nature of the lexical ambiguity approach, and such an entry can therefore spell out their morphological properties just as it would for any other word. In general, the lexical ambiguity approach correctly predicts that inflection will continue to occur as it should, since the words involved in MWEs are fully-fledged words, and thus should behave just like any others. However, on one interpretation it does not follow in the same way as for the semantic mapping approach that words with irregular inflectional paradigms will keep those paradigms in their idiomatic uses. Returning to *come a cropper*, if *come*₁, with the literal meaning, and *come*₂, with the idiomatic meaning, are supposed to be separate words, there seems no reason why they must inflect in the same way. Or at the very least, there is no reason why they should not inflect differently. This may be an argument in favour of the polysemy version of lexical ambiguity, whereby the two meanings share a lexical entry, rather than being two different, merely homophonous, forms. Here the paradigmatic facts follow naturally: there is only one lexical entry under consideration, and so whatever inflectional pattern it follows with one meaning it must also follow with the other.

The lexical ambiguity approach also affords more control than the semantic mapping approach over exactly what kind of syntactic flexibility is allowed in MWEs, since we have direct access to the words bearing the idiomatic meanings in the lexicon. We can thus add specific constraints to their lexical entries if needed, barring passivisation, for example. However, we must recall the collocational challenge facing lexical ambiguity approaches. Since each of the words in a MWE must have its distribution constrained, syntactic flexibility can be a big problem: there is no natural syntactic way of describing the long-distance

relationships that can hold between parts of idioms across e.g. relative clause boundaries, and we must then resort to a hybrid syntactic-semantic account.

Of course, the construction-based approach can encode all such long-distance dependencies locally, using adjunction to account for the separation of the MWE parts. This means that strikingly non-local relationships can still be described locally, and we do not face the same challenges of constraining the distribution of the sub-parts of MWEs: since they cannot appear (with their idiomatic meanings) outside of the elementary tree for the MWE, they are necessarily only licensed when the other parts of the MWE are present.

By treating MWEs as containing multiple terminal nodes, each corresponding to a morphological object, the construction-based approach also naturally accounts for morphosyntactic variation and decomposability. The facts about irregular lexemes do not follow as naturally here as for the other approaches, since there is no sense that the *come* in *come a cropper* is the same as the normal *come*, other than by stipulation. There is certainly nothing stopping us associating the first word in the MWE with the same morphological paradigm as standard *come*, but equally there is nothing motivating us to, either. Perhaps there is a diachronic explanation for these facts to be had: when words enter MWEs they bring their paradigms with them. Indeed, we must assume that historical fossils which exist only in MWEs contribute their own paradigms, even though their independent lexical entries have fallen out of use, and this would follow if the diachronic explanation were true more generally.

In general, then, we can say that the lexical ambiguity and construction-based approaches deal with the phenomena associated with internal complexity the best. The semantic mapping approach also allows for decomposability and for (morpho)syntactic flexibility, but it cannot constrain them, unlike the other two approaches. This is because the lexical ambiguity and construction-based approaches treat MWEs as lexical items rather than simply conventions about meaning, and so they are able to impose constraints on their behaviour in ways the semantic mapping approach cannot.

3.7.6 Psycholinguistic plausibility

On the matter of psycholinguistic plausibility, the lexical ambiguity and semantic mapping approaches make the wrong predictions, while the monolexemic and construction-based approaches make the right ones. Taking a step back, let us see why each of the first two fails.

Since the psycholinguistic evidence points to idiomatic expressions being processed faster than non-idiomatic ones, we must conclude two things:

1. idioms are *different* from non-idioms in some way, since the processing times are different; and
2. whatever the difference is, it cannot involve *additional* computation, since idioms are processed *faster* than non-idioms.

The lexical ambiguity approach fails because of conclusion 1: it assumes that idiom composition is exactly the same process as normal syntactic combination, and we are therefore led to believe there should be no difference in processing times, contrary to fact.

The semantic mapping approach fails because of conclusion 2: this approach claims that normal parsing happens, and then *after* that, subsequent, additional computation is carried out to derive the idiomatic meaning. This ought to make idioms *more* computationally taxing than literal expressions, again contrary to fact.

The other two approaches both claim that idiom processing involves some form of *en bloc* insertion, whereby the idiomatic expression is added to the derivation as a single unit, thereby reducing the number of steps involved, and simplifying the computation. This makes the correct prediction that idiomatic processing times are faster than non-idiomatic ones.

3.7.7 Extended uses

The difficulty in analysing the extended uses of idioms is that they rely simultaneously on the literal and idiomatic meanings of expressions. None of the monolexemic, lexical ambiguity, or construction-based approaches are suitable for this task, because they all take the analysis of sentences like (57) to involve ambiguity at one level or another:

(57) Jadzia kicked the bucket.

We can either parse this literally or idiomatically, but we have to make a choice; and once that choice is made, the other meaning is not available. The semantic mapping approach is different, however, since it requires that the literal meaning be present first, and then maps this onto an idiomatic meaning. Thus, both kinds of meaning are available in the course of the analysis, and it is at least conceivable that a theory could be developed which makes use of this.

I will not have any more to say about extended uses in this thesis, but I include this category for the sake of completeness, and to show that no one theory is able to account for all the analytical challenges posed by MWEs.

3.7.8 Chapter summary

The analytical problem of MWEs is embodied in the tension they exhibit between their word-like and phrase-like statuses, between their idiosyncrasies and their regularities, between their fixedness and their flexibility. The construction-based theories take this tension seriously, and represent MWEs as a kind of hybrid object: they are listed just like normal words, and have the same formal status; yet they contain complex syntactic structure and are made up of multiple morphological objects. For this reason, the construction-based approach succeeds in meeting more of the challenges of MWEs than the other approaches we have examined in this chapter.

The monolexemic approach takes the first of each of the contrasting pairs listed above as primary, and thus totally fails to explain the properties associated

with the other side of the tension. MWEs can vary arbitrarily in their form, but we have no explanation as to why they inflect internally, can have internally-structured meanings, or are subject to normal syntactic processes. The lexical ambiguity approach fares slightly better, but it too attempts to dissolve the tension in favour of one side to the exclusion of the other. This approach has a natural explanation for the facts of decomposability and for syntactic flexibility, but struggles to represent the fact that MWEs are units and not merely conspiracies of individual words. Not only is this a formal failing, but it also means that the approach does not successfully describe our intuitions about MWEs: when I learn the meaning of *pull strings*, it seems as though I have learned one new thing, not two.

The semantic mapping approach offers the unique means to handle the linguistically playful, ‘extended’ uses of idiomatic expressions, but when it comes to the ‘core’ uses it is deficient in a number of ways. Since it is purely a semantic theory, it relies on there being a literal interpretation of the expression available in the first place, which is not always the case. And since it says nothing about the morphological or syntactic properties of MWEs, it cannot explain formal idiosyncrasies like lexical or syntactic idiomaticity, and it has no way of constraining the syntactic flexibility of MWEs.

Going forward, then, we should look for a construction-based account of MWEs. In other words, we want a theory which allows us to represent MWEs as single objects with complex internal structure. As already mentioned, Tree-Adjoining Grammar gives us the formal means to achieve this. The next chapter introduces this theory in more detail, along with two other theories, Lexical Functional Grammar and Glue Semantics, which will form the core of the theory I develop in subsequent chapters.

4

Theoretical background

Contents

4.1	Formal grammars	126
4.1.1	Formalisation	126
4.1.2	Theories of grammar	128
4.2	Tree-Adjoining Grammar	131
4.2.1	Introduction	131
4.2.2	Components of a grammar	133
4.2.3	Elementary structures	134
4.2.4	Combining operations	137
4.2.5	Summary	142
4.3	Lexical Functional Grammar	144
4.3.1	Introduction	144
4.3.2	Two levels of syntax	145
4.3.3	C-structure	147
4.3.4	F-structure	149
4.3.5	Connecting c-structure and f-structure	154
4.3.6	The correspondence architecture	165
4.3.7	Templates	166
4.3.8	Summary	167
4.4	Semantics	168
4.4.1	Meaning language	168
4.4.2	Glue Semantics	175
4.5	Chapter summary	193

Having set the stage in the previous three chapters, we now turn to the theoretical meat of the thesis. In this chapter, I present the formal machinery which will be used in developing my own analysis of MWEs in Chapter 6. We start in Section 4.1 with a brief discussion on what it means to give a formal grammatical analysis of a linguistic phenomenon. In Section 4.2, I then describe the relevant aspects of Tree-Adjoining Grammar, first introduced in the previous chapter. After this, in Section 4.3, we look at Lexical Functional Grammar, the overarching framework within which I will develop my analysis. In the final section, Section 4.4, I present the theory of semantics that I will be working with. In Chapter 5, I show how we can make use of the formal advantages of TAG within an LFG setting, thereby integrating the two frameworks, and in Chapter 6 this is applied to the analysis of MWEs.

4.1 Formal grammars

I have been assuming throughout the discussion that we are interested in giving accounts of linguistic phenomena which are couched in terms of a FORMAL GRAMMAR. But what is a formal grammar? We can answer this in two parts: what makes it formal, and what makes it a grammar?

4.1.1 Formalisation

The most important property of a grammar, just as with any scientific theory, is that it be formalised, or at least be capable of being formalised.¹ Another way of putting this is that the theory should be *explicit*, and should therefore make clear predictions. If I have a theory and I have some data, it should (at least in principle) be totally apparent whether the data falsify the theory or not. There should be no room for debate, hand-waving, appeals to metaphor,

¹It may seem perverse to claim that formalisation is more important for a theory than, say, getting the facts right. But if we do not know what claims a theory is making, we cannot know whether it gets the facts right in the first place. Formalisation is conceptually prior to accuracy; we cannot tell whether a theory is right if we don't know what it says, and formalisation is merely the requirement that we know, explicitly and unambiguously, what a theory is saying.

and the like. For this reason, it is vitally important that the terms in which the theory is expounded have clear, unambiguous, and knowable meanings. This is why science often tends to employ mathematical or logical languages, since they have a universally agreed-upon interpretation—claims made in these languages have meanings which are not open to dispute, and so, if we agree on what the facts of the matter are, i.e. the data the claims are made about, then we ought also to agree on whether such claims are true or false.

It is something of an irony that such formalism can serve as a major source of confusion on the part of the reader, and of obfuscation on the part of the author, since the point of using it is to ensure maximal clarity. This is why it is important that claims be *in principle* translatable into some explicit formal system, but need not always be presented in that manner. This avoids unnecessary use of technical notation which serves no end other than to obscure. In practice, however, the best way to demonstrate that something *could* be made formally explicit is in fact to do so. Nonetheless, in what follows I will always attempt to introduce formalism alongside a ‘prose’ translation which makes clear the point of the formal expression.

As well as a formal theory, we can talk about FORMAL LANGUAGES. Used in this sense, ‘language’ is a term of art from computer science and information theory, and simply refers to a set of strings. For example, (1) is a (very simple) language whose elements are made up of English words, but so is (2), whose alphabet consists only of the letter *a*:

(1) {Julian loves Dax, Dax loves Julian}

(2) {*a, aa, aaa*}

It might be thought that these objects have very little in common with natural languages, but in fact it has proven instructive to consider natural languages as if they were formal languages. Some researchers, like Montague, even deny that there is any meaningful difference between the two (Montague 1970: 222):

There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single natural and mathematically precise theory.²

This is arguably the view taken by Chomskyan linguists also, although Chomsky's recent focus on the biological aspects of human language departs from this view in important respects.³

But just as a semanticist who makes use of possible worlds is not committed to modal realism (cf. Stalnaker 1976), we need not be committed to the identity of natural and formal languages to take advantage of the tools of formal language theory; it is enough to see formal languages as a useful approximation of natural languages. In a similar vein, even though Newtonian physics may have been superseded by the quantum revolution of the twentieth century, it remains a useful tool for addressing issues in classical mechanics.

The power and utility of formal language theory is another instance of the 'unreasonable effectiveness of mathematics' in the sciences (Wigner 1960): poorly understood natural phenomena can be profitably modelled as better understood mathematical objects, from which predictions can be made and more precise theories can be crafted. Natural languages are the unruly naturally-occurring phenomena; formal languages are their mathematically well-behaved kin. This thesis continues the tradition of viewing languages as objects which can profitably be looked at using the tools of logic and mathematics.⁴

4.1.2 Theories of grammar

The task of a grammar is to describe a language, and we want a grammatical theory to be flexible enough to help us write grammars for any human language.

²Égré (2018: 702) notes that Church (1951: 106) expressed a similar idea several years before Montague when he wrote "[a]lthough all the foregoing account has been concerned with the case of a formalized language, I would go on to say that in my opinion there is no difference in principle between this case and that of one of the natural languages".

³Montague himself clearly saw his ideas as compatible with the syntactic theories of the time; the above quotation continues thus: "On this point I differ from a number of philosophers, but agree, I believe, with Chomsky and his associates".

⁴See Partee et al. (1990) for a thoroughgoing overview of the role of logic and mathematics in linguistic theory.

For those in the Chomskyan tradition, a theory of grammar is also a psychological theory: it is a theory about what we know when we know a language. We need not commit to this strongly internalist view, however, and can take a more conservative, externalist view of things, where language is taken as an abstract entity in its own right. In either case, the task of a grammar is to characterise the well-formed expressions of a language. By ‘expressions’ here we generally do not mean simply strings of words, but also some structure(s) associated with them, of e.g. a syntactic, semantic, morphological, or phonological nature.

However we conceive of the formalism, a theory of grammar has two important tasks. Firstly, it should describe all of the linguistically relevant information which is encoded in an utterance. That is, it must tell us *what* is contained in a linguistic expression. This includes the identification of phonological features all the way up to the illocutionary force of a speech act. Writ large, this is perhaps too ambitious for any one theory to manage, and so in practice theories tend to delineate some smaller portion of this information to describe. Nonetheless, it is worthwhile keeping this overall goal in mind.

Secondly, though, a theory should explain *how* we can arrive at this analysis. To a large extent, theoreticians agree on the details of the first task. They may disagree about the status of certain properties (whether grammatical functions are primitive or derived concepts, for instance), but they will generally still agree that the properties must be encoded somehow. We are, ultimately, all trying to explain the same facts. Where theoretical disputes tend to occur is around how we respond to the second task. For instance, does grammatical knowledge amount to being able to carry out a series of computations on a single type of formal object, or does it involve assessing the truth or falsity of various constraints described over (potentially several types of) formal objects? This question characterises the difference between DERIVATIONAL and CONSTRAINT-BASED theories of syntax. A similar distinction is made by Pullum & Scholz (2001) between GENERATIVE-ENUMERATIVE and MODEL-THEORETIC syntax (GES and MTS, respectively): the former views a grammar as a mechanism for producing (or

generating, or enumerating) all of the permitted expressions in a language, while the latter merely gives a statement of the necessary conditions any expression must meet to be licensed. To some extent this corresponds to the division between derivational and constraint-based theories—certainly any derivational theory will be generative-enumerative, but constraint-based theories can be made to fit either conception.

Very often, the two kinds of approach are inter-translatable anyway, although this is not always true. Pullum & Scholz (2001) and Pullum (2013) discuss some significant differences between GES and MTS, which they interpret as arguing in favour of the latter. The theory presented in Chapter 5 is constraint-based, and of the model-theoretic kind, but the two theories introduced in this chapter, Tree-Adjoining Grammar (TAG) and Lexical Functional Grammar (LFG), are not quite so straightforward to categorise. LFG is broadly constraint-based, but in its standard formulation it is not possible to view it entirely in these terms, as we will see. What is more, owing to its origins in generative grammar, it is not clear that Pullum & Scholz would unproblematically characterise it as model-theoretic in its standard form. TAG on the other hand is standardly understood as derivational, and therefore clearly generative-enumerative, but it can be reinterpreted in constraint-based terms, which is exactly what I will do in Chapter 5—this is an example of the potential for translation between GES and MTS formalisms.

* * *

In the remaining three sections of this chapter, I will introduce three theories which will be significant for the proposals introduced in this thesis. In the next two sections, we look at TAG and LFG, introducing theoretical constructs and notation which will be used later. In the last section, I introduce the logical language I will use to represent meanings, and then explain Glue Semantics, a theory of the syntax-semantics interface.

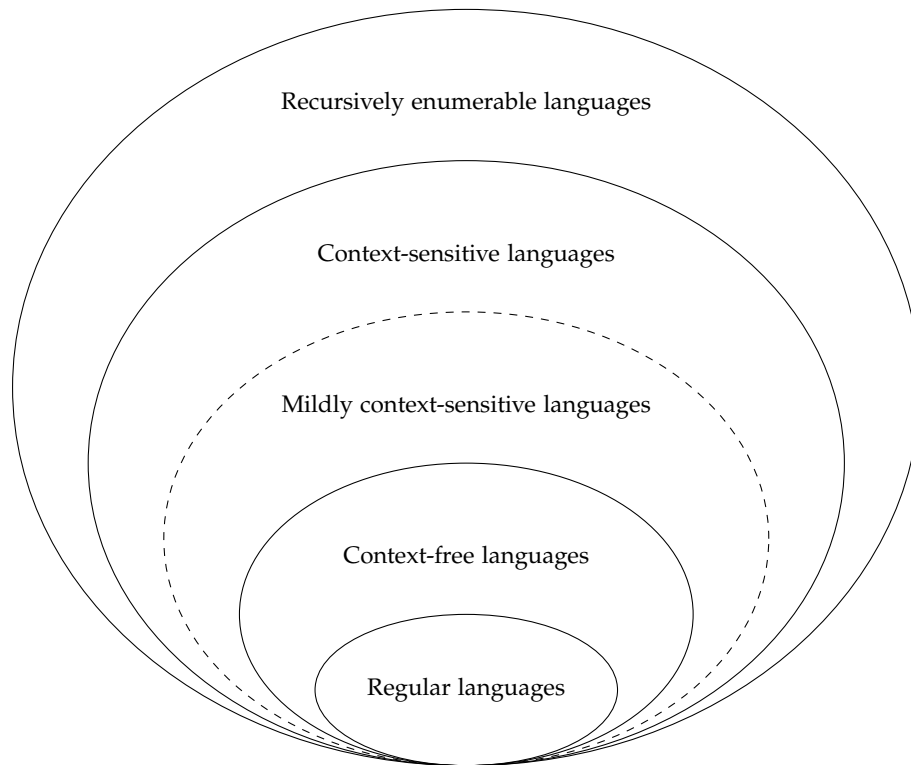


Figure 4.1: The Chomsky hierarchy

4.2 Tree-Adjoining Grammar

4.2.1 Introduction

One important insight from formal language theory is that different kinds of grammar have different expressive powers. For example, the simplest kind of formal grammars, the `REGULAR` grammars, cannot describe languages which have centre-embedding (one constituent appearing inside another constituent of the same type). We can organise the different kinds of grammars into a hierarchy based on the types of language they are able to generate; this is called the `CHOMSKY HIERARCHY` (Chomsky 1959), and is shown in Figure 4.1. It is an inclusion hierarchy, because every regular language can also be described by a context-free grammar, and every context-free language can be described by a context-sensitive grammar, and so on.

As in all scientific endeavours, when it comes to designing grammars for natural languages, we want the most constrained theory which still captures

the facts. If our grammars for natural languages are regular, for example, then we make strong predictions about what kinds of dependency we expect to see in natural languages. This makes the theory falsifiable, which is what makes it of scientific interest. If our grammars are recursively enumerable, on the other hand, then our theories are very powerful indeed, which makes their predictions much less interesting: almost anything goes.

An important question in theoretical linguistics, then, is how complex natural languages are. From the fact that English *does* have centre-embedding, along with various other kinds of dependencies which cannot be expressed by regular grammars, Chomsky (1957) argued that English is not a regular language, and, thus, by extension, that the class of natural languages in general cannot be regular. It is for this reason that context-free grammars (CFGs) have become so central in linguistic theorising since the 1950s. However, in the following decades there was a growing realisation that natural languages may fall even beyond the class of languages which can be described by context-free grammars.

It was Shieber (1985) who delivered the conclusive proof that this was indeed the case, showing that cross-serial dependencies in Swiss German pushed the language into the non-context-free space (see Kallmeyer 2010: 17–20 for discussion).⁵ However, it seems that the full power of a context-sensitive grammar is overkill for natural languages—the areas where greater-than-context-free power is required are relatively few and relatively rare cross-linguistically. Joshi (1985) coined the term MILDLY CONTEXT SENSITIVE to describe precisely the degree of context-sensitivity required to adequately describe natural languages, which falls below the full expressive power of context-sensitive grammars.

Beginning in the 1970s, a number of formalisms have been proposed which take the expressive power of the grammar into (or sometimes beyond) the space of context-sensitivity. These include Head-driven Phrase Structure Grammar

⁵Bresnan et al. (1982) had made the same claim for Dutch, but it turned out that the string language of Dutch *could* be described in context-free terms, even if its constituent structure required a more powerful mechanism (Culy 1983), and so there was still some debate about the impact of the findings.

(HPSG: Pollard & Sag 1994), Combinatory Categorical Grammar (CCG: Steedman 1987), as well as Lexical Functional Grammar and Tree Adjoining Grammar, the theories discussed here. LFG achieves this effect by augmenting the CFG with an additional apparatus, *f*-structure (on which see below, Section 4.3.4). TAG takes a different approach, replacing the CFG wholesale with a new, more powerful formalism.

Unlike CFGs, which (at least as classically conceived) rewrite strings to generate other strings (with trees merely being secondary objects implied by the derivation), TAGs rewrite trees. Trees, not strings, are therefore the basic objects on which the formalism operates, and thus languages generated by TAGs are sets of trees, rather than sets of strings. In this section, I will lay out the basic components of a TAG and how it can be applied to natural languages. The presentation will be somewhat light, since I only want to borrow a few high-level insights from the TAG framework; for more detailed or more mathematically-minded treatments, the reader should consult the foundational text for TAG, Joshi et al. (1975), as well as the collection of papers in Abeillé & Rambow (2000b) and references cited therein.

4.2.2 Components of a grammar

Abeillé & Rambow (2000a: 4) invite us to think about formal grammatical systems in terms of two sub-components: their *ELEMENTARY STRUCTURES* and their *COMBINING OPERATIONS*. The elementary structures are the ‘atoms’ or building blocks which are then put together using the combining operations to build larger structures.⁶ The elementary structures of a CFG, for instance, are the phrase-structure rules, and the only combining operation is the application of a rule to a non-terminal.⁷ TAG, by contrast, treats trees as elementary structures, and combines them via two operations, *SUBSTITUTION* and *ADJUNCTION*. In

⁶Note that this presupposes a derivational/generative understanding of formal grammars.

⁷It may seem rather counterintuitive to talk about phrase-structure rules as building blocks, but the point is that they are the units which make up a CFG. In a certain fundamental sense, a CFG is simply a set of phrase-structure rules; there are no other elementary formal objects in the theory.

the next two sections, I introduce the elementary structures and combining operations of TAG.

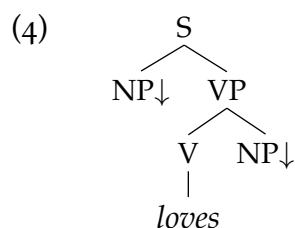
4.2.3 Elementary structures

As mentioned, the elementary structures of TAG are trees. The TAG formalism itself is of course ambivalent about the exact character of these trees, but it is standard to assume a relatively straightforward phrase structure which omits bar levels or complex functional projections, and sticks to the traditional $S \rightarrow NP VP$ view of clause structure. More specifically, each ELEMENTARY TREE is taken to be the extended maximal projection (Grimshaw 2005) of a lexical item—that is, the maximal projection which also encompasses the word’s full syntactic argument structure. In almost all modern implementations of TAG, a grammar will also be LEXICALISED: this means that each elementary structure is associated with at least one lexical item. TAG grammars which have this property are known as Lexicalised TAGs, or LTAGs (Schabes et al. 1988). We will return to the topic of lexicalisation in more detail in Chapter 5.3. For now, we turn to some examples of elementary trees.

Since proper nouns like *Alex* do not take arguments, their extended maximal projection is simply the NP that contains them. The elementary tree for a proper noun like *Alex* is thus as follows:

- (3)
- $$\begin{array}{c} NP \\ | \\ N \\ | \\ Alex \end{array}$$

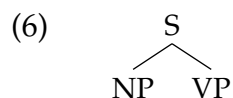
For an argument-taking item, however, the elementary tree must contain slots for each of its arguments. For instance, the elementary tree for active voice, present tense *loves* is as follows, with NP slots in the subject and object positions:



These SUBSTITUTION SITES, gaps where arguments are to be inserted by substitution (see Section 4.2.4), are indicated by arrows (\downarrow).

The rules in a CFG can be thought of in these terms as well, as embodying small trees. For example, the rule in (5) could be thought to stand for the tree in (6):

(5) $S \rightarrow NP VP$



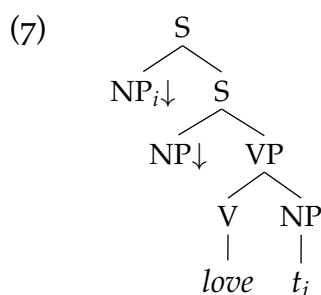
Taking this view makes clear one crucial difference between a CFG and a TAG: the latter has what is called an EXTENDED DOMAIN OF LOCALITY. If we convert all the rules of a CFG into trees as we did in (5)–(6), then each of the trees will have depth 1, since the rules only describe immediate dominance relations.⁸ In a TAG, though, the elementary trees can in principle be of any size. This is linguistically significant, since it means that certain relationships can be described locally in a TAG but only indirectly in a CFG.

For example, in English, the subject of a sentence usually agrees with the verb. But in a CFG, this cannot be expressed locally, in a single elementary structure, since no rule mentions both the subject and the verb. Instead, it must be conceived of as agreement between the subject and the verb *phrase* (with some way of passing agreement information up from the verbal head to its phrasal projection). By contrast, a TAG allows the relationship to be stated directly between the subject and the verb, since both positions appear

⁸The depth of a tree is the longest path from a terminal node to the root node. A tree consisting of a single node thus has a depth of 0.

in the same elementary tree, as in (4), for example.⁹ Exactly how agreement is encoded is not constrained by the TAG formalism *per se*; it is usually achieved by decorating the nodes with feature structures of some kind, which must then unify (Vijay-Shanker & Joshi 1988). This is somewhat similar to the way LFG handles agreement, discussed below (Section 4.3.4); the crucial difference, though, is that the feature structure in question for LFG is non-local, being a projection of the c-structure as a whole, while in Feature Structure based TAG (FTAG) each node bears its own feature matrix.

Perhaps more significantly, it means that the relationship between the filler and gap in a long-distance dependency can also be encoded locally. For example, (7) is an elementary tree which would be used in the derivation of sentences like *Who does Julian love?*, where the object of *love* has been questioned:



The relationship between the extracted object, indicated here by a trace, *t*, and the surface position of the (to-be-inserted) *wh*-word, the first NP, is indicated by coindexation.¹⁰ Whatever this coindexation is taken to actually mean formally, the important thing is that both positions appear in the *same* elementary tree, and so the relationship between them can be encoded in a single lexical entry/elementary structure, rather than having to rely on a subsequent operation of movement or trying to list all possible paths from filler to gap. Of course,

⁹Of course, there may be independent linguistic evidence which leads us to believe that the subject agrees with the verb phrase as a whole rather than just the verb. But a TAG can express this equally well; the point is that the decision is forced on us by the formalism in a CFG, but can be made based on linguistic data in a TAG.

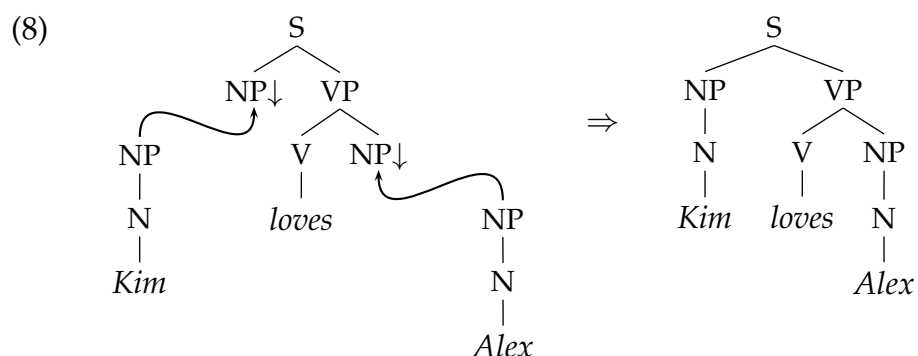
¹⁰The trace in object position is not an essential part of the TAG analysis, but in practice it is usually used. See also Kroch & Joshi (1985) and Kroch (1987), who argue that having a trace here is useful for theoretical reasons, *viz.* that it allows constraints on *wh*-movement to be stated in terms of the topology of elementary trees, rather than requiring any additional kind of specification.

the filler and the gap can appear arbitrarily far apart, and so the elementary tree in (7) is not enough by itself—it only accounts for *wh*-extraction in the same clause. As we saw earlier, what allows us to encode such dependencies locally even though they may ultimately end up separated is the operation of adjunction, to which we turn in the next section.

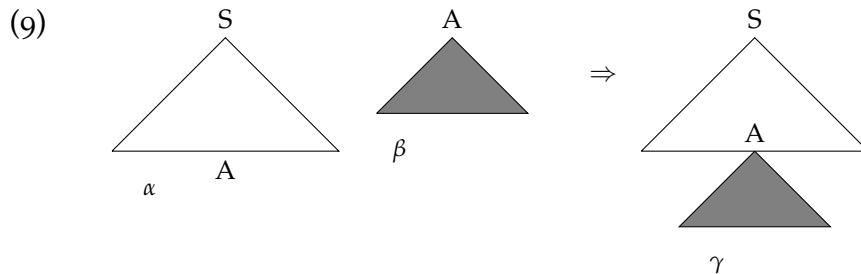
4.2.4 Combining operations

Strictly speaking, TAG employs only a single operation for combining elementary structures: adjunction. However, it is conventional to also include a second operation called substitution in presentations of TAG. Substitution can be simulated by adjunction (it is essentially the trivial case of adjunction), so in terms of formal parsimony, only the latter is required. Nonetheless, for the sake of explanatory simplicity it is often preferable to treat the two as if they were separate, since they serve different purposes in the linguistic applications of TAG, and I will continue this tradition here. I therefore first introduce substitution, before moving on subsequently to adjunction.

Substitution is the process whereby one elementary tree is inserted into another. The node which is the target of substitution, the substitution site, must share a label with the root node of the tree being inserted. For example, the tree for *Alex* given in (3), which has NP as its root, can be substituted into one of the NP nodes in the tree for *loves* given in (4), as shown in (8), a sample derivation for *Kim loves Alex*:



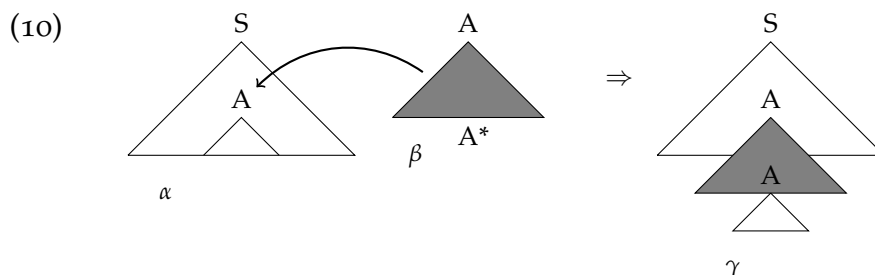
The process of substitution is shown schematically in (9) (after Abeillé & Rambow 2000a: 5, fig. 3):



Tree β , which has root node labelled A , can be substituted into tree α because α has a frontier node also labelled A . The result is the larger tree, γ . Elementary trees where the frontier nodes are all either terminals or substitution sites (all of those we have seen so far) are called **INITIAL TREES** (in contrast to auxiliary trees, to be introduced below).

A tree grammar which only includes the operation of substitution is called a **Tree Substitution Grammar (TSG)**, as opposed to a TAG, and remains within the context-free space of computational complexity. To reach the mild context-sensitivity of TAG, we must add the operation of adjunction (Joshi et al. 1975).

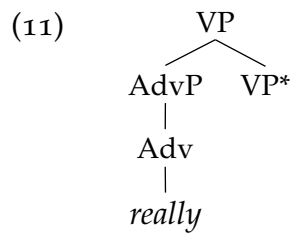
As we have already seen, adjunction can be represented schematically as in (10) (after Abeillé & Rambow 2000a: 9, fig. 6):



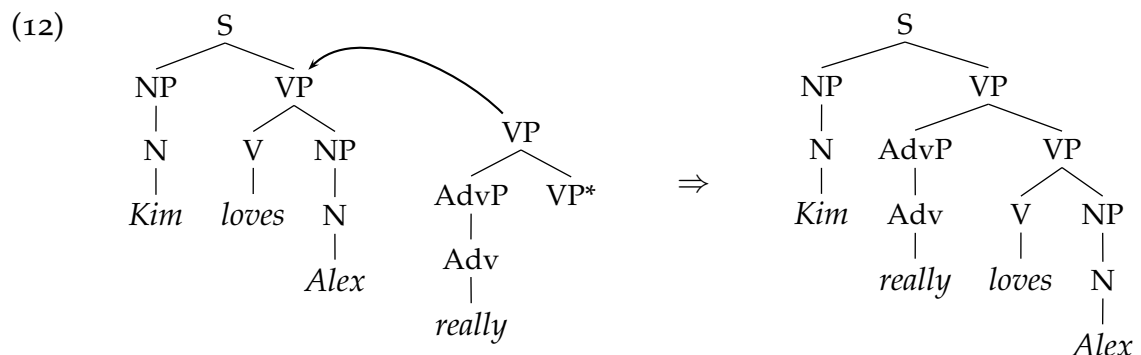
An elementary tree α has a node labelled A somewhere other than on its frontier. Another tree, β , is a special kind of elementary tree called an **AUXILIARY TREE**, which has both a root node and exactly one frontier node with the same label, the latter called the **FOOT NODE** (indicated with a $*$). In this case, β has root and foot nodes labelled A . To adjoin β into α , we remove the subtree rooted in A

from α , replace it with β , and then attach the subtree which we removed to the foot node of β . This produces a larger tree, γ .

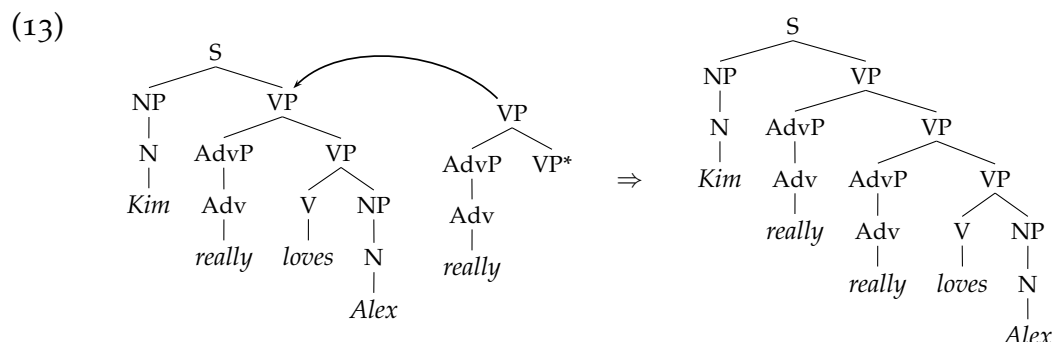
Linguistically, adjunction is used for several purposes, including the insertion of modifiers or functional categories, and sentential embedding. A typical auxiliary tree for an adverb, here *really*, is given in (11):



This has a VP as its root and foot nodes, and so will adjoin to a VP, expanding it by adding the AdvP containing *really* on its left edge. If we adjoin this to the derived tree for *Kim loves Alex* from (8), we get the result given in (12), namely the tree for *Kim really loves Alex*:



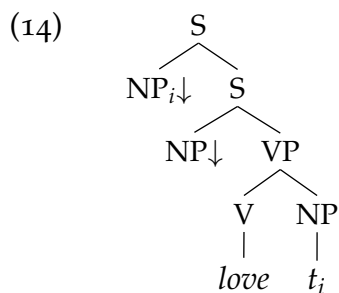
Of course, such modification can be repeated, and adjunction permits this too. The derived tree on the right-hand side of (12) has two VP nodes which can themselves be targets of adjunction for *really*:



And it is clear that this process can be repeated indefinitely, just like a recursive rule in a CFG.

Notice that in (13) either VP node is a potential target for adjunction. Since this complicates parsing (there are two distinct derivations for a single sentence in cases like *Kim really really loves Alex*, and they do not correspond to a real ambiguity), it is standard practice to block adjunction at one of the nodes for an auxiliary tree like that for *really*. We do this by using ADJUNCTION CONSTRAINTS. These are added to the nodes of elementary trees and come in three kinds: selective adjunction constraints (SA), obligatory adjunction constraints (OA), and null adjunction constraints (NA) (Abeillé & Rambow 2000a: 11). An SA constraint is a list of auxiliary trees which are permitted to be adjoined at the node, while an OA constraint is a list of trees one of which *must* be adjoined at the node. An NA constraint prohibits adjunction at the node which bears it; we may think of it equivalently as an SA constraint which consists of the empty list. To avoid having multiple possible parses for adverbially modified sentences, we mark the foot node of adverbial auxiliary trees with an NA constraint (Bleam et al. 2001). This means that the higher VP node is the only target for further adjunction.

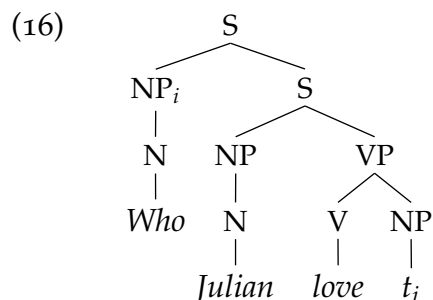
As well as modification, adjunction is also used in the analysis of sentential embedding, which is similarly recursive. Recall that the object-question version of *love* has the elementary tree in (14):



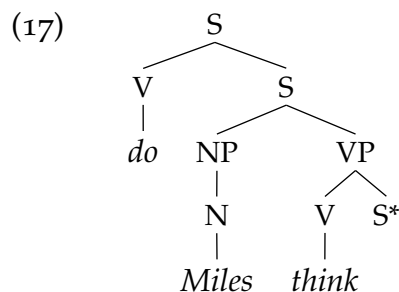
How can we derive a multi-clause question like (15) from (14)?

(15) Who does Miles think Julian loves?

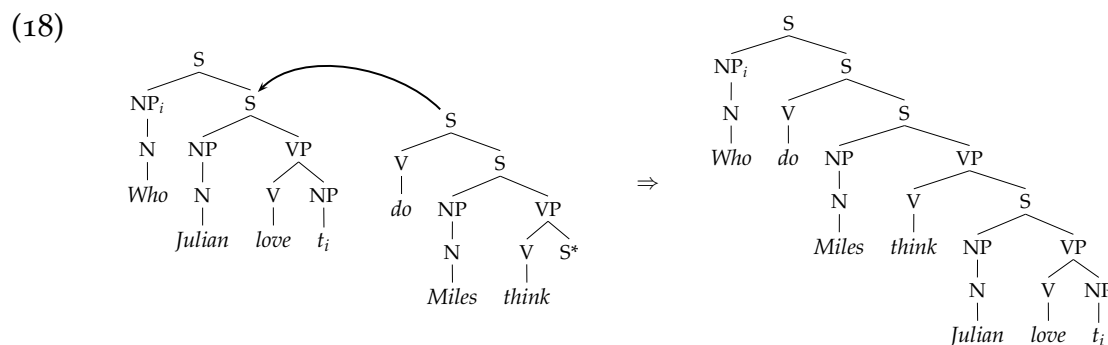
Ignoring morphological variation for the sake of simplicity, we can arrive at the following extended elementary tree representing the subordinate clause and the fronted *wh*-element simply by substitution:



Since the distance between a fronted element and the gap which it fills can be arbitrarily large, we need a way to expand this initial tree ‘from the inside out’. This is exactly what adjunction allows. The main clause of (15) is represented by the auxiliary tree in (17) (once again ignoring morphological variation):¹¹

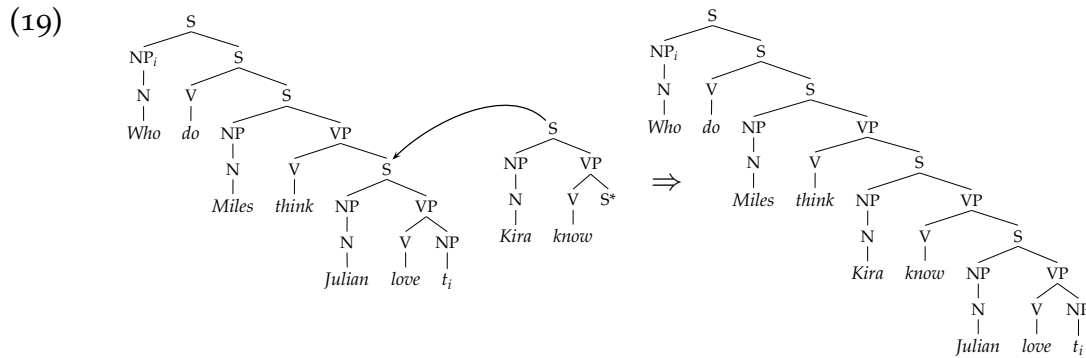


Since this is an auxiliary tree, it can be adjoined at the intermediate S node in (16):



¹¹I omit the details of how *do*-support works, which is orthogonal to the main point here. *Do*-support is one potential application of an OA constraint or of an appropriate feature structure annotation, ensuring that verbs in interrogative clauses are accompanied by an auxiliary. See Bleam et al. (2001) for some discussion in an FTAG setting.

Now the filler and the gap have been separated, even though they are both from the same elementary tree. Crucially, of course, this process can be repeated:



Because of this, the fronted element can end up arbitrarily far from the gap which it fills, even though the relation is represented locally in the grammar. As mentioned above, this avoids the need for other mechanisms such as movement or feature percolation in order to correctly describe long-distance dependencies.

4.2.5 Summary

TAG is a formalism for manipulating trees directly, as opposed to seeing trees as merely side effects of the rewriting of strings, as CFGs were originally conceived. It is of theoretical interest in its own right, and has also played (and continues to play) an important role in research into both computational (e.g. Bleam et al. 2001; Kasai et al. 2017) and psycholinguistic (e.g. Ferreira 2000; Ferreira et al. 2004) models of processing. However, for the purposes of this thesis, there are two properties of TAG in particular which are of interest, since they lend themselves very naturally to a constructional theory of MWEs, as outlined in the previous chapter.

Firstly, TAGs have an extended domain of locality when compared to CFG-based theories. The only way for a CFG to describe a MWE in a single place would be to include multiple terminal nodes on the right-hand side of a phrase-structure rule. However, this would mean that all such terminals would have to be at the same depth within the tree, since PSRs only describe immediate dominance relations. It is impossible to represent the articulated internal syntax

of MWEs using a single context-free rule. By contrast, owing to their extended domain of locality, TAG elementary trees can have anchors which appear at different levels.¹²

The other property of interest is the operation of adjunction. This is what enables structures to grow internally, and thus enables the parts of MWEs to be separated, even though they can still be represented locally in the grammar. This is crucial for explaining the syntactic flexibility of some MWEs while still representing them locally: without adjunction we would be forced to represent MWEs as conspiracies of single-word expressions, in order to permit their parts to be separated by various syntactic processes. With adjunction, it remains possible to represent MWEs as single objects, which can then be manipulated. These strengths of TAG for handling idiomatic expressions were noted by Abeillé (1988, 1995). I highlight them here as they will be the crucial aspects of TAG which I seek to introduce into LFG in the following chapter.

¹²One might imagine a CFG treatment of MWEs which makes use of very specific sub-categories in its rules—for example, a series of rules like (i):

- (i)
- a. $S \rightarrow NP VP_{\text{spill-the-beans}}$
 - b. $VP_{\text{spill-the-beans}} \rightarrow V_{\text{spill}} NP_{\text{the-beans}}$
 - c. $NP_{\text{the-beans}} \rightarrow \text{Det } N_{\text{beans}}$
 - d. $V_{\text{spill}} \rightarrow \textit{spill}$
 - e. $N_{\text{beans}} \rightarrow \textit{beans}$

This essentially simulates the effect of the extended domain of locality provided by a TAG. However, we now have a version of the lexical ambiguity approach, and so are faced with all of the attendant problems discussed in Chapter 3.3. On top of this, we actually lose one of the advantages of that approach, which is that idiom words have normal syntactic categories and so ought to be subject to all the usual rules of the grammar regarding e.g. extraction, modification, etc. Under the approach suggested here, this is not true: *beans* is of category N_{beans} , not N , and so we need additional special rules for e.g. adjective modification. That is, alongside (ii), we also need (iii):

- (ii) $N \rightarrow \text{Adj } N$

- (iii) $N_{\text{beans}} \rightarrow \text{Adj } N_{\text{beans}}$

And the same goes for every construction such expressions can appear in. So each MWE would actually be associated with a family of PSRs, which would largely recapitulate information already in the grammar.

4.3 Lexical Functional Grammar

4.3.1 Introduction

Despite the attractive advantages of TAG when it comes to the analysis of MWEs, the theory developed in this thesis will be couched in terms of Lexical Functional Grammar (LFG), not TAG. There are a number of reasons for this, which will become clear as I present the theory, the philosophy behind it, and what its working assumptions are. In brief, it offers a fuller and more integrated architecture for the description of all levels of linguistic analysis, not just phrase structure; the separation of functional from configurational syntactic information makes for a highly cross-linguistically applicable theory; and its modularity leads to a leaner formal theory over all, which does not commit the framework to a particular analysis of the different levels, and which, by making use of different kinds of data structure for different kinds of phenomena, does not force the data to fit the theory.

The founding work of LFG is Kaplan & Bresnan (1982), but earlier traces can be seen in Joan Bresnan's work on the psychological plausibility of transformational grammars (e.g. Bresnan 1978). Indeed, LFG arose as a response to the transformational theories popular at the time, which were seen as both psychologically and computationally deficient. Outside of linguistic theory itself, it is perhaps in the latter area where LFG has had the most impact, serving as the basis of several important computational implementations, most notably the Xerox Linguistic Environment (XLE: Crouch et al. 2012). But LFG also continues to play a role in psycholinguistic work—for instance, it underpins the influential Bock-Levelt model of language production (Bock & Levelt 1994).¹³

The important properties of LFG which have made it appealing to these two constituencies are the fact that it is well formalised and the fact that it is modular. As discussed above, the first goes without saying for any good

¹³Of course, success in these two domains is not unrelated: if we assume the computational theory of mind (e.g. Putnam 1967; Fodor 1975, 1983), then any computationally-implemented theory could in principle stand as a(n aspect of a) psychological theory.

theory.¹⁴ The second is a more specific LFG advantage over other theories, however. As we will see in more detail below, LFG separates out the various levels of representation which make up a linguistic description, and allows for each to be described using a different kind of data structure, with the different levels connected by correspondence functions. This view is described by Bresnan (1993: 45) as follows:

Semantic argument structure, constituent structure and functional structure are parallel information structures of very different formal character. They are related not by syntactic derivation, but by structural correspondences, as a melody is related to the words of a song. Semantic, structural and functional representations of a sentence can be superimposed, but they are independent planes of grammatical organisation.

In this quotation we can see the rejection of the derivational/transformational approach to grammatical analysis which was prevalent at the time, and a focus instead on mutually constraining, parallel structures. This is the hallmark of LFG.

4.3.2 Two levels of syntax

Historically, LFG started life as a theory of syntax, and so its initial focus was here. Unlike other (so-called MONOSTRATAL) generative approaches, LFG posits two distinct levels of syntactic structure: CONSTITUENT STRUCTURE (c-structure) and FUNCTIONAL STRUCTURE (f-structure). While it may seem obvious that phonology and semantics, say, require the use of different kinds of formal object to describe, it may seem less apparent that syntax is itself divided in this way. However, it is not a unitary phenomenon: what we call 'syntax' refers both to superficial, 'surfacey' phenomena such as word order (which vary widely across the world's languages) and to more abstract, 'deeper' phenomena such as subjecthood (which exhibit many more commonalities cross-linguistically). LFG formalises this distinction by using two different kinds of formal object to represent these

¹⁴This fact makes it all the more striking that so much recent work in the Chomskyan paradigm, e.g. in the Minimalist Program, has tended to eschew proper formalisation (Asudeh & Toivonen 2006); although see Stabler (1997, 2011) and Harkema (2001) for computational work on 'minimalist grammars' which attempts to remedy this.

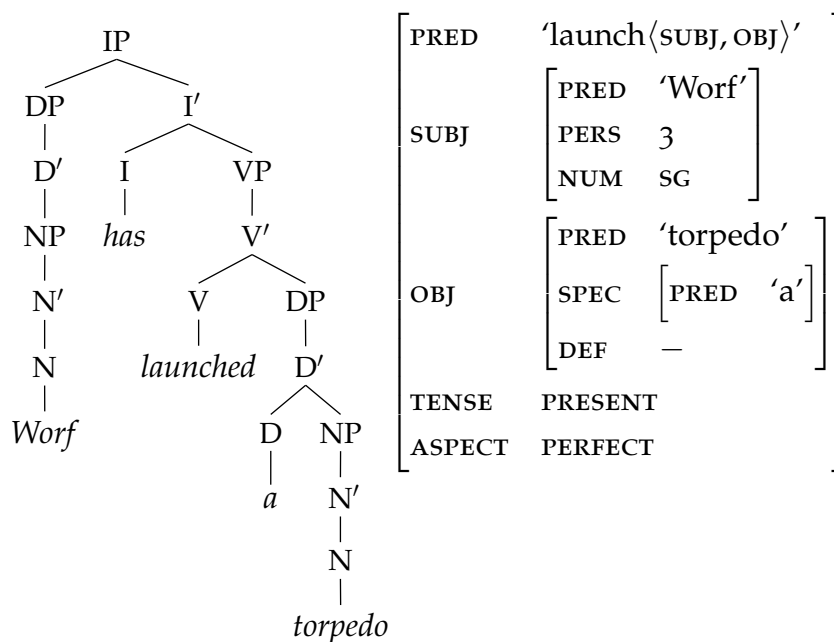


Figure 4.2: C- and f-structures for *Worf has launched a torpedo*

two different kinds of information. C-structure consists of a familiar phrase-structure tree, which is used to encode linear order, constituency, and lexical category information, while f-structure is an attribute-value matrix (AVM), also known as a feature structure, which represents abstract relational information about grammatical functions, agreement, long-distance dependencies, etc. An example of a c- and f-structure for *Worf has launched a torpedo* is given in Figure 4.2. In the next two sections, I will discuss in more detail what it is that these structures represent and how they do so. In Section 4.3.5 we will turn to the question of how they are connected, and how we go about analysing a sentence.

There are several motivations for separating out these two kinds of syntactic information, but one of the most significant is the ability to describe a wide variety of languages as neutrally as possible, without needing to propose (sometimes quite extraordinary) transformations in order to accommodate the various constituent orders exhibited by the world's languages. A subject is as much of a subject in a language where it precedes the verb, like English (SVO) or Urdu (SOV), as it is in a language where it follows it, like Irish (VSO) or Malagasy (VOS). But if subjecthood is seen as merely a reflex of a certain

phrase structure configuration, then whatever configuration is taken to encode subjecthood will be treated as basic, and the other constituent orders have to be derived from it in some way. If, however, subjecthood is represented at a level where order does not matter, then we can represent it in exactly the same way across languages. This is the approach LFG takes: the locus of cross-linguistic similarity is f-structure, not some (underlying/derivationally prior) level of c-structure. At the same time, this also allows c-structure to be a more direct representation of surface constituent order than it is in other, transformational theories, where the phrase structure tree must also accommodate a large amount of functional information, and so necessarily includes terminal nodes which do not correspond to pronounced elements of the string (e.g. traces/copies, pro/PRO, 'operators', etc.). LFG c-structures are much more 'WYSIWYG' ('What You See Is What You Get'), and unpronounced functional elements are represented at f-structure if they are represented at all.¹⁵

4.3.3 C-structure

Trees are natural formal objects with which to represent certain kinds of syntactic information, especially constituency, linear order, and category information: each node in a tree can be seen as the root of a smaller sub-tree, standing for a constituent; the nodes are (partially) ordered;¹⁶ and the nodes can be labelled with grammatical categories (both lexical, e.g. N, and phrasal, e.g. NP). Other kinds of grammatical information, though, such as grammatical relations like 'subject', are not so well served by phrase-structure trees. Such properties are not inherently configurational, and so cannot be reduced to basic properties of trees; rather, if they are to be represented in a tree, they must be seen as derivative

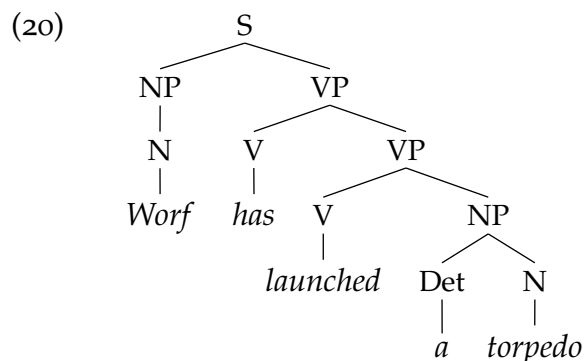
¹⁵Some earlier versions of LFG did continue to make use of traces at c-structure, including Kaplan & Bresnan (1982), along with e.g. Bresnan (1995, 1998); but since the introduction of functional uncertainty (Kaplan et al. 1987), it is not usually seen as part of the mainstream theory—various handbook and textbook presentations of LFG, such as Dalrymple (2001) and Falk (2001) do not employ traces, for instance, although they do still appear in Bresnan et al. (2016), albeit only in marginal cases.

¹⁶Nodes do not precede or follow their ancestors or descendants, hence the rider, but otherwise are fully ordered.

of one of the other properties, like linear order. But, as discussed above, this forces us to make certain potentially uncomfortable theoretical choices simply in order to describe the facts of e.g. subjecthood and objecthood—as Kroeger (2007: 294) puts it, we end up constrained by “dogmatic assertions which make the descriptive task more difficult”. LFG takes this problem seriously, and therefore uses formally appropriate data structures to model different kinds of data. C-structure trees are used only to represent the things which they are good at representing, i.e. constituency, linear order, and category information.

LFG’s c-structures are generated by a context-free grammar (CFG) standardly taken to be grounded in *X’ THEORY* (Jackendoff 1977). However, in keeping with LFG’s focus on descriptive adequacy above theoretical dogmatism, analyses frequently diverge from *X’* orthodoxy. For instance, some languages, especially those with relatively ‘free’ word orders, make extensive use of the exocentric sentential category *S*, which does not obey standard *X’*-theoretic restrictions. In addition, several researchers follow a version of the principle of *ECONOMY OF EXPRESSION*, whereby c-structure nodes are omitted when they do not contribute important functional information (for instance, the *N’* nodes in Figure 4.2—see Bresnan et al. 2016: 89–94 for an overview and Dalrymple et al. 2015 for further discussion). See Dalrymple (2001: Ch. 3) and Dalrymple & Findlay (2019: §3.1) for more on the properties of c-structure as commonly employed in LFG.

The theory of c-structure which I will be assuming is a very pared down one, following the TAG convention of using more traditional categories and avoiding bar levels altogether. In this version of the theory, the c-structure for *Worf has launched a torpedo* looks like (20), rather than as it appears in Figure 4.2:



There are two main motivations for this. The theory developed in the following chapter, which encodes descriptions of whole trees in lexical entries, makes the more articulated c-structure otiose, and so rather than introduce a more complex X'-theoretic c-structure and then abandon it, I have chosen instead to use the simpler version from the start. The other reason is that such a representation is more in keeping with the WYSIWYG goal of c-structure—it reduces the number of unary branching nodes, for example, which are only ever needed for theory-internal rather than empirically-motivated reasons.

4.3.4 F-structure

In order to represent the kinds of syntactic information for which trees are less well suited, LFG uses f-structures. These are represented as feature matrices, and, formally speaking, are functions from their attributes to their values. This means they can be defined in set-theoretic terms: for instance, (21a) could be written instead as the set of ordered pairs shown in (21b):

- (21) a.
$$\begin{bmatrix} \text{ATTRIBUTE1} & \text{VALUE1} \\ \text{ATTRIBUTE2} & \text{VALUE2} \end{bmatrix}$$
- b.
$$\{\langle \text{ATTRIBUTE1}, \text{VALUE1} \rangle, \langle \text{ATTRIBUTE2}, \text{VALUE2} \rangle\}$$

Note that since sets are unordered, f-structures are also unordered; (22) is the same f-structure as (21a):

- (22)
$$\begin{bmatrix} \text{ATTRIBUTE2} & \text{VALUE2} \\ \text{ATTRIBUTE1} & \text{VALUE1} \end{bmatrix}$$

Various other properties follow from the functional nature of f-structure: most notably, that each attribute can only have a single value. This is sometimes referred to as the principle of CONSISTENCY or the UNIQUENESS CONDITION (see Kaplan & Bresnan 1982 and Dalrymple 2001: 39 for discussion). One area where this is linguistically relevant is in agreement. An English present tense verb marked with *-s* indicates that its subject is singular, for example, and results in ungrammaticality if used with a plural subject. This ungrammaticality follows straightforwardly from the functional nature of f-structure, since there will be two conflicting statements regarding the value of the subject's NUM attribute in a sentence like (23)—the verb says it is SG, while the noun says it is PL (we will see how these constraints are implemented formally in the following section):

$$(23) \quad *The\ students\ dances.$$

$$\left[\begin{array}{ll} \text{PRED} & \text{'dances<SUBJ>'} \\ \text{SUBJ} & \left[\begin{array}{ll} \text{PRED} & \text{'students'} \\ \text{NUM} & \text{SG/PL} \end{array} \right] \end{array} \right]$$

Such a structure is impossible, since a function can only have one value for a given argument, and so the sentence is excluded by the grammar.

The values of f-structure attributes can themselves be f-structures, or sets of f-structures, allowing for complex nested structures. This is demonstrated below, where the value of ATTRIBUTE2 is a further f-structure, and the value of ATTRIBUTE5 is a set of two f-structures:

$$(24) \quad \left[\begin{array}{ll} \text{ATTRIBUTE1} & \text{VALUE1} \\ \text{ATTRIBUTE2} & \left[\begin{array}{ll} \text{ATTRIBUTE3} & \text{VALUE3} \\ \text{ATTRIBUTE4} & \text{VALUE4} \end{array} \right] \\ \text{ATTRIBUTE5} & \left\{ \left[\begin{array}{ll} \text{ATTRIBUTE6} & \text{VALUE6} \\ \text{ATTRIBUTE7} & \text{VALUE7} \end{array} \right], \left[\begin{array}{ll} \text{ATTRIBUTE8} & \text{VALUE8} \end{array} \right] \right\} \end{array} \right]$$

What is more, f-structure values can be shared when that value is itself an f-structure; that is, a single f-structure can be the value of multiple attributes. This is conventionally illustrated by drawing a line from one occurrence to another, as below:

$$(25) \quad \left[\begin{array}{l} \text{ATTRIBUTE1} \\ \text{ATTRIBUTE2} \end{array} \left[\begin{array}{ll} \text{A1} & \text{v1} \\ \text{A2} & \text{v2} \end{array} \right] \right]$$

This convention helps to make clear the fact that the values of `ATTRIBUTE1` and `ATTRIBUTE2` are not merely type identical but actually token identical—the very same f-structure is the value of both. Each unique f-structure is only ever represented on the page once.

This structure sharing is generally used in analyses of raising and control (Bresnan 1982a; Dalrymple 2001: Ch. 12; Asudeh 2005), as well as of long-distance dependencies (Dalrymple et al. 1995a; Kaplan & Zaenen 1989; Dalrymple 2001: Ch. 14). Some knowledge of the LFG treatment of long-distance dependencies will be useful in later chapters, but I will not present it here, instead introducing the relevant aspects when they are needed; see the citations given for more information.

Some of the most important f-structure attributes used in LFG are the `GRAMMATICAL FUNCTIONS` (GFs). These are abstract syntactic labels used to characterise the relations between different elements of a sentence, and many of them are familiar from traditional grammars. They are taken to be theoretical primitives, not merely inferred from other properties of the analysis. The complete list is given in Table 4.1 (taken from Lipps 2011: 44; see Dalrymple 2001: 8–28 for further discussion). It is also conventional to define the abbreviation `GF` as a meta-category which represents a disjunction over all the possible attributes listed in Table 4.1, thereby allowing for easy reference to a non-specific attribute:

$$(26) \quad \text{GF} \equiv \{ \text{SUBJ} \mid \text{OBJ} \mid \text{OBJ}_\theta \mid \text{COMP} \mid \text{XCOMP} \mid \text{OBL}_\theta \mid \text{ADJ} \mid \text{XADJ} \}$$

LFG abbreviation	Grammatical function
SUBJ	Subject
OBJ	Object
OBJ _θ	Object (indexed by thematic role)
COMP	Closed sentential complement
XCOMP	Open sentential complement
OBL _θ	Oblique (indexed by thematic role)
ADJ	Adjunct
XADJ	External (open) adjunct

Table 4.1: Grammatical functions in LFG

There are many other attributes which are used in LFG, including `NUMBER`, `PERSON`, `GENDER`, `TENSE`, `ASPECT`, etc. These are represented at f-structure to the extent that they participate in syntactic processes such as agreement. Unlike the `GFS`, which take whole f-structures as their values, many of these other attributes take what are called `ATOMIC` values, such as `SG` or `PL` (potential values of `NUM`, for instance). These are single symbols which have no internal structure. Finally, f-structure values can be `SEMANTIC FORMS`. These are the values of `PRED` features, to which we now turn.

The feature `PRED` was originally used in LFG for two things: to capture the semantic content associated with a particular f-structure, and to describe subcategorisation information. The value of a `PRED` feature, as we have said, is called a semantic form, and is enclosed in single quotes, as in (27):

(27) $\left[\text{PRED } \text{'yawn} \langle \text{SUBJ} \rangle \right]$

The subcategorisation information is described by the `GFS` appearing after the name of the `PRED` value (those appearing inside the angled brackets are thematic, those appearing outside, non-thematic). Example (27) says that this f-structure corresponds to the meaning 'yawn' and that this predicate subcategorises for a (thematic) subject.

Both of these roles for the `PRED` feature are somewhat redundant in modern LFG, since the theory of semantics assumed, Glue Semantics, to which we turn in Section 4.4.2, handles both of them—see Andrews (2008) for some discussion of

this point. Capturing the semantic content of a word is obviously better served by a full theory of semantics rather than being represented at f-structure, which is after all a level of *syntax*, not semantics. Subcategorisation is less purely semantic, but a moment's reflection shows that notions like incoherence (too many arguments) or incompleteness (too few) can easily be captured semantically as well as syntactically. For example, (28a) is, syntactically speaking, missing an oblique (OBL_{GOAL}) argument, but it is also, semantically speaking, missing one of the arguments needed to saturate the predicate. Similarly, (28b) contains an OBJ which is not subcategorised for, but it also contains a semantic expression which will be left over once the predicate has composed with its single argument.

- (28) a. *Trevor put the mug.
 b. *Elizabeth yawned Tracey.

Now, since the semantic theory is independently motivated, and since it offers a straightforward account of the ungrammaticality of (28a)–(28b), it becomes unnecessary to also represent subcategorisation at f-structure. And so, in what follows, I will omit subcategorisation information from semantic forms, e.g. writing (27) as (29), instead:

- (29) [PRED 'yawn']

This is not a new suggestion, and the possibility was already noted by Dalrymple et al. (1999a) and Kuhn (2001). Recently, work by Asudeh (2012), Asudeh & Giorgolo (2012), and my own work on mapping theory (Findlay 2016a) have brought the idea further into the LFG mainstream.¹⁷

One remaining role for PRED values is that of differentiating f-structures. Since f-structures are set-theoretic objects, they are subject to the Axiom of Extension, whereby any two sets which contain the same members are identical.¹⁸ This

¹⁷One *prima facie* problem facing such a move is the existence of expletive arguments, i.e. arguments which are selected for by the syntax but explicitly *not* by the semantics, such as the dummy subjects required by English weather verbs. This problem is not insurmountable, however: see Asudeh (2012: 113) for some suggestions.

¹⁸For more on set theory, see Partee et al. (1990: Part A).

means that two f-structures which contain identical attribute-value pairs are in fact one and the same f-structure. However, in a sentence like (30), we do not want the subject and object f-structures to be identified, just because they both describe 3rd person, singular, masculine nouns; but without some other feature to distinguish them, this is what will happen.

(30) Julian frustrates Benjamin.

However, simply adding an additional feature (PRED) would not in itself be enough: if we had the *same* word in both argument positions (even if they are referentially distinct), then the PRED values would also be the same, and the problem would return.

(31) Julian hates Julian.

Instead, PRED values have the property known as UNIQUE INSTANTIATION: equality between uniquely instantiated values fails. That is, PRED values are distinct even when they have the same form. Asudeh (2012: 113) claims that even this aspect of PRED values can be subsumed under our semantic account, since each word will contribute a unique meaning resource. However, if otherwise distinct f-structures are allowed to be equated, this may cause problems for Glue Semantics, which relies on f-structure (either directly or indirectly, via s-structure—see below) to differentiate semantic arguments. Any obstacles this raises may be surmountable, but until such a theory is fully developed, it seems prudent to retain the PRED feature at least for its uniqueness properties. It also serves the not insignificant function of making f-structures easier to read, since it is clearer which f-structure corresponds to which word in the sentence when they are labelled with their PRED values.

4.3.5 Connecting c-structure and f-structure

C-structure and f-structure are linked by a function, ϕ , whose domain is the set of c-structure nodes, and whose range consists of f-structures. The phrase

structure rules (PSRs) of the context-free grammar which generates c-structures are annotated with constraints which describe how the c-structure configurations are related to f-structures. Lexical entries can be seen as special kinds of phrase structure rules (with a single terminal symbol on their right-hand side), and so can be similarly annotated. The general CFG rules introduce configurational information, while lexical entries give us the idiosyncratic lexical information. For instance, the fact that the left-most NP is the subject in canonical declarative sentences in English can be captured by the following PSR, where $*$ refers to the current node (the one which bears the annotation) and $\hat{*}$ refers to its mother:

$$(32) \quad S \rightarrow \begin{array}{cc} \text{NP} & \text{VP} \\ (\phi(\hat{*}))(\text{SUBJ}) = \phi(*) & \phi(\hat{*}) = \phi(*) \end{array}$$

This says that an S can consist of an NP and a VP, and that the sentential f-structure, when applied to the attribute SUBJ, returns the NP's f-structure (i.e. that the NP is the subject of the sentence), while the VP corresponds to the same f-structure as the whole clause.

Several useful notational shorthands are employed to improve the readability of these annotations. First of all, we define two abbreviations for commonly used expressions; namely, applying ϕ to the current node or its mother:

$$(33) \quad \begin{array}{ll} \text{a.} & \phi(*) \equiv \downarrow \\ \text{b.} & \phi(\hat{*}) \equiv \uparrow \end{array}$$

These arrows were originally intended to be iconic, since annotations were written above the nodes that bore them in early LFG work, so that \downarrow would point to the current node, and \uparrow to its mother. However, it has now become standard practice to write annotations below nodes (this improves readability when nodes bear multiple annotations, especially on terminal (lexical) nodes, for example), and so the iconicity has been broken (\uparrow now points to the current node, but actually refers to its mother's f-structure). This is a regrettable potential source of confusion, but the convention remains.

Secondly, since f-structures are always functions of only one argument, no ambiguity arises from removing the parentheses and taking functional application to be left associative (Kaplan & Bresnan 1982: 183). For the present example this is trivial, since e.g. (\uparrow SUBJ) only contains a single attribute, but there are other cases where we want to be able to talk about more deeply nested attributes, such as the subject's number or person. As mentioned above, English third-person singular verb forms include such constraints: e.g. *loves* states that its subject's number is singular and that its subject's person is 3rd. (34) shows how this would look written out in full, while (35) uses the abbreviations and associativity convention, making it much more readable:

$$(34) \quad V \rightarrow \begin{array}{l} \textit{loves} \\ (\phi(\hat{*}))(\text{PRED}) = \textit{'love' } \\ ((\phi(\hat{*}))(\text{SUBJ}))(\text{NUM}) = \textit{SG} \\ ((\phi(\hat{*}))(\text{SUBJ}))(\text{PERS}) = 3 \end{array}$$

$$(35) \quad V \rightarrow \begin{array}{l} \textit{loves} \\ (\uparrow \text{PRED}) = \textit{'love' } \\ (\uparrow \text{SUBJ NUM}) = \textit{SG} \\ (\uparrow \text{SUBJ PERS}) = 3 \end{array}$$

Although formally speaking lexical entries are just phrase-structure rules, it is common to write them in a slightly different format; (36) shows how this would look for (35), as an example:

$$(36) \quad \textit{loves} \quad V \quad \begin{array}{l} (\uparrow \text{PRED}) = \textit{'love' } \\ (\uparrow \text{SUBJ NUM}) = \textit{SG} \\ (\uparrow \text{SUBJ PERS}) = 3 \end{array}$$

Once a parse has been obtained for a sentence using the context-free grammar rules, the associated constraints are collected into what is called a FUNCTIONAL DESCRIPTION. The f-structure associated with the sentence is the minimal structure which satisfies all of the constraints in the functional description. Let us go through an example to see how this works.

We will give an analysis for the sentence in (37):

$$(37) \quad \textit{Keiko loves the gift.}$$

This relies on the following rules (including lexical entries), which form part of the grammar of English:

(38) **C-structure rules:**

$$\text{a. } S \rightarrow \text{NP} \quad \text{VP} \\ (\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow$$

$$\text{b. } \text{NP} \rightarrow \text{N} \\ \uparrow = \downarrow$$

$$\text{c. } \text{NP} \rightarrow \text{Det} \quad \text{N} \\ \uparrow = \downarrow \quad \uparrow = \downarrow$$

$$\text{d. } \text{VP} \rightarrow \text{V} \quad \text{NP} \\ \uparrow = \downarrow \quad (\uparrow \text{OBJ}) = \downarrow$$

(39) **Lexical entries:**

$$\text{a. } \textit{Keiko} \quad \text{N} \quad \begin{array}{l} (\uparrow \text{PRED}) = \text{'Keiko'} \\ (\uparrow \text{NUM}) = \text{SG} \\ (\uparrow \text{PERS}) = 3 \\ (\uparrow \text{GEN}) = \text{F} \end{array}$$

$$\text{b. } \textit{loves} \quad \text{V} \quad \begin{array}{l} (\uparrow \text{PRED}) = \text{'love'} \\ (\uparrow \text{SUBJ NUM}) = \text{SG} \\ (\uparrow \text{SUBJ PERS}) = 3 \\ (\uparrow \text{TENSE}) = \text{PRES} \end{array}$$

$$\text{c. } \textit{the} \quad \text{D} \quad \begin{array}{l} (\uparrow \text{SPEC PRED}) = \text{'the'} \\ (\uparrow \text{DEF}) = + \end{array}$$

$$\text{d. } \textit{gift} \quad \text{N} \quad \begin{array}{l} (\uparrow \text{PRED}) = \text{'gift'} \\ (\uparrow \text{NUM}) = \text{SG} \\ (\uparrow \text{PERS}) = 3 \end{array}$$

From these, we can produce the c-structure in Figure 4.3. I have labelled each node in this c-structure with the name of the f-structure which corresponds to it (via ϕ). Doing this allows us to instantiate the metavariables \uparrow and \downarrow in the annotations, and from there we can collect up all of the constraints into the following functional description:

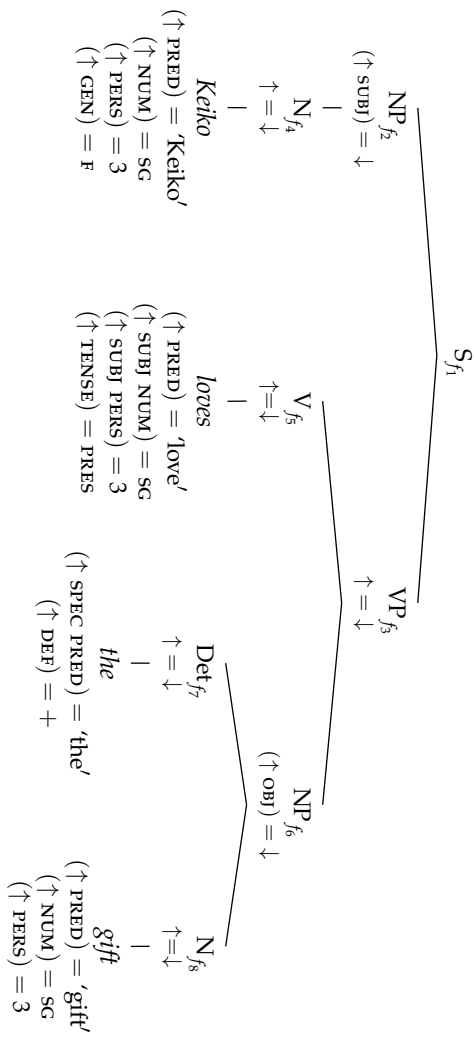


Figure 4.3: Annotated c-structure for *Keiko loves the gift*

$$\begin{array}{ll}
 (40) & (f_1 \text{ SUBJ}) = f_2 & (f_4 \text{ GEN}) = F \\
 & f_1 = f_3 & (f_5 \text{ PRED}) = \text{'love'} \\
 & f_2 = f_4 & (f_5 \text{ SUBJ NUM}) = \text{SG} \\
 & f_3 = f_5 & (f_5 \text{ SUBJ PERS}) = 3 \\
 & (f_3 \text{ OBJ}) = f_6 & (f_5 \text{ TENSE}) = \text{PRES} \\
 & f_6 = f_7 & (f_7 \text{ SPEC PRED}) = \text{'the'} \\
 & f_6 = f_8 & (f_7 \text{ DEF}) = + \\
 & (f_4 \text{ PRED}) = \text{'Keiko'} & (f_8 \text{ PRED}) = \text{'gift'} \\
 & (f_4 \text{ NUM}) = \text{SG} & (f_8 \text{ NUM}) = \text{SG} \\
 & (f_4 \text{ PERS}) = 3 & (f_8 \text{ PERS}) = 3
 \end{array}$$

A functional description like (40) is essentially a set of simultaneous equations to be solved, and the f-structure for the sentence is the smallest f-structure which satisfies all of them. By using substitution, it is possible to convert all of the constraints into statements about the sentential f-structure, f_1 , which makes it clear how we can reach such a structure. (41) shows, for example, how we might determine the PRED value of the subject:

$$\begin{array}{ll}
 (41) & (f_1 \text{ SUBJ PRED}) = (f_2 \text{ PRED}) \quad (\text{since } (f_1 \text{ SUBJ}) = f_2) \\
 & (f_1 \text{ SUBJ PRED}) = (f_4 \text{ PRED}) \quad (\text{since } f_2 = f_4) \\
 & (f_1 \text{ SUBJ PRED}) = \text{'Keiko'} \quad (\text{since } (f_4 \text{ PRED}) = \text{'Keiko'})
 \end{array}$$

By continuing in this fashion for all of the constraints in (40), we arrive at the f-structure in Figure 4.4, where the individual f-structures have been decorated with their labels (for a full account of the solution algorithm for LFG, see Kaplan & Bresnan 1982: 189ff.).

Notice that the f-description in (40) is compatible with other f-structures (an infinite number, in fact), which can be obtained by adding extra features. Figure 4.5 gives an example of one such f-structure. This contains a spurious PERSON feature in the clausal f-structure, and an unmotivated masculine pronominal $\text{OBL}_{\text{COAL}} \text{ GF}$, but nonetheless satisfies all of the equations in (40) (ignoring any semantic issues). However, since there is a smaller f-structure which also satisfies them, namely the one shown in Figure 4.4, the f-structure in Figure 4.5 is rejected by the requirement that we find the *minimal* such f-structure.

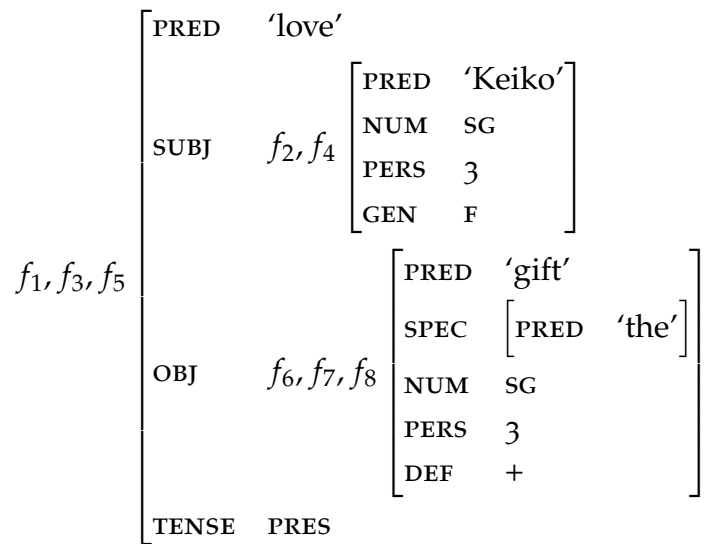


Figure 4.4: F-structure for *Keiko loves the gift*

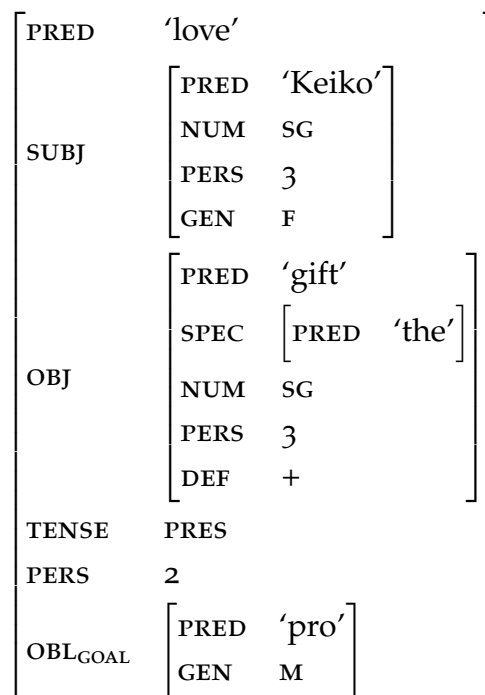


Figure 4.5: Incorrect, non-minimal f-structure for *Keiko loves the gift*

The kinds of equations seen in (40), which state that certain f-structures or their values are equal, are known as DEFINING EQUATIONS. This is because by their presence they define the corresponding f-structure to be the way they say it is; they have a “make it so” character (Bresnan et al. 2016: 59). However, it is also sometimes useful to describe f-structures in a different way, using CONSTRAINING EQUATIONS. Rather than stating that such-and-such a situation *is* the case, these state that it *must* be the case. This difference is necessitated by the fact that, unlike HPSG signs, for example, f-structures need not be fully specified—that is, we do not insist that every f-structure contain an attribute-value pair for every relevant attribute (and, besides, since f-structures aren’t typed, we don’t even have a meaningful notion of ‘relevant’—there are no ‘nominal’ or ‘clausal’ f-structures except as derivative concepts).

Consider what this means for constraints that heads might place on the form of their dependents. An auxiliary like progressive *is*, for example, requires that the verb which heads its open complement clause (xCOMP) be a present participle (the *-ing* form of the verb).¹⁹ For example, we have the contrast in (42):

- (42) a. Miles is winning.
b. *Miles is wins.

However, given the following, apparently sensible, lexical entries for *is*, *winning*, and *wins*, we fail to rule out (42b), for reasons that will become clear.²⁰

¹⁹This example is based on Kaplan & Bresnan (1982: 205–209). Modern LFG treatments of auxiliaries would not necessarily agree with the formal details presented here, being more likely to treat auxiliaries as f-structure co-heads, contributing tense and aspect information but not PREDs of their own (e.g. Falk 2003), and handling the form restrictions via constraints at a different level of structure, e.g. m-structure (Butt et al. 1996a). All of this notwithstanding, I think it is an intuitively clear example with which to make the formal point.

²⁰The second-to-last equation in (43) establishes a structure-sharing relationship, as discussed above, which identifies the subject of the main clause and the subject of the xCOMP, thus ensuring that the auxiliary and the main verb share a common subject.

$$\begin{aligned}
 (43) \quad & is \quad V \quad (\uparrow \text{PRED}) = \text{'is'} \\
 & \quad \quad (\uparrow \text{SUBJ NUM}) = \text{SG} \\
 & \quad \quad (\uparrow \text{SUBJ PERS}) = 3 \\
 & \quad \quad (\uparrow \text{TENSE}) = \text{PRES} \\
 & \quad \quad (\uparrow \text{XCOMP SUBJ}) = (\uparrow \text{SUBJ}) \\
 & \quad \quad (\uparrow \text{XCOMP PARTFORM}) = \text{PRESENT}
 \end{aligned}$$

$$\begin{aligned}
 (44) \quad & winning \quad V \quad (\uparrow \text{PRED}) = \text{'win'} \\
 & \quad \quad (\uparrow \text{PARTFORM}) = \text{PRESENT}
 \end{aligned}$$

$$\begin{aligned}
 (45) \quad & wins \quad V \quad (\uparrow \text{PRED}) = \text{'win'} \\
 & \quad \quad (\uparrow \text{SUBJ NUM}) = \text{SG} \\
 & \quad \quad (\uparrow \text{SUBJ PERS}) = 3 \\
 & \quad \quad (\uparrow \text{TENSE}) = \text{PRES}
 \end{aligned}$$

Assume also that we have the requisite phrase-structure rule like (46), which allows verbs to take VP xCOMP arguments:

$$\begin{aligned}
 (46) \quad & VP \rightarrow V \quad VP \\
 & \quad \quad \uparrow = \downarrow \quad (\uparrow \text{xCOMP}) = \downarrow
 \end{aligned}$$

Now, the last constraint in (43) states that (the head of) its xCOMP has the form of a present participle. This means, for instance, that when *winning* takes this role, there is no clash, since its lexical entry, (44), specifies that it has the correct form—the value of its PARTFORM attribute is PRESENT, as required. Crucially, however, there will only be a clash when the xCOMP verb contains a *conflicting* specification for the value of its PARTFORM feature (e.g. *won* might contribute the pair ⟨PARTFORM, PAST⟩, which would be incompatible with the specification in (43)). If it contains *no* specification, there will be no clash. This means that (42b) will not be ruled out, since *wins* simply contains no specification for the feature PARTFORM, and so *is* is free to define a value for it without provoking a feature clash. As Kaplan & Bresnan (1982: 208) put it, “[w]e have concluded that [*wins*] is a present participle just because *is* would like it to be that way!”

To avoid this problem, we introduce a different kind of equation, the constraining equation, written with a subscript *c* on the equals sign. We thus rewrite the lexical entry for *is* as in (47):

$$\begin{aligned}
 (47) \quad is \quad V \quad & (\uparrow \text{PRED}) = \text{'is'} \\
 & (\uparrow \text{SUBJ NUM}) = \text{SG} \\
 & (\uparrow \text{SUBJ PERS}) = 3 \\
 & (\uparrow \text{TENSE}) = \text{PRES} \\
 & (\uparrow \text{XCOMP SUBJ}) = (\uparrow \text{SUBJ}) \\
 & (\uparrow \text{XCOMP PARTFORM}) =_c \text{PRESENT}
 \end{aligned}$$

The final equation here does not define the attribute-value pair into existence, but instead requires that something else contribute that information. Such equations are initially set aside during the analysis of a sentence, and an f-structure solution is obtained from the defining equations alone. This f-structure is then assessed against the constraining equations; if they are not satisfied, then the f-structure is rejected. Constraining equations thus serve as checks on the well-formedness of f-structures, rather than contributing directly to their construction or verification.

We could avoid the addition of a whole new kind of formal object to our theory by instead insisting that all features must have a value specified (as in HPSG, for example); Kaplan & Bresnan (1982: 208) suggest we could use the value `NONE` for the `PARTFORM` attribute of finite verbal forms which do not appear to merit another value, for instance. However, such full specification is undesirable for a number of reasons. Kaplan & Bresnan (1982: 208) point out that it makes the formal system more cumbersome for linguists using it, and that it makes the theory's representations much less plausible as characterisations of the linguistic knowledge that children acquire.

A related concern is that it makes the intended meaning of the constraint less clear. The claim that we are trying to formalise is that *is* requires that its complement *must* have the form of a present participle. It is better if we can encode that directly in our theory rather than doing so indirectly. Of the two approaches under discussion, the use of constraining equations is markedly more direct. It encodes the constraint in the lexical entry under discussion, while the alternative approach encodes it in the whole of the rest of the lexicon. To see how, imagine we want to constrain a word *A* to be of form *X*, and it is the case that all words of this form contribute a feature *F* with value *V*. Then, to pick out all the *Xs*, we can either require that the f-structure for *A* contain feature *F* with value

V , or we can *state* that it contains such a pair while simultaneously requiring that every lexical entry of the same category as A that is not X contribute an attribute-value pair $\langle F, W \rangle$ where $W \neq V$. It seems to me that the former is considerably more perspicuous: it refers only to the relevant attribute-value pair, and places no additional constraints on the form of the rest of the lexicon.

There are a number of other forms of constraining equation which are used in LFG. For instance, we can require that an attribute have *some* value, without specifying what it is; (48) is an example of such an EXISTENTIAL constraining equation, requiring the presence of a subject:

$$(48) \quad (\uparrow \text{SUBJ})$$

We can also require that a certain attribute-value pair *not* be present, using a NEGATIVE constraining equation:

$$(49) \quad (\uparrow \text{SUBJ PERS}) \neq 3$$

And we can combine these to require that an attribute not be present at all, using a negative existential constraining equation:

$$(50) \quad \neg(\uparrow \text{SUBJ})$$

Bresnan et al. (2016: 60–61) also discuss the use of CONDITIONAL constraining equations, written as (51), although these are less commonly used:

$$(51) \quad (\downarrow \text{CASE}) = \text{NOM} \Rightarrow (\uparrow \text{SUBJ}) = \downarrow$$

Such constraints have the force of their right-hand side, provided that their left-hand side is true (when checked against the minimal f-structure being considered). (51) says that the f-structure of the node bearing the annotation is the subject of its mother's f-structure, provided that it also bears nominative case.²¹

One consequence of the inclusion of constraining equations to the formalism is that LFG is not wholly constraint-based in the usual sense of the term.

²¹See Bresnan et al. (2016: 60–61) for some discussion of the precise interpretation such conditional constraining equations should receive.

Grammatical analysis is not simply a question of simultaneous constraint satisfaction; instead, it is constraint satisfaction followed by an additional step of verification against a further set of constraints. Another way of seeing this is as taking constraining equations to be higher-order constraints—they are essentially constraints on constraints, requiring (or barring) the presence or absence of certain defining equations, rather than of certain structures. This clearly adds to the complexity of the theory when compared with simply having a first-order constraint system. As always, there is a trade-off: as discussed above, the cost of avoiding such tools is added complexity in the substance of the theory; the cost of including them, however, is added complexity in the form of the theory.

4.3.6 The correspondence architecture

The LFG grammatical architecture originally consisted of just the two syntactic levels c-structure and f-structure, but Kaplan (1987) suggested generalising it to other levels of analysis. This is known as the CORRESPONDENCE, OR PARALLEL PROJECTION, ARCHITECTURE (Kaplan 1987, 1995; Halvorsen & Kaplan 1988). This modular approach to the grammar expands the approach outlined above, and treats different components of grammatical analysis as their own encapsulated levels, with their own formal properties, which are related to one another via correspondence functions. Several other levels of analysis have been proposed and developed, including a(rgument)-structure (e.g. Butt et al. 1997; Kibort 2007), m(orphosyntactic)-structure (e.g. Butt et al. 1996a), s(emantic)-structure (e.g. Dalrymple 1999; Lowe 2014), i(nformation)-structure (e.g. Dalrymple & Nikolaeva 2011), and p(rosodic)-structure (e.g. Mycock & Lowe 2013). This is one of the great strengths of LFG as a framework: it provides wide coverage of diverse linguistic phenomena in a modular fashion. However, in this thesis we will only be considering c- and f-structure, so I will say no more about the correspondence architecture here.

4.3.7 Templates

It can be convenient to abbreviate long descriptions such as we might find in lexical entries. LFG does this using `TEMPLATES` (Dalrymple et al. 2004; Crouch et al. 2012; Asudeh et al. 2013). A template is just a short-hand way of referring to a functional description. They can contain single equations, such as (52) or (53):²²

$$(52) \quad \text{SG-SUBJ} := \\ (\uparrow \text{SUBJ NUMBER}) = \text{SG}$$

$$(53) \quad \text{1-SUBJ} := \\ (\uparrow \text{SUBJ PERSON}) = \text{1}$$

But they can contain any number of equations, corresponding to any level of the correspondence architecture (along with meaning constructors, to be introduced in Section 4.4.2). For example, we can define a more complex template `1SG-SUBJ`:

$$(54) \quad \text{1SG-SUBJ} := \\ (\uparrow \text{SUBJ PERSON}) = \text{1} \\ (\uparrow \text{SUBJ NUMBER}) = \text{SG}$$

Of course, we can abbreviate this template itself using the simpler templates. When templates are used in descriptions, they are preceded by the '@' symbol, which represents a 'call' of the template:

$$(55) \quad \text{1SG-SUBJ} := \\ @\text{1-SUBJ} \\ @\text{SG-SUBJ}$$

Template invocation is simple substitution, so (55) is equivalent to (54).

In addition, we can parametrise templates to make them more flexible. For example, we can define a template `SUBJ-PERSON`, which takes a single argument, so that (57a) can be rewritten as (57b):

$$(56) \quad \text{SUBJ-PERSON}(X) := \\ (\uparrow \text{SUBJ PERSON}) = X$$

$$(57) \quad \text{a. } @\text{1-SUBJ} \\ \text{b. } @\text{SUBJ-PERSON}(1)$$

²²This presentation is drawn in large part from Findlay (2016a: 310–311).

We can do something similar for SUBJ-NUMBER, and then define a general SUBJECT template which takes two arguments, the person and the number of the predicate's subject:

$$(58) \quad \text{SUBJ-NUMBER}(X) := \\ (\uparrow \text{SUBJ NUMBER}) = X$$

$$(59) \quad \text{SUBJECT}(P, N) := \\ @\text{SUBJ-PERSON}(P) \\ @\text{SUBJ-NUMBER}(N)$$

Now (60a) can be rewritten as (60b):

$$(60) \quad \begin{array}{ll} \text{a.} & @1\text{SG-SUBJ} \\ \text{b.} & @\text{SUBJECT}(1, \text{SG}) \end{array}$$

4.3.8 Summary

Lexical Functional Grammar has a number of desirable properties. Its constraint-based nature means it is both readily computationally implementable and at the same time psychologically plausible. Its modular architecture has both theoretical and practical advantages: it directly encodes the fact that different aspects of human language differ in their formal properties, and it makes it easier for the researcher to work on one aspect of the grammar while holding the other components constant. And its descriptive flexibility makes it possible to analyse data from a variety of typologically diverse languages without forcing them into a framework which does not fit them. This last point, taken together with the relative accessibility of the formalism to linguists from a variety of backgrounds, and “a tradition of taking grammatical details seriously” (Kroeger 2007: 294), has meant that LFG is the tool of choice for a variety of fieldworkers, bringing much fruitful data to the attention of linguistic theory (e.g. Dahlstrom 1991, 2009; Austin & Musgrave 2008; Hemmings 2016).

However, the strictly lexical nature of the theory means that it does not have a natural means of describing MWEs—since lexical entries can only describe single terminal nodes in c-structure, LFG is forced into some version of the

lexical ambiguity approach described in the previous chapter (as in e.g. the proposal of Arnold 2015). As I argued there, however, this kind of approach is inadequate. What we need is some way to integrate the advantages of the construction-based approach which TAG provides into the LFG framework. The next chapter will demonstrate how we can accomplish this. Before that, however, I must say a little about the semantic framework I will be using, and how this connects to the LFG grammar.

4.4 Semantics

This section has two subsections, devoted to each of these questions. In the first, I outline the meaning language I will use, which is a version of neo-Davidsonian event semantics. In the second, I introduce *GLUE SEMANTICS*, a theory of the syntax-semantics interface, and show how we can use LFG f-structures to control semantic composition.

4.4.1 Meaning language

The semantic theory I am assuming takes its role to be determining the truth conditions of a sentence; that is, describing how the world would have to be for the sentence to be true. These conditions are given in the form of statements in a higher-order predicate logic. Logical languages have a number of advantages over natural languages for expressing meanings, not least of which is their unambiguous nature—each well-formed statement in an appropriately designed logic has one and only one interpretation. Another important component of the semantic theory is that it should derive the truth conditions of a sentence from the semantic contributions of its constituents and the manner in which they are combined syntactically; i.e. it should adhere to some version of the Principle of Compositionality.²³ The standard approach to compositionality is

²³Sometimes called Frege's Principle on the assumption that it has its origins with 19th-century mathematician and logician Gottlob Frege (cf. Dummett 1973), although e.g. Janssen (1997: 420ff.) disputes this origin. See Szabó (2017) for extended discussion and references.

that it involves functional application. The semantic contributions of words and phrases are therefore given in terms of the typed lambda calculus. I will not defend the merits of such a system any further here, since it is well-established practice in the ‘formal’ or ‘model-theoretic’ semantics literature. For a concise and easily readable motivation of such a theory, see Bach (1989), and for textbook introductions, see Heim & Kratzer (1998) or Chierchia & McConnell-Ginet (2000).

There are only two noteworthy differences to be found between the meaning language used here and the more canonical textbook presentations. The first is that although I assume an intensional logic, I do not make use of the intensional type *s* as in e.g. Montague (1973). Instead, following van Benthem (1988, 1991) (cf. also Asudeh & Toivonen 2012), I take the type of sentences, *t*, to stand for propositions (sets of possible worlds) rather than truth values. In other words, its domain consists of the power set of the set of worlds, rather than of the two elements 1 and 0 (or T and F). Full details of the types I am assuming are given in Appendix B.

The second point of departure is that I will be using a neo-Davidsonian event semantics, whereby verbs are seen as predicates which take an EVENTUALITY argument—where the set of eventualities is defined as the union of the set of EVENTS and STATES; cf. Bach (1981)—and which contribute a number of other semantic role predicates which specify the various participants in the event or state.²⁴ The semantic representation in (61) for the verb *repair* illustrates the approach:²⁵

$$(61) \quad \lambda x \lambda y \lambda e. \mathbf{repair}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y$$

²⁴Davidson (1967) is the original work which proposes to treat events as implicit arguments of certain predicates. See Parsons (1990) for a more thoroughgoing implementation, including the separation of semantic roles from the main predicate; Parsons (1995) and Pietroski (2002: 97–102) give further motivation, especially for this latter move.

²⁵By convention, I use **boldface** for predicates and individual constants, Roman type for connectives, brackets, etc., and *italics* for variables. To avoid clutter in meaning representations, I also employ the following abbreviations for commonly used semantic role names: **ag** = agent, **pt** = patient, **th** = theme, **goal** = goal, **ben** = beneficiary, **exp** = experiencer.

Such an expression takes three arguments: the two nominal arguments corresponding to the agent and the patient, and the event argument. Once it has combined with its nominal arguments, the verb's denotation is a set of events; in this case it is the set of events where x repairs y .

In many standard compositional realisations of event semantics, the event argument is existentially closed at the sentence level by some kind of closure operator, e.g. (62), with the result that the full meaning for a sentence like *Miles repaired the computer* is (63):^{26,27}

$$(62) \quad \llbracket[\text{closure}]\rrbracket = \lambda F_{\langle vt \rangle} . \exists e [F(e)]$$

$$(63) \quad \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)]]$$

However, this approach causes problems when it comes to other kinds of quantificational expressions. As Champollion (2015: 33) explains, “[i]n many implementations of event semantics in compositional frameworks, accounts of scope-taking expressions such as quantifiers, negation, and conjunctions need to be complicated compared with the more standard treatments that would be available if events were not present”. Part of the reason for this is due to the lack of scopal interactions between things like quantifiers or negation and the existential binding of the event variable.

In general, quantificational elements can interact scopally with one another, so that when a sentence contains more than one of them it becomes ambiguous, corresponding to multiple different translations in our meaning language. Some examples are given below (omitting events from consideration for the time being):

²⁶ v is the type of events; F is thus a function from events to propositions, i.e. the type of verbal expressions once they have combined with all of their arguments.

²⁷I abstract away from the semantics of definite descriptions and simply give an unanalysed meaning for NPs like *the computer* using the iota-operator, a formalism first introduced by Whitehead & Russell (1910: *14) (although notice that these authors use a turned iota, ι , instead). The expression $\iota x [\mathbf{computer}(x)]$ is to be read as “the unique (contextually salient) x such that x is a computer”.

- (64) \llbracket Every student saw a badger. $\rrbracket =$
 $\forall x[\mathbf{student}(x) \rightarrow \exists y[\mathbf{badger}(y) \wedge \mathbf{see}(x, y)]]$ $\forall \gg \exists$
 (every student saw some badger, but not necessarily the same one)
 OR
 $\exists y[\mathbf{badger}(y) \wedge \forall x[\mathbf{student}(x) \rightarrow \mathbf{see}(x, y)]]$ $\exists \gg \forall$
 (there is at least one badger that every student saw)
- (65) \llbracket Kirk didn't kiss every alien. $\rrbracket =$
 $\neg[\forall x[\mathbf{alien}(x) \rightarrow \mathbf{kiss}(\mathbf{kirk}, x)]]$ $\neg \gg \forall$
 (Kirk didn't kiss all the aliens, but he might have kissed some)
 OR
 $\forall x[\mathbf{alien}(x) \rightarrow \neg[\mathbf{kiss}(\mathbf{kirk}, x)]]$ $\forall \gg \neg$
 (every alien is such that Kirk didn't kiss them—he kissed *no* alien)

But the event quantification does not interact in this way; instead, it always scopes as low as possible (cf. Champollion 2015: 36–37):

- (66) \llbracket Kirk kissed every alien. $\rrbracket =$
 $\forall x[\mathbf{alien}(x) \rightarrow \exists e[\mathbf{kiss}(e) \wedge \mathbf{ag}(e) = \mathbf{kirk} \wedge \mathbf{pt}(e) = x]]$ $\forall \gg \exists e$
 (for every alien, there is a different kissing event)
 NOT
 $\exists e[\mathbf{kiss}(e) \wedge \mathbf{ag}(e) = \mathbf{kirk} \wedge \forall x[\mathbf{alien}(x) \rightarrow \mathbf{pt}(e) = x]]$ $*\exists e \gg \forall$
 (there is a single kissing event with every alien as the theme)
- (67) \llbracket Spock didn't smile. $\rrbracket =$
 $\neg[\exists e[\mathbf{smile}(e) \wedge \mathbf{ag}(e) = \mathbf{spock}]]$ $\neg \gg \exists e$
 (there is no smiling event of which Spock is the agent)
 NOT
 $\exists e[\neg[\mathbf{smile}(e) \wedge \mathbf{ag}(e) = \mathbf{spock}]]$ $*\exists e \gg \neg$
 (there is an event in which Spock does not smile)

The second translation in (66) is problematic because intuitively we think of each kissing as its own event, but this translation only makes available a single,

macro-level kissing event.²⁸ Another concern is that semantic role predicates are generally thought of as functional: that is, a single event can only have one theme. But in the inappropriate translation, the expression $\forall x[\mathbf{alien}(x) \rightarrow \mathbf{th}(e) = x]$ is under the scope of the event quantification, which means that every alien is said to be the theme of the *same* event. On the assumption that **th** is functional, then, this translation will be false in any model with more than one alien.

Similar concerns arise with (67): in the inappropriate second translation, the sentence means that there is some event in which Spock doesn't smile. But almost all events are of this kind! Assuming our model contains at least one event in which Spock doesn't smile, then this sentence will be trivially true.²⁹

So, if the existential quantification of the event variable does not interact with other scope-taking expressions, it seems inappropriate to introduce it only at the sentence level—without further stipulation this implies the existence of readings which are not attested. Various proposals have been made to remedy this, but, as noted above, they usually involve additional complexification of other components of the semantics, namely the scope-taking expressions like quantifiers or negation. A different solution is to change how we introduce the existential quantification of the event variable; we could, for instance, include it in the lexical entry for the verb, which would ensure that it scopes below any arguments or other modifiers of the verb like negation:

$$(68) \quad \lambda x \lambda y. \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y]$$

This correctly rules out the unwanted readings, but is now *too* restrictive—we need the event argument to still be available to adverbial expressions, for instance, which are taken to also be predicates of events:

²⁸There might well be a need for some complex event corresponding to the higher-scope quantification here, but this must be as well as the lower-level quantification which scopes under the universal quantifier, not instead of it. See Champollion (2010) for discussion of some of the issues.

²⁹Here we are ignoring the role of tense in semantics—obviously sentence (67) does not mean the same thing as 'Spock never smiles', i.e. that there is no such event *at any time*; rather, it means that at the one arbitrarily fixed time (in the past) which the model represents there is no such event.

$$(69) \quad \llbracket \text{Miles repaired the computer quickly.} \rrbracket = \\ \exists e[\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x[\mathbf{computer}(x)] \wedge \mathbf{quick}(e)]$$

But if *quickly* contributes a meaning like (70), it will be unable to compose with (68) (or, rather, with the result of (68) composing with its two arguments):

$$(70) \quad \llbracket \text{quickly} \rrbracket = \\ \lambda F \lambda e. F(e) \wedge \mathbf{quick}(e)$$

Once (68) has composed with its arguments it is of type t , not the type $\langle vt \rangle$ which (70) is looking for. Even if we made (70) take a type t argument instead, it would still be unable to apply the **quick** predicate to the same event argument as the verbal predicate, since that argument is not left open in the verbal meaning contribution. Fortunately, Champollion (2015) proposes an ingenious solution.

Rather than seeing verbal predicates as holding of events, we see them as holding of *sets* of events: “[i]nstead of denoting the set of all raining events, think of a predicate like ‘rain’ as being true of any set that contains a raining event” (Champollion 2015: 38–39). Formally, this means maintaining the intuition that the existential quantification of the event variable is encoded lexically, but also adding an additional argument to the verbal predicate:

$$(71) \quad \llbracket \text{repair} \rrbracket = \\ \lambda x \lambda y \lambda F. \exists e[\mathbf{repair}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y \wedge F(e)]$$

Once this is combined with its nominal arguments, we arrive at an expression like (72):

$$(72) \quad \lambda F. \exists e[\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x[\mathbf{computer}(x)] \wedge F(e)]$$

This is true of any set of events which contains an event of Miles repairing a (uniquely contextually salient) computer. We can think of the actual world as the set of all events whatsoever, i.e. $\lambda e. \top$, a characteristic function which returns True for any event, and so we can assess the truth of (72) by applying it to this set. When we do, we arrive at the more familiar translation of the sentence:

$$\begin{aligned}
(73) \quad \llbracket \text{Miles repaired the computer} \rrbracket &= \\
&\quad \vdots \\
&\lambda F. \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)] \wedge F(e)] (\lambda e. \top) = \\
&\exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)] \wedge (\lambda e. \top)(e)] = \\
&\exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)] \wedge \top] = \\
&\exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)]]
\end{aligned}$$

Importantly, though, until we apply this new kind of closure operation, the event argument remains open—not because the variable itself is open, but because there is an open slot for a predicate which takes it as an argument.

With suitable amendments, then, we can easily ensure that adverbs are able to apply to the event argument appropriately. Because verbal expressions are now of type $\langle vt, t \rangle$, rather than the lower $\langle vt \rangle$, adverbs must also have a higher type: $\langle \langle vt, t \rangle, \langle vt, t \rangle \rangle$ instead of the $\langle vt, vt \rangle$ type they would have in a standard event semantics:

$$\begin{aligned}
(74) \quad \llbracket \text{quickly} \rrbracket &= \\
&\lambda \mathcal{V}_{\langle vt, t \rangle} \lambda G_{\langle vt, t \rangle}. \mathcal{V}(\lambda e. \mathbf{quick}(e) \wedge G(e))
\end{aligned}$$

Syntactically, the adverb takes the verb as an argument, but it then feeds the adverbial meaning to the verb in place of its own argument (the predicate of events F in the verb's meaning). Because the meaning in (74) is a symmetrical, modifier type (i.e. of the form $\langle x, x \rangle$, for any type x), it leaves the verb with an open predicate of events, exactly as it found it, so that modification can continue indefinitely. (75) shows how composition proceeds:

$$\begin{aligned}
(75) \quad \llbracket \text{quickly} \rrbracket (\llbracket \text{Miles repaired the computer} \rrbracket) &= \\
&\lambda G. \llbracket \text{Miles repaired the computer} \rrbracket (\lambda e. \mathbf{quick}(e) \wedge G(e)) = \\
&\lambda G. \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)] \\
&\quad \wedge (\lambda e. \mathbf{quick}(e) \wedge G(e))(e)] = \\
&\lambda G. \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota x [\mathbf{computer}(x)] \\
&\quad \wedge \mathbf{quick}(e) \wedge G(e)]
\end{aligned}$$

We end up with the same $\langle vt, t \rangle$ type we started with, simply with the addition of an extra predicate describing the event, just as desired.

This concludes our overview of the meaning language to be used in this thesis. We now show how we can connect our semantics to an LFG syntactic analysis.

4.4.2 Glue Semantics

Traditional approaches to semantic composition make use of phrase-structure trees to determine the order of application of their meaning language terms. This means they are reliant on linear order, and so variations in word order (and therefore in phrase-structure configuration), both between languages and within the same language, can potentially pose problems (for more on which see Dalrymple et al. 1995b). The solution which Glue Semantics (Glue) offers to these problems is to not rely on surface constituent order and hierarchical structure, but rather on a more abstract level of syntactic representation, namely, LFG's f-structures.³⁰

Glue sees lexical entries (and potentially also syntactic structure) as contributing premises, known as MEANING CONSTRUCTORS, to a logical deduction. The question then arises: which logic should we use? Neither classical nor intuitionistic logic, it turns out, is appropriate for this task. This is because lexical contributions to meaning are not “context-independent assertions that may be used or not in the derivation of the meaning of the sentence depending on the course of the derivation”; rather, they are “*occurrences* of information which are generated and used exactly once” (Dalrymple et al. 1995b: 8; emphasis in original). In classical logic, premises are either true or not true, and they continue to be so whatever we use them for. What is more, we can choose whether to use premises in a deduction or not, and we can use the same one multiple times. But this is not how natural language works—as Dalrymple (2001:

³⁰Although Glue was developed for LFG and is most prevalent in the LFG community, it has also been integrated with a number of other syntactic frameworks, including LTAG (Frank & van Genabith 2001) and HPSG (Asudeh & Crouch 2002), along with Categorical Grammar and Context-Free Grammar (Asudeh & Crouch 2001).

242) points out, “it would be incorrect to deduce the same meaning for the sentence if words or phrases are added or subtracted”. This claim is discussed in some detail by Asudeh (2012: Ch5), where he argues for what he calls the RESOURCE SENSITIVITY HYPOTHESIS, namely, that natural languages are resource sensitive in that “elements of combination in grammars cannot be freely *reused* or *discarded*” (Asudeh 2012: 97; emphasis in original). Expressions in natural language utterances contribute their meaning exactly once, no more and no less; what we need therefore is a logic which is sensitive to this conception of meaning contributions as *resources*. We achieve this with LINEAR LOGIC.

4.4.2.1 Linear logic

It is possible to analyse different kinds of logic based on what sorts of rules are valid in their deductions. Three important rules in the typology of logics are WEAKENING, CONTRACTION, and COMMUTATIVITY (Restall 2000; see also Asudeh 2012: 96ff, from which this section draws liberally):

1. **Weakening:** Premises can be *freely added*.
2. **Contraction:** Additional occurrences of a premise can be *freely discarded*.
3. **Commutativity:** Premises can be *freely reordered*.

A schematic representation of Weakening is given in (76):

$$(76) \quad \frac{\Gamma \vdash B}{\Gamma, A \vdash B}$$

That is, if we can prove B from the set of premises Γ , we can also prove it from Γ and A (we can freely add A). Example (77) gives a representation of Contraction:

$$(77) \quad \frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}$$

In other words, if we can prove B from Γ and two instances of A , we can also prove it from Γ and a single instance of A (we can freely discard A). Finally, Commutativity is as in (78):

$$(78) \quad \frac{\Gamma, A, B \vdash C}{\Gamma, B, A \vdash C}$$

If we can prove C from Γ, A , and B in that order, then we can also prove it from Γ, B , and A in *that* order (we can freely reorder our premises).

Classical logic possesses all of these properties, which makes it inappropriate for modelling linguistic meaning. This is because, as alluded to above, semantics is RESOURCE SENSITIVE; in other words, we cannot freely add or discard premises, and so we must not include Weakening or Contraction in our logic. As for Commutativity, we can make the following observations (Asudeh 2012: 99–100). In natural language, functors can equally well combine with arguments to their left as to their right. This is true cross-linguistically (e.g. languages where the direct object precedes as opposed to follows the verb) but also within a single language: if the subject of a verb is of a lower type than the verb (e.g. it is a proper name), then the verb will be the functor and the name the argument; but if the subject is a quantified expression, it will be of a higher type than the verb and so the subject will be the functor and the verb the argument. There may be syntactic constraints, but as the quantifier example shows, this is semantically irrelevant. Really, “it is the types of the expressions that determine functor-argument combination, not their order” (Asudeh 2012: 100). For this reason, it seems sensible to maintain Commutativity in our logic.

If we remove Weakening and Contraction from our logic, while maintaining Commutativity, we obtain a RESOURCE LOGIC, whereby logical formulae become dynamic resources which require “strict accounting” (Asudeh 2012: 97). The particular resource logic used in Glue is linear logic (Girard 1987), which has the desired properties. In fact, only a restricted fragment of linear logic is needed; in this thesis I make use of the first-order ($-\circ, \forall$) fragment of predicate linear logic. It has first-order universal quantification, but not existential quantification, and no negation. The symbol $-\circ$ stands for linear implication, the linear logic counterpart of classical logic’s material conditional \rightarrow or \supset (the crucial difference being that application of *modus ponens* uses up the resource on the left-hand side

of an implication in linear logic; see below). Other works in Glue sometimes additionally make use of multiplicative conjunction \otimes (the counterpart of classical logic's conjunction \wedge), e.g. in treatments of anaphora (Dalrymple et al. 1999b; Asudeh 2012), but this is not essential (cf. Kokkonidis 2005: , fn. 7). For the sake of simplicity, and since anaphora will not be treated in this thesis, we can stick to the simpler fragment.

The introduction and elimination rules for \multimap are shown below (after Asudeh 2012: 78–9):³¹

(79) **Implication Elimination** (*/modus ponens*):

$$\frac{\begin{array}{c} \vdots \\ A \end{array} \quad \begin{array}{c} \vdots \\ A \multimap B \end{array}}{B} \multimap\mathcal{E}$$

(80) **Implication Introduction** (*/abstraction/hypothetical reasoning*):

$$\frac{\begin{array}{c} [A]^1 \\ \vdots \\ B \end{array}}{A \multimap B} \multimap\mathcal{I},1$$

The rule of Implication Elimination in (79) tells us that if we have a resource $A \multimap B$, it can consume a resource A to produce a resource B (note that both A and $A \multimap B$ are used up in the process). Implication Introduction, (80), tells us that if we can assume a resource A and produce a resource B , then we can consume B to produce $A \multimap B$, while discharging our assumption of A .

It turns out that there is an exact mapping between proof rules in constructive logics (such as our linear logic fragment) and operations in the lambda calculus, discovered over several decades during the 20th century, and known as the CURRY-HOWARD CORRESPONDENCE or the CURRY-HOWARD ISOMORPHISM (CHI). We will not discuss the formal details here (on which see Crouch & van Genabith 2000), but the fundamental insight of the CHI is that implications in the

³¹The following conventions are used in writing proofs. Proof steps are annotated by the rule they correspond to: for the most part, this means implication elimination, $\multimap\mathcal{E}$, or implication introduction, $\multimap\mathcal{I}$. Assumptions are enclosed in square brackets and annotated with a number, which is used to indicate when the assumption is discharged: $\multimap\mathcal{I},1$ means that we are using the rule of implication introduction and, in so doing, discharging assumption 1. Vertical dots indicate an unspecified sequence of proof steps.

linear logic correspond to functional types in the lambda calculus. Specifically, implication elimination corresponds to functional application, while implication introduction corresponds to lambda-abstraction. Meaning constructors in Glue are thus expressions in some meaning language paired with expressions in linear logic which control their composition. The following expansions of the introduction and elimination rules for \multimap show how the CHI applies:

(81) **Implication Elimination** (= functional application):

$$\frac{\begin{array}{c} \vdots \\ a : A \end{array} \quad \begin{array}{c} \vdots \\ f : A \multimap B \end{array}}{f(a) : B} \multimap\mathcal{E}$$

(82) **Implication Introduction** (= lambda-abstraction):

$$\frac{\begin{array}{c} [x : A]^1 \\ \vdots \\ f : B \end{array}}{\lambda x.f : A \multimap B} \multimap\mathcal{I},1$$

In addition to constraining the ways in which meanings can be combined, the linear logic component of a meaning constructor is also used to connect the semantic resources to the syntax. It is because of this latter use, of ‘holding things together’, that the theory bears its ‘glue’ moniker. To see how this works, consider an analysis of a simple sentence like *Benjamin likes Jadzia*. Ignoring the complexities of tense, etc., the f-structure is given in (83):

$$(83) \quad \left[\begin{array}{l} \text{PRED} \quad \text{'like'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'Benjamin'} \right] \\ \text{OBJ} \quad \left[\text{PRED} \quad \text{'Jadzia'} \right] \end{array} \right]$$

If we simply take the linear logic component of the meaning constructor to represent typing information, and to impose the resource sensitivity constraint, then the meaning constructors contributed by the words of this sentence could be as follows (ignoring tense and events for the sake of illustrative simplicity):

$$(84) \quad \text{a.} \quad \llbracket \text{like} \rrbracket = \lambda x \lambda y. \mathbf{like}(x, y) : e \multimap e \multimap t$$

- b. $\llbracket \text{Benjamin} \rrbracket = \mathbf{benjamin} : e$
 c. $\llbracket \text{Jadzia} \rrbracket = \mathbf{jadzia} : e$

From this, we can construct the correct proof for the sentence, (85):³²

$$(85) \quad \frac{\frac{\llbracket \text{like} \rrbracket \quad \llbracket \text{Benjamin} \rrbracket}{\lambda x \lambda y. \mathbf{like}(x, y) : \mathbf{benjamin} : e \multimap e \multimap t} \quad e}{\lambda y. \mathbf{like}(\mathbf{benjamin}, y) : \mathbf{jadzia} : e \multimap t} \quad \llbracket \text{Jadzia} \rrbracket}{\mathbf{like}(\mathbf{benjamin}, \mathbf{jadzia}) : t}$$

However, since we are simply matching types, we can also construct the incorrect proof, (86), which flips the arguments around:

$$(86) \quad \frac{\frac{\llbracket \text{like} \rrbracket \quad \llbracket \text{Jadzia} \rrbracket}{\lambda x \lambda y. \mathbf{like}(x, y) : \mathbf{jadzia} : e \multimap e \multimap t} \quad e}{\lambda y. \mathbf{like}(\mathbf{jadzia}, y) : \mathbf{benjamin} : e \multimap t} \quad \llbracket \text{Benjamin} \rrbracket}{\mathbf{like}(\mathbf{jadzia}, \mathbf{benjamin}) : t}$$

Although we have constrained the proofs to be resource sensitive, owing to the linear logic—so that we can't derive $\mathbf{like}(\mathbf{jadzia}, \mathbf{jadzia})$, for instance—we have failed to constrain the proofs functionally-speaking. What is needed is some way to add functional information to the linear logic, so that we make it clear that it is the *subject* which likes the *object*, and not the other way around. To aid in this, we can add labels to the f-structure from (83):

$$(87) \quad l \left[\begin{array}{l} \text{PRED} \quad \text{'like'} \\ \text{SUBJ} \quad b \left[\text{PRED} \quad \text{'Benjamin'} \right] \\ \text{OBJ} \quad j \left[\text{PRED} \quad \text{'Jadzia'} \right] \end{array} \right]$$

Using labels gives us a concrete way of referring to specific f-structures, which is important since e.g. a sentence might contain multiple embedded clauses, each with their own subject, and so simply referring to SUBJ will not uniquely

³²I include the sources of meaning constructors in proofs to aid readability, but they are not a formal part of the proofs. If otherwise unannotated, proof steps should be taken as \multimap -elimination.

identify a particular f-structure. With these labels in place, there are at least two ways of proceeding (as Kokkonidis 2008: 51–52 discusses):

1. We can use f-structure labels directly as types in the linear logic. That is, the meaning constructors from before would become the following:

$$(88) \quad \begin{array}{l} \text{a. } \llbracket \text{like} \rrbracket = \lambda x \lambda y. \text{like}(x, y) : b \multimap j \multimap l \\ \text{b. } \llbracket \text{Benjamin} \rrbracket = \mathbf{benjamin} : b \\ \text{c. } \llbracket \text{Jadzia} \rrbracket = \mathbf{jadzia} : j \end{array}$$

Now there is only one way to compose the meaning, as desired. However, we have lost the distinction between e and t in the base types. There are also other formal complications—e.g. the treatment of quantifiers now becomes more complex (see discussion in Kokkonidis 2008: 52, 6off.).

2. We can instead treat the linear logic as a predicate logic, and take the base types e and t to instead be type-constructor predicates which take a single argument, rather than types in their own right. In this format, the meaning constructors would be as follows:

$$(89) \quad \begin{array}{l} \text{a. } \llbracket \text{like} \rrbracket = \lambda x \lambda y. \text{like}(x, y) : e(b) \multimap e(j) \multimap t(l) \\ \text{b. } \llbracket \text{Benjamin} \rrbracket = \mathbf{benjamin} : e(b) \\ \text{c. } \llbracket \text{Jadzia} \rrbracket = \mathbf{jadzia} : e(j) \end{array}$$

Most work in Glue takes the first approach. However, in this thesis I depart from this trend and instead follow the second approach, which is known as first-order Glue (after Kokkonidis 2008). There are a number of reasons for my choice. Firstly, as pointed out by Kokkonidis (2008: 52, 6off.), scope-taking expressions require quantification in the linear logic, and this pleads in favour of the second approach, since quantification over arguments is natural in a predicate logic, whereas quantification over propositions in a propositional logic is not. Secondly, and more significantly, first-order Glue obviates much of the need for a separate semantic structure projected from f-structure, since

it permits multiple Glue resources to be associated with the same f-structure, by passing them as arguments to different type constructors. This makes for a major architectural saving. Let us see how this comes about.

In mainstream work in Glue, the linear logic takes s(ematic)-structures, not f-structures, as base types, mostly because of mismatches between types and f-structures. For instance, since common nouns have the functional type $\langle et \rangle$, the CHI means that the linear logic side of the corresponding meaning constructor must be an implication of the form $A \multimap B$. We therefore need to refer to two base types, A and B . Now, since in the mainstream approach base types are structures, that means we need to have two distinct structures associated with a common noun. But common nouns project just a single f-structure, so what do we do? The standard Glue treatment (e.g. in Dalrymple 2001: 250ff.) is to assume the existence of two ‘dummy’ attributes in the corresponding s-structure (projected from the f-structure by the σ function):³³

$$(90) \quad c \left[\begin{array}{l} \text{PRED} \\ \text{'computer'} \end{array} \right] \xrightarrow{\sigma} c_{\sigma} \left[\begin{array}{l} \text{VAR} \\ \text{RESTR} \end{array} \right]$$

The meaning constructor for a common noun like *computer* is then as follows:

$$(91) \quad \lambda x. \mathbf{computer}(x) : (c_{\sigma} \text{ VAR}) \multimap (c_{\sigma} \text{ RESTR})$$

However, in first-order Glue the problem does not arise. We can simply use the two different base type constructors to form two distinct base types from one and the same f-structure:³⁴

$$(92) \quad \lambda x. \mathbf{computer}(x) : e_c \multimap t_c$$

We do not need to propose otherwise unmotivated structures like the values of VAR and RESTR which are only present for theory-internal reasons, and

³³VAR stands mnemonically for the ‘variable’ and RESTR for the ‘restriction’, gesturing to a view of NPs as quantifiers; but in the kind of implementation described in the text the two structures do not actually correspond to any meanings individually.

³⁴From hereon out, I abbreviate types of the form *constructor(argument)* to the slightly more compact *constructor_{argument}*; e.g. e_c in (92) stands for $e(c)$.

instead can base semantic composition on the independently syntactically motivated f-structure.³⁵ Thus, first-order Glue leads to a more parsimonious representation, and a cleaner analysis. This does not necessarily mean we can do without semantic structures entirely, however, since they have been used for other purposes, e.g. for mediating the interface with i(nformation)-structure (Dalrymple & Nikolaeva 2011; Lowe 2014). But it does at least mean that many of the redundant features which serve only to make the Glue derivations work can be removed.

Finally, though not insignificantly, I also believe it is a strength of first-order Glue that the semantic types of meaning language expressions are represented more prominently in meaning constructors; this aids clarity when discussing some semantic issues.

4.4.2.2 Meaning constructors in lexical entries

The meaning constructors we have seen so far make reference to specific f-structures by name. But this is inappropriate for a lexical entry, for example, where an expression only knows which f-structures it and its arguments (if any) will correspond to *relatively*, rather than *absolutely*. To remedy this, we use the same conventions as for f-descriptions, describing paths through f-structure anchored at the node which bears the annotation: in this case the terminal node for the lexical item which carries the meaning (although meaning constructors can also be introduced constructionally by phrase-structure rules—see e.g. Dalrymple 2001: 416ff.). Using this procedure, the (event-free) meaning constructor for a transitive verb such as *like* is as follows:

$$(93) \quad \lambda x \lambda y. \mathbf{like}(x, y) : e_{(\uparrow_{\text{SUBJ}})} \multimap e_{(\uparrow_{\text{OBJ}})} \multimap t_{\uparrow}$$

³⁵And this is not to mention the formal problems of the s-structure approach, which requires multiple empty AVMs to nonetheless be interpreted as distinct; given a set-theoretic understanding of s-structure, two empty structures should in fact be identified (Kokkonidis 2008: 63, fn. 13). There are perhaps straightforward ways to remedy this, e.g. by giving each structure a unique index (see Dalrymple et al. 2018 for a suggestion along these lines in a different domain), but the fact remains that it must be remedied.

Note that it does not matter which order the arguments of the verb are presented in, since, without using any additional premises, it is always possible to change the order in which the verb consumes its arguments during a proof:

$$(94) \quad \frac{\frac{\lambda x \lambda y. \mathbf{like}(x, y) : e_{(\uparrow \text{SUBJ})} \multimap e_{(\uparrow \text{OBJ})} \multimap t_{\uparrow} \quad [\mathbf{a} : e_{(\uparrow \text{SUBJ})}]^1}{\lambda y. \mathbf{like}(\mathbf{a}, y) : e_{(\uparrow \text{OBJ})} \multimap t_{\uparrow}} \quad [\mathbf{b} : e_{(\uparrow \text{OBJ})}]^2}{\mathbf{like}(\mathbf{a}, \mathbf{b}) : t_{\uparrow}} \multimap_{\mathcal{I},1}}{\lambda y \lambda x. \mathbf{like}(x, y) : e_{(\uparrow \text{OBJ})} \multimap e_{(\uparrow \text{SUBJ})} \multimap t_{\uparrow}} \multimap_{\mathcal{I},2}}$$

Given the generality of this procedure, I will make use of whichever ordering of arguments is convenient for the situation at hand, leaving the reordering implicit.

Of course, the verbal lexical entries we will be making use of are somewhat more complex, but the underlying principles are the same. (95) gives the full form of the meaning constructor for *like*:

$$(95) \quad \lambda x \lambda y \lambda F. \exists e [\mathbf{like}(e) \wedge \mathbf{exp}(e) = x \wedge \mathbf{th}(e) = y \wedge F(e)] : \\ e_{(\uparrow \text{SUBJ})} \multimap e_{(\uparrow \text{OBJ})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$$

Note again the parsimony of first-order Glue: instead of requiring an additional EVENT semantic structure to correspond to the event argument (as in e.g. Asudeh & Giorgolo 2012; Findlay 2016a), we simply use v_{\uparrow} , the result of applying the event-type function to the clause's f-structure, which is distinct from t_{\uparrow} , the propositional type associated with the clause.

The verbal meaning constructors make reference to grammatical functions like SUBJ and OBJ in order to anchor their composition in the syntax. However, argument alternations disrupt this mapping. Consider a more prototypical (agent-patient) verb, like *repair*. The meaning constructor below claims that it is always the SUBJ which repairs the OBJ:

$$(96) \quad \lambda x \lambda y \lambda F. \exists e [\mathbf{repair}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y \wedge F(e)] : \\ e_{(\uparrow \text{SUBJ})} \multimap e_{(\uparrow \text{OBJ})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$$

But this isn't true in the passive voice, e.g. *the computer was repaired by Miles*; here, the subject becomes the patient, and the agent is realised (if at all) by an OBL GF.

Starting with Asudeh & Giorgolo (2012), therefore, some researchers have preferred not to use f-structure GFS in meaning constructors, but instead to use numbered ARG attributes at s-structure, which correspond to the specific semantic argument involved, and so do not vary across diathesis alternations.³⁶ This enables a single meaning constructor to be used for both the active and passive cases, but comes at a cost. The first cost is that we lose the advantage brought by first-order Glue of not needing to refer to s-structures in semantic composition. The second is that we require a theory of how the GFS and the ARG slots are connected. This last problem has been partially resolved by Findlay (2016a), who provides a formalisation of the mapping between grammatical functions and semantic structure arguments (building on work by Kibort 2001, 2007). However, since most researchers working in Glue do not use first-order Glue, they do not even face the first problem, and so no solutions have been discussed.

I propose a simple workaround, whereby the semantic contribution of passive is precisely to reorganise the arguments of a verb in the appropriate way:³⁷

$$(97) \quad \lambda P \lambda y. \exists x [P(x)(y)] : (e_{(\uparrow \text{SUBJ})} \multimap e_{(\uparrow \text{OBJ})} \multimap t_{\uparrow}) \multimap e_{(\uparrow \text{SUBJ})} \multimap t_{\uparrow}$$

This takes the meaning of any verb once it has combined with all of its arguments except its SUBJ and OBJ, then reorders those last two arguments and existentially closes the argument which corresponds to the active voice SUBJ. Figure 4.6 shows

³⁶See e.g. Asudeh et al. (2014) and Lowe (2015a); Przepiórkowski (2017) follows the intuition of referring to s-structure attributes instead of GFS, but uses thematic role names instead of numbered ARG labels.

³⁷This is in some respects an alternative way of formalising the concept of LEXICAL RULE used by some (especially earlier) practitioners in LFG to describe such diathesis alternations (e.g. Bresnan 1982b). One advantage of my proposal over the traditional LFG lexical rules, however, is that it makes clear what the relationship between the *meaning* of the active and passive is. Lexical rules as conceived of by e.g. Bresnan (1982b) manipulate argument structure (by saying, for instance, that a predicate which takes a SUBJ and an OBJ argument in the active takes an OBL/∅ and a SUBJ in the passive), but rely on some additional theory to tell us how this corresponds to meaning (see e.g. Butt et al. 1997 for such a proposal). Such theories assume a separate level of a(rgument)-structure, which runs counter to our goal, facilitated by first-order Glue, of doing without a separate level of structure on which semantic composition depends, be it a-structure or s-structure. The approach suggested here shows that we still do not need to resort to such additions to our theoretical ontology.

how it works in the Glue proof for the passive sentence *Janak was poked*, using the labels from the f-structure in (98):³⁸

$$(98) \quad p \left[\begin{array}{ll} \text{PRED} & \text{'poke'} \\ \text{SUBJ} & j \left[\begin{array}{ll} \text{PRED} & \text{'Janak'} \end{array} \right] \\ \text{VOICE} & \text{PASSIVE} \end{array} \right]$$

(Notice that in Figure 4.6 the verb has had its arguments reordered, using the technique described above, so that it consumes the event predicate first.) This approach enables us to retain the intuitive, active-voice meaning of the verb as basic, and then to derive the passive version when necessary.

4.4.2.3 Completeness and coherence

One strength of treating our semantic theory as resource sensitive is that we gain a natural account of completeness and coherence, as alluded to above. We can now show more explicitly what this means. We restrict Glue proofs so that a valid proof for a sentence must terminate in a single premise of type t (i.e. the type of sentences). This ensures that the proof must use up all resources available (ending in the conjunction of type t with something else is not acceptable), and that it must not leave any out (ending in a functional type $X \multimap t$ is also not acceptable).

Under this view, an incomplete sentence corresponds to a RESOURCE DEFICIT, where there are not enough resources to reach the goal type, and we will be left with some functional type, while an incoherent sentence leads to RESOURCE SURPLUS, since there will be additional resources left over once the goal type is reached (Asudeh 2012). For example, an attempted proof for **Benjamin likes*, an incomplete sentence, fails since it terminates in a type $e \multimap t$ resource:

³⁸I assume that it does not matter that there is no f-structure (\uparrow_{OBJ}), and so no determinate value for $e_{(\uparrow_{\text{OBJ}})}$ in the passive meaning constructor and the meaning constructor for the verb *poked*, since no meaning constructor consumes this resource directly. It is only mentioned in an implicational type which is itself consumed, and so identity of form is sufficient.

$$\begin{array}{c}
\llbracket \text{poked} \rrbracket \\
\lambda F \lambda x \lambda y. \exists e [\mathbf{poke}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y \wedge F(e)] : \llbracket \text{closure} \rrbracket \\
\hline
(v_p \multimap t_p) \multimap e_j \multimap e_{obj} \multimap t_p \quad \lambda e. \top : v_p \multimap t_p \quad \llbracket \text{passive} \rrbracket \\
\lambda x \lambda y. \exists e [\mathbf{poke}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y] : \lambda P \lambda y. \exists x [P(x)(y)] : \\
e_j \multimap e_{obj} \multimap t_p \quad (e_j \multimap e_{obj} \multimap t_p) \multimap e_j \multimap t_p \quad \llbracket \text{Janak} \rrbracket \\
\lambda y. \exists x [\exists e [\mathbf{poke}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y]] : \mathbf{Janak} : \\
e_j \multimap t_p \quad e_j \\
\hline
\exists x [\exists e [\mathbf{poke}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = \mathbf{Janak}]] : t_p
\end{array}$$

Figure 4.6: Glue proof for *Janak was poked*

$$(99) \quad \frac{\begin{array}{l} \llbracket \text{likes} \rrbracket \quad \llbracket \text{Benjamin} \rrbracket \\ \lambda x \lambda y. \text{like}(x, y) : \mathbf{benjamin} : \\ e_b \multimap e_{obj} \multimap t_l \quad e_b \end{array}}{\lambda y. \text{like}(\mathbf{benjamin}, y) : \\ e_{obj} \multimap t_l}$$

There are no further steps available to be taken here, but we have not reached our goal type, so this sentence is unacceptable. On the other hand, an attempted proof for **Julian yawns Worf*, an incoherent sentence, fails because there is an extra type e meaning left over:

$$(100) \quad \frac{\begin{array}{l} \llbracket \text{yawns} \rrbracket \quad \llbracket \text{Julian} \rrbracket \\ \lambda x. \text{yawn}(x) : \mathbf{julian} : \\ e_j \multimap t_y \quad e_j \end{array} \quad \llbracket \text{Worf} \rrbracket}{\mathbf{yawn}(\mathbf{julian}) : \quad \mathbf{worf} : \\ t_y \quad e_w} \\ !$$

Once again, there are no more proof steps available, but we have not terminated with a single type t resource; instead we also have an additional type e resource.

This approach can therefore subsume syntactic subcategorisation restrictions where these apply to semantically contentful arguments. Obviously, as mentioned in footnote 17 above, it cannot help with expletive arguments, since, not making a semantic contribution, these do not enter into the resource accounting in the first place. For those such cases, we cannot resort to a purely semantic solution.

This section has explained how a resource-sensitive semantics not only constrains semantic composition, but also extends into domains often thought to be properly syntactic, such as valency. We might, in addition, think of this as offering some insight into the argument-adjunct distinction: arguments are obligatory in the syntax because their absence leads to a resource deficit, and cannot be iterated indefinitely because any repetition leads to a resource surplus.³⁹ Adjuncts, of course, since they consume other resources rather

³⁹Optionally transitive verbs, like *eat* (*We ate* and *We ate supper* are both grammatical), and the existence of ‘cognate’ objects (*Dukat laughed a cruel laugh*) are *prima facie* counterexamples to both these claims, since in one there is an apparent resource deficit, and in the other a

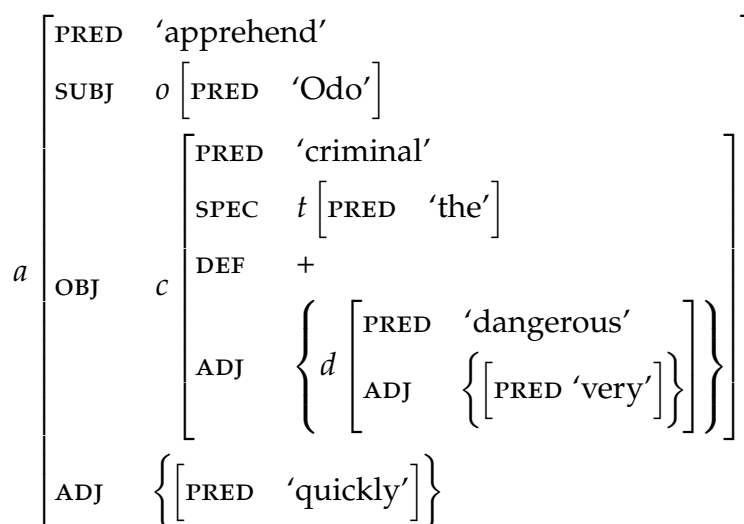


Figure 4.7: F-structure for *Odo apprehended the very dangerous criminal*

than being consumed themselves, do not lead to such issues, and so are in principle unrestricted in their number. This is in many ways consonant with Przepiórkowski's (1999) view that there is no meaningful distinction to be made in the syntax between arguments and adjuncts, and that theories like LFG would do well to abandon the idea that there is, however minimally such a distinction is actually recognised in the theory (Przepiórkowski 2016).

4.4.2.4 Sample analysis

In order to summarise the presentation above, as well as to introduce the basic treatment of modifiers, in this section I give an analysis of the sentence in (101), which has the f-structure shown in Figure 4.7

(101) Odo quickly apprehended the very dangerous criminal.

surplus, but without any adverse effects on grammaticality. Both cases, however, show specific constraints on their use, be they lexical or meaning-related. In the first case, even if it is not such a rare occurrence for (English) verbs to exhibit optional transitivity, it is still a matter of lexical variability whether they do so: for example, in contrast to *We ate*, **We devoured* is ungrammatical. In the second case, there are semantic constraints on the cognate object (e.g. the object of *laugh* must be a kind of laugh). This means we can retain as our default position the idea that resource sensitivity explains the basic transitivity facts, with the understanding that lexically-specified and -constrained exceptions to it are also possible. See Asudeh & Giorgolo (2012) and Asudeh et al. (2014) for proposals for tackling both these kinds of data in a resource-sensitive framework.

We begin by introducing the relevant lexically-contributed meaning constructors, before giving the proof in full.

The verb has the usual kind of meaning constructor we have already seen, given in (102). I first give the general meaning constructor, and then the instantiated version relative to the f-structure in (101).

- (102) a. $\lambda x \lambda y \lambda F. \exists e [\mathbf{apprehend}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{th}(e) = y \wedge F(e)] :$
 $e_{(\uparrow_{\text{SUBJ}})} \multimap e_{(\uparrow_{\text{OBJ}})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$
- b. $\lambda x \lambda y \lambda F. \exists e [\mathbf{apprehend}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{th}(e) = y \wedge F(e)] :$
 $e_o \multimap e_c \multimap (v_a \multimap t_a) \multimap t_a$

It also contributes the [closure] meaning constructor discussed in Section 4.4.1:

- (103) a. $\lambda e. \top :$
 $v_{\uparrow} \multimap t_{\uparrow}$
- b. $\lambda e. \top :$
 $v_a \multimap t_a$

Proper names like *Odo* have a very simple meaning constructor:

- (104) a. **odo** : e_{\uparrow}
- b. **odo** : e_o

And common nouns have the type discussed earlier:

- (105) a. $\lambda x. \mathbf{criminal}(x) : e_{\uparrow} \multimap t_{\uparrow}$
- b. $\lambda x. \mathbf{criminal}(x) : e_c \multimap t_c$

The definite article introduces the iota operator mentioned in footnote 27:

- (106) a. $\lambda P. \iota x [P(x)] : (e_{(\text{SPEC}\uparrow)} \multimap t_{(\text{SPEC}\uparrow)}) \multimap e_{(\text{SPEC}\uparrow)}$
- b. $\lambda P. \iota x [P(x)] : (e_c \multimap t_c) \multimap e_c$

The functional descriptions here use so-called INSIDE-OUT specifications: instead of going deeper into the f-structure from the f-structure corresponding to the node which bears the annotation, we work our way outwards. (SPEC \uparrow)

refers to the f-structure which contains the SPEC attribute having the current f-structure as its value. If \uparrow in (106) corresponds to t in Figure 4.7, then (SPEC \uparrow) corresponds to c .

We now turn to how to analyse the modifiers. On the linear logic side of their meaning constructors, modifiers must have an implication which has the same, potentially complex, type on both sides:

$$(107) \quad X \multimap X$$

This is because modification should not affect the combinatorial possibilities of the thing it modifies: a noun modified by an adjective is still, combinatorially speaking, a noun, and should still be able to serve as the argument of a determiner, for example. Also, modification is in general recursive: being modified by one adjective should not stop a noun being modified by another, for example. This means that modification should be type-preserving, which is precisely what the type in (107) achieves. Anything which takes something of type X as its argument (be it a predicate or a modifier) will still be able to take the modified expression as its argument, since that expression retains the same type afterwards.

As discussed above, adverbial expressions are of type $\langle\langle vt, t \rangle, \langle vt, t \rangle\rangle$. This is because they modify verbal meanings, which are of type $\langle vt, t \rangle$. With this in mind, the meaning constructor for *quickly* is given in (108):

$$(108) \quad \begin{array}{l} \text{a. } \lambda\mathcal{V}\lambda G.\mathcal{V}(\lambda e.\mathbf{quick}(e) \wedge G(e)) : \\ \quad ((v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap \\ \quad (v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)} \\ \text{b. } \lambda\mathcal{V}\lambda G.\mathcal{V}(\lambda e.\mathbf{quick}(e) \wedge G(e)) : \\ \quad ((v_a \multimap t_a) \multimap t_a) \multimap (v_a \multimap t_a) \multimap t_a \end{array}$$

Once again we see an inside-out f-description, and additionally note the use of the set membership symbol, ' \in ', as an attribute. This is a convenient shorthand that enables us to refer to members of sets in functional descriptions. See Dalrymple (2001: 154–155) for more on this use. (ADJ $\in \uparrow$) refers to the f-structure

which contains an ADJ attribute having as its value a set which contains the current f-structure as one of its members.

Nominal modifiers like *dangerous* ultimately must be of type $\langle et, et \rangle$, since they modify type $\langle et \rangle$ common nouns. This meaning is given in (109):

- (109) a. $\lambda P \lambda x. P(x) \wedge \mathbf{dangerous}(x) :$
 $(e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$
 b. $\lambda P \lambda x. P(x) \wedge \mathbf{dangerous}(x) :$
 $(e_c \multimap t_c) \multimap e_c \multimap t_c$

This combines straightforwardly with a common noun to give the intersective meaning:

- (110) $\llbracket \text{dangerous} \rrbracket, \llbracket \text{criminal} \rrbracket \vdash \lambda x. \mathbf{criminal}(x) \wedge \mathbf{dangerous}(x) : e_c \multimap t_c$

However, modifiers like *dangerous* can themselves be modified, e.g. by *very*, as we have here. But the meaning constructor in (109) does not make the meaning of *dangerous* itself available for modification—it is embedded in a conjunction. Instead, we need to split the meaning of *dangerous* into two parts, one of which gives the lexical meaning of *dangerous* and one of which handles the composition with common nouns (Dalrymple 2001: 265–269):

- (111) a. $\lambda x. \mathbf{dangerous}(x) : e_\uparrow \multimap t_\uparrow$
 b. $\lambda x. \mathbf{dangerous}(x) : e_d \multimap t_d$
- (112) a. $\lambda Q \lambda P \lambda x. P(x) \wedge Q(x) :$
 $(e_\uparrow \multimap t_\uparrow) \multimap (e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$
 b. $\lambda Q \lambda P \lambda x. P(x) \wedge Q(x) :$
 $(e_d \multimap t_d) \multimap (e_c \multimap t_c) \multimap e_c \multimap t_c$

Andrews (2010) calls these the LEXICAL and GRAMMATICAL meaning constructors, respectively. When composed, they give the same meaning constructor as in (109), but since the meaning of *dangerous* is available independently of its intersective compositional meaning, it can be modified before it combines with (112). The meaning constructor for *very* which achieves this is given in (113):

- (113) a. $\lambda P \lambda x. \mathbf{very}(P(x)) :$
 $(e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$
- b. $\lambda P \lambda x. \mathbf{very}(P(x)) :$
 $(e_d \multimap t_d) \multimap e_d \multimap t_d$

In a full account, the meaning constructor for all modifiers, including *very* and *quickly*, should also be broken down in the same fashion. I omit this here for simplicity, however, since the principle is clear.

In fact, it turns out to be frequently useful to decompose meaning constructors in this way. We may compare the semantics of verbs, for instance, where the verb makes two contributions to the meaning, although only one contains any lexical information (the other being the closure meaning constructor).

We now have all the pieces in place to give the full proof for this sentence. I present this in two pieces, for readability. Figure 4.8 shows the proof for *the very dangerous criminal*, and Figure 4.9 incorporates this into the proof for the whole sentence. The final step in Figure 4.9 simply eliminates the redundant *verum* conjunct, relying on the equivalence in (114):

$$(114) \quad A \wedge \top \equiv A$$

In general, this step will not be performed explicitly in proofs, so that the result of composing the verb with the [closure] meaning constructor will just be to eliminate the open predicate in the verb's meaning.

4.5 Chapter summary

In this chapter I have introduced three theories which will form the basis of the analysis in this thesis. LFG has a number of theoretical and practical strengths: it is well-formalised, constraint-based, and descriptively flexible. Its modular architecture lends itself well to the description of the complex phenomenon that is natural language. Glue Semantics is a natural accompaniment; it too offers a modular approach to semantics, being compatible with a number of meaning languages and syntactic frameworks. It is also grounded in a plausible

$$\begin{array}{c}
\text{[[very]]} \\
\lambda P \lambda x. \text{very}(P(x)) : \quad \text{[[dangerous]]-1} \\
\hline
(e_d \multimap t_d) \multimap e_d \multimap t_d \quad e_d \multimap t_d \\
\lambda x. \text{very}(\text{dangerous}(x)) :
\end{array}$$

$$\begin{array}{c}
\text{[[dangerous]]-2} \\
\lambda Q \lambda P \lambda x. P(x) \wedge Q(x) : \\
\hline
(e_d \multimap t_d) \multimap (e_c \multimap t_c) \multimap e_c \multimap t_c \\
\lambda P \lambda x. P(x) \wedge \text{very}(\text{dangerous}(x)) : \\
\hline
e_c \multimap t_c
\end{array}$$

$$\begin{array}{c}
\text{[[criminal]]} \\
\lambda x. \text{criminal}(x) : \\
\hline
e_c \multimap t_c \\
\lambda x. \text{criminal}(x) \wedge \text{very}(\text{dangerous}(x)) : \\
\hline
e_c \multimap t_c
\end{array}$$

$$\begin{array}{c}
\text{[[the]]} \\
\lambda P. l_x[P(x)] : \\
\hline
(e_c \multimap t_c) \multimap e_c \\
l_x[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))]: \\
\hline
e_c
\end{array}$$

Figure 4.8: Glue proof for the very dangerous criminal

$$\begin{array}{c}
\llbracket \text{Odo} \rrbracket \llbracket \text{apprehended} \rrbracket \\
\text{odo} : \frac{\lambda x \lambda y \lambda F. \exists e [\text{apprehend}(e) \wedge \text{ag}(e) = x \wedge \text{th}(e) = y \wedge F(e)] : \llbracket \text{the very dangerous criminal} \rrbracket}}{e_0 \rightarrow e_c \rightarrow (v_a \rightarrow t_a) \rightarrow t_a} \\
\frac{\lambda y \lambda F. \exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = y \wedge F(e)] : \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))]} :}{e_c \rightarrow (v_a \rightarrow t_a) \rightarrow t_a} \text{quickly} \\
\frac{\lambda F. \exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))]] \wedge F(e)] :}{(v_a \rightarrow t_a) \rightarrow t_a} \lambda \mathcal{V} \lambda G. \mathcal{V}(\lambda e. \text{quick}(e) \wedge G(e)) : \\
\frac{\lambda G. \exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))] \wedge \text{quick}(e) \wedge G(e)] :}{(v_a \rightarrow t_a) \rightarrow t_a} \llbracket \text{closure} \rrbracket \\
\frac{\exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))] \wedge \text{quick}(e) \wedge \top]} :}{t_a} \lambda e. \top : \\
\frac{\exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))] \wedge \text{quick}(e)] : t_a}{\exists e [\text{apprehend}(e) \wedge \text{ag}(e) = \text{odo} \wedge \text{th}(e) = \text{ix}[\text{criminal}(x) \wedge \text{very}(\text{dangerous}(x))] \wedge \text{quick}(e)] : t_a} \top \varepsilon
\end{array}$$

Figure 4.9: Glue proof for Odo quickly apprehended the very dangerous criminal

theory about natural language, the Resource Sensitivity Hypothesis. We saw in the previous chapter, however, that from a formal perspective a TAG has a number of advantages when it comes to the analysis of MWEs. In the next chapter, I show how we can integrate some of those advantages into the LFG architecture. Not only will this improve our analysis of MWEs, but it will also lead to a more elegant and formally coherent theory overall.

5

Lexicalised LFG

Contents

5.1	Incorporating TAG into LFG	198
5.1.1	TAG-LFG	199
5.1.2	Structures and descriptions of structures	201
5.2	A language for tree descriptions	202
5.2.1	Phrase structure rules and lexical entries	203
5.2.2	A description-based TSG	205
5.2.3	Some linguistic constraints	207
5.3	Lexicalisation	209
5.3.1	Lexicalised vs. non-lexicalised	209
5.3.2	Lexicalising a CFG	211
5.3.3	Advantages of lexicalised grammars	214
5.3.4	<i>Lexical</i> Functional Grammar	216
5.4	Description-based TAG	217
5.4.1	Motivation	217
5.4.2	Partial descriptions of trees	219
5.4.3	Constraining the formalism	222
5.5	Lexicalised LFG	226
5.5.1	Noun phrases	229
5.5.2	Verbs	231
5.5.3	Modification	234
5.5.4	Long-distance dependencies	238
5.5.5	Lexical rules	243
5.6	Chapter summary	259

In the previous chapter, I explained the key features of LFG and considered its advantages as a theory of grammar. In this chapter, I show how the insights of a TAG-based approach can be incorporated into LFG, by including descriptions of trees in lexical entries (or, rather, quasi-trees, as we shall see). In doing this, we not only provide an adequate analysis of MWEs but also arrive at a neater, more theoretically parsimonious theory in general terms, which also possesses the property of being *LEXICALISED*.

We begin in Section 5.1.1 by outlining a previous attempt to incorporate TAG into LFG, and identifying a fundamental weakness which needs to be addressed: a confusion between structures and *descriptions* of structures, a distinction which we discuss in Section 5.1.2. In Section 5.2, I give a straightforward account of how one can give descriptions of phrase-structure trees, and show how this can be used as the basis of a Tree Substitution Grammar. Section 5.3 explores the notion of lexicalisation, and its role in the grammar. I argue that this is another reason to prefer a TAG-style approach, in that it more closely encapsulates the lexicalist position. Given this, Section 5.4 then shows how the formal language introduced in Section 5.2 can be extended to accommodate a lexicalised TAG. Section 5.5 shows how this can be embedded in an LFG grammar, and gives examples of how this new theory works for the simple, single-word cases, before the next chapter shows how it can handle MWEs.

5.1 Incorporating TAG into LFG

We noted in Chapter 3 that an approach to MWEs that took their hybrid nature seriously was the most promising, and that TAG offered a good formal basis for such an approach. We also noted, in Chapter 4, that there are reasons to like LFG as a general theory of the architecture of the grammar. The formal challenge, then, is to integrate the insights of the TAG-style approach, i.e. that lexical entries are associated with trees which can then be manipulated, into the LFG formalism. I have elsewhere argued for just such a combination (Findlay 2017a,b), but I believe that these earlier proposals are deficient in a formally

important way, and in this section I want to explore what that is. I begin by briefly introducing these earlier proposals.

5.1.1 TAG-LFG

An LFG lexical entry can be thought of as a triple (W, C, F) , where W is a word form, i.e. the terminal node in the phrase-structure tree, C is a c-structure category, i.e. the pre-terminal node, and F is a functional description, i.e. a set of expressions spelling out additional linguistic information via the correspondence architecture. The obvious first step in incorporating TAG into LFG is therefore to replace the category C with a whole tree instead, and to replace the context-free grammar of c-structure with a TAG. This gives us what Findlay (2017a,b) calls TAG-LFG. In TAG-LFG, a lexical entry is instead a triple $(\langle W \rangle, T, F)$, where the word forms are now provided as a list, to accommodate MWEs, and the category information C is replaced by a tree, T .

An example of a lexical entry in TAG-LFG for the word *kisses* is given in (1):

$$(1) \quad \begin{array}{l} \langle W \rangle = \langle \text{kisses} \rangle \\ T = \end{array}$$

```

graph TD
    S --> NP1[NP↓]
    S --> VP[VP]
    NP1 --- S1["(↑ SUBJ) = ↓"]
    VP --> V[V◇0]
    VP --> NP2[NP↓]
    NP2 --- S2["(↑ OBJ) = ↓"]
  
```

$$F = \begin{array}{l} (S_\phi \text{ PRED}) = \text{'kiss'} \\ (S_\phi \text{ TENSE}) = \text{PRESENT} \\ \vdots \\ \lambda x \lambda y \lambda F. \exists e [\mathbf{kiss}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y \wedge F(e)] : \\ e_{(S_\phi \text{ SUBJ})} \multimap e_{(S_\phi \text{ OBJ})} \multimap (v_{S_\phi} \multimap t_{S_\phi}) \multimap t_{S_\phi} \end{array}$$

The standard TAG down arrows used to mark substitution sites are replaced with \downarrow s so as to avoid confusion with the LFG metavariable \downarrow , and, in the functional description, categories are used as shorthand ways to refer to the nodes they

label.¹ As in TAG treatments of abstracted tree schemata, the lozenge (\diamond) marks the anchor of the tree, the node under which the head word is to appear. In addition, such anchors are annotated with the list index of the word form in $\langle W \rangle$ which is to appear there. In a simple case like *kisses*, the list $\langle W \rangle$ is a singleton, and so the only list index that appears is 0.

Multiply-anchored trees are only slightly different. (2) gives the entry for *kicked the bucket* (adapted from Findlay 2017a: 222, fig. 2):

$$\begin{array}{l}
 (2) \quad \langle W \rangle = \langle \text{kicked, the, bucket} \rangle \\
 \quad \quad T = \\
 \quad \quad \quad \begin{array}{c}
 \text{S} \\
 \swarrow \quad \searrow \\
 \text{NP}\downarrow \quad \text{VP} \\
 (\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow \\
 \quad \quad \quad \swarrow \quad \searrow \\
 \quad \quad \quad \text{V}\diamond 0 \quad \text{NP} \\
 \quad \quad \quad \quad \quad \quad (\uparrow \text{OBJ}) = \downarrow \\
 \quad \quad \quad \quad \quad \quad \swarrow \quad \searrow \\
 \quad \quad \quad \quad \quad \quad \text{Det}\diamond 1 \quad \text{N}\diamond 2
 \end{array} \\
 \\
 F = \begin{array}{l}
 (\text{S}_\phi \text{ PRED}) = \text{'kick-the-bucket'} \\
 (\text{S}_\phi \text{ TENSE}) = \text{PAST} \\
 \vdots \\
 \lambda x \lambda F. \exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = x \wedge F(e)] : \\
 e_{(\text{S}_\phi \text{SUBJ})} \multimap (v_{\text{S}_\phi} \multimap t_{\text{S}_\phi}) \multimap t_{\text{S}_\phi}
 \end{array}
 \end{array}$$

In this case, we have multiple word forms in the list $\langle W \rangle$, and so the various anchors are annotated accordingly. Otherwise the entry looks just the same as the lexical entry for a non-MWE. The assumption is that the tree schema in question will be listed in some metagrammar (Candito 1996; Crabbé et al. 2013) as inheriting the properties of regular transitive verb trees, like in (1), but with further specification of the contents of the object NP.²

¹Of course, the labels are not the nodes, but merely the result of applying some node labelling function to the nodes; I give the specifics of this in Section 5.2. Notice also that this will cause problems when there are two nodes with the same label in a single elementary tree (e.g. the two NPs in (1)); in this case, we can follow the TAG convention of using different labels like NP0 and NP1, even though really such nodes share the ‘same’ label, *viz.* NP.

²There is a problem here in that the VP-internal NP is annotated so as to be the OBJ of the clause (just like with any transitive verb), but in this case contributes no further information since no lexical item will be substituted there. This is an issue I did not notice when writing Findlay (2017a), but which will be addressed in Chapter 6.2.2.

We might admire the simplicity of the TAG-LFG approach, but unfortunately this is also a stumbling block for the theory. The first step in incorporating TAG into LFG is, as noted, associating extended c-structure information with lexical entries. In standard LFG, lexical entries only contain category information. In TAG-LFG, trees (or tree families) are associated with lexical entries directly. But this confuses the important distinction between structures and *descriptions* of those structures (Kaplan 1995), to which we now turn.

5.1.2 Structures and descriptions of structures

Grammatical analysis is concerned with associating structures with sentences. This can be achieved in (at least) two different ways, which Kaplan (1995: 11) calls the CONSTRUCTIVE OR PROCEDURAL approach and the DESCRIPTIVE, DECLARATIVE OR MODEL-BASED approach. In the former, a set of operations builds or manipulates structures, either starting from the string or with the string as a goal. In the latter, a formal description obtained from the string constrains which structures are permissible—any structures which satisfy the description are admitted; they are said to be MODELS of the description. There is of course considerable overlap here with the derivational vs. constraint-based/generative-enumerative vs. model-theoretic distinctions mentioned in Chapter 4.

TAG, as standardly conceived, is firmly in the first camp: structures—trees—are manipulated directly to arrive at the final analysis. On the other hand, LFG is very much of the latter type: Kaplan (1995: 11) calls the descriptive approach the “hallmark of LFG”. TAG-LFG ignores this mismatch, and hence, without further elaboration, is not formally coherent. If we want to make use of a TAG in LFG, we must associate lexical entries not with trees, but with descriptions of trees. In the simplest case, there is a one-to-one correspondence between description and structure; in this case, the mismatch described above is only apparent, and can be resolved by a trivial translation procedure. However, if descriptions involve disjunction or negation, or other operations beyond simple conjunction of propositions, then the relation between structure and

description is no longer isomorphic, and so it becomes important not to confuse the two (Kaplan 1995: 14).

Describing trees is only half the challenge, however; we also need a way to implement the TAG operations of substitution and adjunction on descriptions of trees rather than on trees directly. The former is straightforward, and simply amounts to unification: substituting a tree rooted in node A at node B is just the identification of nodes A and B. Adjunction is more complicated, however, and relies on partial descriptions of so-called *quasi-trees* (Vijay-Shanker 1992: 486ff.). We will have more to say about this in Section 5.4. But first, we explain in more detail how to describe trees.

5.2 A language for tree descriptions

As pointed out by Kaplan (1995: 9ff.), any mathematical structure, like a c-structure or f-structure, can be described by giving “a listing of its defining properties and relations” (*ibid.*: 12). Different kinds of structure will be made up of different kinds of formal objects and different relations between them. LFG makes extensive use of this approach in its functional annotations. These provide a set of properties which various structures in the correspondence architecture, most prominently f-structures, must satisfy. As noted, we say that the structures which correspond to the analysis of the sentence in question are the minimal ones which satisfy the description.

But as well as giving descriptions of f-structures, we can do the same thing for c-structures. Kaplan (1995: 10) gives the following formal language for describing c-structures:

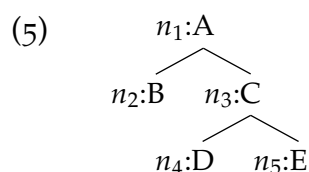
- (3) N: set of nodes, L: set of category labels
 $\mathcal{M}: N \rightarrow N$
 $\prec \subseteq N \times N$
 $\lambda: N \rightarrow L$

In words, this language consists of a set of nodes (call them n_1, n_2, \dots, n_i), a set of category labels (e.g. S, NP, V, etc.) and three relations between them: firstly,

the mother relation \mathcal{M} , which maps nodes to the node which immediately dominates them (i.e. their mother);³ secondly, a partial ordering between nodes (\prec indicates linear precedence; it is partial because e.g. nodes do not precede or follow their daughters); and lastly, the labelling function λ , which maps nodes to their labels.

As an example, (4) gives the description which corresponds to the tree in (5) (Kaplan 1995: 12):

$$(4) \quad \begin{array}{llll} \mathcal{M}(n_2) = n_1 & \lambda(n_1) = A & n_2 \prec n_3 & \\ \mathcal{M}(n_3) = n_1 & \lambda(n_2) = B & n_4 \prec n_5 & \\ \mathcal{M}(n_4) = n_3 & \lambda(n_3) = C & & \\ \mathcal{M}(n_5) = n_3 & \lambda(n_4) = D & & \\ & \lambda(n_5) = E & & \end{array}$$



Notice that once again the description would also be satisfied by indefinitely many larger trees, and so we specify that we need the minimal tree satisfying the constraints.

5.2.1 Phrase structure rules and lexical entries

Since phrase-structure rules essentially describe small trees,⁴ we might also consider them as abbreviations for descriptions in such a formal language, e.g. (6) could be rewritten as (7):

$$(6) \quad S \rightarrow NP VP$$

³This relation is what is sometimes called a partial function; that is, elements in the domain appear a maximum of one time in the left-hand side of a pair, but, unlike in a total function, some elements appear zero times (Partee et al. 1990: 31f.). So no node is mapped to more than one mother, but some are mapped to none—in this case, root nodes, which have no mother.

⁴At least when they are viewed as node admissibility conditions rather than as rewrite rules (McCawley 1968).

- (7) $\mathcal{M}(n_2) = n_1$
 $\mathcal{M}(n_3) = n_1$
 $\lambda(n_1) = \text{S}$
 $\lambda(n_2) = \text{NP}$
 $\lambda(n_3) = \text{VP}$
 $n_2 \prec n_3$

However, we have committed a formal fudge at this point, since phrase-structure rules do not describe one specific tree, but rather a particular *type* of tree. The description in (7), on the other hand, describes a specific tree, since it mentions nodes by name (n_1, n_2, n_3). In describing a type of tree (as opposed to a particular instance or token of this type), we should really speak in terms of variables over node names rather than node names *per se*. A simple way to get around this is to keep the same description as we have in (7), but to assume that the symbols n_1, n_2, n_3 are actually variables over nodes, not names of nodes, and that any variables mentioned in a description are implicitly existentially quantified. This is the convention we will stick to going forward.

Now, since lexical entries in LFG are just special phrase structure rules, they too can be rewritten in this form. For example, the partial lexical entry for *cat* in (8) can be rewritten as (9):

- (8) $cat \quad \text{N} \quad (\uparrow \text{PRED}) = \text{'cat'}$
 $\quad \quad \quad (\uparrow \text{NUM}) = \text{SG}$

- (9) $\mathcal{M}(n_2) = n_1$
 $\lambda(n_1) = \text{N}$
 $\lambda(n_2) = cat$
 $(\phi(n_1) \text{ PRED}) = \text{'cat'}$
 $(\phi(n_1) \text{ NUM}) = \text{SG}$

What this means is that lexical entries do not have to be hybrid formal objects—phrase structure rules with associated annotations in a description language, or triples of word form, category, and description—but can instead be of a single formal kind, namely the functional description, which now covers all levels of the correspondence architecture, including c-structure.

With the tools we have so far, we are now in a position to implement a simple description-based Tree Substitution Grammar (a tree-based grammar which

uses only substitution, not adjunction: see Chapter 4.2.4). A full TAG requires additional considerations, which we return to in Section 5.4.

5.2.2 A description-based TSG

In order to have a point of comparison, I will assume that we are trying to convert a CFG into a description-based TSG. In order to do this, we assign a description of an elementary tree to each of the phrase structure rules, including the lexical entries. In this simple example we ignore the LFG annotations and stick to a plain CFG; the rest of this chapter develops the theory until, in Section 5.5, we reach the point of having full LFG annotations associated with lexical entries and syntactic structures as well. For now, though, we will consider the simple CFG in (10):

- (10) a. $S \rightarrow NP VP$
 b. $VP \rightarrow V NP$
 c. $NP \rightarrow Kim$
 d. $NP \rightarrow Sandy$
 e. $V \rightarrow kisses$

This can be rewritten as a set of descriptions as in (11):

- (11) a. $\mathcal{M}(n_2) = n_1$
 $\mathcal{M}(n_3) = n_1$
 $\lambda(n_1) = S$
 $\lambda(n_2) = NP$
 $\lambda(n_3) = VP$
 $n_2 \prec n_3$
- b. $\mathcal{M}(n_2) = n_1$
 $\mathcal{M}(n_3) = n_1$
 $\lambda(n_1) = VP$
 $\lambda(n_2) = V$
 $\lambda(n_3) = NP$
 $n_2 \prec n_3$
- c. $\mathcal{M}(n_2) = n_1$
 $\lambda(n_1) = NP$
 $\lambda(n_2) = Kim$
- d. $\mathcal{M}(n_2) = n_1$
 $\lambda(n_1) = NP$
 $\lambda(n_2) = Sandy$
- e. $\mathcal{M}(n_2) = n_1$
 $\lambda(n_1) = V$
 $\lambda(n_2) = kisses$

The traditional view of CFGs as rewrite rules sees a derivation as involving repeated replacement of a non-terminal according to one of the phrase structure

rules in the grammar. In our description-based grammar, on the other hand, we simply add together the appropriate sets of equations and find the minimal structure which satisfies the new, combined set. This will involve inferring the identity of certain nodes, which is what gives us the effect of substitution.⁵

Let us consider an example from our toy TSG, namely the substitution of the subject NP, using descriptions (11a) and (11c), repeated below:

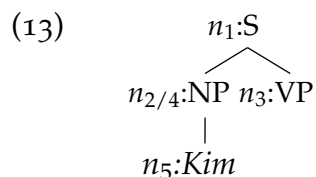
- (11) a. $\mathcal{M}(n_2) = n_1$
 $\mathcal{M}(n_3) = n_1$
 $\lambda(n_1) = S$
 $\lambda(n_2) = NP$
 $\lambda(n_3) = VP$
 $n_2 \prec n_3$
- c. $\mathcal{M}(n_2) = n_1$
 $\lambda(n_1) = NP$
 $\lambda(n_2) = Kim$

Firstly, we notice that variable names are being reused—in this case, n_1 and n_2 are used in both descriptions. We can rename node variables in different descriptions harmlessly (in the manner of α -conversion in the lambda calculus), so let us relabel the variables in (11c), changing n_1 to n_4 and n_2 to n_5 . We can then combine the two descriptions to obtain the following set of equations (the contributions of (11c) are enclosed in boxes, for the sake of illustration):

- (12) $\mathcal{M}(n_2) = n_1$ $\lambda(n_1) = S$ $n_2 \prec n_3$
 $\mathcal{M}(n_3) = n_1$ $\lambda(n_2) = NP$
 $\mathcal{M}(n_5) = n_4$ $\lambda(n_3) = VP$
 $\lambda(n_4) = NP$
 $\lambda(n_5) = Kim$

Our target is to find the minimal tree which satisfies this description. The only way to achieve this is to assume that $n_2 = n_4$. Doing this gives us the following tree:

⁵The formal properties of the trees described will take us a long way, but will not always guarantee that we arrive at a linguistically sensible solution, and so we also employ certain linguistically-motivated constraints, to be discussed in the next section.



Other trees will satisfy the description, including those which do not identify n_2 and n_4 , but they will include additional nodes, and thus will be ruled out by the requirement of minimality.

Although I have singled out the substitution of the subject NP, this was purely for expository convenience: the process does not proceed step-by-step as this might imply. The correct analysis for a sentence is the one which *simultaneously* satisfies all of the constraints imposed by all of its parts: for example, to produce *Kim kisses Sandy*, we would take the union of (11a)–(11e) and find the minimal tree which satisfies the resulting description. This is what makes such a formalism declarative rather than derivational: a single description provides a single minimal tree; there are no intermediary steps.

5.2.3 Some linguistic constraints

Alongside satisfying the description provided by the grammar, and doing so minimally, trees must satisfy three more basic axioms. Firstly, since each sentence must only contain one root node, we prevent parses which posit multiple, unconnected trees:

(14) **Single Root (informal)**

A description can only be satisfied by a single tree (i.e. the solution only involves a single root node).⁶

⁶This constraint will in general fall under the requirement of minimality. If minimality is determined by counting nodes, a single tree which collapses potentially identical nodes will always have fewer nodes than multiple unconnected trees which do not. However, we want to rule out the case where a single tree *cannot* be arrived at from the description; in this case a parse which produces more than one tree should still be invalid, even if it truly is the minimal solution.

Secondly, trees must be non-tangled:

(15) **Non-tangled Trees (informal)**

Node connections in a tree cannot cross.

Thirdly, we want to prevent trees like (13), which have non-terminals at the frontier, from being accepted as the final parse of a sentence:

(16) **Terminal Frontiers (informal)**

The frontier nodes of a tree must be made up entirely of terminal nodes.

To facilitate this last condition, we must, as is standard in a CFG, divide up the labels into terminals and non-terminals. At this point we can update the definition of our description language as follows, incorporating the fact that N is a set of node variables rather than nodes, and that the labels are divided into two subsets:

- (17) N : set of variables over nodes
 C : set of category labels, T : set of terminal labels
 $L: C \cup T$
 $\mathcal{M}: N \rightarrow N$
 $\prec \subseteq N \times N$
 $\lambda: N \rightarrow L$

This also allows us to state our axioms in formal terms, as shown below (where $w, x, y, z \in N$):

(18) **Single Root (formal)**

$$\exists x[\neg\exists y[\langle x, y \rangle \in \mathcal{M}] \wedge \forall z[(\neg\exists w[\langle z, w \rangle \in \mathcal{M}]) \rightarrow z = x]]$$

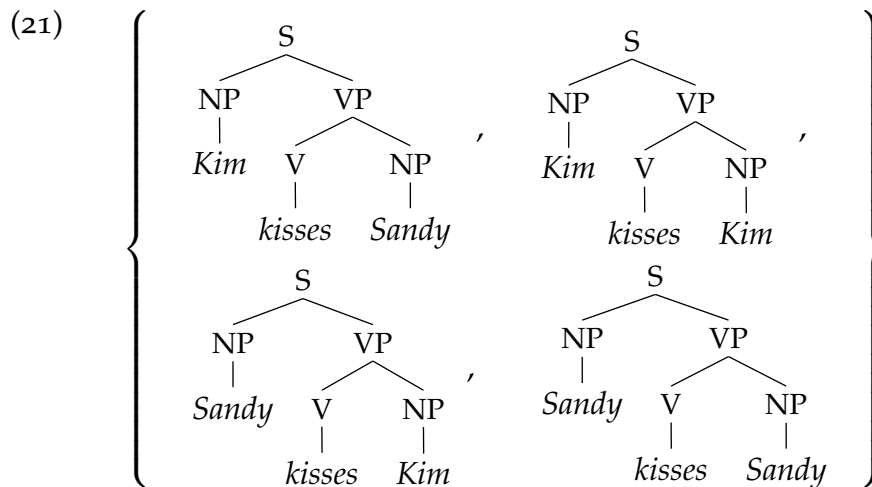
(19) **Non-tangled Trees (formal)**

$$\forall x, y[\mathcal{M}(x) \prec \mathcal{M}(y) \rightarrow x \prec y]$$

(20) **Terminal Frontier (formal)**

$$\forall x[(\neg\exists y[\langle y, x \rangle \in \mathcal{M}]) \rightarrow \lambda(x) \in T]$$

With these assumptions, the grammar in (11) is indeed equivalent to the CFG in (10), and generates the tree language in (21):



Before we extend this description language to cover TAGs rather than merely TSGs, and show how it can be incorporated into LFG, we turn to a related issue, that of lexicalisation, which will bear on the kinds of description that we want to associate with our lexical entries.

5.3 Lexicalisation

One important property that grammars may possess is that of being *LEXICALISED* (Schabes et al. 1988), described by Abeillé & Rambow (2000a: 7) as follows (emphasis in original):

We will call a grammar *lexicalised* if every elementary structure is associated with exactly one lexical item (which can consist of several words), and if every lexical item of the language is associated with a finite set of elementary structures in the grammar.

Most linguistic research using TAG makes use of grammars which possess this property, and this variant is known as Lexicalised TAG (LTAG). By contrast, LFG grammars, despite having ‘Lexical’ in the name, are not in general lexicalised. In this section, we examine the topic of lexicalisation a little further, and ask why we might prefer a lexicalised grammar to a non-lexicalised one.

5.3.1 Lexicalised vs. non-lexicalised

Grammars need not be lexicalised. For example, CFGs are in general not lexicalised. The simple CFG from (10) is repeated below:

- (22)
- a. $S \rightarrow NP VP$
 - b. $VP \rightarrow V NP$
 - c. $NP \rightarrow Kim$
 - d. $NP \rightarrow Sandy$
 - e. $V \rightarrow kisses$

This grammar contains rules (elementary structures) like (22a) or (22b), which consist entirely of non-terminals, and therefore are not associated with a lexical item. Consequently, the grammar does not meet the definition given above, and so does not count as lexicalised.

Tree-based grammars need not be lexicalised either; for instance, the following grammar is the TSG equivalent of (22), and is not lexicalised for the same reason, because the first two elementary structures, in this case trees, are not associated with a lexical item:

- (23)
-
- S VP V NP NP
 \swarrow \swarrow $|$ $|$ $|$
 $NP \downarrow$ $VP \downarrow$ $kisses$ Kim $Sandy$

We can speak of one grammar **LEXICALISING** another if the former is (weakly or strongly) equivalent to the latter, except that the former is lexicalised whereas the latter is not. The Tree Substitution Grammar in (24), for example, **STRONGLY LEXICALISES** both the grammar in (23) and the grammar in (22), since not only does it lexicalise them but it is also strongly equivalent to them—that is, as well as describing the same string language, it also produces the same structures (trees) corresponding to those strings, and so describes the same tree language as well.

- (24)
-
- S NP NP
 \swarrow $|$ $|$
 $NP \downarrow$ VP Kim $Sandy$
 \swarrow \swarrow
 V $NP \downarrow$
 $|$
 $kisses$

Do we need to use a TSG to lexicalise a CFG like (22), though? Is it possible to lexicalise a CFG with another CFG instead? That is, can we keep the

same formalism but simply add the property of lexicalisation? If we require a lexicalising grammar to be *strongly* equivalent to the grammar it lexicalises, it turns out that the answer is no (Schabes et al. 1988). For, although there is a way of converting any CFG into so-called GREIBACH NORMAL FORM (Greibach 1965), where the right-hand side of each rule begins with a terminal symbol—thereby lexicalising the grammar—such grammars do not in general generate the same set of trees as the grammars they normalise, since they will include different (and many more) rules. This means they are only weakly, not strongly, equivalent. To strongly lexicalise a CFG in fact requires not only a TSG, but actually a TAG, as we will now see.

5.3.2 Lexicalising a CFG

We saw in (24) an example of a TSG which strongly lexicalises a CFG, and so we might wonder whether actually TSGs are already sufficient for lexicalising CFGs in general. Certainly, they have one property which makes them well-suited to the task: their extended domain of locality. This is the difference between the lexicalised TSG in (24) and the non-lexicalised one in (23). Since elementary structures are permitted to be trees with depth greater than one, (24) can combine the trees not associated with a lexical item in the non-lexicalised TSG into larger trees which are. Thus, an extended domain of locality is a natural step in the lexicalisation of a CFG.

Unfortunately, this is not on its own sufficient, and so it turns out that TSGs cannot in general lexicalise CFGs. And even in the cases where they *can*, the results can be linguistically unsatisfying. We consider both these cases in turn.

For the formal proof that they cannot fulfill this function, consider the artificial language $\{a^n | n > 0\}$ (that is, any number of *as* except zero).⁷ This can be straightforwardly described by the following CFG:

$$(25) \quad \begin{array}{l} S \rightarrow S S \\ S \rightarrow a \end{array}$$

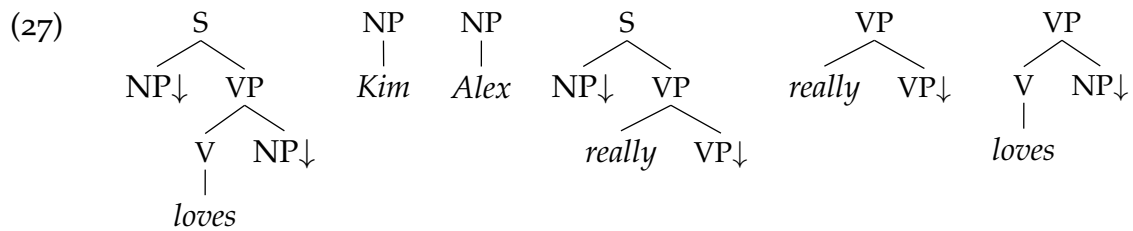
⁷This discussion is drawn from Kallmeyer (2010: 22–23).

However, we cannot describe the same language with a strongly equivalent lexicalised TSG. We can show this by contradiction. Call the CFG in (25) G . Assume that there is a TSG G' which can strongly lexicalise G . Each member of the tree language of G' , $\mathcal{T}(G')$, is derived from some elementary tree anchored in a . Call the length of the path from a to the root of this tree n . Since we cannot substitute at the anchor, any tree derived from this tree will also contain this same path from a to the root. Call the length of the longest such path contained in any elementary tree n_{max} . Then, because the length of this path will not increase during the derivation, there is no derived tree in the language in question in which all of the paths are of length $> n_{max}$. But such trees do exist in the tree language of G : because of the recursive rule $S \rightarrow S S$, the path from the root to each of the terminals can be arbitrarily long. Thus, assuming a finite grammar for G' , there are trees in $\mathcal{T}(G)$ not in $\mathcal{T}(G')$. But we assumed that G' was strongly equivalent to G , so we have reached a contradiction, and our assumption of equivalence must be false. This means that TSGs cannot always lexicalise CFGs.

On the other hand, even when it is formally possible to lexicalise a CFG with a TSG, the results may be quite linguistically unappealing. Let us assume that we want a TSG to abide by the same kinds of linguistic constraints standardly applied to TAG. For example, elementary trees should constitute extended maximal projections (Grimshaw 2005) of their anchors, including their full argument structure. Now, consider the CFG in (26):

- (26) $S \rightarrow NP VP$
 $VP \rightarrow V NP$
 $NP \rightarrow Alex$
 $NP \rightarrow Kim$
 $VP \rightarrow really VP$
 $V \rightarrow loves$

As Schabes et al. (1988: 579) point out, a grammar like this can be lexicalised with the following TSG:



The first three trees are as we have seen before. But to capture the recursive nature of the adverb placement, we need three additional trees, one of which must be a sentential tree with the adverb as its anchor. Such a tree is clearly anomalous, since it is standardly assumed that S is the extended maximal projection of verbs, not of adverbs. Note also that we now require two different elementary trees for both the verb and the adverb. Although it is common in TAG grammars to associate lexical entries with multiple trees, these trees correspond to different argument realisations (e.g. passive vs. active) or extraction constructions (relative clauses, *wh*-questions, etc.), which is not the case here. Finally, one of the elementary trees for *loves* does not encode its full argument structure (the last tree in (27) does not include the subject NP slot), again going against one of the linguistic, if not formal, principles discussed above.

Just like in the formal case discussed above, the problem is once again a recursive rule, this time the fifth rule of the CFG (26), $VP \rightarrow \textit{really} VP$. Again, because in a TSG we cannot extend the path between two nodes in an elementary tree, we need all three of the fourth, fifth, and sixth trees in (27) to capture the fact that the head V of the VP can appear arbitrarily deeply down the tree. But this seems to be missing a generalisation, captured by the recursive rule in the CFG, that this is simply one operation—adjoining more adverbs at the VP node. So in order to satisfactorily lexicalise a CFG, we need to add the operation of adjunction. And, of course, once we add adjunction to a TSG, we arrive at a TAG. Thus, in a certain sense, the quest to lexicalise a CFG leads naturally to a TAG (Schabes et al. 1988).

As we saw earlier, LFG is based on a CFG (augmented with additional structures), and this means that LFG grammars are not, in general, lexicalised.

If we did want to lexicalise LFG, we have seen that we would be naturally led to replacing the CFG with a TAG. However, at this point one might reasonably be wondering whether we *do* want to lexicalise the grammar: after all, I have so far offered no motivation for doing so other than formal interest. Let us, then, take a moment to examine some of the advantages of lexicalised grammars.

5.3.3 Advantages of lexicalised grammars

One early advantage touted for lexicalised grammars was based on parsing. In a lexicalised grammar, a given sentence can contain at most as many elementary structures as there are words in the sentence. If we assume that each lexical item is associated with a finite number of elementary structures, then this also means that the number of analyses for the sentence is finite, and so the recognition problem is decidable (Schabes et al. 1988: 581–582). As Kallmeyer (2010: 21; emphasis in original) puts it, “[l]exicalized grammars are *finitely ambiguous*, i.e., no sequence of finite length can be analyzed in an infinite number of ways”. However, in practice, the dangers of non-terminating parses are virtually non-existent in sensibly-written natural-language grammars, and so this advantage is not so great as it may seem.⁸

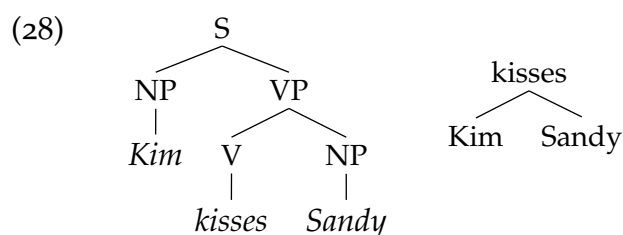
A related claim is that lexicalised grammars assist parsing because “parsing need consider only those trees of the grammar that are associated with the lexical symbols in the input string” (Eisner & Satta 2000: 79–80), rather than searching the whole grammar, and so the specific words used in a sentence “help to restrict the search space during parsing” (Kallmeyer 2010: 20). Once again, however, this argument carries less practical weight than it might seem, since parsing times for TAG grammars are actually rather slow: the best parsing algorithms for TAGs have a worst-case time complexity of $\mathcal{O}(n^6)$ (Joshi & Schabes 1997; Kallmeyer 2010: Ch. 5), as opposed to $\mathcal{O}(n^3)$ in the case of CFGs, for example.⁹ And this is not to mention the fact that state-of-the-art parsers tend to be purely data-driven

⁸My thanks to Adam Przepiórkowski and Timm Lichte for discussion of this point.

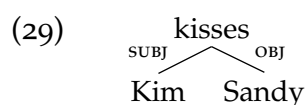
⁹Although Joshi & Schabes (1997) report that the *average* time complexity of the Earley parsing algorithm for TAGs is actually better, e.g. $\mathcal{O}(n^4)$ for unambiguous TAGs.

anyway, and therefore don't even use 'grammars' *per se*. It seems to me, therefore, that the parsing arguments in favour of lexicalisation, which may have once been significant (e.g. when Schabes et al. 1988 were writing), no longer carry so much clout, and their continued repetition in certain textbook presentations of TAG parsing might well be more of a sociological than a scientific phenomenon.

Fortunately, however, they are not the only reasons to prefer a lexicalised grammar: there are also linguistic motivations. The first of these is that lexicalised grammars facilitate a straightforward connection to dependency structures (Kuhlmann 2007), since "combinations during derivation can be interpreted as dependencies" (Kallmeyer 2010: 20). A TAG parse produces not only the final phrase-structure tree, called the *DERIVED TREE*, but also a record of how it was constructed, called a *DERIVATION TREE*. This latter kind of tree has the units which are combined as its nodes, with the substituting or adjoining unit dominated by the unit it is substituting or adjoining to. For example, the derived tree and derivation tree for *Kim kisses Sandy*, using the lexicalised TSG from (24), are given in (28):



Crucially, because the grammar in (24) is lexicalised, the nodes in the derivation tree correspond to lexical items, and so the derivation tree corresponds to a dependency structure. This is made even clearer if we annotate the node connections with e.g. grammatical functions:



Especially given the recent interest in dependency grammars prompted by the Universal Dependencies project (Nivre et al. 2016), it is clearly advantageous if

our formalism has a transparent connection to dependency structures. Using a lexicalised grammar makes this that much easier.

Another linguistic argument is that a lexicalised grammar very strongly instantiates the lexicalist position. Although, as discussed in Chapter 2.1.2, the status of the Lexicalist Hypothesis as an empirical claim may be open to question, what is perhaps less controversial in modern syntactic theories is the broader definition of lexicalism, whereby the lexicon is placed centre stage in the grammar, treated as a “richly structured” object, and whereby the overall syntactic theory assumes “an articulated theory of complex lexical structure” (Dalrymple 2001: 3). This is in stark contrast to the historically more prominent theory of the lexicon as a mere “collection of the lawless”, to use Di Sciullo & Williams’s (1987: 4) term, where it is simply a repository of exceptions, “incredibly boring by its very nature”, about which “there neither can nor should be a theory” (*ibid.*: 3–4). Most modern syntactic theories, including LFG, are lexicalist in this methodological sense of taking the lexicon to be a richly structured object at the heart of the theory, and for them lexicalisation is a sensible step. A lexicalist syntactic theory already needs a richly detailed lexicon, and in the most parsimonious theory, this is *all* it would need, the syntactic component being encoded in the lexical entries themselves. This is just what lexicalisation provides: in LTAG, for example, aside from the basic operations of adjunction and substitution, any other grammatical constraints are described in the elementary trees of lexical items; that is, in the lexicon. In a lexicalised grammar, the lexicon essentially *is* the grammar.¹⁰

5.3.4 *Lexical Functional Grammar*

Ever since Chomsky’s seminal ‘Remarks on nominalization’ (Chomsky 1970), several scholars have pursued a research programme which attempts to treat as lexical many phenomena which were previously treated as syntactic (see

¹⁰This would also be in line with what Baker (2008: 353) calls the BORER-CHOMSKY CONJECTURE, after Borer’s (1984) proposal and Chomsky’s (1995) later adoption of it, whereby parametric variation is restricted to the lexicon.

e.g. Halle 1973; Jackendoff 1975 for some early theories of the lexicon along these lines). Joan Bresnan has been one leading proponent of this approach (cf. e.g. Bresnan 1978, 1982b), and LFG is in many respects the culmination of this line of thought. Yet the formalism underpinning it, the CFG, does not permit the ultimate expression of this philosophy, namely the lexicalisation of the grammar. I believe it is time to change that.

In the standard formulation of LFG, the ‘building blocks’ of the theory are phrase structure rules, some of which correspond to lexical items, and some of which don’t. In the proposal we are developing in this chapter, on the other hand, lexical entries, which are just descriptions of various structures at all levels including c-structure, are themselves the building blocks, being combined by simply taking their union. It seems to me that this is much more in keeping with the descriptive approach that is the “hallmark” of LFG, and places far more importance on a richly detailed and structured lexicon, as is good and proper for a lexicalist theory. Let us call this version of the theory *LEXICALISED LFG*.

We started the chapter with one motivation for integrating TAG into LFG: giving a satisfactory account of MWEs. We now have another: the possibility of lexicalising LFG. However, in order to implement this change, the simple procedure for describing trees outlined in Section 5.2 is inadequate. As Vijay-Shanker (1992) points out, in order to make sense of the TAG operation of adjunction from the point of view of tree descriptions, we must think in terms of quasi-trees, not trees. It is to this distinction and the changes required which we now turn.

5.4 Description-based TAG

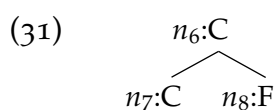
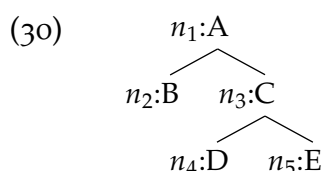
5.4.1 Motivation

As we have seen, the full power of a Tree-Adjoining Grammar, and not just a Tree Substitution Grammar, is required in order to successfully localise MWE dependencies, and in order to lexicalise a CFG. However, the description language

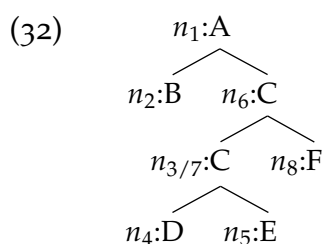
outlined in Section 5.2 is only adequate for a TSG, and not, unfortunately, for a TAG. This is because, as Vijay-Shanker (1992: 486) explains (my emphasis), such a formalism

manipulates (composes) *fully specified* structures (trees in this case) rather than *partially specified* structures. The composition operation of adjoining creates a new structure that does not maintain all of the properties that held in the original (fully specified) structures of which it is composed.

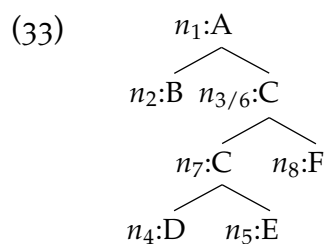
Or, to look at it from the other direction, the operation of substitution in a TSG does not overwrite information contained in a full description of the structures involved, but only ever adds new information. We saw this above in the sample derivation in Section 5.2: resolving the constraints involved positing an equality between two nodes, but this just added an extra equation to the set of equations already present; we did not have to remove any of the existing equations. By contrast, adjunction changes the immediate dominance relationships which the \mathcal{M} function describes. For instance, while a description of (30) will contain the equations $\mathcal{M}(n_3) = n_1$ and $\mathcal{M}(n_4) = n_3$, if we adjoin the auxiliary tree in (31) at node n_3 , at least one of these will cease to be true.



Depending on how we conceive of adjunction, we get one of two trees. In the standard TAG definition of adjunction, the target of adjunction and the foot node of the auxiliary tree are identified, and so we obtain (32):



In this case, $\mathcal{M}(n_3) = n_1$ is now false, since the mother of n_3 is n_6 . If we wanted to preserve this particular immediate dominance relation, perhaps we could assume that the target of adjunction is somehow identified with the root of the auxiliary tree instead of the foot:



But in this case, $\mathcal{M}(n_4) = n_3$ is false instead, since the mother of n_4 is now n_7 . Thus, however we conceive of adjunction, the derived tree which results from it will invalidate a fully-specified description of the tree prior to that adjunction taking place.

This is not a bad thing: it is how TAGs avoid the problems facing TSGs in handling recursive rules, discussed above, for example. Recall that the issue there was precisely that we could *not* change the immediate dominance relations between nodes, in that we could not extend the length of the path between them. Adjunction allows us to do precisely that, but it renders our description language too rigid. In order to allow adjunction, we would have to non-monotonically remove information from our lexically-contributed descriptions.

5.4.2 Partial descriptions of trees

In order to avoid this non-monotonicity, what is needed instead is a looser description which remains true even after adjunction has taken place. We thus give underspecified descriptions of what Vijay-Shanker (1992) calls *QUASI-TREES*, making use of the looser relation of dominance rather than immediate dominance.¹¹ Following Vijay-Shanker (1992: 487), we define dominance as reflexive, so that a node always dominates itself. Beyond this, a node n_1 dominates a node n_2 if n_1 is the mother of n_2 or the mother of the mother of n_2 , or ...—in other

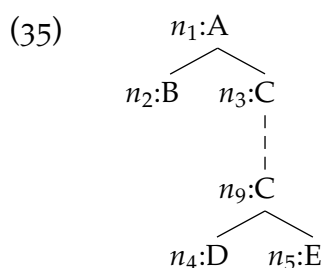
¹¹See Marcus et al. (1983) for more on the advantages of this looser notion of description, especially with respect to parsing.

words, if it appears somewhere in the ancestry of n_2 . In formal terms, dominance is the converse of the transitive, reflexive closure of the mother function:

$$(34) \quad \mathcal{D} := \{\langle x, y \rangle \mid \langle y, x \rangle \in \mathcal{M}^*\}$$

By describing the relationship between some nodes in terms of dominance rather than immediate dominance, we allow for the possibility that other nodes may be added along the path from dominator to dominated, e.g. via adjunction. Unlike immediate dominance, elongating the path between the two nodes does not change the truth or falsity of a statement involving simple dominance: a node dominates its great-granddaughter just as much as it dominates its daughter, or, in the reflexive definition we are working with, itself.

Making use of this looser relation enables us to split potential adjunction sites into two QUASI-NODES, which will be collapsed into a single node in the case where adjunction does not occur. We can replace our working example with the quasi-tree in (35), which is described by (36). The dashed line represents the dominance relation, while the solid lines continue to stand for immediate dominance (Vijay-Shanker 1992: 488).



$$(36) \quad \begin{array}{lll} \mathcal{M}(n_2) = n_1 & \lambda(n_1) = A & n_4 \prec n_5 \\ \mathcal{M}(n_3) = n_1 & \lambda(n_2) = B & n_2 \prec n_3 \\ \mathcal{M}(n_4) = n_9 & \lambda(n_3) = C & \\ \mathcal{M}(n_5) = n_9 & \lambda(n_4) = D & \\ \mathcal{D}(n_3, n_9) & \lambda(n_5) = E & \\ & \lambda(n_9) = C & \end{array}$$

The important thing to note at this point is that nothing in the description in (36) would be incompatible with the assumption that $n_3 = n_9$. And so, since we desire the minimal structure which satisfies our description, we will assume

that the two quasi-nodes are identified unless we have some reason not to. Adjunction at that site would provide just such a reason.

The schematic auxiliary tree introduced in (31) and repeated below as (37), has the description given in (38):

$$(37) \quad \begin{array}{c} n_6:C \\ \swarrow \quad \searrow \\ n_7:C \quad n_8:F \end{array}$$

$$(38) \quad \begin{array}{l} \mathcal{M}(n_7) = n_6 \\ \mathcal{M}(n_8) = n_6 \\ \lambda(n_6) = C \\ \lambda(n_7) = C \\ \lambda(n_8) = F \\ n_7 \prec n_8 \end{array}$$

Combining these two (quasi)-trees, we obtain the full description in (39):

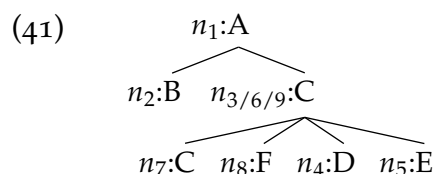
$$(39) \quad \begin{array}{lll} \mathcal{M}(n_2) = n_1 & \lambda(n_1) = A & n_4 \prec n_5 \\ \mathcal{M}(n_3) = n_1 & \lambda(n_2) = B & n_7 \prec n_8 \\ \mathcal{M}(n_4) = n_9 & \lambda(n_3) = C & n_2 \prec n_3 \\ \mathcal{M}(n_5) = n_9 & \lambda(n_4) = D & \\ \mathcal{M}(n_7) = n_6 & \lambda(n_5) = E & \\ \mathcal{M}(n_8) = n_6 & \lambda(n_6) = C & \\ \mathcal{D}(n_3, n_9) & \lambda(n_7) = C & \\ & \lambda(n_8) = F & \\ & \lambda(n_9) = C & \end{array}$$

Since our elementary tree now has a pair of (quasi-)nodes available to be identified with the two nodes in the auxiliary tree, this does accurately describe our desired output, namely (40), where n_3 and n_9 remain separated:

$$(40) \quad \begin{array}{c} n_1:A \\ \swarrow \quad \searrow \\ n_2:B \quad n_{3/6}:C \\ \swarrow \quad \searrow \\ n_{9/7}:C \quad n_8:F \\ \swarrow \quad \searrow \\ n_4:D \quad n_5:E \end{array}$$

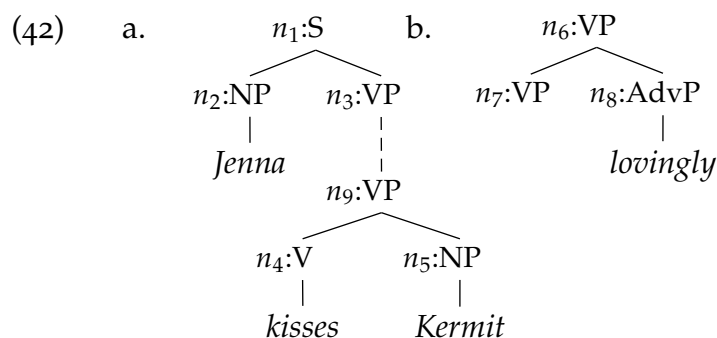
5.4.3 Constraining the formalism

Unfortunately, (40) is not the only structure which (39) describes. If we wanted to maintain the identification of n_3 and n_9 , we could also propose the following tree, which equally well matches the description in (39):

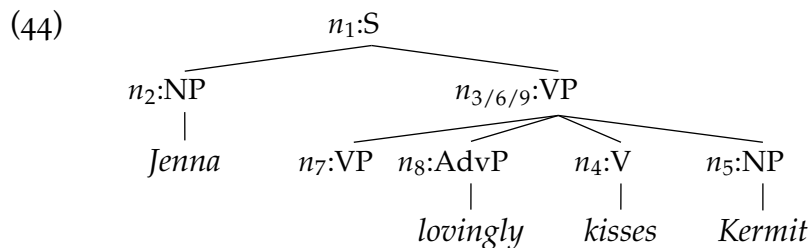
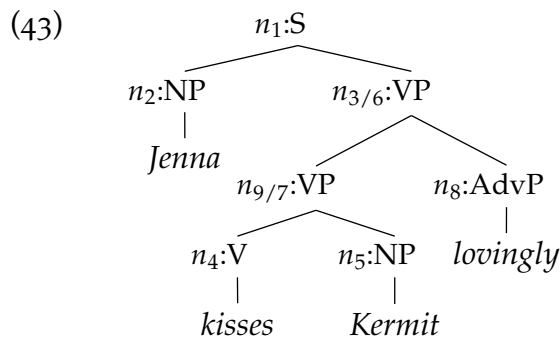


However, this is not the desired result. What rules this out? In this section, we will address the potential for problematic over-generation caused by relying solely on constraint satisfaction.

In the case of example (41), the linguistically motivated constraints can help us. In the present example, we have not distinguished between terminal and non-terminal nodes. If we do this, and replace the trees in question with linguistic examples, then we can successfully rule out the unwanted solution, (41).



I will not present the descriptions, as they will be equivalent to those already given, simply with the labels changed and with additional statements giving the labels and immediate dominance relations of the new terminal nodes. To aid readability, I have therefore not given node names to the terminal nodes in the diagrams either. We can consider how the two contenders given in (40) and (41) fare when translated into these new, complete trees, however. Their equivalents are as follows:

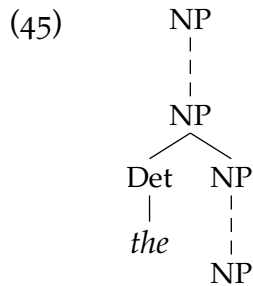


The tree in (44) falls foul of the condition named Terminal Frontier, which requires that all frontier nodes be terminal nodes. In (44), n_7 is a frontier node but it is a non-terminal, therefore this is not a tree which the grammar will generate, either in production or comprehension. In comprehension, of course, (44) would not be a valid parse of *Jenna kisses Kermit lovingly* anyway, since it misorders the words in the string. But *Jenna lovingly kisses Kermit* is also a valid string in English, and so (44) must still be ruled out as a valid parse for *this* string.

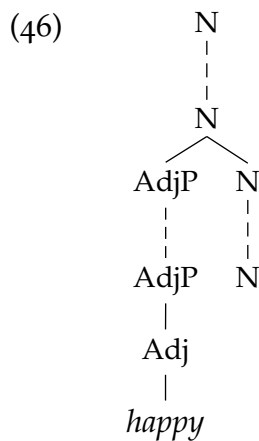
Terminal Frontier requires that candidates for the full parse of a sentence do not have non-terminals at their frontiers. However, what is stopping us deriving (44) as an intermediate step, and then substituting another tree at the VP node n_7 ? In this case, the worry does not arise, since there will be no elementary trees rooted in VP (the maximal extended projection of a verb is S, not VP). The only trees rooted in VP will be auxiliary trees, and inserting another auxiliary tree at n_7 will not help us escape the strictures of Terminal Frontier, since all auxiliary trees include a non-terminal at one of their frontier nodes.

There are cases where the concern might be relevant, however: with NP nodes in the same situation, for example, since there *are* elementary trees rooted in NP, and therefore Terminal Frontier would not be enough on its own to rule out the unwanted analysis. However, the number of auxiliary trees with NP

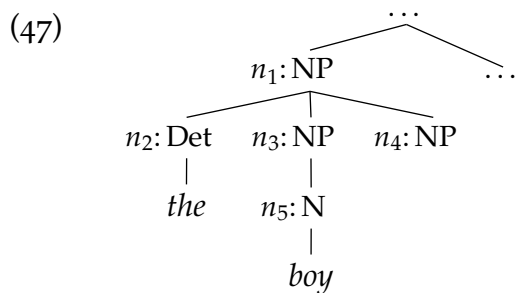
feet/roots is also limited, in the usual case to determiners (e.g. in Bleam et al. 2001). In our terms, determiners will have quasi-trees like the following:



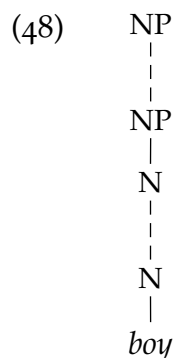
Other nominal modifiers, like adjectives, are restricted to N, which does not root any elementary trees:



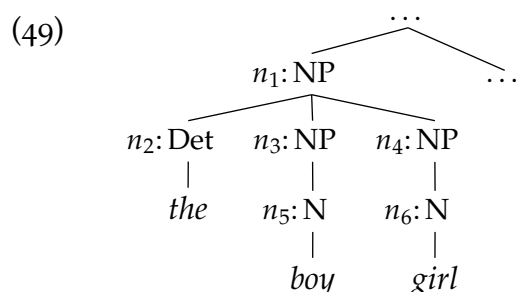
The potentially problematic case, then, is exemplified in (47):



Here we have adjoined the tree in (45) to the elementary tree for *boy*, given in (48), and then collapsed the various quasi-nodes.



Although (47) would be ruled out by Terminal Frontier, the problem is that we could substitute another tree like (48) at the NP node n_4 :



But this yields sentences like (50), which are ungrammatical:

(50) *The boy girl is sleeping.

In this case, we cannot rely on the tree-description language alone to save us. But fortunately this is not the only means we have at our disposal. Once we have integrated our TAG into the LFG framework, we can also draw on other components of the architecture in constraining the grammar. In this case, (50) would clearly result in a resource surplus in the semantics, for example, since the determiner *the* will only take one of the nouns as an argument, leaving the other left over. This is quite in keeping with the standard approach in LFG: it is the co-constraining nature of the correspondence architecture which correctly characterises all and only the possible expressions in a language; any one component taken alone will be inadequate. With this in mind, let us now turn to the question of how we can incorporate our new description-based understanding of TAG into the LFG architecture.

5.5 Lexicalised LFG

LFG lexical entries include a functional description which gives details of all levels of structure except c-structure. Now armed with our description-based TAG, we can also include c-structure in this description, and reduce lexical entries to functional descriptions alone. Thus, in Lexicalised LFG a lexical entry is a description, i.e. a set of propositions, which covers all of the different levels of structure included in the correspondence architecture. When several lexical entries are combined, we simply take the union of their descriptions; the structure they describe is the minimal structure which satisfies the combined set of statements. This section will give extensive exemplification in order to show how this works. In it, we will cover the basics of the grammar as it applies to normal, single-word expressions; Chapter 6 extends the coverage to MWEs. We begin by giving examples of lexical entries for various major categories, starting with the nominal domain, before moving onto verbs and then to modifiers of various kinds. For the most part, this involves simply porting existing LFG analyses and adapting them to the TAG setting, often borrowing the TAG syntax from the XTAG grammar of English (Bleam et al. 2001). We then look at how long-distance dependencies can be analysed, and how syntactic flexibility can be captured using `LEXICAL RULES`, which serve the same purpose as tree families in TAG. The linguistic analysis presented in this section will not be very detailed, but I hope it will serve as a proof of concept and lay the framework for the more detailed discussion of MWEs in the following chapter. Integrating many LFG and TAG analyses of the various complex corners of the grammar into the mixed framework of Lexicalised LFG remains a task for future work.

Before embarking on this task, however, there are a few points of house-keeping we must cover. Firstly, rather than writing out descriptions for each lexical entry, I will for the most part represent tree descriptions as the trees they describe. This is both for the sake of readability and in order to maintain some continuity and cross-compatibility with previous work in LFG and TAG.

However, it must be borne in mind that the *actual* grammar is made up of descriptions, not trees: sometimes this distinction can be important, as when trees are underspecified, for example. When this distinction becomes relevant, I will either introduce a notational convention for handling it, or revert to speaking directly about descriptions.

Secondly, because of the extended domain of locality afforded to us by a TAG, when we add LFG annotations to a tree, we are faced with a choice about where to put them. For example, in a verbal tree, the V, VP, and S nodes will all correspond to the same, clausal, f-structure, and so any functional annotations referring to that f-structure could appear under any of the three. I adopt the convention of maintaining the same position for annotations as would be found in standard LFG analyses, where the choice is restricted by the lack of an extended domain of locality, once again maintaining continuity with existing LFG work where possible.

Thirdly, when talking about descriptions, it will sometimes be useful to talk about them as sets, and sometimes as conjunctions of propositions. Sets can be combined and manipulated easily, but they are not the sort of things which are true or false, and that is what we need in order to talk about a structure or structures *satisfying* a description. In the base case, then, we treat descriptions as sets of equations, but I provide a predicate Pr which converts its set argument into a conjunction of its members:

$$(51) \quad \text{Pr}(D) := \bigwedge_{x \in D} x$$

Fourthly, constraining equations necessitate a slight complexification of the simple picture painted so far. I will treat sets of constraining equations as descriptions over defining equations; that is, as higher-order descriptions. This means that lexical entries are in fact pairs, $\langle D, C \rangle$, where D is the set of defining equations and C the set of constraining equations. We can then define the different kinds of constraining equations in the following way:

$$\begin{array}{ll}
 (52) & \text{Constraining: } F =_c V \equiv \exists d \in D[d \rightarrow F = V] \\
 & \text{Existential: } F \equiv \exists d \in D[d \rightarrow \exists V[F = V]] \\
 & \text{Negative existential: } \neg F \equiv \neg \exists d \in D[d \rightarrow \exists V[F = V]]
 \end{array}$$

This way, constraining equations use a more complex logic, which includes quantification, while defining equations can make do with just negation. They could do without even that, except that it makes sense to keep simple negative defining equations in D , since they can be assessed straightforwardly against a proposed analysis without recourse to quantification over descriptions. They state that some relation does not hold, and we can inspect the proposed structures to see whether that is true in just the same way we can inspect them to see if a relation does hold in the case of a positive defining equation. The only way a negative defining equation can be false is if the relation it bars *does* hold, and the only reason this would be the case is if some other equation said that it was. But in that case the two defining equations would conflict and there would be no analysis which satisfied the description, so the constraining equation will have done its job.

Notice that by seeing constraining equations as descriptions over defining equations, we remove the requirement to assess them only *after* the defining equations have proposed a model. Instead, the two sets of equations can be checked simultaneously, since we want to know whether the constraining equations hold of the description, rather than of the structure.

Finally, when it comes to writing lexical entries in Lexicalised LFG, I will use standard LFG notation, even though this may be otiose or slightly inaccurate. For example, I will continue to write constraining equations and defining equations together in a single lexical entry, even though strictly speaking they are separated. Also, I will use the metavariables \uparrow and \downarrow when annotating trees, despite the fact that the actual descriptions will make reference to specific nodes directly rather than relatively. This continues the trend of both aiding readability and avoiding too great a rupture with other work in LFG.

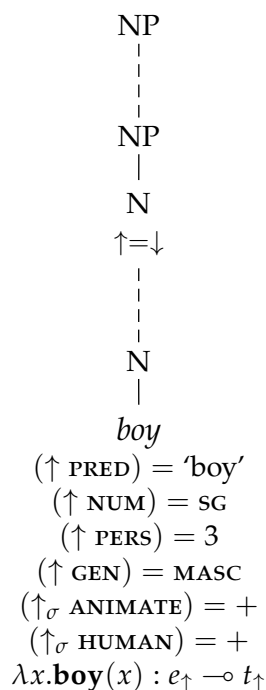


Figure 5.1: Lexical entry for *boy*

We are now in a position to illustrate Lexicalised LFG. We begin in the next section by looking at the NP, before moving onto verbs, modifiers, and finally the question of syntactic flexibility and how we can use lexical rules to model it.

5.5.1 Noun phrases

Nouns are perhaps the simplest lexical entries in terms of syntax: they project NPs, with appropriate adjunction sites to allow both determiners and adjectives to adjoin (at the NP and N levels respectively). We have already seen one example of the kind of tree needed for *boy*; Figure 5.1 shows it with the appropriate LFG annotations. This entry includes a large amount of lexical information, including f-structure and s-structure features. In general, I will not include all of this information in lexical entries, unless it is relevant to the question at hand. Of course, in the full grammar, all of this detail will be present, but it is convenient to retain only the crucial components for our high level discussions.

Normally, only the upper half of a pair of quasi-nodes will be annotated, since the lower half will always unify with some other node (the upper half in

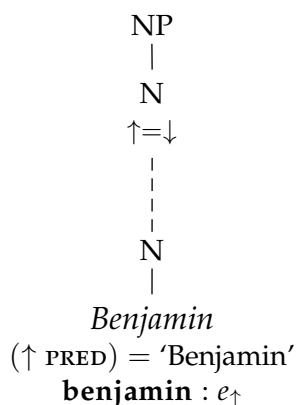


Figure 5.2: Possible lexical entry for *Benjamin*

the simple case), which will bear its own annotations. That is, the lower quasi-node merely indicates adjunction possibilities, and so only makes an important contribution to the configurational part of the syntax (namely, c-structure). We can see this here with the pairs of NP and N quasi-nodes. The root node, too, is unannotated. This is another general convention: in the case of NPs, the tree will ultimately be substituted into some argument position, and so will bear the annotation of the substitution site (see the following section for some examples). Taken together, these two facts mean that only one of the non-terminal nodes in Figure 5.1 bears an annotation—in this case, the upper N, which is annotated $\uparrow=\downarrow$.

We can control what adjunction is allowed by altering how many quasi-nodes a lexical entry contains. For example, we might decide that proper nouns in English cannot host determiners, and so they should have a single NP node rather than two quasi-nodes, as shown in Figure 5.2. However, on the basis of expressions like *the Benjamin that I know*, I will leave open the possibility of determiners adjoining here, and assume that in all but the appropriate constructions it will be ruled out by the semantics.¹² Nevertheless, this case illustrates the flexibility we have regarding our descriptions: modification can

¹²This is also what is assumed in standard LFG: as discussed above, the various levels of the grammar are crucially co-constraining, and any single level of representation on its own will likely be inadequate to fully rule out everything we want to exclude.

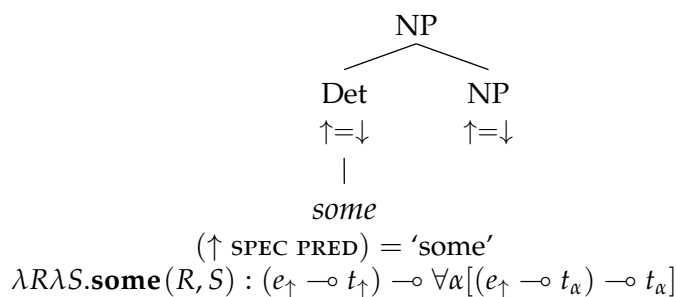


Figure 5.3: Lexical entry for *some*

be permitted or blocked explicitly on the syntactic level, as well as at other levels of representation.

Now that we have a basic NP tree, we turn to the determiners. Following XTAG (Bleam et al. 2001), I represent these as auxiliary trees which adjoin at an NP node (allowing for the possibility of determiner stacking, e.g. *all these bottles*).¹³ Figure 5.3 shows the lexical entry for *some*; other determiners will have the same structure. As is standard in LFG, determiners and nouns are coheads of the same f-structure (see e.g. Bresnan et al. 2016: Ch. 6 on this COHEAD PRINCIPLE), and so the Det node is annotated with $\uparrow=\downarrow$.

While determiners adjoin at NP, the N quasi-nodes in nominal lexical entries host modifiers such as adjectives, relative clauses, and PPs. We will discuss these separately, in Section 5.5.3.

5.5.2 Verbs

Verbal elementary trees are rather more complex than nominal ones, since, following standard TAG practice, they also incorporate the argument structure of the verb. What is more, verbs and their arguments can appear in a number of different configurations beyond the simple active voice declarative sentence—passives, relative clauses, *wh*-questions, etc.—and these all need to be associated with the verb somehow. In many TAG grammars, this is achieved via tree families: associating a lexical item with multiple different syntactic frames. In

¹³Of course, something extra must be said to constrain the relative ordering of such expressions, in order to rule out **these all bottles*, etc. But this is a general problem (cf. the relative ordering of adjectives in the NP) which I do not propose to solve here.

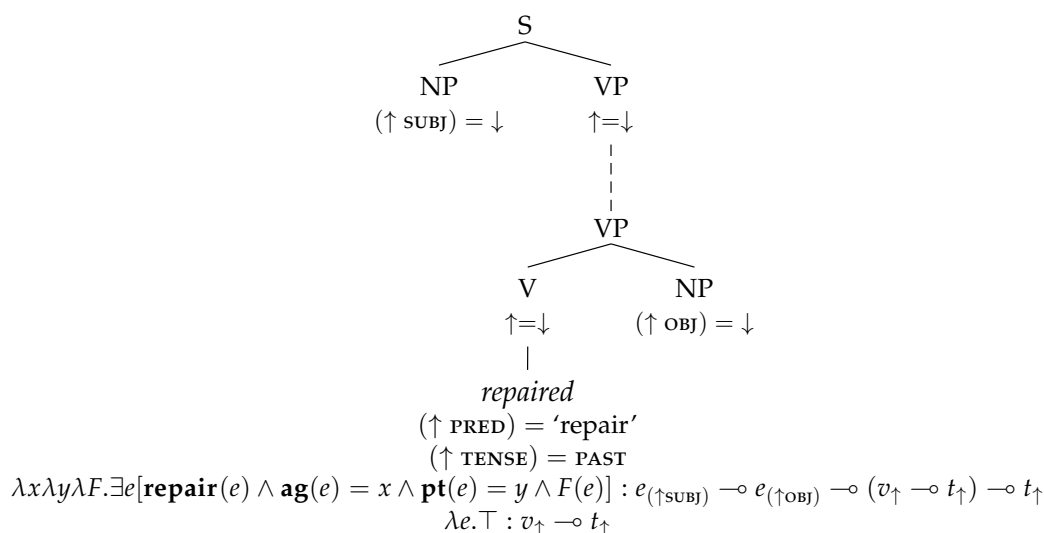


Figure 5.4: Lexical entry for *repaired*

Lexicalised LFG, we instead use lexical rules to associate a basic lexical entry with other structures. The details of this will be presented in Section 5.5.5; in this section, I will just give the simple active voice, declarative forms of verbs.

We begin with an example. Figure 5.4 shows the lexical entry for *repaired*. There is nothing particularly surprising here. The c-structure description contains slots for the subject and object NPs, exactly as we would expect from a TAG-style representation of a transitive verb, and it allows adjunction at the VP node to account for the possibility of adverbial expressions appearing here:

- (53) a. We [_{VP} [_{VP} repaired the hole] carefully/in a rush/wearing galoshes].
 b. We [_{VP} carefully/quickly/nearly [_{VP} repaired the hole]].

Similarly, the meaning constructors have the standard form we would expect from a transitive verb, one with dependencies on both its subject and object, and a separate meaning constructor for Champollion's (2015) closure meaning. Other verbs will have similar lexical entries, *modulo* different categories or grammatical functions for their arguments. Verbs like *say*, for instance, have an S node instead of an NP node as the c-structure complement of V. Figure 5.5 shows the lexical entry for *said*.

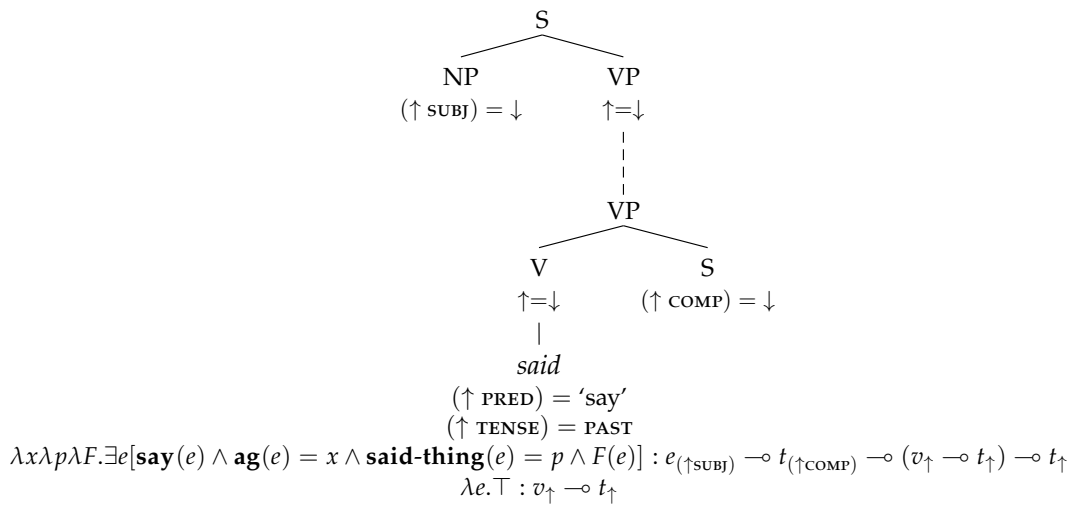


Figure 5.5: Lexical entry for *said*

This particular configuration has an important additional consequence beyond selectional restrictions, however: since the root node and one of the frontier nodes share a label, namely *S*, this makes the tree in Figure 5.5 an auxiliary tree. This means it can adjoin to a pair of *S* quasi-nodes, potentially therefore expanding a tree from the inside rather than simply stacking ‘on top’ of an existing clausal tree. As we saw in Chapter 4.2.4, this is of particular importance in accounting for long-distance dependencies in a TAG setting. We will see an example of a long-distance dependency involving *said* in Section 5.5.4.

One final topic to discuss before we move onto modifiers is that of auxiliary verbs. Following the approach of XTAG (Bleam et al. 2001), auxiliary verbs are represented as auxiliary trees which adjoin to VP. Figure 5.6 provides the lexical entry for perfective *has*, by way of illustration. As argued by Falk (2003), I assume that this auxiliary is an f-structure cohead with the main verb, and simply contributes present perfect feature information; thus, the foot node VP is annotated with $\uparrow = \downarrow$. I ignore the question of the meaning contribution of tense and aspect markers here (instead obscuring the present perfect meaning in a template), but see Haug (2008) and Lowe (2015b) for some suggestions as to how this can be implemented in a Glue semantics setting. Similarly, I do not explain how it is that auxiliaries constrain the form of their verbal c-structure

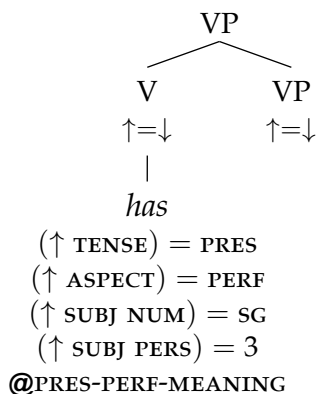


Figure 5.6: Lexical entry for perfective *has*

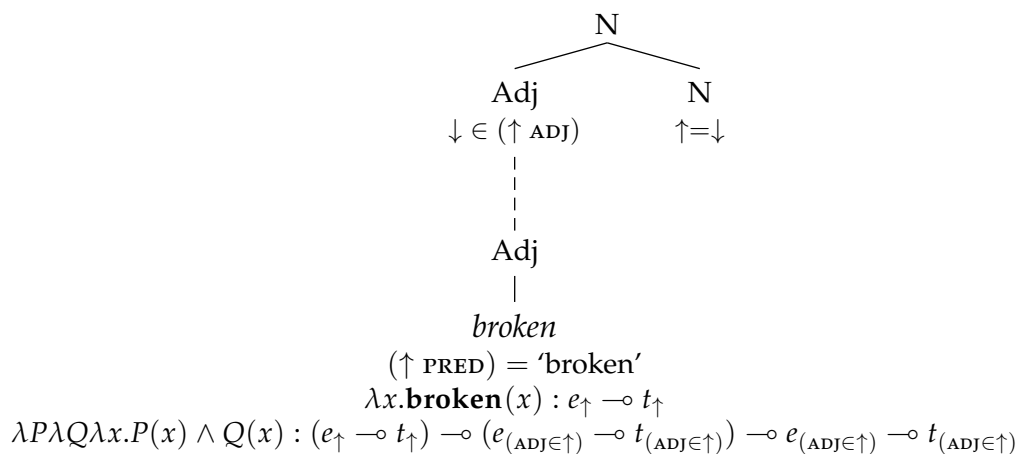


Figure 5.7: Lexical entry for *broken*

complements (e.g. *has worked* vs. **has work/works/working*), assuming that this can be handled through any of the proposed methods in standard LFG (e.g. via a separate level of m(orphosyntactic)-structure, as in Butt et al. 1996b; see also Dalrymple 2001: 178–180 for some discussion).

5.5.3 Modification

The canonical classes of modifiers are unproblematic: as we would expect from TAG, they anchor auxiliary trees, and as we would expect from LFG, they contribute to the ADJ set of the node they adjoin to. Figure 5.7 shows the lexical entry for the adjective *broken* as a pre-nominal modifier. The Adj node is split into a pair of quasi-nodes because of the possibility of adverbial adjunction here:

(54) the [_{Adj} completely/barely/very [_{Adj} broken]] computer

Also, note that there is no AdjP: because pre-N adjectives in English cannot take complements, Sadler & Arnold (1994) argue that they should not project full AdjPs. Post-nominal Adjs do project full phrases, however, since they can (and sometimes must) take complements:

- (55) a. the proud (*of his pupil) teacher
 b. the teacher proud *(of her pupil)

Another motivation for a full phrasal projection for pre-nominal adjectives would be the existence of different adjunction possibilities at AdjP and Adj, as we have in the nominal domain, for instance, where determiners adjoin at NP but adjectives adjoin at N. But there is only one class of tree which can adjoin to Adjs, namely the adverbs, and so no such distinction is warranted on these grounds either. We therefore omit unnecessary structure, and merely include the Adj node. This is also how such trees are implemented in XTAG.

Adverbs are rather more variable than adjectives when it comes to where they can appear. As well as the canonical S and VP modifying adverbs, they can also modify adjectives, other adverbs, PPs, NPs, and determiners (Bleam et al. 2001):

- (56) a. **extremely** good
 b. rich **enough**

- (57) a. oddly **enough**
 b. **very** well

- (58) **right** through the wall

- (59) **quite** some time

- (60) **exactly** five men

Each of these uses requires a different elementary tree, since the category of root and foot will change; what is more, where there is the possibility of both pre- and post-modification, separate trees are required for each. This is not to say multiple descriptions are needed, however: underspecification allows us to

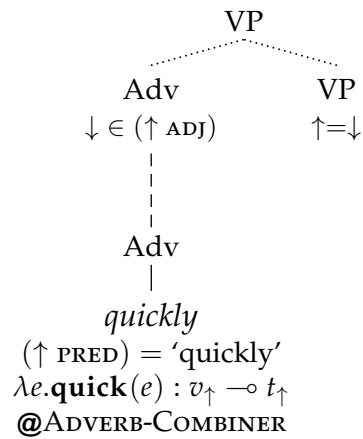


Figure 5.8: Preliminary lexical entry for *quickly*

capture the ordering possibilities, for instance. A classic manner adverb like *quickly* can appear both before and after the VP:

(61) Kira (quickly) ran to the weapons locker (quickly).

When we describe this tree, therefore, we can simply omit any statement about the ordering of the two daughters of the root node. I represent this in diagrams using a dotted line (distinct from the dashed line which represents the dominance relation between a pair of quasi-nodes), which indicates that the ordering of the nodes is unspecified. Figure 5.8 shows this in a preliminary lexical entry for *quickly*. For readability, I abbreviate the combinatorial meaning constructor in a template. The contents of ADVERB-COMBINER is (62):

(62) $\lambda G \lambda \mathcal{V} \lambda F. \mathcal{V}(\lambda e. F(e) \wedge G(e)) :$
 $(v_{\uparrow} -\circ t_{\uparrow}) -\circ ((v_{(\text{ADJ}\uparrow)} -\circ t_{(\text{ADJ}\uparrow)}) -\circ t_{(\text{ADJ}\uparrow)}) -\circ$
 $(v_{(\text{ADJ}\uparrow)} -\circ t_{(\text{ADJ}\uparrow)}) -\circ t_{(\text{ADJ}\uparrow)}$

When it comes to the category of the root and foot nodes, however, underspecification is not enough: I am not aware of any one adverb which can modify *all* of the categories exemplified in (56)–(60); rather, each must specify some subclass. For instance, *quickly* can modify S and VP, but not NP, PP, or Det:

(63) a. [S Quickly, [S Worf stood up]].
 b. Worf [VP quickly [VP stood up]].

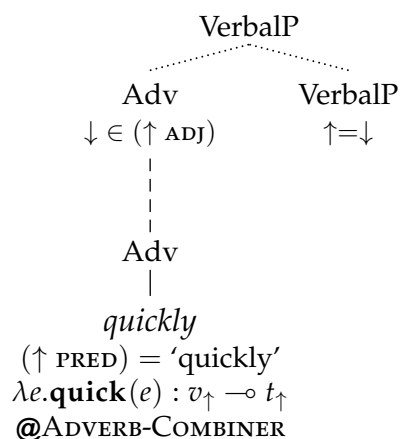


Figure 5.9: Lexical entry for *quickly*

- (64) a. *[NP quickly [NP some time]]
 b. *[PP quickly [PP through the wall]]
 c. *[Det quickly [Det five]] men

I therefore propose a meta-category VerbalP, defined as the disjunction of S and VP, with which to label the root and foot nodes of such manner adverbs.

- (65) VerbalP \equiv {S|VP}

We ensure that the foot and root nodes have the same category (S or NP) by equating them in the description:

- (66) $\lambda(\text{root}) = \lambda(\text{foot}) = \text{VerbalP}$

Figure 5.9 shows the final lexical entry for *quickly*; other manner adverbs will be similar. Of course, adverbs like *very*, which can only modify adjectives or other adverbs, will have different entries, capturing their restrictions. In the case of *very*, for instance, it always pre-modifies, so no underspecification of ordering is needed, and it modifies either Adj or Adv, so we can define a new meta-category A which covers both of these.

- (67) A \equiv {Adj|Adv}

Figure 5.10 shows the lexical entry for *very*. The contents of VERY-MEANING is given in (68):

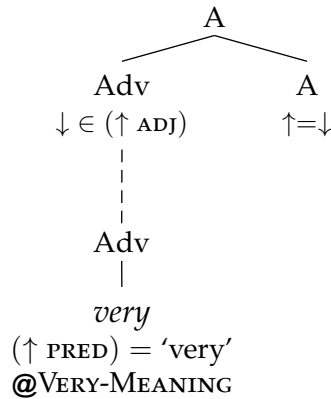


Figure 5.10: Lexical entry for *very*

$$\begin{aligned}
 (68) \quad & \lambda p.\mathbf{very}(p) : t_{\uparrow} \multimap t_{\uparrow} \\
 & \lambda Q\lambda P\lambda x.Q(P(x)) : \\
 & (t_{\uparrow} \multimap t_{\uparrow}) \multimap (E_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}) \multimap E_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}
 \end{aligned}$$

Because *very* can modify both adjectives, whose lexical meaning constructors are of type $\langle et \rangle$, and adverbs, whose lexical meaning constructors are of type $\langle vt \rangle$ or $\langle st \rangle$, we make use of a supertype E in its combinatorial meaning constructor. This supertype's domain is defined as the union of the domain of entities and eventualities (events and states), i.e. $D_E = D_e \cup D_e$, which therefore loosens *very*'s combinatorial possibilities.¹⁴

5.5.4 Long-distance dependencies

One of the hallmarks of TAG is that long-distance dependencies can be handled 'locally': because of the extended domain of locality coupled with the operation of adjunction, we can encode the dependency between a displaced element and the verb of which it is an argument in a single elementary tree. In this section I show how this works in Lexicalised LFG by using relative clauses as a case study, before turning in the next section to the question of how the different syntactic realisations of a predicate are related by lexical rules.

¹⁴Using a looser typing like this might be thought to pose problems: for instance, modification of a strict type $\langle et \rangle$ resource by *very* will turn it into the more permissive type $\langle Et \rangle$. However, the functional information will ensure that this is not a problem: *very* will never be in the ADJ set of something which contributes both a type v and type e resource, for instance.

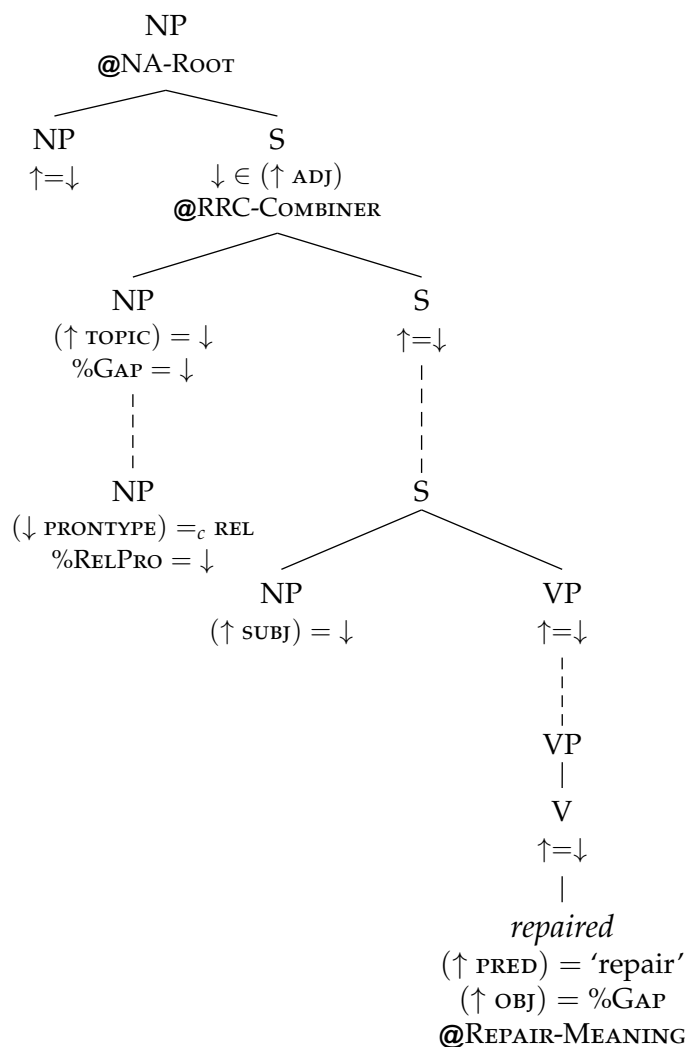


Figure 5.11: Lexical entry for a relative clause headed by the object of *repaired*

Figure 5.11 shows the lexical entry for a relative clause headed by the object of *repaired*. As is usual for a modifier, this is an auxiliary tree; following the XTAG grammar (Bleam et al. 2001), I represent it as adjoining at NP (and since relative clauses are postnominal in English, it attaches to the right).

Without further stipulation, we would now expect at least two parses for NPs containing both a relative clause and a determiner: one where the determiner attaches above the relative clause, and one where it attaches below:

- (69) a. [NP the [NP computer [S which Miles repaired]]]
 b. [NP [NP the computer] [S which Miles repaired]]

On the basis of relative clauses involving collective predicates or reciprocals like those in (70)–(71), Blear et al. (2001) argue for the second analysis, whereby the determiner attaches below the relative clause.

- (70) the man and the woman who met on the bus
- a. ??[_{NP} the [_{NP} man and the woman [_S who met on the bus]]]
 - b. [_{NP} [_{NP} the man and the woman] [_S who met on the bus]]
- (71) the man and the woman who like each other
- a. ??[_{NP} the [_{NP} man and the woman [_S who like each other]]]
 - b. [_{NP} [_{NP} the man and the woman] [_S who like each other]]

This is in contrast to other authors such as Dalrymple (2001: 402) who assume the first analysis, with determiners attaching high. I do not take a principled stand on this matter here, but for the sake of concreteness I assume the analysis of Blear et al. (2001). In order to implement this, we need to prevent adjunction above the top NP node of the relative clause. This is what the template NA-ROOT accomplishes, our version of TAG's null adjoining constraint, which prevents adjunction at a specified node. In our description-based TAG, we normally don't need such a constraint, because, unlike standard TAG, we must specify where adjunction *can* occur, by using pairs of quasi-nodes; the unmarked case is for there to be no adjunction permitted, since explicitly specified immediate dominance relations make it impossible. However, this is not enough for root and frontier nodes, since they, by definition, do not have any further dominance relations specified (up or down, respectively), and so we cannot ensure that they will not be unified with one of a pair of quasi-nodes. In this case, then, we must explicitly rule out adjunction. For root nodes, we need the constraint in NA-ROOT:¹⁵

¹⁵The equivalent constraint for frontier nodes is given in (i):

- (i) NA-FRONTIER :=
 $\neg[\exists n[\langle n, * \rangle \in \mathcal{M} \wedge \lambda(n) = \lambda(*)]]$

This is slightly more complex because we must quantify over all of the (possible) daughter nodes.

$$(72) \quad \text{NA-Root} := \lambda(\mathcal{M}(*)) \neq \lambda(*)$$

In words, this says that the current node does not have the same label as its mother. Since adjunction always involves a mother and daughter node sharing a label, the constraint in (72) essentially blocks adjunction above the node it annotates.¹⁶

Notice that once again both NP arguments of the verb are present in the elementary tree, albeit in rather a different configuration from the basic case. The subject of *repaired* appears in the same place as usual; the object, though, does not appear inside the VP—instead, an annotation on the verb assigns it a local name (Crouch et al. 2012), %GAP. A local name is simply a name assigned to an f-structure that allows us to refer to it elsewhere in the same description; here we specify that the f-structure of the object of *repaired* is the same as the f-structure of the fronted NP.¹⁷ Of course, there is an intervening pair of S quasi-nodes between this NP slot and the verb, which means that we can adjoin sentential embedding verbs here, and thereby split up the verb from its fronted object. This is how we obtain the unbounded nature of such dependencies: adjunction can be applied over and over again, resulting in a tree of arbitrarily large depth. However, the dependency between the relative pronoun and the verb need not be described in terms of such arbitrarily long paths; rather, because of TAG's extended domain of locality, it can be encoded locally in the elementary tree.¹⁸

¹⁶One problem with assuming the analysis of Bleam et al. (2001) over that of Dalrymple (2001) is that we rule out stacked relative clauses like (i), which seem perfectly grammatical:

(i) the [_{NP} [_{NP} [_{NP} meal [_S that Miles cooked]] [_S that Julian knew he wouldn't like]]]

This is because the null adjoining constraint on the top of relative clause trees will prevent them appearing under one another, meaning there can only be one relative clause at a time. Clearly, there are downsides to both approaches.

¹⁷I thank Dag Haug for useful discussion on this point; the idea to use a local name to identify the f-structure of the gapped argument was his.

¹⁸This use of local names therefore means that we no longer have to use functional uncertainty in describing long-distance dependencies. This is a major strength of the extended domain of locality offered by a TAG. An obvious question to pose is whether we now need functional uncertainty at all, since it was invented precisely to handle unbounded dependencies (Kaplan & Zaenen 1989). However, it does now play a role in many other parts of the grammar, perhaps most notably in LFG's binding theory (Dalrymple 1993; Dalrymple et al. 2018), and these cannot so easily be recast in terms of an extended domain of locality.

In addition to the local name %GAP for the gapped argument of the verb, we also use the name %RELPRO to refer to the f-structure of the relative pronoun itself. In the simple case, these two will be identical. However, in cases of pied piping (Ross 1967), they need not be:

(73) the computer [%GAP the processor of [%RELPRO which]] Miles repaired

The meaning of %GAP is the meaning which the verb is looking for (here, the OBJ of *repaired*), while the meaning of %RELPRO is the argument which is left open in the meaning of the relative clause (here, the OBJ of *of*). These may be identical, of course (*the computer Miles repaired*), but they need not be. In the final meaning for the NP modified by the relative clause, %RELPRO will be identified with the head noun, while %GAP may simply contain that meaning as an argument. For example, the meaning of (73) is (74):

(74) $\iota x[\mathbf{computer}(x) \wedge \exists e[\mathbf{repair}(e) \wedge \mathbf{ag}(e) = \mathbf{miles} \wedge \mathbf{pt}(e) = \iota y[\mathbf{processor}(y) \wedge \mathbf{part-of}(y, x)]]]]$

The meaning of the gapped argument in this case is the patient of *repaired*, namely $\iota y[\mathbf{processor}(y) \wedge \mathbf{part-of}(y, x)]$, while the meaning of the relative pronoun corresponds to the variable x . In order to see how this works out compositionally, we need the meaning given by RRC-COMBINER (where ‘RRC’ is mnemonic for ‘restrictive relative clause’), which controls the combination of the relative clause and the noun it modifies. This is shown in (75) (based on Dalrymple 2001: 417):

(75) RRC-COMBINER :=
 $\lambda P \lambda Q \lambda x. Q(x) \wedge P(x) :$
 $(e_{\%RELPRO} \multimap t_{\downarrow}) \multimap (e_{(ADJ \in \downarrow)} \multimap t_{(ADJ \in \downarrow)}) \multimap e_{(ADJ \in \downarrow)} \multimap t_{(ADJ \in \downarrow)}$

The meaning constructor in (75) consumes the unsaturated meaning of the relative clause to produce a modifier type which conjoins the clausal meaning with the nominal meaning, while predicating both of the same variable.

Putting this all together, Figures 5.12 and 5.13 give the c- and f-structure, and Figures 5.14 and 5.15 the Glue proof (separated for readability), for the

NP *the computer the processor of which Julian said Miles repaired*, assuming that *which* is semantically vacuous.¹⁹

In English, the relative pronoun is sometimes optional for object relative clauses, namely when there is no pied piping:

- (76) a. the computer (which) Miles repaired
 b. the computer the processor of *(which) Miles repaired

We therefore also need an alternative tree which does not have a slot for the relative pronoun, but encodes the information it carries elsewhere. Figure 5.16 provides the appropriate lexical entry. Here, there is no NP node for any complex *wh*-phrases to attach, and so we ensure that this tree cannot be used in cases of pied piping; instead, %RELPRO and %GAP are identified and the pronominal PRED value of the relative pronoun is supplied by the construction. We can encode these two types of tree disjunctively in our basic description for the object relative clause tree: either there is an NP slot for the gapped argument pronoun, or the top S node carries the additional annotations in Figure 5.16.

Subject relative clauses are much the same, except that the gapped argument is identified with the verb's SUBJ. Figure 5.17 shows the relevant lexical entry for a relative clause headed by the subject of *repaired*.

5.5.5 Lexical rules

Of course, many other transitive verbs will also be able to appear in exactly the same configurations we have just exemplified with *repair*. The trees for

¹⁹We should perhaps encode the information that *which* refers to non-humans and *who(m)* refers to humans in the meanings of the relative pronouns (this is what Dalrymple 2001: 418 does for *who*, for instance). This meaning contribution does nothing to affect composition, however, and simply adds an extra conjunct to the meaning, so I omit it here for simplicity's sake. The meaning of the possessive *whose*, however, is compositionally crucial, since it encodes both possession and definiteness (Partee & Borschev 1998), and so it must be included. The meaning constructor for *whose* is as follows (adapted from Dalrymple 2001: 421), where **poss** represents whatever relation holds between a possessor and a possessum:

- (i) $\lambda P \lambda Q \lambda x . \lambda y [\mathbf{poss}(x, y) \wedge P(y) \wedge Q(y)] :$
 $(e_{(\text{SPEC}\uparrow)} \multimap t_{(\text{SPEC}\uparrow)}) \multimap (e_{(\text{SPEC}\uparrow)} \multimap t_{(\text{TOPIC SPEC}\uparrow)}) \multimap e_{\uparrow} \multimap t_{(\text{TOPIC SPEC}\uparrow)}$

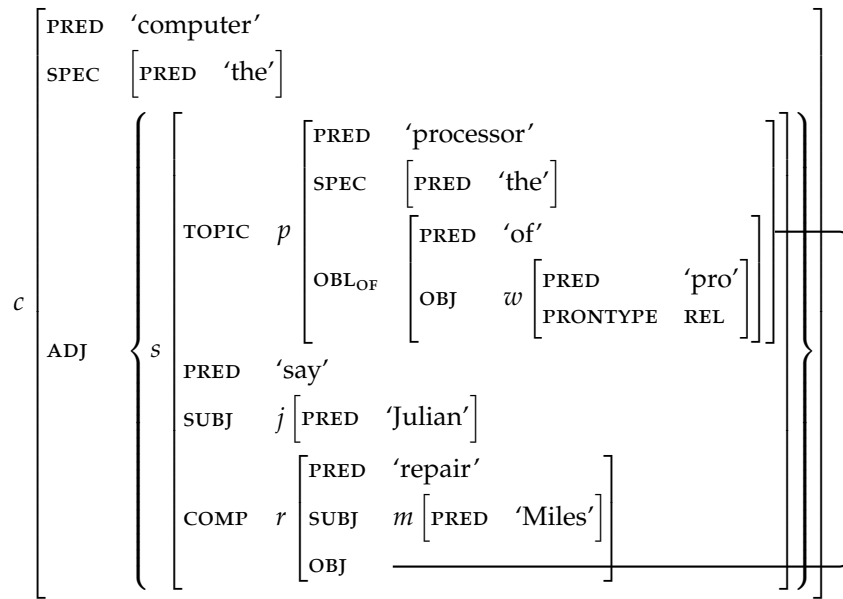


Figure 5.13: F-structure for *the computer the processor of which Julian said Miles repaired*

relative clauses headed by the objects of *hit*, *ate*, *met*, or any other transitive verb that permits relativisation on its object will be identical to that in Figure 5.11 except for the relevant lexical information, and likewise for many other kinds of extraction structures like *wh*-questions, clefts, topicalisations, etc. If we are forced to repeat this every time for every verb, we are surely missing a generalisation. In TAG, this generalisation is usually captured by the notion of tree family: we define a set of trees which a particular class of verbs all have access to—perhaps (nominal object) transitive verbs. In Lexicalised LFG, we will use lexical rules instead.

Lexical rules have been used in LFG since the very start, especially to capture diathesis alternations. For instance, (77), based on Bresnan (1982b: 9), is a passive rule which modifies the semantic form of a verb to change its argument structure:

$$(77) \quad 'F\langle \text{SUBJ}, \text{OBJ} \rangle' \rightarrow 'F\langle \emptyset, \text{SUBJ} \rangle'$$

This can be seen as a (non-monotonic) rewrite rule which modifies lexical entries, or, more perspicuously, and in keeping with the constraint-based philosophy of LFG, as a generalisation over the lexicon which states that if there is any lexical entry of the form $'F\langle \text{SUBJ}, \text{OBJ} \rangle'$, then there is also a lexical entry of the form

$$\begin{array}{c}
\boxed{\mathbf{I}} \quad : \\
\frac{\boxed{\text{RRC-Combiner}} \quad \lambda P \lambda Q \lambda x.Q(x) \wedge P(x) : \quad \lambda x.\exists e[\text{say}(e) \wedge \text{ag}(e) = \text{julian} \wedge \text{said-thing}(e) = \exists e'[\text{repair}(e') \wedge \text{ag}(e') = \text{miles} \wedge \text{pt}(e') = \text{iy}[\text{processor}(y) \wedge \text{part-of}(y,x)]]]] :}{(e_w \rightarrow t_s) \rightarrow (e_c \rightarrow t_c) \rightarrow e_c \rightarrow t_c \quad e_w \rightarrow t_s} \\
\frac{\lambda Q \lambda x.Q(x) \wedge \exists e[\text{say}(e) \wedge \text{ag}(e) = \text{julian} \wedge \text{said-thing}(e) = \exists e'[\text{repair}(e') \wedge \text{ag}(e') = \text{miles} \wedge \text{pt}(e') = \text{iy}[\text{processor}(y) \wedge \text{part-of}(y,x)]]]] :}{(e_c \rightarrow t_c) \rightarrow e_c \rightarrow t_c} \\
\boxed{\text{computer}} \quad \lambda x.\text{computer}(x) : \quad \lambda x.\text{computer}(x) \wedge \exists e[\text{say}(e) \wedge \text{ag}(e) = \text{julian} \wedge \text{said-thing}(e) = \exists e'[\text{repair}(e') \wedge \text{ag}(e') = \text{miles} \wedge \text{pt}(e') = \text{iy}[\text{processor}(y) \wedge \text{part-of}(y,x)]]]] :}{e_c \rightarrow t_c} \\
\frac{\lambda x.\text{computer}(x) \wedge \exists e[\text{say}(e) \wedge \text{ag}(e) = \text{julian} \wedge \text{said-thing}(e) = \exists e'[\text{repair}(e') \wedge \text{ag}(e') = \text{miles} \wedge \text{pt}(e') = \text{iy}[\text{processor}(y) \wedge \text{part-of}(y,x)]]]] :}{e_c \rightarrow t_c} \\
\boxed{\text{the}} \quad \lambda P.\text{ix}[P(x)] : \quad \text{ix}[\text{computer}(x) \wedge \exists e[\text{say}(e) \wedge \text{ag}(e) = \text{julian} \wedge \text{said-thing}(e) = \exists e'[\text{repair}(e') \wedge \text{ag}(e') = \text{miles} \wedge \text{pt}(e') = \text{iy}[\text{processor}(y) \wedge \text{part-of}(y,x)]]]]] : e_c}{(e_c \rightarrow t_c) \rightarrow e_c}
\end{array}$$

Figure 5.15: Final Glue proof for the computer the processor of which Julian said Miles repaired

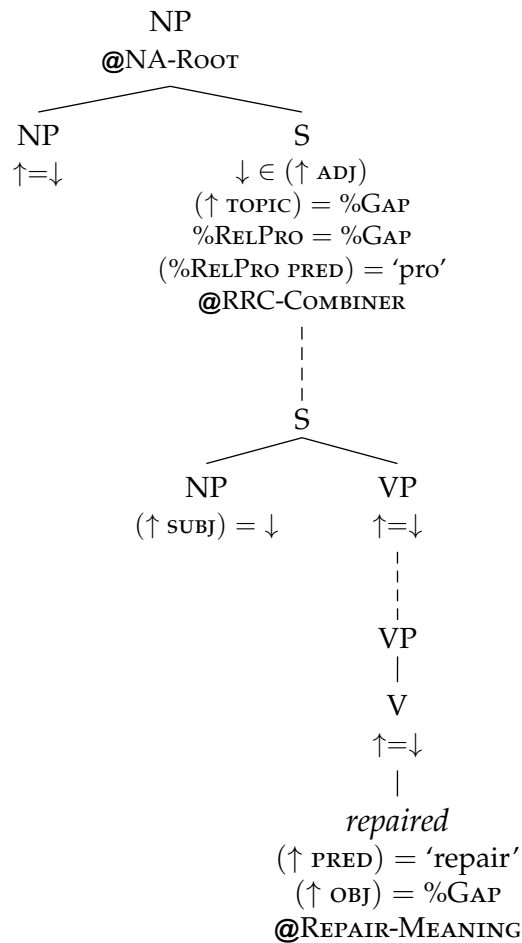


Figure 5.16: Lexical entry for a relative clause headed by the object of *repaired*, without a relative pronoun

' $F\langle \emptyset, \text{SUBJ} \rangle$ '. In this way, lexical rules capture much of the same information as tree families. We will understand lexical rules in this sense, as describing generalisations over the lexicon.

In standard LFG, there is no need for lexical rules to describe extraction structures like relative clauses, since these high-level structures are handled by the context-free grammar component; in other words, the definition of what is *lexical* is more restricted. In Lexicalised LFG, though, such constructions *are* lexical, in that each of them corresponds to an elementary tree for some lexical item. I therefore propose to extend the notion of lexical rule to cover these cases as well.

A first-pass approximation for the definition of a lexical rule is given in (78):

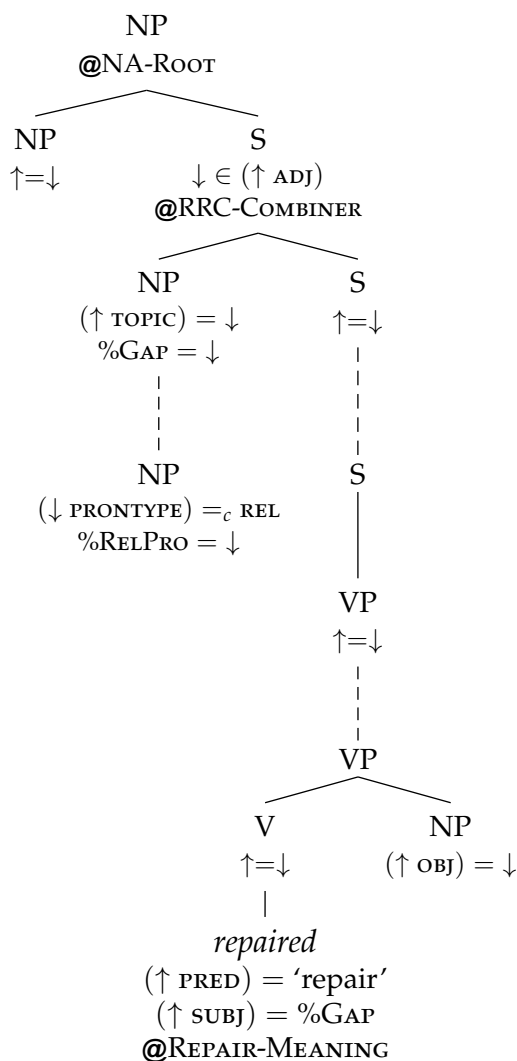


Figure 5.17: Lexical entry for a relative clause headed by the subject of *repaired*

(78) **Lexical Rule (first pass)**

A lexical rule $A \Rightarrow B$ is to be read as follows: the existence of a lexical entry L , of which the description A is a subset, implies the existence of a lexical entry L' , which is identical to L except that A has been replaced by B .

Seen from the theoretician's point of view, what is useful about lexical rules of this kind is that they capture generalisations which would otherwise go undescribed. For example, it might be that all transitive verbs have corresponding object relative clauses, but without a rule to describe this generalisation, a lexicalised grammar will simply have to list each of them in turn. Similarly, from

the grammar writer's perspective, lexical rules represent a saving in manual encoding: rather than having to spell out L' each time, we merely have to give L and the lexical rule. And since there will be multiple lexical entries L which match the left-hand side of the rule, a single rule will provide us with multiple L' s. Finally, from a psychological perspective, such rules might be thought to model the productivity of human language: if we know that when certain properties are present in a particular lexical entry, e.g. a VP containing an NP object, there is also another lexical entry with related but different properties, e.g. the passive voice version of the verb, then this would explain why we do not have to learn the active and passive forms of new verbs separately.

However, as presented in (78), the definition is both too strict and too permissive. On the one hand, it is too much to require that L has A as a subset, since we are not only interested in cases where the descriptions are identical, but also where the contents of one implies the other: for instance, if our lexical rule requires that $\mathcal{D}(a, b)$ hold, it should match lexical entries which include $\mathcal{M}(b) = c \wedge \mathcal{M}(c) = a$, since the latter entails the former. In other words, lexical rules should operate on the content of descriptions rather than merely their form, and so we must talk about $\text{Pr}(A)$ and not merely A . On the other hand, we do not want rules to apply too freely, since lexical exceptions do exist. Verbs like *weigh* do not passivise, despite taking NP objects (and so presumably matching the left-hand side of the relevant rule):

- (79) a. The baby weighed nine pounds.
 b. *Nine pounds were weighed (by the baby).

We therefore need some way of blocking the rule's application on a case-by-base basis. (80) provides a fuller definition which accounts for both these requirements:

(80) **Lexical Rule (final)**

A lexical rule $A \Rightarrow B$ is to be read as follows: the existence of a lexical entry L , of which there is some subset A' , the smallest subset of L such

that $\text{Pr}(A') \rightarrow \text{Pr}(A)$, implies the existence of a lexical entry L' , which is identical to L except that A' has been replaced by B , provided that $\text{Pr}((L - A') \cup B)$ does not express a contradiction.

Firstly, we specify that there must be *some* subset of L whose meaning entails $\text{Pr}(A)$, thus accounting for the content-matching rather than form-matching. Secondly, the final clause provides for the possibility of there being times when the rule cannot apply, due to some feature clash, etc. This allows us to selectively exclude rules from applying to certain lexical items (e.g. unpassivisable verbs) by specifying in their basic lexical entry something which contradicts the description on the right-hand side of the relevant rule (we will see an example of this below).

Once again, we will continue to use trees as shorthand for the (minimal) descriptions they satisfy, in order to aid readability. This means that lexical rules will be written as mappings between trees. Note that what appears on either side of a lexical rule need not be (and usually will not be) a full lexical entry, but rather descriptions which form *part* of these entries; this means the trees on both sides of the rule will often be partial, and so will usually not satisfy the requirements laid out earlier on well-formed final trees.

Lexical rules have to have something to operate on, which means that some structures must be seen as more 'basic' than others. That is, since the structures on the right-hand sides of rules are in a sense derived from those on the left-hand side, the latter are more basic than the former. For verbs, I have chosen the active voice, declarative form as the base case. This is somewhat arbitrary, but does have some justification: it is the one construction which the vast majority of verbs can appear in. Several verbs cannot passivise (e.g. measure verbs like *weigh* or *cost*), and some are restricted in what other kinds of constructions they can appear in. It therefore makes sense to choose the most generally applicable class as the basic one.

Let us now look at some examples of lexical rules. Figure 5.18 gives the object relative clause lexical rule. I use the \bullet symbol to denote the terminal node which hosts the verb. As we can see, this rule states that whenever there is a

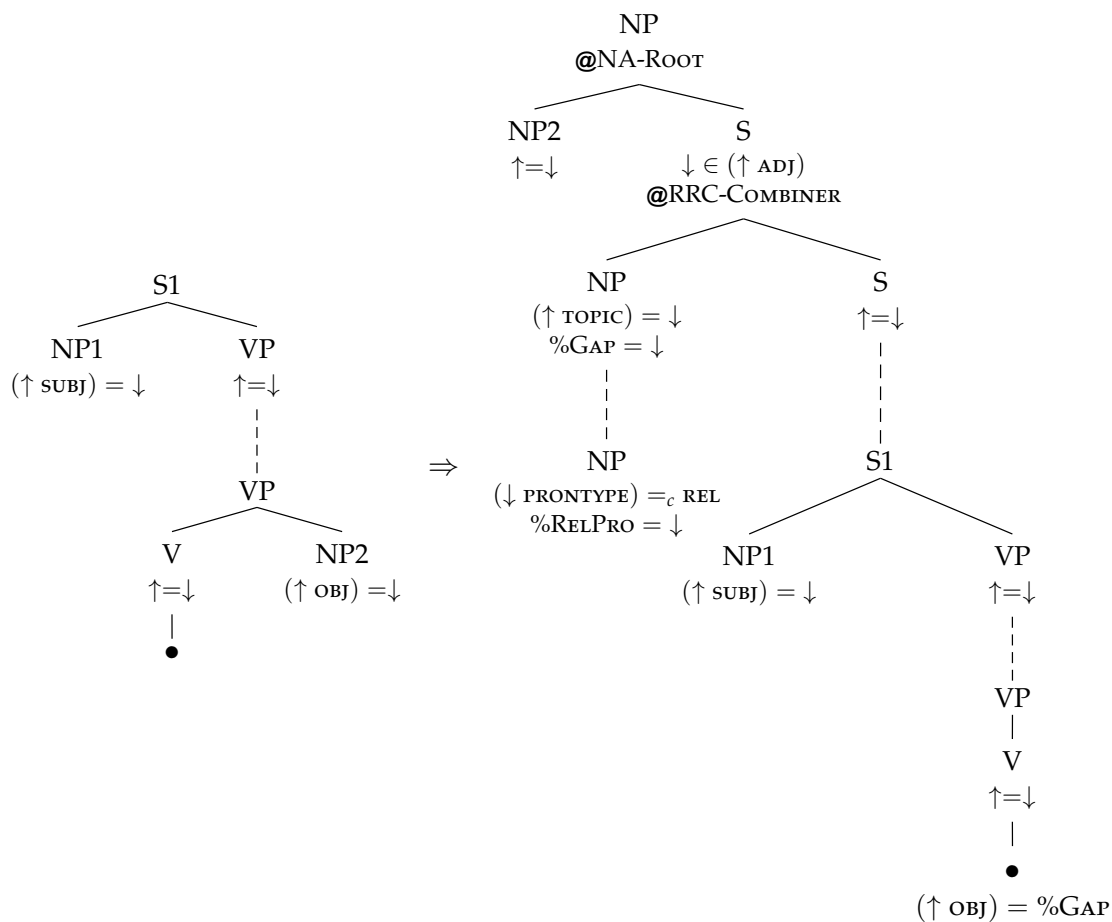


Figure 5.18: Lexical rule for object relative clause (with relative pronoun)

basic transitive verb elementary tree in the lexicon, there will also be a tree for a relative clause headed by that verb's object. I also use the convention of suffixing node labels with numbers where it is important to keep track of which nodes correspond to each other across lexical rules. The rule actually makes reference to the nodes directly, not their labels, since the description language allows us to talk about nodes independently of their labels, but it is convenient to refer to node labels as if they were the nodes. Of course, the numbers are not part of the grammar, and therefore not part of the rule: NP1 will match any node labelled NP, for example. In this rule, the object NP on the left corresponds to the head noun's NP on the right, and the subject NP remains in the same place. The root node on the left-hand side is identified with the lower S node on the right-hand side, so that anything which is stipulated to be true of the clause that

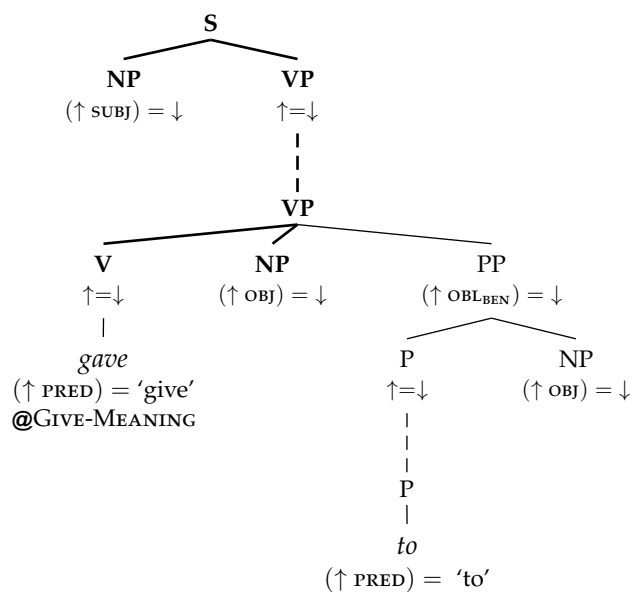


Figure 5.19: Lexical entry for *gave*

contains the verb in its lexical entry is also required to be true of the (lowest sub-clause in the) relative clause. In fact, every node on the left-hand side which also appears on the right-hand side is taken to be the ‘same’ node in the relevant sense, so that anything asserted of it in the basic lexical entry outside of the scope of the rule continues to be true on the right-hand side.

This rule only describes a relative clause with a relative pronoun, and, of course, this is optional in English. The proper rule would have a disjunctive description on the right-hand side. Equivalently, we could simply have two different rules.

Notice that because the left-hand side of a rule must merely be contained in a lexical entry for a rule to apply, the rule in 5.18 correctly applies to ditransitive verbs like *give* as well, since their descriptions contain the description of a simple transitive verb as a sub-part. Figure 5.19 shows the lexical entry for *gave*, with the sub-part which matches the left-hand side of Figure 5.18 highlighted. Recall that a lexical rule says that everything else in the lexical entry it applies to besides whatever matches the left-hand side of the rule remains the same. This means that whatever descriptions were stated to hold of VP in the basic entry,

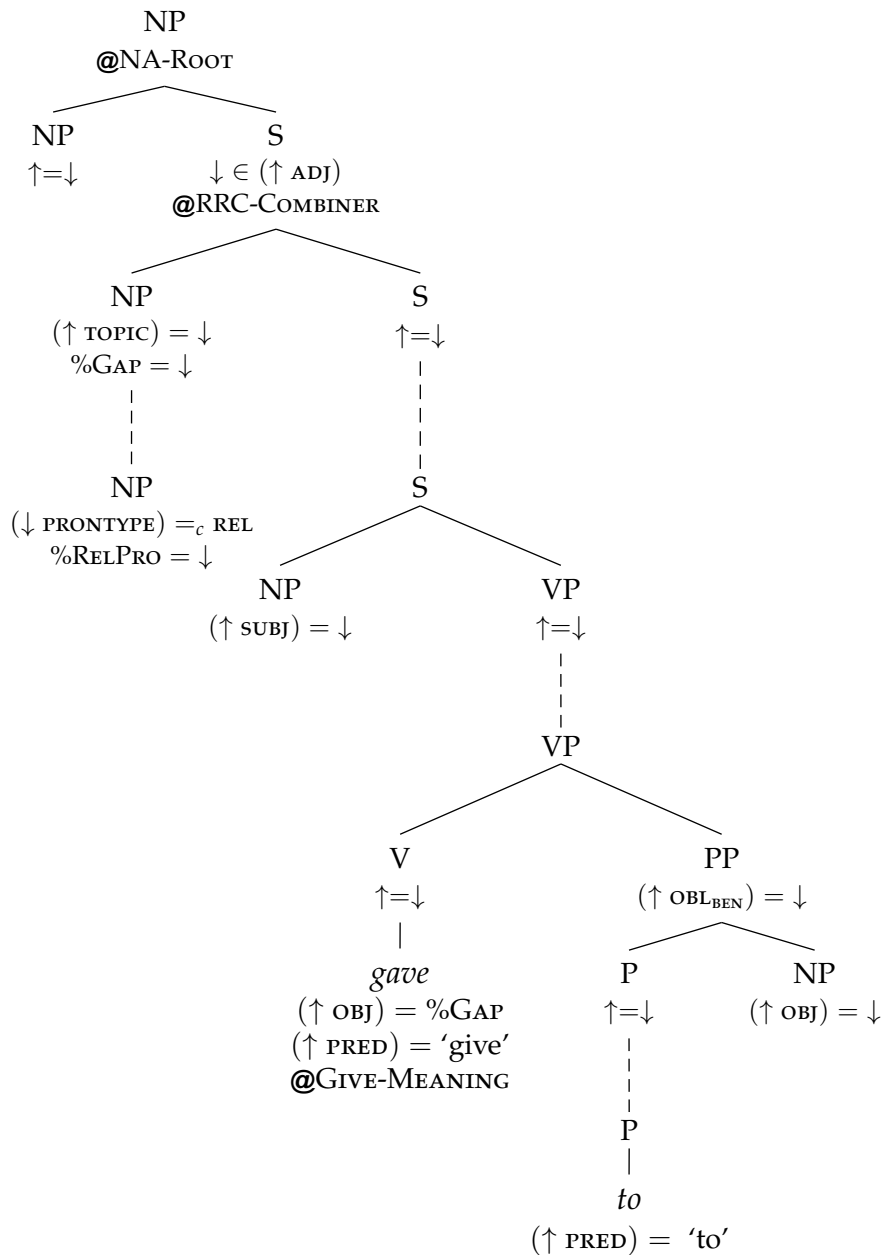


Figure 5.20: Lexical entry for a relative clause headed by the object of *gave*

for example that it dominates a PP headed by *to*, remain in the new lexical entry for the relative clause. Figure 5.20 shows the result for *gave*.

We use this partial matching to our advantage in the subject relative clause rule as well. The rule is given in Figure 5.21. This rule will not only match intransitive verbs, which fit the description exactly, but in fact every verb which has a subject, since they will contain the description as a sub-part. Any other arguments under VP remain in the relative clause, as required:

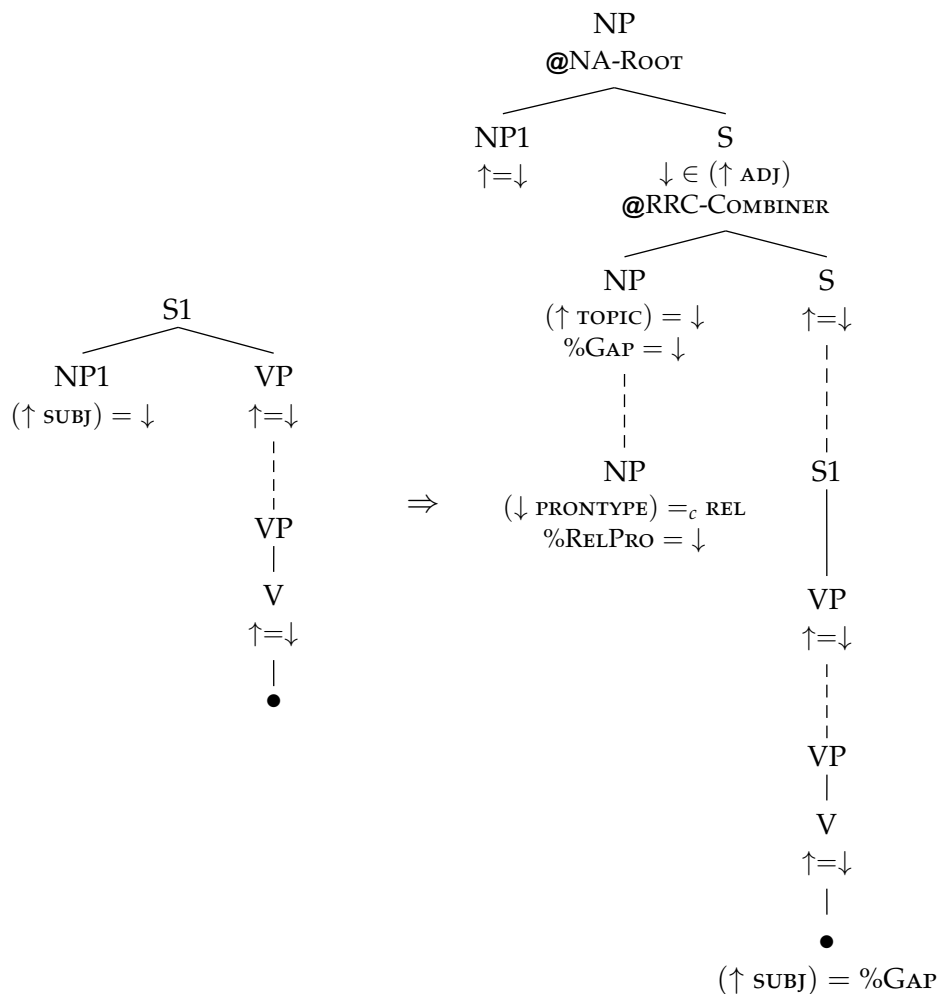


Figure 5.21: Lexical rule for a subject relative clause

- (81) a. the ensign who [VP [V came]]
 b. the ensign who [VP [V wore] [NP a red shirt]]
 c. the ensign who [VP [V gave] [NP his life] [PP to the cause]]

Now that we have the two rules in Figures 5.18 and 5.21, we only need to stipulate the basic elementary tree for each verb in order to also add elementary trees for subject and, where appropriate, object relative clauses to the lexicon. Similar rules can be constructed for *wh*-questions, topicalisation, clefting, etc. I do not present a full grammar here, and so will not discuss these further, but it is clear how such rules could be quite easily formulated along the same lines as those for relative clauses.

One rule that I will discuss is the passive rule, since this will be relevant in our

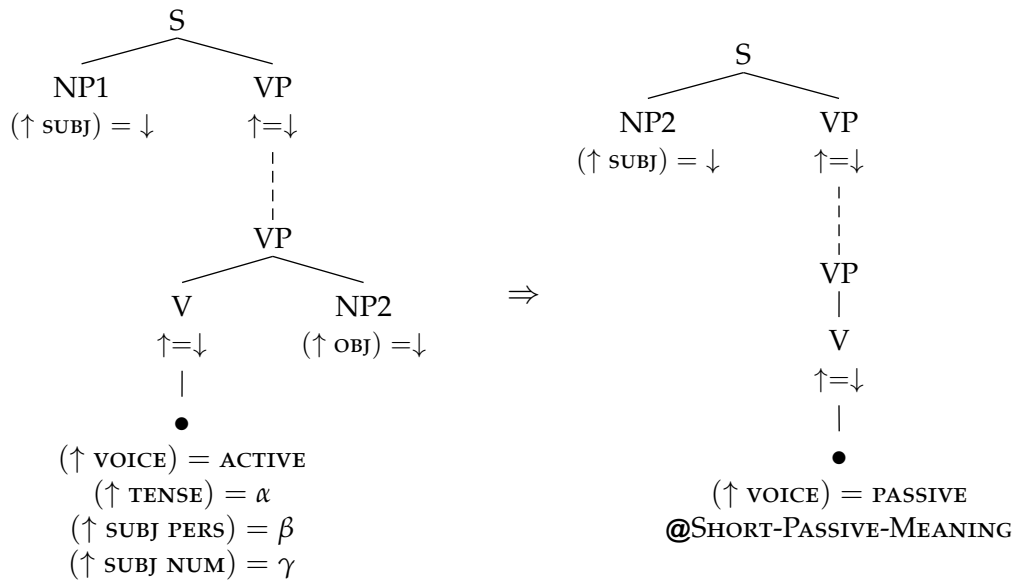


Figure 5.22: Lexical rule for the short passive

discussion of MWEs. The rule for the short passive is given in Figure 5.22. This rule says that for every elementary tree for an active voice transitive verb, there is a corresponding passive tree. What is more, that passive tree neutralises whatever person and number restrictions the verb might place on its subject, and also does not have a tense feature (α , β , and γ are variables which range over whatever the values of TENSE, PERS, and NUM are taken to be). This represents the fact that we do not want separate elementary trees for third-person singular passive forms in English, even though there will be a separate entry for the corresponding active form, and that passive participles are tenseless: passive sentences in English take their tense from whatever auxiliary adjoins to the verb.²⁰

In terms of semantics, the passive rule adds the passive meaning constructor discussed in Chapter 4.4.2.2, abbreviated here as `SHORT-PASSIVE-MEANING`, and shown in (82):

$$(82) \quad \lambda P \lambda y. \exists x [P(x)(y)] : (e_{(\uparrow \text{SUBJ})} \multimap e_{(\uparrow \text{OBJ})} \multimap t_{\uparrow}) \multimap e_{(\uparrow \text{SUBJ})} \multimap t_{\uparrow}$$

²⁰This lack of tense information is also what forces an auxiliary to adjoin in the passive. At the moment, however, there is nothing ill-formed about a sentence without a value for TENSE. I assume that in a fully developed theory, which incorporates tense information into its meaning language, this would result in semantic issues, but I do not consider the details here.

There may be other semantic constraints on the passive, e.g. that its subject must be “a true patient”, i.e. be “genuinely affected by the action of the verb” (Bolinger 1975: 67), and these can be added to the right-hand side of the rule as required. In a full account we would also include whatever morphological process is taken to realise the participial form of the verb which instantiates the •.

Once again, this lexical rule will apply very widely: to every verb that has both a subject and an object NP, in fact. As with relative clauses, this correctly includes verbs that have additional complements:

- (83) a. The bandage was given to the ensign.
b. The ensign was given the bandage.

As it stands, the rule will not allow the secondary object in a ditransitive to be promoted to subject, since this bears the GF OBJ_{THEME}, not OBJ. For many dialects of English, this is the correct result, and (84) is judged ungrammatical; this is not the case for all, however:

- (84) %The bandage was given the ensign.

For those languages, and those dialects of English, which do permit this kind of alternation, an additional rule or rules will be required. Similarly, we require an additional rule to permit the prepositional passive in the appropriate cases:²¹

- (85) a. The ensign was spoken to.
b. *The ensign was given the bandage to.

As we have already noted, certain transitive verbs, like *weigh* or *cost*, cannot passivise. In this case, we must lexically mark these verbs as being unable to undergo passivisation. Figure 5.23 shows the required lexical entry for *weighed*. The important equation is (\uparrow VOICE) \neq PASSIVE. This is not included in the left-hand side of the passive rule, and so will carry over to the right-hand side. But

²¹Correctly constraining the prepositional passive rule is a complex problem—see Tseng (2007), Findlay (2016b), and references therein for some overview.

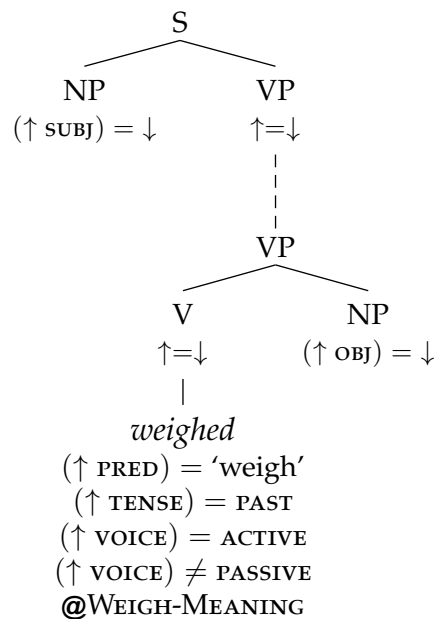


Figure 5.23: Lexical entry for *weighed*

then we would have incompatible constraints in the derived lexical entry: the base entry gives us (86), but the passive rule gives us (87):

(86) (↑ VOICE) ≠ PASSIVE

(87) (↑ VOICE) = PASSIVE

These cannot both be true, and so this case falls under the final clause of the lexical rule definition given in (80), where the rule is blocked if application of it would result in a contradiction. Passivisation is thus correctly ruled out for *weighed*. Note that it would not have been sufficient just to include (↑ VOICE) = ACTIVE in the lexical entry, since this is precisely what the passive lexical rule checks for and replaces.

* * *

This concludes our overview of the machinery of Lexicalised LFG. The proposals shown here are somewhat schematic, but serve to illustrate the viability of the approach, and lay the groundwork for the more detailed analysis of MWEs in the next chapter, which is the focus of the thesis.

5.6 Chapter summary

We started this chapter by recognising that incorporating a TAG into the LFG architecture required recasting the TAG formalism in description-based terms. This means seeing it as involving constraint satisfaction rather than tree manipulation. We provided a language for describing trees, and showed how this could be used for a TSG. In order to apply it to a fully-fledged TAG, we needed to expand our description language to be able to describe simple dominance instead of only immediate dominance. The former is a looser relation between nodes, where a node dominates itself and all of its descendants, but only immediately dominates its daughters. This freedom is what allows us to simulate adjunction in description-based terms.

At the same time, we introduced the notion of lexicalisation, and I argued that lexicalised grammars are desirable from the point of view of lexicalist theories like LFG. We showed that the path to lexicalising a CFG (which LFG c-structures are based on) leads naturally to a TAG, and so found another reason to want to incorporate a TAG into the LFG framework, in addition to its utility in describing MWEs. By lexicalising LFG, we also arrive at a neater theory, whereby lexical entries are unitary objects (descriptions), rather than hybrid phrase-structure rules which also have descriptions attached.

In the final section, I illustrated what a Lexicalised LFG treatment of various components of the grammar looks like. We also discussed the role of lexical rules in capturing generalisations, reducing manual labour in grammar building, and describing the productivity of human language. Now that we have these theoretical tools in hand, we turn to the task of showing how they can be used in the analysis of MWEs. This is the focus of the next chapter.

6

MWEs in Lexicalised LFG

Contents

6.1	Non-verbal MWEs	262
6.1.1	Syntax	263
6.1.2	Semantics	272
6.1.3	Non-verbal MWEs: a summary	295
6.2	Verbal MWEs	295
6.2.1	Decomposable expressions	296
6.2.2	Non-decomposable expressions	308
6.2.3	Verbal MWEs: a summary	319
6.3	Lexical flexibility	321
6.3.1	<i>Grist to the mill</i>	322
6.3.2	Other kinds of lexical flexibility	327
6.3.3	Summary	334
6.4	Chapter summary	335

This chapter discusses how MWEs can be analysed using the tools of Lexicalised LFG introduced in the previous chapter. I take a number of specific examples of MWEs as case studies, and explore various aspects of their syntax and semantics in detail, showing how Lexicalised LFG can account for them. This serves both to broaden our understanding of the empirical phenomenon and to show the efficacy of the proposed theoretical framework. As discussed, my focus will be on idioms (along with certain fixed expressions), since these illustrate a large number of the challenging properties which an account of

MWEs must make sense of. This is not to say that other kinds of MWE, such as prepositional verbs, do not pose their own unique challenges, but considerations of space mean that I cannot provide an explicit treatment of the whole gamut of MWE types here. My hope is that the discussion presented in this chapter will nonetheless suggest the direction such analyses might take.

When presenting analyses, I will continue to use trees as stand-ins for the descriptions which really make up the lexical entries in Lexicalised LFG. For the most part, this is harmless, since each tree has a unique minimal description, and so the task of translating from description to tree is trivial. The fact that trees are easier to understand at a glance than long descriptions simply makes it more reader-friendly to use the graphical representation whenever possible. It should continue to be borne in mind though that the trees are *not* objects in the formal theory, unlike standard TAG; instead, the only objects in the grammar are descriptions. It will not be possible to maintain this simplification at all times—for instance, where a description is underspecified, leaving several different minimal trees which satisfy it, it may be easier to talk about the description rather than list all of the trees which it describes.

We divide MWEs into two categories: non-verbal and verbal. Section 6.1 presents the analysis of the non-verbal MWEs, while Section 6.2 handles the verbal ones. After this, Section 6.3 addresses the question of lexical flexibility, and Section 6.4 offers summary and conclusion.

6.1 Non-verbal MWEs

Although the focus of much work on MWEs, and idioms in particular, is on the verbal domain, there exist a very large number of non-verbal MWEs, including idiomatic NPs, AdjPs, PPs, and AdvPs.¹ Unlike their verbal counterparts, these expressions do not exhibit very much syntactic flexibility, and so it will usually suffice for our purposes merely to describe their syntactic structure and account

¹There are also MWEs which are not constituents, such as *out of the frying pan into the fire*, or *the cat... out of the bag* (cf. *let the cat out of the bag* vs. *the cat is out of the bag*). We will have less to say about these, but will consider one such case, *grist to the mill*, in Section 6.3.1.

for how they combine semantically with other parts of the sentence. Thus, to give an account of a non-verbal MWE, we must at least decide *what* and *how much* syntactic structure it possesses, and also give a compositionally appropriate meaning. Another question is whether or not the expression permits modification of any of its parts: this will determine where we permit adjunction to take place (where we assign quasi-nodes). In this section, we first address the syntax of non-verbal MWEs, and then their semantics.

6.1.1 Syntax

I included the question ‘how much syntactic structure?’ above because many authors claim that certain non-verbal MWEs are truly frozen expressions, and thus amenable to analysis via a version of the monolexemic approach, where they are treated as single words. Kay et al. (2015: 4), for example, claim that Sag et al.’s (2002) category of ‘fixed expressions’, which includes non-verbal MWEs such as *by and large*, *all of a sudden*, *right away*, and *first off*, “can, without loss of generality, be listed as single words in the lexicon, despite their spelling suggesting a multiword history”. Indeed, all things being equal, we should not treat these expressions as phrasal if there is no motivation in the synchronic grammar for doing so. But some such expressions are not as inflexible as these authors give them credit for. We have already seen examples with *by and large* in Chapter 3.1, repeated below in (1); and other ‘fixed’ expressions are even more permissive, as shown for *all of a sudden* in (2):

- (1) a. A: By and large, the economy seems to be doing well.
 B: By but not so large: have you seen the latest unemployment figures? (Glucksberg 1993: 7)
- b. But by and indeed large, the Space-Sim has long since perished. (iWeb)
- (2) a. Weirdest thing yesterday, I got into my car, and put the key in the ignition and the car starts beeping like crazy [...] Anyone have this

experience or know what caused that all of a random sudden?

(iWeb)

- b. 'Twas christmas night in the workhouse, and all of a terrible sudden
there was a dreadful accident. (WWW)

Now, the example in (1a) is perhaps an extended use (somewhat playfully taking the expression to have some literal, decomposable meaning), but (1b) is a straightforward insertion of an adverbial, and the examples in (2) straightforward insertions of adjectives. This shows that speakers are aware of the phrasal status of these expressions, even if they remain relatively infrequent in their modified forms. Frequency of modification, though, is not necessarily a question for the grammar proper: if it is possible, the grammar should permit it; how frequently it happens may be down to all manner of other factors. It seems prudent, therefore, to provide such MWEs with at least the minimal syntactic structure they appear to realise. I therefore propose the syntactic analyses of *by and large* and *all of a sudden* shown in Figures 6.1 and 6.2.² I represent the meaning schematically using templates at this stage; we will unpack the contents of these templates in the following section.

There are a number of things to be said about these representations. Firstly, they are auxiliary trees, since they are modifiers. Secondly, I have given the root and foot nodes of these auxiliary trees the label VerbalP; as discussed in Chapter 5.5.3, this is a metacategory which includes S and VP, accounting for the fact that such adverbials can modify both sentential and verbal elements:

- (3) Adjunction to S:
- a. By and large, we succeeded in our goals.
 - b. All of a sudden, we reached the end.

²The word *sudden* appears, atypically, as a noun in Figure 6.2. This is motivated both by the fact that it hosts a determiner and that adjectives can adjoin to it (e.g. *all of an unforeseen sudden*).

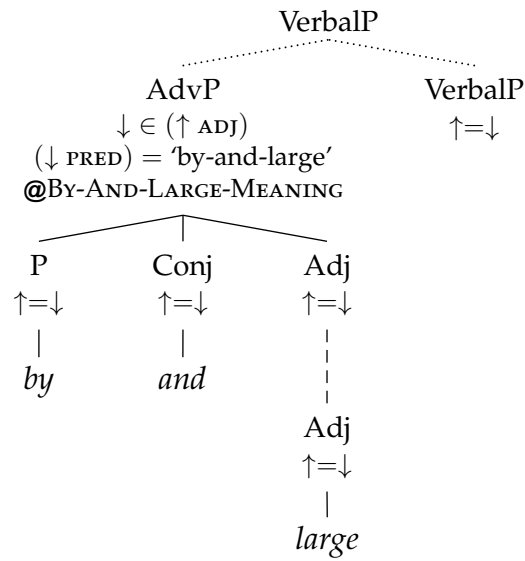


Figure 6.1: Lexical entry for *by and large*

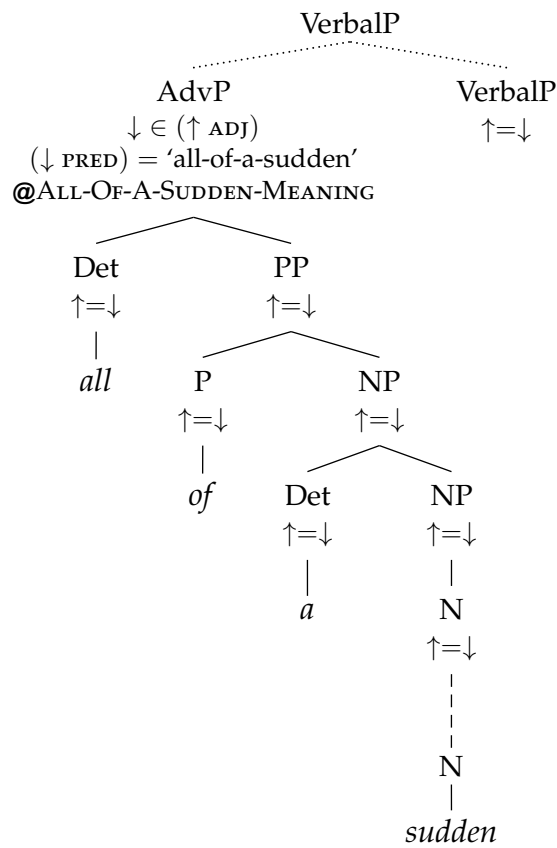


Figure 6.2: Lexical entry for *all of a sudden*

- (4) Adjunction to VP:
- a. We have by and large succeeded in our goals.
 - b. We all of a sudden reached the end.

We might also want to include other categories in this group, such as NP, depending on our analysis of sentences like (5):

- (5) They sold {basically|by and large} our whole stock.

Here the adverbial appears inside the VP, so cannot be adjoined to it.³ At any rate, we can describe as many different elementary trees for each expression as needed, by including a disjunction over as many categories as needed in the description of the root and foot nodes for the tree. As discussed in the previous chapter, we ensure the root and foot are identical by equating them in the description:

- (6) $\lambda(\text{root}) = \lambda(\text{foot}) = \{S|VP|NP|\dots\}$

Given the possibility of post-sentential uses of these expressions, as in (7) (insofar as they are non-parenthetical, which seems more clearly the case for (7b) than for (7a), for instance), we also underspecify the ordering of the foot node with respect to its sister AdvP node, so that the expression can appear to the right as well as to the left of the modified phrase.

- (7) a. We succeeded by and large.
b. Julian arrived all of a sudden.

Thirdly, for the category labelling the MWE itself, I have chosen AdvP, following the convention in XTAG (Bleam et al. 2001) to use AdvP for any multiword adverbial category, even when the expression itself contains no adverbs (examples discussed in the XTAG grammar include *a little* or *sort of*).

Fourthly, the AdvP node also hosts all of the interesting functional annotations and meaning constructors for the expression, the bottom half of the

³There is a variant of (5) where the adverbial is flanked by comma intonation. I do not discuss this potential parenthetical use here. Regardless of the status of parentheticals, examples such as (5), which lack such intonation, still demand an analysis.

tree being devoid of annotations except for $\uparrow=\downarrow$. Since these expressions are not decomposable, and since any modification of the subparts seems to only scope over the whole expression, I propose a simple, atomic f-structure for this kind of MWE, as in (8):

- (8) a. $\left[\text{PRED } \text{'by-and-large'} \right]$
 b. $\left[\text{PRED } \text{'all-of-a-sudden'} \right]$

This means that the expression's phrasal nature is a purely c-structural phenomenon; the phrase as a whole makes an f-structure contribution, but each of its parts does not. This is perhaps the reflex of other authors' judgements that expressions like this can be treated atomically, as 'words with spaces'.

Notice that, because the annotations appear on the maximal projection AdvP and not on its daughters, they use the down arrow \downarrow , not the up arrow \uparrow , since the f-structure they refer to is their own, not their mother's. (But also recall that the \uparrow/\downarrow notation is in fact otiose in Lexicalised LFG, owing to the extended domain of locality.)

Finally, there is one pair of quasi-nodes in each expression, allowing for adjunction at this point. For the modifications of *all of a sudden* seen in (2), it is clear that the noun is the target of adjunction, and so this is where the quasi-node is located. For the insertion of *indeed* in *by and large*, however, the choice of target is less obvious. I have chosen to locate the adjunction site on the Adj node, even though *indeed* is not really modifying the adjective, on the basis that *indeed* seems to be able to attach to the left of various categories, including adjectives (along with e.g. nouns):⁴

- (9) a. It is **indeed true** that diversity promotes creativity. (iWeb)
 b. The mere fact that litigation was in progress did not alter the right and **indeed responsibility** of the mass media to impart information of public interest. (BNC)

⁴See Ernst (1983) and Collins (1988) for evidence that these adverbs are not parenthetical and really do form part of the phrasal projection of the head word.

Notice, however, that, in both MWE examples, because the Adj/N node corresponds to the f-structure of the whole MWE, any modification will, in the semantics, scope over the whole expression, not just the target of adjunction—as indeed it should do, since the parts do not bear any meaning themselves. We will see how this works in detail in the following section.

There do also exist non-verbal MWEs which don't allow any phrase-internal modification, such as *kith and kin*. We do not find examples like **kith and business kin*, for instance, even though an intended meaning of 'close work colleagues (who are like family)' is not implausible.⁵ If these are really completely fixed, then they could be represented as single words. However, we may remark that the prosody of *kith and kin* argues for a phrasal analysis: the expression has final stress, as would be intonationally appropriate for a phrase, but much less so for a single word. Thus, by analogy to the other 'fixed' expressions, I think it is no great loss to represent the apparent syntactic structure of these unmodifiable non-verbal MWEs, while ensuring that they are not open to modification.⁶ These expressions therefore have a similar structure to those discussed above, except that they contain no internal pairs of quasi-nodes—Figure 6.3 gives the relevant lexical entry for *kith and kin*. Notice that the top-most nodes *are* potential adjunction sites here, however, since adjectives, which adjoin to N, can attach to the whole phrase, as can determiners, which adjoin to NP:

- (10) a. I should be at the table with [NP all of my [NP kith and kin]].
(COCA)

⁵One can in fact find cases of modified *kin* inside *kith and kin*:

- (i) a. Finally, the Tsimane' rely on gifts from kith and close kin to tide them over lean times. (journal article)
b. In a city of crooked alleys, // [...] // Suspicious kith and unkind kin [...]
(song lyrics)

This is likely a matter of dialectal variation: where *kith* meaning 'friends' survives independently (e.g. in some northern English and Scottish varieties), the expression as a whole can be seen as compositional, being, in Sag et al.'s (2002) terms, an institutionalised phrase rather than a lexicalised one (i.e. possessing statistical but not semantic idiomaticity).

⁶This also leaves open the possibility of extended or playful uses which make reference to the internal structure—it is easier to see how such uses might reanalyse a node as a pair of quasi-nodes than add phrasal structure where the lexical entry has none.

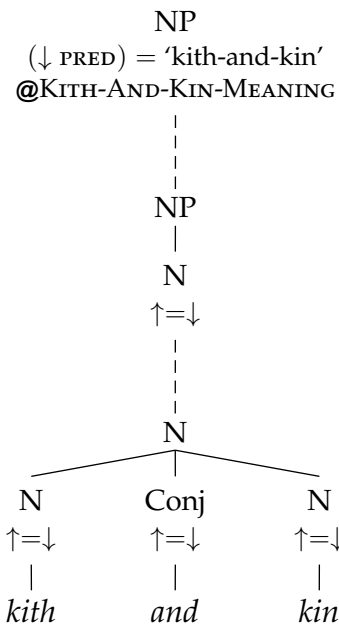


Figure 6.3: Lexical entry for *kith and kin*

- b. Of course, we did have [_{NP} some [_{NP} [_N real [_N kith and kin]]]] up in Scotland. (novel)
- c. Several people have asked me for a voucher they can give [_{NP} their [_{NP} [_N dear [_N donkey/adventure/Wales-loving [_N kith and kin]]]], so I've fashioned one out of the notebook doodles of five weeks of running the Kickstarter! (WWW)

Lastly, there are some cases of decomposable non-verbal MWEs, which therefore require a more articulated f-structure, to allow for the parts to be modified individually (see next section). One such example is *fly in the ointment*, meaning 'flaw (in something)'. The fact that it is decomposable can be seen from examples like (11):

- (11) a. There is one fairly massive fly in the ointment, though—the potential of a “no deal” Brexit. (NOW)
- b. The latest fly in the bilateral ointment is an Australian move to stop counting New Zealand citizens as domestic students in Australian universities and polytechs. (NOW)

- c. The fly in the Labour ointment is “Odd Ed”. Miliband’s personal ratings are poor, with just one in four voters thinking he does well as party leader. (NOW)
- d. I will just have to live with this empirical fly in my theoretical ointment until some additional cases are manifest.⁷

Both *fly* and *ointment* can be modified separately. It might be unclear in example (11a) whether *fairly massive* modifies *fly* alone or the whole phrase (and perhaps semantically this makes no difference: a massive flaw in something might well still count as a massive flaw *tout court*). However, (11b)–(11c) show that *ointment* can be modified independently of the whole expression. (11b) does not mean that the flaw was bilateral; quite the opposite—it refers to a decision made by Australia against the wishes of New Zealand. Rather, it is a flaw in the bilateral nature of the relationship between the two countries. Similarly, (11c) refers to a flaw in the Labour Party (namely, its leader), not a flaw which has Labour-like qualities. The fact that both parts can be modified independently is shown further in (11d), where both are modified simultaneously: we have an empirical flaw in a theoretical enterprise.⁸

Because this is real internal modification (the modification applies to subparts of the expression), we must assign a meaning to each of the parts, and in the syntax this means each part must correspond to its own f-structure. Figure 6.4 gives the appropriate lexical entry: notice that the equation $(\uparrow \text{OBJ}) = \downarrow$ will give an articulated rather than flat f-structure. The schematic f-structure which corresponds to the contribution of the MWE alone is shown in Figure 6.5.

There are two important things to note regarding this analysis. Firstly, I have not represented the determiner as part of the MWE. There are two reasons for this decision. Firstly, when it is present, it has its usual meaning (i.e. a fly in the

⁷I have been unable to relocate the origin of this attested example, which I noted down some months ago without recording the source. Even if it were not an attested example, though, it strikes me as grammatically impeccable, and so certainly in need of explanation.

⁸The precise meaning of idiomatic *ointment* in these examples is not altogether clear; I will return to this question when we discuss the semantics of these expressions in the next section.

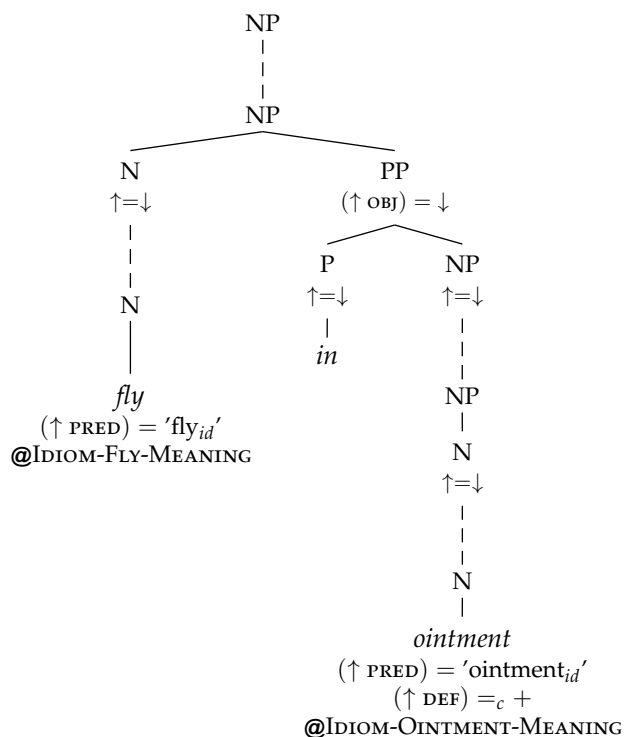


Figure 6.4: Lexical entry for *fly in the ointment*

$$\left[\begin{array}{l} \text{PRED} \quad \text{'fly}_{id}' \\ \text{OBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'ointment}_{id}' \end{array} \right] \end{array} \right]$$

Figure 6.5: Schematic f-structure for *fly in the ointment*

theoretical ointment is a flaw in *the* theory); and, secondly, it can be replaced by a possessive phrase instead, as we saw in (11d). Thus, the MWE places no constraint on what form the specifier of *ointment* takes, but does require that it make the noun definite (something which I take possessor phrases to also accomplish), as encoded in the constraining equation $(\uparrow \text{DEF}) =_c +$.

Secondly, although *in the ointment* looks like a locative modifier, I have represented it syntactically as introducing an OBJ instead. Semantically speaking, the preposition is like *with* in *a problem with* or, indeed, like *in* in *a flaw in*: it marks the argument of a noun. The word *flaw* denotes a two-place relation: if there is a flaw, there is a flaw *in something*. Since *ointment* in *fly in the ointment* corresponds to this something, it makes sense for the *in* to serve the same purpose here as it does in the literal expression. We will see how this is spelled

out in the semantics in the next section.

* * *

We now have three kinds of non-verbal MWE to consider:

1. expressions which permit no phrase-internal modifiers (*kith and kin*),
2. expressions which permit phrase-internal modifiers which apply to sub-parts of the MWE (*fly in the ointment*), and
3. expressions which permit phrase-internal modifiers, but where these modifiers scope over the whole expression (*all of a sudden, by and large*).

In the next section, we turn to the question of how these different modification possibilities are cashed out semantically.

6.1.2 Semantics

As discussed in Chapter 3, we are not treating MWE meanings as derived in any way from their literal ones. Instead, I treat certain strings as ambiguous between literal and idiomatic parses, leaving it to historical linguists and lexicographers to explore the origins of such links. For the synchronic grammar, it is sufficient to treat MWEs as contributing to the semantics of a sentence in exactly the same way as any other lexical item. A brief note on terminology is in order, however. I will represent idiomatic meanings with their own constants in the meaning language, e.g. **kick-the-bucket** instead of **die**, or **fly-id** rather than **flaw** (I use the **id** suffix when the label for the constant would otherwise be the same as that of an existing single-word expression). It might be the case that the idiomatic meanings and their literal translations really are extensionally equivalent, so that **fly-id** = **flaw**, but we have seen some reasons to doubt this (e.g. *kick the bucket* retains the punctual aspect of *kick*, which makes it different from *die*), and so it does not hurt to reserve judgement for the time being. If it turns out they are equivalent, all we have done is introduce some unnecessary extra labels; the alternative would be to collapse important distinctions. The former potentially

makes the theory excessively verbose, but the latter makes it inaccurate, and verbosity is clearly to be preferred over inaccuracy. With this in mind, let us now turn to the analysis of our three classes of non-verbal MWEs.

6.1.2.1 *Kith and kin*

The analysis of the first class is trivial. MWEs like *kith and kin* which do not permit any phrase-internal modification have no special semantics; they just behave like any other common noun. The contents of the template KITH-AND-KIN-MEANING introduced in Figure 6.3 is simply (12), where **k&k** is the meaning constant associated with whatever lexical concept ‘kith and kin’ stands for:

$$(12) \quad \lambda x.\mathbf{k\&k}(x) : e_{\downarrow} \multimap t_{\downarrow}$$

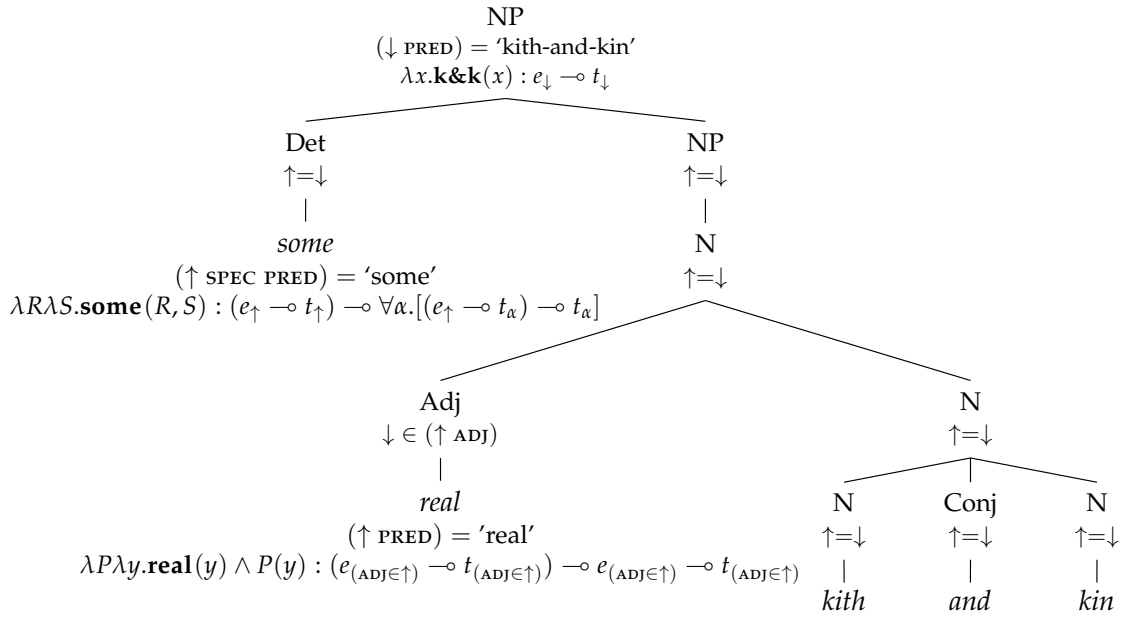
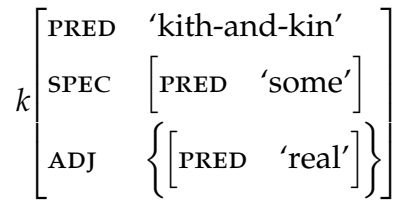
This has the usual form of a nominal meaning constructor, the only difference being the use of \downarrow rather than \uparrow , owing to the fact that the annotation appears at the top of the NP rather than on a terminal node.⁹

Figure 6.6 shows the c-structure for the noun phrase *some real kith and kin* from example (10b). This gives us the f-structure in Figure 6.7, and the Glue proof in Figure 6.8. Notice that this is identical in the relevant respects to the proof for any similar noun phrase without a MWE, such as *some real family*. Semantically speaking, these expressions really do behave just like big words.

6.1.2.2 *Fly in the ointment*

The second class, the decomposable MWEs, are also relatively straightforward analytically speaking. They just require one additional move; namely, the association of multiple meaning constructors with the same lexical entry. This is already commonplace in standard Glue—for example, we saw in Chapter 4.4.2.4 that adjectives are standardly assumed to contribute both a ‘lexical’ and ‘grammatical’

⁹Recall that where we write the meaning constructor, just like any annotation, is purely a matter of convention. In the tree diagrams which we are using as a shorthand, I have chosen to place annotations on the maximal projection of the relevant phrase when they are not obviously associated with a specific word (terminal node), as is the case for non-decomposable MWEs.

Figure 6.6: Annotated c-structure for *some real kith and kin*Figure 6.7: F-structure for *some real kith and kin*

$$\begin{array}{l}
 \llbracket \text{some} \rrbracket \\
 \lambda R \lambda S. \text{some}(R, S) : \\
 (e_k \multimap t_k) \multimap \forall \alpha [(e_k \multimap t_\alpha) \multimap t_\alpha] \\
 \hline
 \llbracket \text{real} \rrbracket \\
 \lambda P \lambda y. \text{real}(y) \wedge P(y) : \\
 (e_k \multimap t_k) \multimap e_k \multimap t_k \\
 \hline
 \llbracket \text{kith-and-kin} \rrbracket \\
 \lambda x. \mathbf{k\&k}(x) : \\
 e_k \multimap t_k \\
 \hline
 \lambda S. \text{some}(\text{real} \wedge \mathbf{k\&k}, S) : \\
 \forall \alpha [(e_k \multimap t_\alpha) \multimap t_\alpha]
 \end{array}$$

Figure 6.8: Glue proof for *some real kith and kin*

meaning constructor, the former providing the core lexical meaning, and the latter controlling the combinatorics:

- (13) a. $\lambda x.\mathbf{swedish}(x) : e_{\uparrow} \multimap t_{\uparrow}$
 b. $\lambda Q\lambda P\lambda x.P(x) \wedge Q(x) :$
 $(e_{\uparrow} \multimap t_{\uparrow}) \multimap (e_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}) \multimap e_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}$
 (adapted from Dalrymple 2001: 266)

Standardly, each word will contribute at most a single lexical meaning, but it could contribute any number of grammatical meaning constructors. The only innovation in dealing with MWEs is that we now allow multiple lexical meaning constructors as well. For example, *fly in the ointment* contributes the two meaning constructors indicated in Figure 6.4; these are expanded below:

- (14) a. **IDIOM-FLY-MEANING** :=
 $\lambda y\lambda x.\mathbf{fly-id}(x, y) : e_{(\uparrow\text{OBJ})} \multimap e_{\uparrow} \multimap t_{\uparrow}$
 b. **IDIOM-OINTMENT-MEANING** :=
 $\lambda x.\mathbf{ointment-id}(x) : e_{\uparrow} \multimap t_{\uparrow}$

Idiomatic *fly* is a two-place relation like *flaw*: **fly-id**(x, y) means that x is a flaw in y . Its internal argument will be filled by the meaning for the NP containing idiomatic *ointment*, its object. Essentially, *fly in the ointment* is a normal NP, but one which has been pre-assembled in the lexicon. In other words, each of the nouns has a normal nominal meaning, allowing modifiers and determiners to apply as they would to any other noun. I illustrate both these possibilities in the Glue proof given in Figure 6.10 for the phrase *the empirical fly in my theoretical ointment*, assuming the f-structure in Figure 6.9.

6.1.2.3 All of a sudden

The third class of non-verbal MWEs are like the first class in only contributing a single lexical meaning, but unlike it in allowing modifiers to appear syntactically ‘inside’ them. Such modifiers sometimes apply to the expression that contains them, but sometimes they seem to scope higher. In this section, I will first

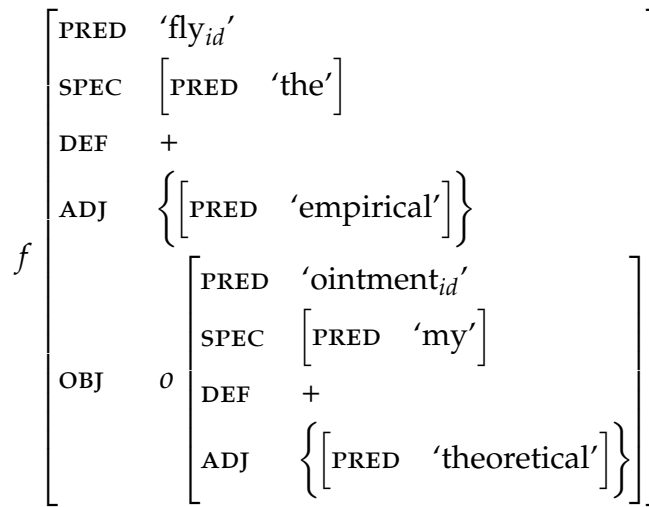


Figure 6.9: F-structure for *the empirical fly in my theoretical ointment*

present the basic meaning of *all of a sudden*, and then consider three kinds of adjectival modification which it can undergo.

For the basic case, let us take (15) as our running example, which has the f-structure given in Figure 6.11.

(15) All of a sudden, Jadzia died.

Setting aside whatever pragmatic effects arise from using an idiomatic expression, this has much the same (truth-conditional) meaning as (16):

(16) Suddenly, Jadzia died.

The semantic representation of (16) is as in (17).¹⁰ Let us assume, then, that the target meaning representation for (15) is parallel to this, as shown in (18):

¹⁰For simplicity I am treating **sudden** (and, analogously, **all-of-a-sudden**) as a one-place predicate on events, and thus abstracting away from many complicated questions about its lexical semantics. For instance, it presumably also takes a time argument which it says occurs unexpectedly quickly after some contextually specified reference time. Of course then we must ask what 'unexpectedly' means, and this leads us onto the question of gradability. Certainly, for something to be 'sudden' it has to happen fast with respect to some comparison class, and so a full analysis might also assign a degree argument to **sudden** and specify how this relates to the relevant comparison class. My aim in the text is not to give a detailed account of the lexical semantics of *suddenly*, or of the MWEs under discussion, but rather to show how they can be made to behave compositionally. To this end, I omit these details from consideration here, trusting that any needed complexification could be accomplished without invalidating my proposals.

$$d \left[\begin{array}{l} \text{PRED} \quad \text{'die'} \\ \text{SUBJ} \quad j \left[\text{PRED} \quad \text{'Jadzia'} \right] \\ \text{ADJ} \quad \left\{ \left[\text{PRED} \quad \text{'all-of-a-sudden'} \right] \right\} \end{array} \right]$$

Figure 6.11: F-structure for *All of a sudden, Jadzia died*

$$(17) \quad \exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{sudden}(e)]$$

$$(18) \quad \exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e)]$$

In other words, our approach is to treat *all of a sudden* as if it just means *suddenly*. So what is the meaning contribution of *suddenly*?

Like all modifiers, *suddenly* provides two meaning constructors: one lexical, the other controlling its combinatorics. These are given in (19):

$$(19) \quad \lambda e.\mathbf{sudden}(e) : v_{\uparrow} \multimap t_{\uparrow} \\ \lambda G \lambda \mathcal{V} \lambda F. \mathcal{V}(\lambda e.F(e) \wedge G(e)) : \\ (v_{\uparrow} \multimap t_{\uparrow}) \multimap ((v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap \\ (v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)}$$

The standard compositional meaning constructor for the adverb combines these, leaving us with an expression which modifies verbal types ($\langle vt, t \rangle$):

$$(20) \quad \lambda \mathcal{V} \lambda F. \mathcal{V}(\lambda e.F(e) \wedge \mathbf{sudden}(e)) : \\ ((v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap (v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap t_{(\text{ADJ} \in \uparrow)}$$

The basic meaning contribution of *all of a sudden* is exactly parallel, but, like *kith and kin*, makes use of \downarrow rather than \uparrow , since the constructor is introduced at a maximal projection. The contents of ALL-OF-A-SUDDEN-MEANING is thus as follows:

$$(21) \quad \lambda e.\mathbf{all-of-a-sudden}(e) : v_{\downarrow} \multimap t_{\downarrow} \\ \lambda G \lambda \mathcal{V} \lambda F. \mathcal{V}(\lambda e.F(e) \wedge G(e)) : \\ (v_{\downarrow} \multimap t_{\downarrow}) \multimap ((v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap \\ (v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}$$

These combine to give (22):

$$\begin{array}{c}
\begin{array}{c}
\text{[[Jadzia]]} \quad \text{[[died]]} \\
\mathbf{jadzia} : \lambda x \lambda G. \exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = x \wedge G(e)] : \\
\frac{e_j \quad e_j \multimap ((v_d \multimap t_d) \multimap t_d)}{\lambda F. \exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge F(e) \wedge \mathbf{all-of-a-sudden}(e)] :} \\
\frac{\text{[[all of a sudden]]} \quad \lambda \mathcal{V} \lambda F. \mathcal{V} (\lambda e. F(e) \wedge \mathbf{all-of-a-sudden}(e)) :}{((v_d \multimap t_d) \multimap t_d) \multimap (v_d \multimap t_d) \multimap t_d} \quad \frac{\lambda G. \exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge G(e)] :}{(v_d \multimap t_d) \multimap t_d} \\
\frac{\lambda F. \exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge F(e) \wedge \mathbf{all-of-a-sudden}(e)] :}{(v_d \multimap t_d) \multimap t_d} \quad \frac{\text{[[closure]]}}{\lambda e. \top : v_d \multimap t_d} \\
\frac{}{\exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e)] : t_d}
\end{array}
\end{array}$$

Figure 6.12: Glue proof for *All of a sudden, Jadzia died*

$$(22) \quad \lambda \mathcal{V} \lambda F. \mathcal{V} (\lambda e. F(e) \wedge \mathbf{all-of-a-sudden}(e)) : \\
((v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap (v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}$$

Figure 6.12 shows the derivation for the meaning of (15), using this meaning constructor (and the f-structure labels from Figure 6.11). The proof for (16) is formally identical.

Now that we know what the basic meaning contribution of *all of a sudden* is, we can turn to the question of how to handle modifiers which appear inside the expression. There are three crucial cases, which I illustrate with variations on our running example:

- (23) a. All of a complete sudden, Jadzia died.
(≈ Completely suddenly, Jadzia died.)
 $\exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{complete}(\mathbf{all-of-a-sudden}(e))]$
- b. All of a random sudden, Jadzia died.
(≈ Suddenly and randomly, Jadzia died.)
 $\exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e) \wedge \mathbf{random}(e)]$
- c. All of a terrible sudden, Jadzia died.¹¹
(≈ Terribly, Jadzia suddenly died.)
 $\mathbf{terrible}(\exists e [\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e)])$

In the first of these, (23a), the adjective modifies the MWE: it is the suddenness which is complete, and not, for instance, the dying. In the other two cases, though, the modifier ‘reaches out’ of the MWE to apparently modify some

¹¹This sentence also has a reading analogous to (23a), where *terrible* simply has an intensifying meaning: ‘Jadzia died terribly suddenly’.

element of the clause as a whole—either the event argument, as in (23b), or the entire proposition expressed, as in (23c).

Each of these can be given a relatively straightforward analysis in our approach, the first two requiring no additional mechanisms to be posited, and the third requiring only a minor addition. We begin with the more expected case, exemplified by (23a), where *complete* modifies the meaning of *all of a sudden*, before turning to the other, more ‘unusual’ cases.

Sentence (23a) has the meaning representation in (24):¹²

$$(24) \quad \exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{complete}(\mathbf{all-of-a-sudden}(e))]$$

In other words, *complete* contributes a predicate modifier **complete**, such that **complete**($P(x)$) means that P applies completely to x (which could be spelled out in terms of degrees in a more articulated semantic theory).¹³

In order to reach our target representation, there are two potential issues that must be dealt with: one syntactic, one semantic. Firstly, *complete* appears next to *sudden*, and yet it is interpreted as modifying the whole expression *all of a sudden*, not just *sudden*. This, though, follows from the analysis given

¹²Once again, I abstract away from some orthogonal questions of lexical semantics. Obviously *complete* is precisely the kind of expression which takes a degree argument, so that the meaning of (23a) might be better represented as (i), i.e. that the dying event e is sudden to some degree d , which is in fact complete:

$$(i) \quad \exists e \exists d[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(d, e) \wedge \mathbf{complete}(d)]$$

However, this once again adds nothing to the compositional analysis, so I stick with the simplified account for the sake of clarity.

¹³There is another meaning of *complete* on which it is not a predicate modifier. This is exemplified in (i):

- (i) The house is finally complete.

The attributive equivalent of this is not (iia) but (iib):

- (ii) a. the complete house
b. the completed house

In other words, it is not that its house-ness is total, but rather that some process has finally finished (the construction or the decorating or something else contextually retrievable). However, as demonstrated in (ii), such a reading is not available for *complete* in attributive position, and since the adjectival modification of *all of a sudden* is necessarily attributive, this meaning of *complete* is not available in the cases we are considering.

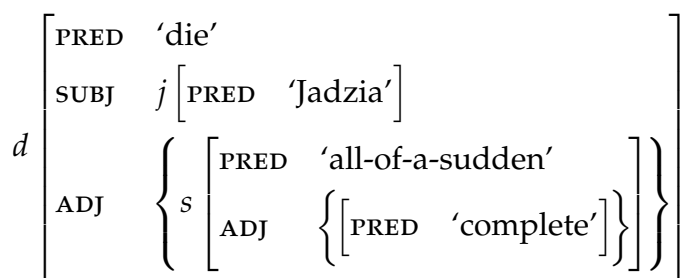


Figure 6.13: F-structure for *All of a complete sudden, Jadzia died*

above whereby all of the c-structure nodes in *all of a sudden* project to the same f-structure. Since meanings are computed from f-structure and not c-structure, *complete* still appears as an adjunct to the whole expression, not just to the noun. This is shown in the f-structure in Figure 6.13. In other words, since *all of a sudden* only contributes a single meaning, it does not matter where the modifier appears inside the expression—it will still modify the same thing. Where the modifier appears is just a matter of syntax: in this case, the noun is the only pair of quasi-nodes available, and so this is the only place adjunction can take place. This is also why the modifier takes the form of an adjective, even though semantically it behaves like an adverb: only adjectival auxiliary trees have root and foot nodes of category N.

The ability to separate out the functional from the configurational in this way is one of the strengths of LFG as a theory. Other frameworks might struggle to account for situations like this where the facts about phrasal position do not line up with the facts about modification, but LFG expects such mismatches to be possible.

Syntactically, then, things are running smoothly. When it comes to the semantics, however, there is a potential problem regarding types. Meaning constructors for adjectives normally take a predicate of individuals as an argument, not a predicate of events, so the (composed) meaning constructor for *complete* might ordinarily be thought to be (25):

$$(25) \quad \lambda P \lambda x. \mathbf{complete}(P(x)) : (e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$$

That is, its first argument is of type $\langle et \rangle$, not the $\langle vt \rangle$ type required to modify an adverbial like *all of a sudden*. An example of this ‘normal’ use would be in a sentence like (26a), which has the representation given in (26b), where **failure** is a normal type $\langle et \rangle$ predicate over individuals:

- (26) a. I see a complete failure.
 b. $\exists x[\exists e[\mathbf{see}(e) \wedge \mathbf{exp} = \mathbf{SPKR} \wedge \mathbf{th}(e) = x \wedge \mathbf{complete}(\mathbf{failure}(x))]]$

But if we want *complete* to modify the meaning of *all of a sudden*, it must have a different type:

- (27) $\lambda P \lambda e. \mathbf{complete}(P(e)) : (v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap v_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$

The most straightforward solution to this clash is simply to relax the type restrictions on the **complete** predicate, so that predicates of both individuals and eventualities are included in its domain. We can once again use the supertype E , where $D_E = D_e \cup D_{e'}$, which includes both types. The meaning constructor for *complete* is then as in (28):

- (28) $\lambda P \lambda X. \mathbf{complete}(P(X)) : (E_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap E_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$

This meaning constructor will do double duty as both an adverbial and adjectival modifier. Figure 6.14 shows how it can be used to derive the correct meaning for *All of a complete sudden, Jadzia died*, using the f-structure labels from Figure 6.13.

One question which should immediately arise is how *ad hoc* such a move is. If we are tinkering with the semantics of *complete* (and, by extension, any other adjective that can appear in this position) solely to get it to compose correctly in the present case, then this is a rather heavy-handed and inelegant solution. However, there is actually independent motivation for such a change. For instance, if we follow Parsons’s (1990) analysis of state sentences, then (29a) has the representation in (29b):

- (29) a. Dukat is a failure.
 b. $\exists s[\mathbf{failure}(s) \wedge \mathbf{th}(s) = \mathbf{dukat}]$

$$\begin{array}{c}
\frac{\text{[[all of a sudden]]-1} \quad \text{[[complete]]}}{\lambda e.\text{all-of-a-sudden}(e) : \lambda P\lambda X.\text{complete}(P(X)) :} \\
\frac{(E_s \multimap t_s) \multimap E_s \multimap t_s}{\lambda X.\text{complete}(\text{all-of-a-sudden}(X)) :} \\
\frac{\lambda \forall \lambda G.\forall(\lambda e.\text{complete}(\text{all-of-a-sudden}(e)) \wedge G(e)) :}{((v_d \multimap t_d) \multimap t_d) \multimap t_d} \\
\frac{\lambda F\lambda \forall \lambda G.\forall(\lambda e.F(e) \wedge G(e)) :}{(v_s \multimap t_s) \multimap ((v_d \multimap t_d) \multimap t_d) \multimap} \\
\frac{\text{[[all of a sudden]]-2}}{\lambda F\lambda \forall \lambda G.\forall(\lambda e.F(e) \wedge G(e)) :} \\
\frac{\text{[[Jadzia]]} \quad \text{[[died]]}}{\text{Jadzia} : \lambda x\lambda F.\exists e[\text{die}(e) \wedge \text{pt}(e) = x \wedge F(e)] :} \\
\frac{e_j \quad e_j \multimap ((v_d \multimap t_d) \multimap t_d)}{\lambda F.\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{Jadzia} \wedge F(e)] :} \\
\frac{(v_d \multimap t_d) \multimap t_d}{\lambda G.\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{Jadzia} \wedge \text{complete}(\text{all-of-a-sudden}(e)) \wedge G(e)] :} \\
\frac{\text{[[closure]]}}{\lambda e.\top : v_d \multimap t_d} \\
\frac{\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{Jadzia} \wedge \text{complete}(\text{all-of-a-sudden}(e))] : t_d}{\lambda e.\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{Jadzia} \wedge \text{complete}(\text{all-of-a-sudden}(e))] : t_d}
\end{array}$$

Figure 6.14: Glue proof for *All of a complete sudden, Jadzia died*

In this theory, a noun in predicative position contributes a predicate of states. But nouns can be modified by adjectives like *complete*, even in predicative position. And this means that such adjectives must take a type $\langle st \rangle$ argument here:¹⁴

- (30) a. Dukat is a complete failure.
 b. $\exists s[\mathbf{complete}(\mathbf{failure}(s)) \wedge \mathbf{th}(s) = \mathbf{dukat}]$

States are one kind of eventuality, but what about events, the case in point? It has been proposed that some nouns denote sets of events or relations between events and individuals rather than merely sets of individuals (Grimshaw 1992; Larson 1998). This includes nominalisations such as *destruction* or *transformation*, which can, of course, also be modified by adjectives like *complete*:

- (31) a. the complete destruction of the city
 b. $\iota e[\mathbf{complete}(\mathbf{destroy}(e)) \wedge \mathbf{pt}(e) = \iota x[\mathbf{city}(x)]]$

Here, the adjective must have taken a type $\langle vt \rangle$ argument, just as we need it to in the case of *all of a complete sudden*. In other words, even in its regular adjectival use, *complete* does sometimes take predicates of eventualities as arguments, and thus our proposal to loosen its typing is needed independently of the unusual MWE modification cases. In fact, the only reason such uses seem strange to us is that, syntactically speaking, adjectives are not normally in a position to modify adverbial expressions in this way, since they can only adjoin to nouns. It is the peculiar *syntax* of the MWE which gives rise to this unusual, but independently motivated, semantic possibility.

Let us turn now to the intersective modifiers like *random*. I repeat our example sentence and target meaning representation below:

- (32) a. All of a random sudden, Jadzia died.
 b. $\exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e) \wedge \mathbf{random}(e)]$

¹⁴Note that s is the type of states, not the intensional type used by e.g. Montague (1973). This means $\langle st \rangle$ is the type of predicates of states, not the type of propositions, which in our framework is just t .

$$d \left[\begin{array}{l} \text{PRED} \quad \text{'die'} \\ \text{SUBJ} \quad j \left[\text{PRED} \quad \text{'Jadzia'} \right] \\ \text{ADJ} \quad \left\{ s \left[\begin{array}{l} \text{PRED} \quad \text{'all-of-a-sudden'} \\ \text{ADJ} \quad \left\{ \left[\text{PRED} \quad \text{'random'} \right] \right\} \right] \right\} \right\} \end{array} \right]$$

Figure 6.15: F-structure for *All of a random sudden, Jadzia died*

Sentence (32a) has the f-structure given in Figure 6.15. This time, the modifier is in the ‘wrong’ place in the f-structure: *random* modifies the event argument of the main clause, but appears in the ADJ set of *all of a sudden*. However, despite this syntactic quirk, composition proceeds quite straightforwardly. First of all, we again assume that the meaning constant provided by *random* takes both eventuality and individual arguments, on the basis of equivalent examples to those we saw above, e.g. (33):

- (33) a. the random destruction of the city
 b. $\iota e[\mathbf{destroy}(e) \wedge \mathbf{pt}(e) = \iota x[\mathbf{city}(x)] \wedge \mathbf{random}(e)]$

Since *random* can apply to event nouns as much as non-event nouns, its lexical meaning constructor therefore makes use of the supertype E :

$$(34) \quad \lambda X.\mathbf{random}(X) : E_{\uparrow} \multimap t_{\uparrow}$$

Now, *random* also contributes the usual combinatorial meaning constructor for adjectives, which we also modify in the appropriate way:

$$(35) \quad \lambda Q\lambda P\lambda X.P(X) \wedge Q(X) : \\ (E_{\uparrow} \multimap t_{\uparrow}) \multimap (E_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap E_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$$

Notice that this will *not* lift *random* into an adverbial type: the output of this meaning constructor is a type $\langle Et, Et \rangle$ meaning, not the type $\langle \langle vt, t \rangle, \langle vt, t \rangle \rangle$ needed to modify verbal meanings. But this is as it should be: *random* cannot normally function as an adverbial (for which we use the adverb *randomly*). It

is the special context of *all of a sudden* which permits it to have this function (for reasons which we turn to presently).

Using the f-structure labels in Figure 6.15, we can give the result of combining (34) and (35) as (36):

$$(36) \quad \lambda P \lambda X. P(X) \wedge \mathbf{random}(X) : (E_s \multimap t_s) \multimap E_s \multimap t_s$$

This is a modifier of a predicative meaning associated with f-structure s —the f-structure of *all of a sudden*. And in fact there is such a predicative meaning, since the lexical meaning of *all of a sudden* is of the correct type:

$$(37) \quad \lambda e. \mathbf{all-of-a-sudden}(e) : v_s \multimap t_s$$

Since *random* is intersective, all (36) does is to conjoin the property of being random with whatever its argument is: in the current case, that is the property of happening all of a sudden. We are left with a modified version of the lexical meaning constructor for *all of a sudden*, (38), which can then be consumed by its combinatorial meaning constructor to produce the usual adverbial type, (39):

$$(38) \quad \lambda X. \mathbf{all-of-a-sudden}(X) \wedge \mathbf{random}(X) : E_s \multimap t_s$$

$$(39) \quad \lambda \mathcal{V} \lambda G. \mathcal{V}(\lambda e. \mathbf{all-of-a-sudden}(e) \wedge \mathbf{random}(e) \wedge G(e)) : \\ ((v_d \multimap t_d) \multimap t_d) \multimap (v_d \multimap t_d) \multimap t_d$$

Figure 6.16 shows the full Glue proof for *All of a random sudden, Jadzia died*.

How did our semantic analysis work out, when *random* was in the ‘wrong’ place in the syntax? The answer lies in the structure of (Neo-)Davidsonian meaning representations: all modifiers of the event argument are conjoined at the same level, which flattens out the articulated structure of the syntax. This means that *random* can ‘modify’ *all of a sudden* in the sense of conjunctively adding an additional predicate which applies to the same argument—even though ultimately that argument also serves as the argument of the main verbal predicate, and is bound by a quantifier that the verb’s meaning introduces: it is in this sense that *random* modifies a part of the verbal meaning too. This

$$\begin{array}{c}
\boxed{\text{random}} \\
\lambda P \lambda X.P(X) \wedge \text{random}(X) : \lambda e.\text{all-of-a-sudden}(e) : \\
(E_s \rightarrow t_s) \rightarrow E_s \rightarrow t_s \quad v_s \rightarrow t_s \\
\hline
\lambda X.\text{all-of-a-sudden}(X) \wedge \text{random}(X) : \\
E_s \rightarrow t_s \\
\hline
\boxed{\text{all of a sudden}}\text{-1} \\
\lambda F \lambda \lambda G.\mathcal{V}(\lambda e.F(e) \wedge G(e)) : \\
(v_s \rightarrow t_s) \rightarrow ((v_d \rightarrow t_d) \rightarrow t_d) \rightarrow (v_d \rightarrow t_d) \rightarrow t_d \\
\boxed{\text{all of a sudden}}\text{-2} \\
\lambda F \lambda \lambda G.\mathcal{V}(\lambda e.F(e) \wedge G(e)) : \\
(v_s \rightarrow t_s) \rightarrow ((v_d \rightarrow t_d) \rightarrow t_d) \rightarrow (v_d \rightarrow t_d) \rightarrow t_d \\
\vdots \\
\boxed{\text{jadzia died}} \\
\lambda F.\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge F(e)] : \\
(v_d \rightarrow t_d) \rightarrow t_d \\
\hline
\lambda G.\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge \text{random}(e) \wedge G(e)] : \\
(v_d \rightarrow t_d) \rightarrow t_d \\
\boxed{\text{closure}} \\
\lambda e.\top : v_d \rightarrow t_d \\
\hline
\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge \text{random}(e)] : t_d
\end{array}$$

Figure 6.16: Glue proof for *All of a random sudden, Jadzia died*

situation is only possible here because *all of a sudden*, despite having an adverbial meaning, contains a quasi-N node, thus allowing an adjective to appear inside it. Once again, the MWE provides an unusual syntax which in turn permits the adjective meaning to apply to something it would not normally have access to.

Finally, we come to (23c), repeated as (40a), and with the proposed meaning given in (40b):

- (40) a. All of a terrible sudden, Jadzia died.
 b. **terrible**($\exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{all-of-a-sudden}(e)]$)

I take this reading of the sentence to be approximately paraphrasable as (41):

- (41) Terribly, Jadzia suddenly died.

In this context, *terribly* is a so-called EVALUATIVE adverb. These express some attitude of the speaker towards the proposition they modify. (41) means, essentially, (42):

- (42) Jadzia suddenly died and I think that's terrible.

Other canonical evaluative adverbs include *remarkably*, *fortunately*, or *oddly*. One thing they have in common is that sentences modified by them can very often be paraphrased using the corresponding adjective, as in (42). Thus, the following pairs of sentences are approximately equivalent:

- (43) a. Fortunately, the police came.
 b. It is fortunate that the police came.
- (44) a. Remarkably, nobody minded.
 b. It is remarkable that nobody minded.

One important difference between the (a) and (b) sentences, however, is in the information status of the modifier. In the (b) sentences, the adjective is part of what is asserted, “part of the main content of the sentence” as Bonami & Godard (2008: 282) put it; in the (a) sentences, this is not the case—the adverbial

contribution is backgrounded, or set aside. This becomes clearer when we consider the sentences in (45)–(46) (Bonami & Godard 2008: 284):

- (45) a. If Paul goes away on vacation, we will, strangely, not know about it.
 ⇔
 b. If Paul goes away on vacation, we will not know about it.
- (46) a. If Paul is going away on vacation, it is strange that we don't know about it.
 ⇏
 b. If Paul is going away on vacation, we don't know about it.

The sentences in (45) are mutually entailing, but those in (46) are not: removing the adjective changes the meaning of the sentence in a way that removing the adverb does not.

This and other considerations have led various authors to suggest that evaluatives introduce conventional implicatures, (CIs: Jayez & Rossari 2004; Potts 2005—Bonami & Godard 2008 offer a slightly more complex analysis, but working from the same basic idea). This is undoubtedly right, but in the interests of formal simplicity, I refrain from integrating a full-blown Pottsonian multidimensionality into the semantics here.¹⁵ Instead, I will follow Ernst (2001: Ch. 2) and take speaker-oriented evaluative adverbs like *terribly* to simply denote predicates of propositions. Compositionally, this will amount to the same thing, since one crucial aspect of CI-introducing meanings is that they leave their at-issue content untouched; so a sentence-level adverb like *terribly* would be of type $\langle t^a t^c \rangle$, taking an at-issue proposition to a CI-dimension proposition. For our purposes, we simply ignore the dimensions and treat the modifier as a type $\langle tt \rangle$.

We therefore have the following meaning constructor for *terribly*, and also for *terrible* in its use in *all of a terrible sudden*:

¹⁵See Arnold & Sadler (2010), and work by Ash Asudeh and Gianluca Giorgolo (Giorgolo & Asudeh 2011; Asudeh & Giorgolo 2016) for two different proposals on how to go about achieving this in LFG+Glue.

$$(47) \quad \lambda p.\mathbf{terrible}(p) : t_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}$$

In fact, we will make use of the equivalent but higher typed (48), for reasons which will be explained shortly:

$$(48) \quad \lambda \mathcal{V} \lambda F.\mathbf{terrible}(\mathcal{V}(\lambda e.F(e))) : \\ ((v_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}) \multimap t_{(\text{ADJ}\in\uparrow)}) \multimap (v_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}) \multimap t_{(\text{ADJ}\in\uparrow)}$$

One natural question to ask is whether the adjective can have this evaluative meaning independently, or whether we need to propose some special mechanism for deriving it. The answer is that it can, and so we don't. *Terrible* has many different meanings, some of which are illustrated below with the relevant part of the meaning language translation:

- (49) a. Throw that [terrible chair] away!
b. $\mathbf{terrible}_1(\mathbf{chair}(x))$

- (50) a. Will suffered a [terrible beating].
b. $\mathbf{terrible}_2(\mathbf{beat}(e))$

- (51) a. Julian is a terrible driver.
b. $\exists e[\mathbf{drive}(e) \wedge \mathbf{ag}(e) = \mathbf{julian} \wedge \mathbf{terrible}_3(e)]$

- (52) a. We're only wasting a small amount, but [it is a terrible waste] because it could so easily be avoided.
b. $\mathbf{terrible}_4(\exists s[\mathbf{waste}(s)])$

Sentence (49) illustrates perhaps the canonical sense of *terrible*: $\mathbf{terrible}_1(P(x))$ means 'x is a P, but a very bad example of one'. What we understand from (49) is that the chair in question is not very good in some respect—most plausibly, it is not very good as a chair. Contrast this with (50), however, where the beating is not said to be of poor quality—if anything, its quality *qua* beating is very high!—but rather to be an extreme example of a beating. In other words, although $\mathbf{terrible}_2(P(x))$ has the same form as $\mathbf{terrible}_1(P(x))$, it has almost the opposite denotation: $\mathbf{terrible}_2(P(x))$ means that x is an extreme example of

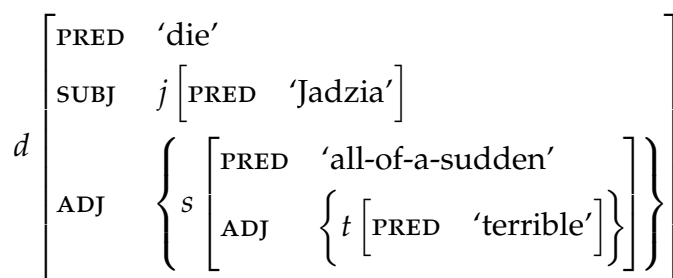


Figure 6.17: F-structure for *All of a terrible sudden, Jadzia died*

a *P*.¹⁶ Sentence (51) illustrates another standard meaning for *terrible*, modifying the event argument of a participant noun like *driver*. Here we understand that Julian's driving is terrible: he is terrible as a driver.¹⁷ Lastly, we come to (52). On the most salient reading, the speaker is not claiming that the waste is a deficient example of a waste (the **terrible**₁ reading), nor, certainly given the context, that it is an extreme example of waste (the **terrible**₂ reading). Rather, the claim is that it is terrible that there even is any waste at all (given that it could easily be avoided). In other words, we have the evaluative meaning described above.

Assuming the adjective itself can provide the right kind of meaning, the question of types is resolved. However, this time the syntax is more of an obstacle. The f-structure for *All of a terrible sudden, Jadzia died* is given in Figure 6.17. Once again, the f-structure of the adjective, *t* is contained inside the f-structure of the MWE, *s*. However, since **terrible**₄ needs to take a verbal type, $\langle vt, t \rangle$, as an argument, such a configuration is more problematic than it was for *random*.¹⁸ Although the event argument is available at both *d* and *s*, the verbal meaning constructor only corresponds to *d*. Unfortunately, the meaning constructor for

¹⁶The lexical semantics are a little bit more complicated than this: the extreme thing must be a bad thing, for instance. This is why (i) cannot mean that Julian had a very fast time:

(i) Julian ran a terrible marathon.

My thanks to Mary Dalrymple (pers. comm.) for pointing this out to me.

¹⁷It seems to me that the so-called 'intersective' reading of this sentence (Larson 1998) is marginally available too, whereby Julian is terrible (as a person) and is also a driver.

¹⁸Since *terrible* can also apply to stative predicates like *waste*, as seen in (52), it should probably refer to eventualities rather than events in its meaning constructor, i.e. it should be looking for a type $\langle et, t \rangle$ argument rather than a type $\langle vt, t \rangle$ one. However, I ignore this complexity in the text.

terrible refers to $(\text{ADJ} \in \uparrow)$, which instantiates to s in Figure 6.17—this means there are no meaning constructors of the proper type available for it to consume.

To resolve this issue, we resort to a pair of special meaning constructors introduced by the MWE. We add the following (jointly optional) meaning constructors to the template ALL-OF-A-SUDDEN-MEANING:

$$(53) \quad \left(\begin{array}{l} \lambda \mathcal{V}. \mathcal{V} : ((v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap (v_{\downarrow} \multimap t_{\downarrow}) \multimap t_{\downarrow} \\ \lambda \mathcal{V}. \mathcal{V} : ((v_{\downarrow} \multimap t_{\downarrow}) \multimap t_{\downarrow}) \multimap ((v_{(\text{ADJ} \in \downarrow)} \multimap t_{(\text{ADJ} \in \downarrow)}) \multimap t_{(\text{ADJ} \in \downarrow)}) \end{array} \right)$$

Let us call these two meaning constructors [lower] and [raise], respectively. The idea is this: [lower] changes the type of the main clause’s verbal meaning so that it appears to be part of the same f-structure as *all of a sudden*, thus ‘flattening’ the f-structure; once the other modification has applied, [raise] then returns it to the appropriate type so that composition can continue as usual. Since they are jointly optional, if one is used, both must be—this ensures that if the f-structure is flattened, it is also restored subsequently. Figure 6.18 shows the Glue proof for *All of a terrible sudden, Jadzia died* using this technique.

Special meaning constructors like this are not the most elegant of solutions—indeed, they appear to be something of a hack. However, they do have the advantage of allowing us to associate all of the ‘special’ meanings with the MWE itself, which is already idiosyncratic in a number of ways. We should not underplay this virtue. It seems desirable to associate any anomalous combinatorial possibilities that arise in relation to MWEs with the lexical entries of the MWEs themselves, and in this way to leave the rest of the grammar untouched. Here the meaning of *terrible* remains unchanged, and its unusual combinatorial possibilities in the context of the MWE are explained by the MWE itself. It may be that there is a higher generalisation being missed, but it could well be that these expressions simply behave in unusual ways.

Before we conclude, I offer a brief note on why we could not use the lower type $\langle tt \rangle$ meaning for *terrible*. Essentially, it leads to an unattested reading for the sentence. Since it would consume a type t_s argument, and since the lexical meaning constructor for *all of a sudden* is of type $\langle v_s t_s \rangle$, we can, by hypothetical

$$\begin{array}{c}
\llbracket \text{jadzia died} \rrbracket \\
\vdots \\
\frac{\llbracket \text{all of a sudden} \rrbracket \quad \lambda \lambda \lambda G, \mathcal{V}(\lambda e. \text{all-of-a-sudden}(e) \wedge G(e)) : \lambda F. \exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge F(e)] : ((v_d \multimap t_d) \multimap t_d)}{\lambda G. \exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge G(e)] : ((v_d \multimap t_d) \multimap t_d)} \llbracket \text{lower} \rrbracket \\
\frac{\lambda G. \exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge G(e)] : ((v_d \multimap t_d) \multimap t_d)}{\lambda \mathcal{V}. \mathcal{V} : ((v_d \multimap t_d) \multimap t_d) \multimap t_s} \llbracket \text{terrible} \rrbracket \\
\frac{\lambda \mathcal{V}. \mathcal{V} : ((v_s \multimap f_s) \multimap f_s)}{\lambda \mathcal{V}. \mathcal{V} : ((v_s \multimap f_s) \multimap f_s) \multimap f_s} \llbracket \text{raise} \rrbracket \\
\frac{\lambda F. \text{terrible}(\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge F(e)]) : ((v_s \multimap f_s) \multimap f_s)}{\lambda F. \text{terrible}(\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e) \wedge F(e)]) : t_d} \llbracket \text{closure} \rrbracket \\
\frac{\lambda e. \top : v_d \multimap t_d}{\text{terrible}(\exists e[\text{die}(e) \wedge \text{pt}(e) = \text{jadzia} \wedge \text{all-of-a-sudden}(e)]) : t_d}
\end{array}$$

Figure 6.18: Glue proof for *All of a terrible sudden, Jadzia died.*

$$\begin{array}{c}
\llbracket \text{all of a sudden} \rrbracket\text{-1} \\
\lambda e.\text{all-of-a-sudden}(e) : \\
\frac{v_s \multimap t_s \quad [e' : v_s]^1 \quad \llbracket \text{terrible} \rrbracket}{\text{all-of-a-sudden}(e') : \quad \lambda p.\text{terrible}(p) :} \\
\frac{t_s \quad \text{terrible}(\text{all-of-a-sudden}(e')) :}{t_s} \multimap_{\mathcal{I},1} \frac{\llbracket \text{all of a sudden} \rrbracket\text{-2} \quad \lambda F\lambda\mathcal{V}\lambda G.\mathcal{V}(\lambda e.F(e) \wedge G(e)) :}{\lambda e.\text{terrible}(\text{all-of-a-sudden}(e)) :} \\
\frac{v_s \multimap t_s \quad ((v_d \multimap t_d) \multimap t_d) \multimap (v_d \multimap t_d) \multimap t_d}{\lambda\mathcal{V}\lambda G.\mathcal{V}(\lambda e.\text{terrible}(\text{all-of-a-sudden}(e)) \wedge G(e)) :} \\
((v_d \multimap t_d) \multimap t_d) \multimap (v_d \multimap t_d) \multimap t_d
\end{array}$$

Figure 6.19: Glue proof for the unattested reading of *all of a terrible sudden*

reasoning, derive a meaning whereby **terrible** scopes only over the adverbial meaning of *all of a sudden*, and not the whole sentence. Figure 6.19 shows the relevant portion of the proof. We derive a standard adverbial meaning constructor, but with **terrible** embedded under the verbal meaning. This would leave us with the meaning representation in (54a), paraphrasable perhaps as (54b) (with appropriate stress on *sudden*):

- (54) a. $\exists e[\mathbf{die}(e) \wedge \mathbf{pt}(e) = \mathbf{jadzia} \wedge \mathbf{terrible}(\mathbf{all-of-a-sudden}(e))]$
b. Terribly, Jadzia died all of a sudden.

Here it is the fact of the dying being sudden which is evaluated as terrible by the speaker, rather than the whole proposition. Since this is not an available reading of *All of a terrible sudden, Jadzia died*, and since assigning *terrible* the lower $\langle tt \rangle$ type allows it to be derived, we must give *terrible* the higher type shown above instead. This higher type does not lead to the unwanted reading, because it applies **terrible** to the whole verbal complex directly, and any further modification applies within the scope of the \mathcal{V} predicate, which corresponds to the verbal meaning constructor. Thus, **terrible** cannot apply any lower, and any other (straightforward adverbial) modification also occurs within its scope, just as we want.

6.1.3 Non-verbal MWEs: a summary

In this section, we have shown how our framework can be used to analyse non-verbal MWEs. I argued that we should represent these expressions as having an articulated phrasal syntax, even when they appear to be relatively fixed. In general, though, the syntax of these expressions is fairly straightforward: our main question has been one of representation rather than variability. The only alternations involve modification, which is simply a question of where adjunction is possible, and therefore which nodes are represented by pairs of quasi-nodes. When addressing the question of modification, we divided up the non-verbal MWEs into three types: those which permit no modification, those which permit direct modification of their parts, and those which allow unusual modifications which apply not just to the expression itself, but to elements external to it. I showed how each of these could be accounted for semantically. The first type simply behave like a single word would; the second type include multiple lexical meaning constructors which can then be modified independently; the third type have a special syntax that allows modifiers to appear in places they otherwise would not, with semantic results which often follow from independent considerations (but which in other cases require that the MWE contribute special meaning constructors which allow for more unusual modification possibilities).

In the next section, we turn to verbal MWEs. In many respects our approach will be the same, but the verbal domain offers more complexity owing to the increased syntactic flexibility of verbal MWEs when compared to non-verbal ones.

6.2 Verbal MWEs

There is rather more to discuss when it comes to verbal MWEs: they share with non-verbal MWEs the possibility for their parts to be modified, and for the meaning of the expression to be distributed among those parts; but because verbs usually have more complex argument structures than other parts of speech, we also must address the question of voice alternations, and the possibility

of unbounded dependencies existing between the sub-parts. We begin this section by looking at decomposable verbal MWEs, before moving onto non-decomposable ones. Throughout, I continue to follow the principle that any peculiarity surrounding a MWE should be represented in the MWE itself, and leave the rest of the grammar untouched.

6.2.1 Decomposable expressions: *spill the beans*

The basic facts about decomposability can be captured in the same way they were for *fly in the ointment*: we associate certain lexical entries with multiple lexical meaning constructors. In this section, we look at such decomposable expressions, and show how the grammar accounts for their behaviour. The overall claim will be that decomposable expressions behave like pre-assembled versions of normal phrases, and that this goes hand in hand with their more ‘normal’ syntactic behaviour.

The schematic lexical entry for a canonical decomposable idiom, *spill the beans*, is given in Figure 6.20. This has the form of the entry for a standard transitive verb, except that the contents of the object NP are also specified. But this NP itself has the form of a normal NP, with adjunction sites at both NP and N. Notice that the determiner is not pre-specified, because it is (a) variable, as shown in (55), (b) optional, as shown in (56), and (c) carries its usual non-idiomatic meaning when it is present (*spill the beans* presupposes that there are contextually salient secrets, whereas *spill some beans* does not, for instance).

- (55) a. They keep us actors totally in the dark so we can't spill **any** beans.
(iWeb)
- b. They do have a podcast though so hopefully they'll spill **some** beans there.
(iWeb)
- c. Without spilling **all of the** beans, here is the idea in a nutshell: [...] (GloWbE)

der the head verb, $(\uparrow \text{VOICE}) = \text{ACTIVE} \Rightarrow \neg(\uparrow \text{OBJ ARG})$. ARG is an abbreviation for all of the argument GFS, i.e. not ADJ or XADJ:

$$(58) \quad \text{ARG} \equiv \{\text{SUBJ} \mid \text{OBJ} \mid \text{OBJ}_\theta \mid \text{COMP} \mid \text{XCOMP} \mid \text{OBL}_\theta\}$$

The force of this constraint is thus to prevent the object of active voice *spill the beans* from having any embedded argument f-structures. This essentially ensures that *beans* is itself the object of *spill*, and not merely embedded inside the object. We still correctly allow determiners and modifiers, since SPEC and ADJ are not prohibited, but we rule out cases like (57), because these involve an embedded OBJ:

$$(59) \quad * \left[\begin{array}{l} \text{PRED} \quad \text{'spill}_{id}' \\ \text{SUBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'Quark'} \end{array} \right] \\ \text{OBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'importance'} \\ \text{SPEC} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'the'} \end{array} \right] \\ \text{OBJ} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'beans}_{id}' \\ \text{SPEC} \quad \left[\begin{array}{l} \text{PRED} \quad \text{'the'} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

The second conditional expression places the same constraint on the passive subject, thus ruling out sentences like (60):

$$(60) \quad \# \text{The importance of the beans was spilled by Quark.}$$

From the VP node upwards, the entry is identical to the lexical entry for any verb, and so the VP is also split into a pair of quasi-nodes, since adverbial expressions and auxiliaries can adjoin here as usual:

$$(61) \quad \text{a.} \quad \text{When an official [}_{VP} \text{ inadvertently [}_{VP} \text{ [}_{VP} \text{ spills the beans] during an interview]], the smart reporter suppresses his excitement and caps his pen in hopes that the official will dig himself in deeper.}$$

(GloWbE)

- b. Since stepping down as head of Mossad in January, after eight years in the post, Dagan [VP has [VP been [VP spilling the beans]]] in a series of unusual public statements and appearances. (iWeb)

The decomposability of the expression is encoded in the fact that the lexical entry contains multiple meaning templates which each introduce a lexical meaning:

- (62) a. **IDIOM-SPILL-MEANING** :=
 $\lambda x \lambda y \lambda F. \exists e [\mathbf{spill-id}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{pt}(e) = y \wedge F(e)] :$
 $e_{(\uparrow\text{SUBJ})} \multimap e_{(\uparrow\text{OBJ})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$
 $\lambda e. \top : v_{\uparrow} \multimap t_{\uparrow}$
- b. **IDIOM-BEANS-MEANING** :=
 $\lambda x. * \mathbf{bean-id}(x) : e_{\uparrow} \multimap t_{\uparrow}$

The * in (62b) is Link's (1983) pluralising operator, which closes the following predicate under the sum operation. I say no more about the semantics of plurality here, since it is orthogonal to the main discussion. This is the normal denotation of a plural noun.¹⁹ The meaning constructors in (62a) are, similarly, just the normal meaning contribution of a transitive verb: a lexical meaning constructor with dependencies on its two arguments, and the closure constructor. In this case, since one of those two arguments is mentioned in the same lexical entry, it might seem logical to 'pre-compose' the two meanings, but we leave them separate since idiomatic *beans* can be modified, and, compositionally speaking, this has to take place before its meaning is consumed by the verb or by a quantifier.²⁰

With these meanings and structures in place, composition proceeds as it would for any other sentence involving a transitive verb, with modification of

¹⁹To account for the bare plural uses in (56), this entry must also include a meaning constructor which modifies the type appropriately (e.g. in the manner described by Chierchia 1982); I leave the details of this to one side here also.

²⁰Note that this has an interesting parallel with the motivation which led Dalrymple (2001: 264ff.) to split the meaning contribution of modifiers in two. In that case, the lexical meaning of an adjective like *Swedish* had to be of a different type to its final, compositional type in order to account for the possibility of modification; here, we have two lexical meanings rather than one lexical and one grammatical.

the object *beans* possible as it would be for any other noun. As an example, Figure 6.21 shows the annotated c-structure for *Julian spilled some important beans*, and Figure 6.22 shows the f-structure. Figure 6.23 gives the Glue proof which derives the meaning.

The parallelism with the analysis of a regular transitive sentence highlights the fact that decomposable MWEs are really just pre-assembled chunks of language. Because they are listed, they can give their parts idiosyncratic meanings which are not available outside the expression, but grammatically they behave (largely) like non-idiomatic expressions. This goes some way to accounting for their tendency to be more open to the usual syntactic processes of the grammar, leading to their increased syntactic flexibility (Nunberg et al. 1994). Since they look, structurally, like ‘normal’ language, the normal rules of the grammar will automatically apply to them unless something is done to prevent it. For instance, the lexical entry in Figure 6.20 matches the left-hand side of both the passive rule and the relative clause formation rules given in Chapter 5.5.5, repeated here as Figures 6.24–6.26. This means these rules will apply to *spill the beans* just as much as they apply to a regular transitive verb like *repair*, correctly generating passive forms and relative clauses headed by both the object and subject of *spill the beans*. Crucially, since any descriptions not included in the left-hand side of a lexical rule carry over, all of the specifications about the idiomatic NP complement, *beans*, carry over to the corresponding node on the right-hand side of the lexical rules, namely NP2. This means that the subject in the passive and the noun which heads the object relative clause are both constrained to have the form *beans*, as required. In other words, we correctly predict the existence of the lexical entries shown in Figure 6.27.²¹

²¹Notice that the conditional constraining equations under the verb *spilled* serve an additional purpose in the case of relative clauses, now also correctly ruling out NPs like (i), which lack an idiomatic interpretation:

- (i) a. #beans the importance of which Quark spilled
- b. #beans the importance of which was spilled

The pied piped phrase in each of these expressions corresponds to a complex object in the active, or complex subject in the passive, both of which are ruled out by the conditional constraints.

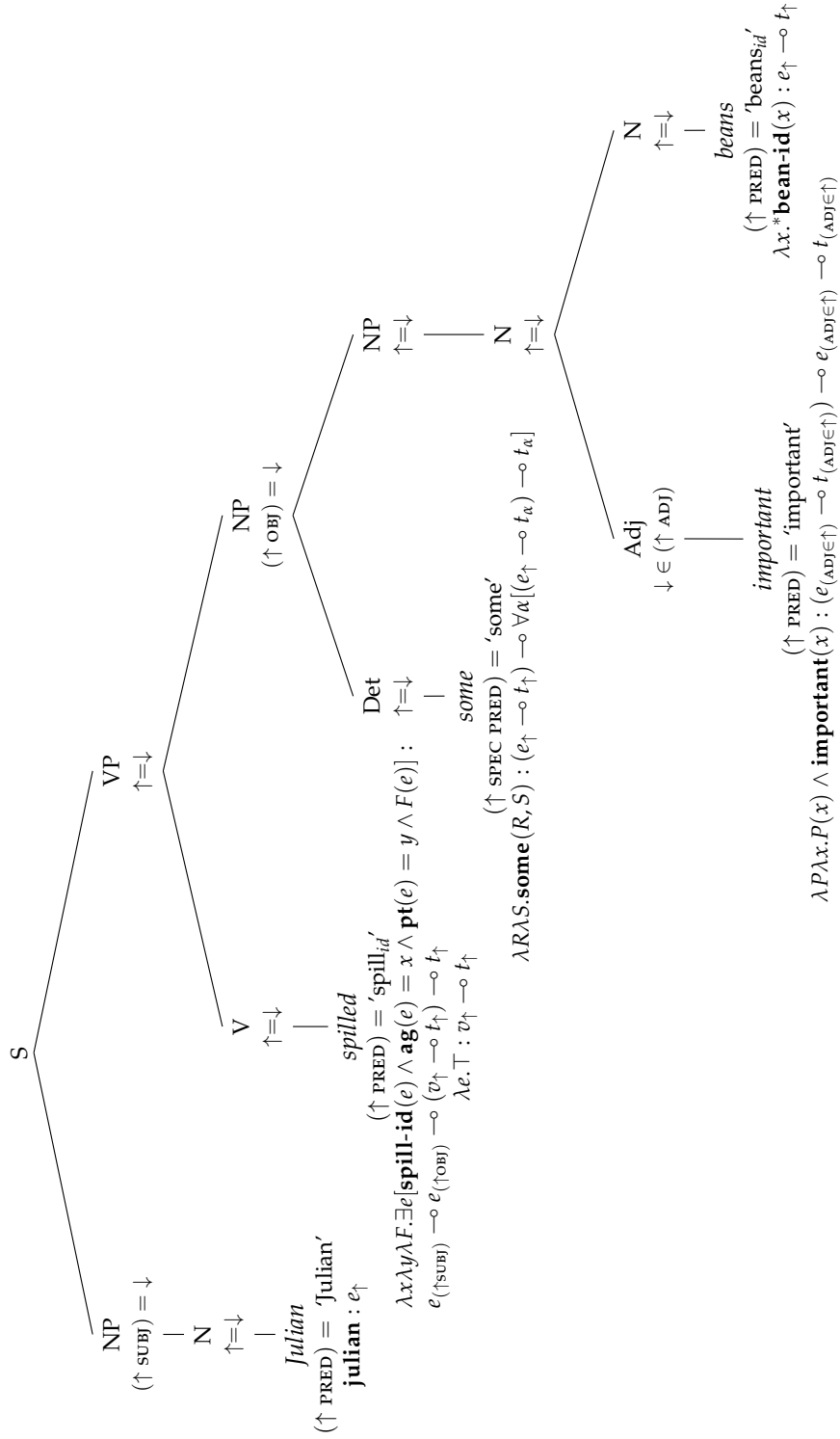


Figure 6.21: Annotated c-structure for *Julian spilled some important beans*

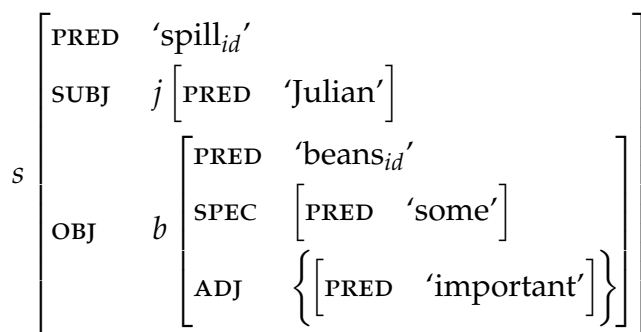


Figure 6.22: F-structure for *Julian spilled some important beans*

Insofar as decomposable expressions look like regular expressions of the language, then, they will be susceptible to the same grammatical alternations. For the most part, this is what we want. But we noted earlier that some decomposable expressions are not as flexible as others. Indeed, for some speakers, relative clauses headed by idiomatic *beans* are not acceptable, in contrast to, say, idiomatic *strings*, from *pull strings*:

(63) #The beans that Joe spilled caused us a lot of trouble.

(64) The strings that Joe pulled got us out of trouble.

Such predictions do not follow from our theory without further stipulation, but this is perhaps as it should be: unless there is some underlying (synchronic) difference between *spill the beans* and *pull strings* which accounts for this fact independently, it truly is a matter of lexical idiosyncrasy. And lexical idiosyncrasies must simply be listed.²² In order to prevent *spill the beans* appearing in object relative clauses, we could annotate the V node in its basic lexical entry with the following equation:

(65) $\neg(\text{TOPIC} (\uparrow \text{OBJ}))$

The object relative clause identifies the %GAP f-structure with the verb's object and with the relative clause's TOPIC. The negative constraining equation in (65) explicitly states that the verb's object is *not* also a topic, thus blocking the

²²At least this is true as far as syntax is concerned; there may well, of course, be historical reasons, or the irregularity may be motivated by another level of the grammar.

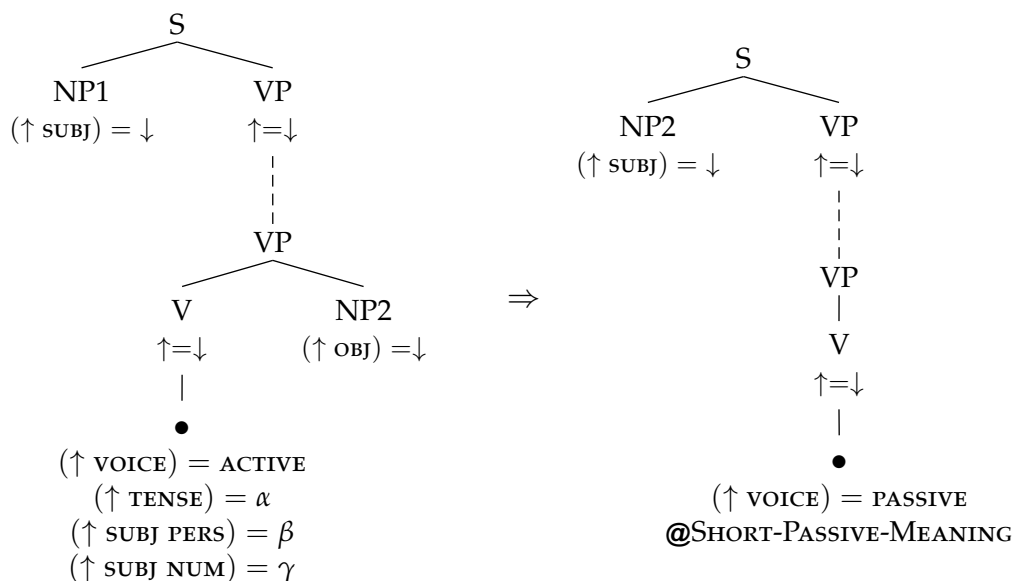


Figure 6.24: Lexical rule for the short passive

relative clause construction where it is. If we agree with Horn (2003) that any construction where idiomatic *beans* is extracted is out, then we can generalise this constraint as in (66), where DF is a disjunction of discourse functions (TOPIC and FOCUS):

$$(66) \quad \neg(\text{DF}(\uparrow \text{OBJ}))$$

$$(67) \quad \text{DF} \equiv \{\text{TOPIC}|\text{FOCUS}\}$$

To sum up, the decomposability of decomposable MWEs is accounted for once again by including multiple lexical meaning constructors in the same lexical entry. In addition, these MWEs have a more ordinary syntactic structure than non-decomposable expressions (as we saw above for non-verbal MWEs, and will see again in the next section for verbal MWEs). This, in turn, means that they are made available to the lexical rules in the grammar, and so they naturally participate in a variety of syntactic alternations. If required, such alternations can be restricted on a lexical basis. Of course, such restrictions may ultimately follow from higher-order properties of the expressions, and it is to be hoped that any such generalisations would fit naturally into the current framework.

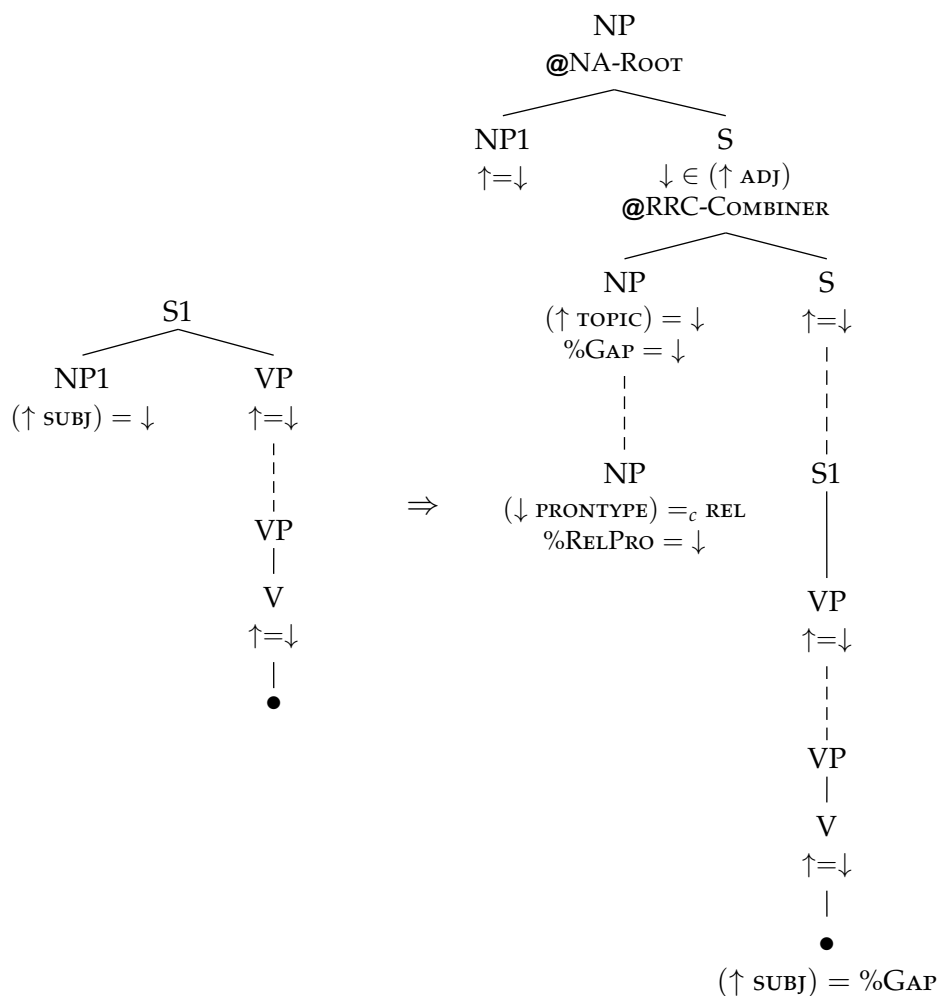


Figure 6.25: Lexical rule for a subject relative clause

Before we move on to non-decomposable expressions, it is worth pointing out that the approach proposed here for decomposable MWEs also provides a plausible analysis for Sag et al.'s (2002) institutionalised phrases, which are syntactically normal and semantically compositional, but which have a fixed form. In this case, we would list an expression like *salt and pepper* in the lexicon as a MWE, but give each word the same meaning it has outside of the expression. We thus give the impression that the expression is assembled normally, even though it is actually listed as a whole. We could then appeal to a blocking rule which excludes any meaning arrived at through composition of single word expressions that is already listed in the lexicon as a pre-assembled phrase. For instance, because $[[\text{pepper and salt}]] = [[\text{salt and pepper}]]$, and because *salt and*

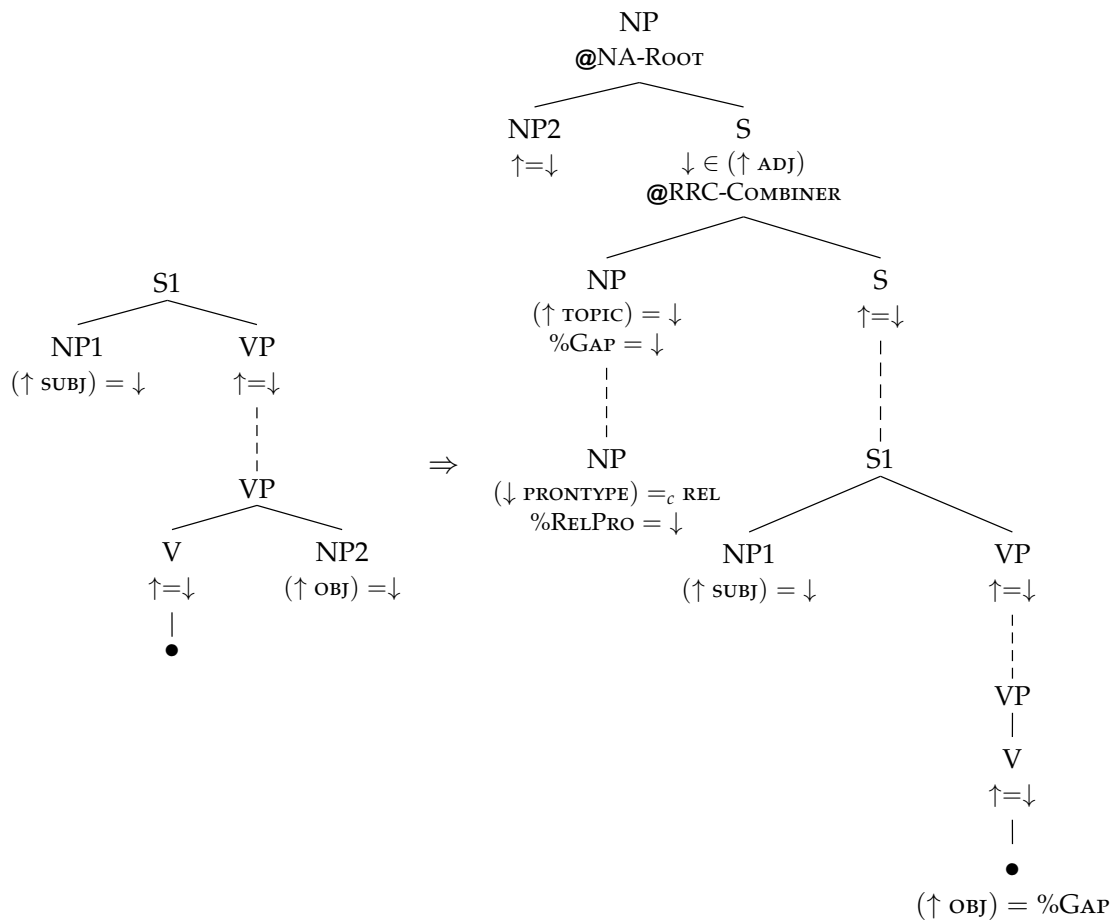


Figure 6.26: Lexical rule for an object relative clause (with a relative pronoun)

pepper is listed, *pepper and salt* is blocked.²³ This would give an account of the heavy preference for the ‘institutionalised’ version of such expressions, even though their syntax and semantics are apparently transparent. Of course, if the meaning is not strictly identical, it will not be blocked—so *black pepper and sea salt*, for instance, is correctly predicted to be possible. This approach also explains how such institutionalised phrases could come to have more idiosyncratic or restricted meanings: once they are listed separately from their parts, nothing stops the meanings diverging independently.

²³Perhaps we don’t want to completely *block* the unusual order: *pepper and salt* is, after all, perfectly grammatical. Rather, our rule might trigger some Gricean inference based on Manner (Grice 1975): the speaker is using a marked (non-canonical) form, so we should infer some additional meaning based on this.

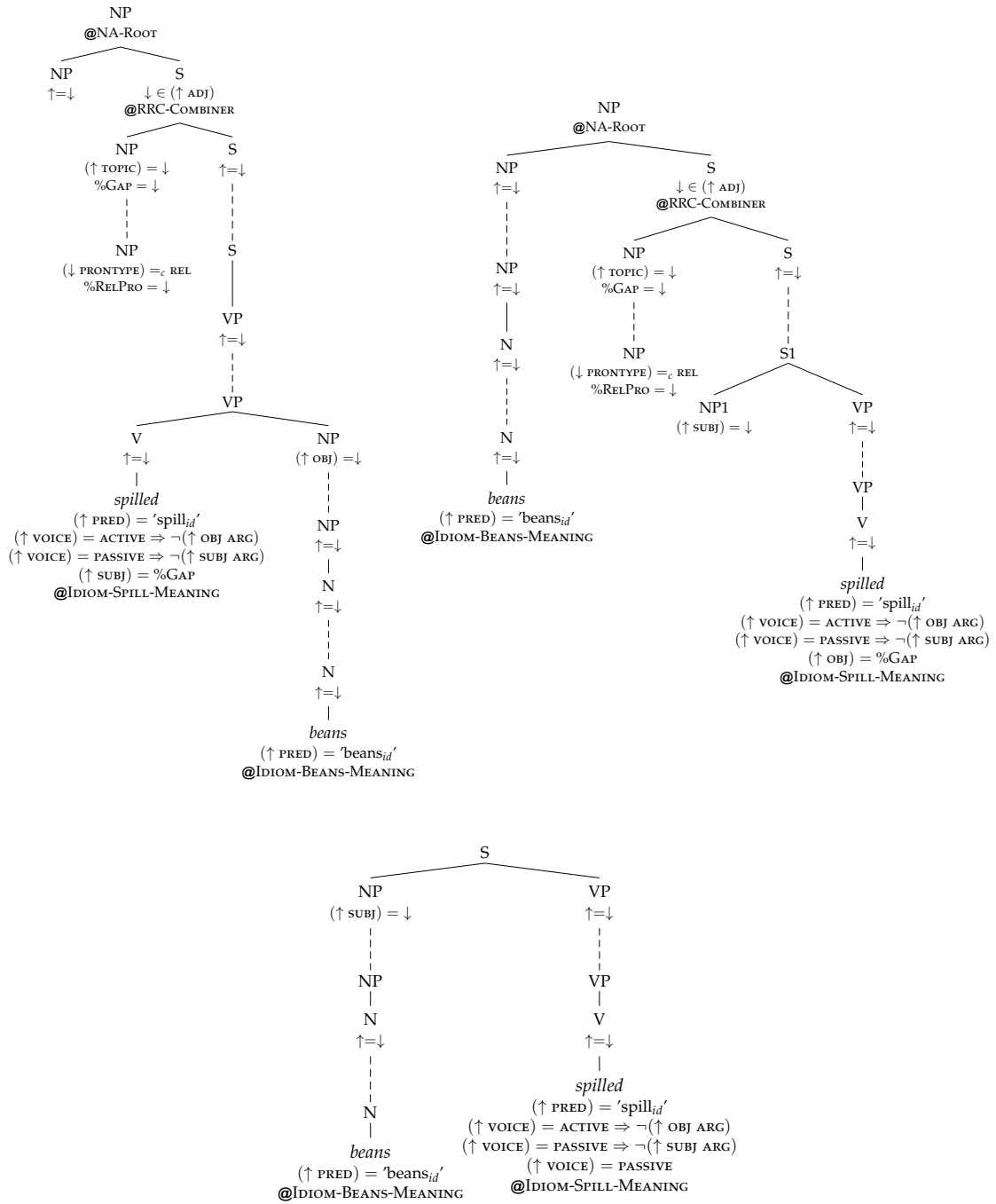


Figure 6.27: Lexical entries for a subject relative clause, object relative clause, and passive clause derived from *spill the beans*

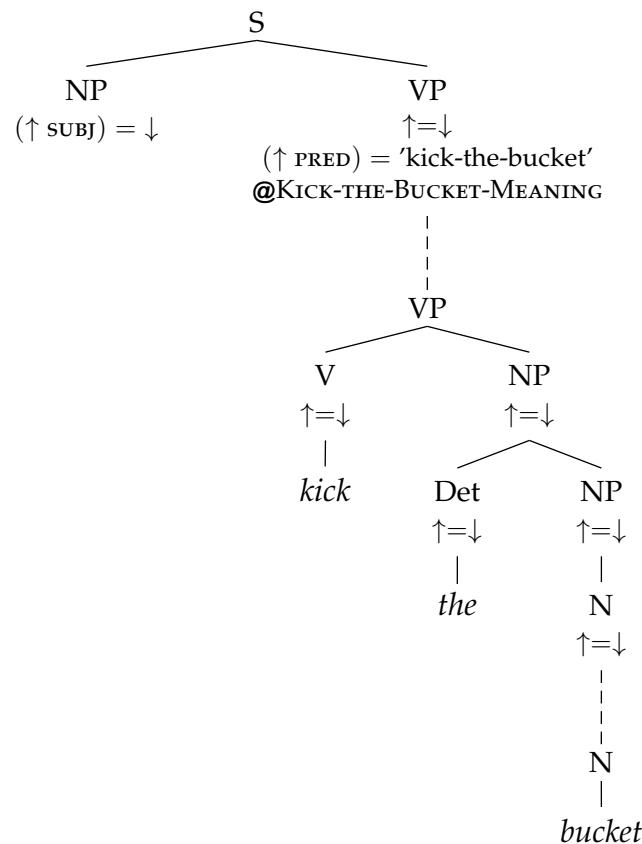


Figure 6.28: Lexical entry for *kick the bucket*

6.2.2 Non-decomposable expressions

While decomposable expressions have a lot in common with their surface-equivalent non-MWEs, non-decomposable expressions are rather different. For instance, expressions like *kick the bucket* and *shoot the breeze*, despite having the surface form of transitive VPs, semantically behave like intransitives, contributing one-place predicates meaning ‘die’ and ‘chat’. This kind of mismatch between levels of representation is something which LFG can describe very neatly: I propose, therefore, that such expressions have exactly the structure they appear to have at the level of c-structure, but contribute only a single PRED at f-structure (and, correspondingly, a single lexical meaning). *Kick the bucket*, for instance, has the lexical entry given in Figure 6.28. Instead of being annotated $(\uparrow \text{OBJ}) = \downarrow$, as it would be in a normal transitive verb, the ‘object’ NP is given the annotation $\uparrow = \downarrow$, which means that it corresponds to the same f-structure

$$\left[\begin{array}{l} \text{PRED} \quad \text{'kick-the-bucket'} \\ \text{SUBJ} \quad \left[\text{PRED} \quad \text{'Jadzia'} \right] \end{array} \right]$$

Figure 6.29: F-structure for *Jadzia kicked the bucket*

as the rest of the MWE. For instance, the sentence *Jadzia kicked the bucket* has the ‘flat’ f-structure shown in Figure 6.29.

Semantically speaking, *kick the bucket* makes the contribution we would expect from an intransitive verb. The template KICK-THE-BUCKET-MEANING contains the meaning constructors in (68):

$$(68) \quad \lambda x \lambda F. \exists e [\mathbf{kick-the-bucket}(e) \wedge \mathbf{pt}(e) = x \wedge F(e)] : \\ e_{(\uparrow \text{SUBJ})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow} \\ \lambda e. \top : v_{\uparrow} \multimap t_{\uparrow}$$

Because of its configurational position, *the bucket* retains the c-structure properties of an object. For instance, adverbs cannot intervene between it and the verb, because there is no adjunction site available:

- (69) a. Jadzia suddenly kicked the bucket.
 b. Jadzia kicked the bucket suddenly.
 c. *Jadzia kicked suddenly the bucket.

However, it lacks one of the important diagnostic properties of OBJs in LFG: that of being able to be promoted to SUBJ through passivisation. This makes it less appealing to represent *the bucket* as an f-structure OBJ, even if it does appear as a c-structure complement (*viz.*, immediately to the right of the verb).

Markantonatou & Samaridi (2018) consider the question of what GF to assign the nominal component of (non-decomposable) V+NP idioms, and conclude that a new GF ‘FIX’ is needed. They correctly observe that the passivisation facts are the crucial ones when it comes to deciding what GF to assign to the relevant NPs. For, although such NPs equally well cannot be the head of a relative clause (70a) or questioned (70b), these facts can be explained on semantic/pragmatic

grounds (a point also made by Kay et al. 2015): it makes no sense to predicate something of a noun that has no meaning, nor to apply a *wh*-operator to it.

- (70) a. #The bucket that she kicked was sudden.
 b. #Which bucket did she kick?

The passivisation facts, though, are not so easily explained. To repeat some points we have already seen: although passive subjects in English generally have some informationally prominent role, it is not true that all passive subjects are contentful, since expletives can appear here:

- (71) a. It is considered polite to remove your shoes when inside.
 b. It was hoped that this indulgence would help bring back the people to their parish churches and draw them away from the more radical dissenters and ‘mad preachers’. (COCA)

At the same time, not all contentful NPs can appear as passive subjects; some predicates resist passivisation, such as measure verbs:

- (72) a. This book weighs 12 oz.
 b. *12 oz. are weighed by this book.
- (73) a. The Norwegian coastline measures over 20,000 km. (WWW)
 b. *Over 20,000 km are measured by the Norwegian coastline.

All of this leads one to the conclusion that passivisation is controlled syntactically rather more than semantically, and so the semantic emptiness of *the bucket* is not enough to explain why *kick the bucket* cannot be passivised.

Since Markantonatou & Samaridi’s (2018) underlying theory is a version of the lexical ambiguity approach, they do not consider the possibility of describing the whole expression in a single lexical entry. Thus, while they are right to note that the best way to prevent the passivisation of expressions like *kick the bucket* is not to assign *the bucket* the GF OBJ, they fail to observe that one solution to this problem is to assign it *no* GF, rather than having to invent a new one (their FIX).

Such a move also correlates with the fact that the whole expression corresponds to a single semantic predicate—it is appropriate therefore that it should contribute a single PRED value. I believe, therefore, that the flat f-structure analysis is the proper one for such non-decomposable expressions. It also accords with the modification possibilities of such expressions, which we turn to momentarily.

Under the present analysis, both the ‘semantic’ phenomena—relativisation and *wh*-questioning—and the ‘syntactic’ phenomenon of passivisation are in fact ruled out on syntactic grounds. Specifically, since the lexical rules for passive, object relativisation, and object questioning will all specify that the verbal complement NP bears the GF OBJ, the lexical entry for *kick the bucket* fails to trigger such rules. Subject relativisation, of course, still correctly applies. In short, because of their ‘anomalous’ syntax, such expressions are excluded from the scope of the normal rules of the grammar.²⁴

This approach does raise some questions about cross-linguistic application, however. As we noted earlier, non-decomposable idioms appear to be more syntactically flexible in German, for instance. The non-decomposable idiom *den Löffel abgeben* ‘die’ (literally ‘pass on the spoon’) can passivise, for example:

- (74) Hier wurde der Löffel ab-gegeben.
 here was the spoon on-passed
 ‘Someone died here.’

This might make us more inclined to view the c-structure complement here as a genuine OBJECT, since it now passes the passivisation test. But if we do this we go against the semantic facts, since the MWE still only contributes

²⁴I claim, therefore, that the apparent exceptions to this generalisation, such as (i), are in fact ‘extended’ uses of the idioms in question (Chapter 2.2.9).

- (i) When you are dead, you don’t have to worry about death anymore. [...] The bucket will be kicked.

In their extended uses, idioms are treated as metaphors, and therefore interpreted as entirely compositional expressions. This means that they are subject to all of the normal rules of the grammar once more: *kick* as a separate lexical entry can passivise as well as any other transitive verb.

a single, unary predicate, and so we would still like it to correspond to a single, flat f-structure. I do not have a solution to this quandary at present. One possibility is that, since functional information is less bound to syntactic position in German, owing to its case system and reflected in its more flexible word order, the passive rule is also less sensitive to functional information and is instead more sensitive to c-structure position. Then what is important for allowing passivisation is that the MWE look configurationally like a transitive verb, even if functionally it is different.

A naïve version of this theory, however, would predict that German should be ‘symmetric’ with respect to double-object constructions; i.e. that for a ditransitive verb which takes two accusative complements, both should be potential passive subjects, since both will match the configurational description of following the head V inside the VP. This is not the case, however. In fact, in the double-accusative construction, illustrated in (75), *neither* ACC-marked complement can be promoted to subject using the standard passive with *werden*; in the more colloquial *kriegen*- or *bekommen*-passive, only the first one can be (Lee-Schoenfeld & Diewald 2017).

- (75) Jemand lehrt ihn den Seiltrick.
someone.NOM teaches him.ACC the.ACC rope-trick
 ‘Someone is teaching him the rope trick.’ (Lee-Schoenfeld & Diewald 2017: 1)

I regret that I do not have the time to explore this interesting puzzle further, but I am confident that a suitably well-articulated passive rule for German, which will be independently needed anyway, and therefore also independently motivated, will be able to account for the extra flexibility of non-decomposable MWEs in German. In general, it seems probable that cross-linguistic variation in MWE behaviour like this is due to higher-level variation between grammars more generally, rather than to differences in how particular types of MWE are represented. For example, between German and English, German is generally more permissive in terms of word order flexibility; it is not surprising, then,

that it is also more permissive in terms of the active-passive alternation in non-decomposable MWEs. What would instead cry out for explanation would be if it turned out that English non-decomposable MWEs were more flexible than German ones, even though, in general, German word order is more flexible.

* * *

Besides their restricted syntactic behaviour, non-decomposable expressions also exhibit special behaviour with regard to modification, to which we now turn. Firstly, we correctly predict that normal nominal modifiers cannot appear with *the bucket*, since there is no type $\langle et \rangle$ resource to consume:

(76) #Jadzia kicked the purple/dangerous/former bucket.

Simple intersective adjectives like *purple* have the modifier type $\langle et, et \rangle$, and apply to the f-structure of which they are an ADJunct:

(77) $\lambda P \lambda x. P(x) \wedge \mathbf{purple}(x) : (e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}) \multimap e_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$

However, in this case, $(\text{ADJ} \in \uparrow)$ refers not to a separate f-structure for the noun, but to the outer f-structure for the whole clause. And there is no type $\langle et \rangle$ meaning associated with this f-structure.

Of course, it is not true that no modification is possible with these MWEs. We will mention three kinds here: domain-delimiters, expressives, and meta-linguistic modifiers.

We saw in Chapter 2.2.7 that domain-delimiting ‘external’ modification is possible (Ernst 1981); some further examples are given below:

- (78) a. With that dumb remark at the party last night, I really kicked the social bucket. (Ernst 1981: 51)
- b. Thank heavens that soul draining Mr. Morgan has finally kicked the political bucket at least. (NOW)

What is the type of these adjectives? The meaning of the relevant part of (78a) can be paraphrased as (79):

(79) Socially/In the social domain, I kicked the bucket.

In other words, it is a sentential modifier, of type $\langle tt \rangle$. Note that *social* can have this meaning outside of MWEs, as in (80):²⁵

(80) He is a social pariah.
 [≈ Socially, he is a pariah.]

In this adnominal use, *social* seems to be restricted in its scope, however. Consider the difference between (81) and (82), given alongside their proposed meaning representations:

- (81) a. He is a social pariah.
 b. **socially**($\exists s[\text{pariah}(s) \wedge \text{th}(s) = \text{he}]$)
- (82) a. We spat at the social pariah. [$\not\approx$ Socially, we spat at the pariah.]
 b. $\exists e[\text{spit-at}(e) \wedge \text{ag}(e) = \text{we} \wedge \text{pt}(e) = \iota x[\text{socially}(\text{pariah}(x))]]$

In (82), unlike (81), the modification does not scope over the whole sentence, but only over the proposition expressed by the nominal, that something is a pariah. Interestingly, this corresponds to the unattested reading for *all of a terrible sudden* discussed above, given in (54a), where the modifier scopes low over a sub-expression of the whole clausal proposition. The non-existence of such a reading there was used as an argument not to give *terrible* the simple type $\langle tt \rangle$. The fact that it *is* attested in the case of *social* suggests that such domain delimiters *should* have this type. We therefore assign domain-delimiting *social* the meaning constructor in (83):²⁶

²⁵Of course, it does not always have this meaning, and there are cases where it just has a simple intersective meaning:

- (i) This is a social problem.
 [≈ This is a problem of a social character.]
 [$\not\approx$ Socially, this is a problem.]

Clearly there is more to be said about how such adjectives come to have their domain-delimiting meanings, but I will not go into detail about the lexical semantics of this particular use of *social* here.

²⁶If the domain-delimiting meaning is to be related directly to the intersective one, then the meaning must be more complicated than this. Perhaps we need a **domain** predicate, where

$$(83) \quad \lambda p.\mathbf{socially}(p) : t_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}$$

In (81a), $(\text{ADJ}\in\uparrow)$ corresponds to the outer f-structure, since the noun is being used as the main predicate. In (82a), on the other hand, $(\text{ADJ}\in\uparrow)$ refers to the nominal f-structure for *pariah*. This means that the adjective only scopes over the proposition which that noun expresses (namely that its argument is a pariah).

Now, in *Julian kicked the social bucket*, the *social* scopes over the whole clause, even though superficially the sentence resembles (82a). This follows straightforwardly from our flat analysis for the idiom, however: *bucket* makes no independent meaning contribution, and corresponds to the outer, clausal f-structure; therefore, the expression $(\text{ADJ}\in\uparrow)$ in the adjective's meaning constructor also refers to this f-structure, and *social* scopes correctly over the whole clause. Figure 6.31 shows the proof which derives the meaning of *Julian kicked the social bucket*, based on the f-structure in Figure 6.30.²⁷

$\mathbf{domain}(p, Q)$ means that p holds in the domain of Q . For instance, *Julian kicked the social bucket* might have the meaning representation in (i):

$$(i) \quad \mathbf{domain}(\exists e[\mathbf{kick-the-bucket}(e) \wedge \mathbf{pt}(e) = \mathbf{julian}], \lambda x.\mathbf{social}(x))$$

Then the combinatorial meaning constructor associated with the adjective can be given as in (ii), which combines with its lexical meaning to produce the new domain-delimiting one:

$$(ii) \quad \lambda Q\lambda p.\mathbf{domain}(p, Q) : (e_{\uparrow} \multimap t_{\uparrow}) \multimap t_{(\text{ADJ}\in\uparrow)} \multimap t_{(\text{ADJ}\in\uparrow)}$$

²⁷Note that under this analysis, the phenomenon under discussion is distinct from that of HYPALLAGE or EPITHET TRANSFER (Hall 1973; Zwicky 2007), which is a generalisation of this process to all manner of adjectives in other contexts:

- (i) a. I balanced a **thoughtful** lump of sugar on the teaspoon.
 [≈ I thoughtfully balanced... /Thoughtful, I balanced...]
 (Hall 1973: 92, quoting P. G. Wodehouse)
- b. As British political scandals go, this one is not particularly juicy. No honours seem to have been sold, no politician's Parisian hotel bills picked up, no **extramarital** toes sucked.
 [≈ no toes sucked extramaritally] (Zwicky 2007, quoting the *The Economist*)

Notice that, in non-MWE contexts, the classic domain-delimiting adjectives, such as *social*, *financial*, *political*, etc., do not scope widely when used to modify a noun in an argument position (cf. (81a) vs. (82a) above), whereas the transferred epithets in (i) do. This suggests that the phenomena are different. Impressionistically, too, they seem distinct to me: the transferred epithet examples have a much more affected and/or literary feel than the idiom examples. Perhaps this is because the former involve some kind of extra-lexical type coercion, whereas the latter just use the normal lexical entries for the component parts.

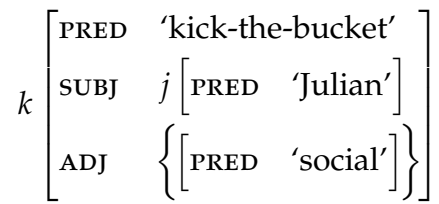


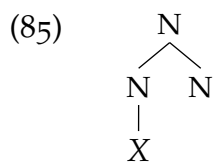
Figure 6.30: F-structure for *Julian kicked the social bucket*

If nouns are also possible domain delimiters, as in (84), then their analysis will be similar.

- (84) a. Our team is not as good as last year's, but we aren't going to drop out of the **soccer** picture.
- b. He denied that the Saudis, angry over *Death of a Princess*, were seeking some **celluloid** revenge.

(Ernst 1981: 55)

Presumably, these expressions must also include some combinatorial meaning constructor which takes a type $\langle et \rangle$ noun and turns it into a type $\langle tt \rangle$ domain delimiter (see footnote 26 for a suggestion as to how this might work). This might be represented as a constructional meaning associated with auxiliary trees of the form shown in (85), i.e. nouns which pre-modify other nouns, as the domain-delimiting meaning is only available when the noun is used as a modifier, not when it is used as a predicate or an argument.



We turn now to expressive adjectives, which can also appear inside non-decomposable idioms:

- (86) a. [Manchester] Utd have been my team for over 50 years and will be till I kick the fucking bucket. (GloWbE)
- b. Same symptoms, it detects the drive at first, then after I try to burn a CD, it kicks the bloody bucket. (WWW)

$$\begin{array}{c}
\begin{array}{c}
\llbracket \text{Julian} \rrbracket \quad \llbracket \text{kick the bucket} \rrbracket \\
\text{julian} : e_j \quad \lambda x \lambda F. \exists e [\text{kick-the-bucket}(e) \wedge \text{pt}(e) = x \wedge F(e)] : e_j \multimap (\vartheta_k \multimap t_k) \multimap t_k
\end{array} \\
\hline
\begin{array}{c}
\llbracket \text{closure} \rrbracket \\
\lambda e. \top : \vartheta_k \multimap t_k
\end{array}
\end{array}
\quad
\begin{array}{c}
\lambda F. \exists e [\text{kick-the-bucket}(e) \wedge \text{pt}(e) = \text{julian} \wedge F(e)] : \\
(\vartheta_k \multimap t_k) \multimap t_k \\
\hline
\exists e [\text{kick-the-bucket}(e) \wedge \text{pt}(e) = \text{julian}] : \\
t_k \\
\hline
\begin{array}{c}
\llbracket \text{social} \rrbracket \\
\lambda p. \text{socially}(p) : t_k \multimap t_k
\end{array}
\end{array}
\quad
\begin{array}{c}
\text{socially}(\exists e [\text{kick-the-bucket}(e) \wedge \text{pt}(e) = \text{julian}]) : t_k
\end{array}$$

Figure 6.31: Glue proof for *Julian kicked the social bucket*

Expressives include adjectives like *damn*, *fucking* and *bloody*, notable for the fact that their meaning contribution is entirely connotative: they do not affect the truth conditions of the expressions they modify (Potts 2005: Ch. 5; Potts 2007; Potts et al. 2009). Given this property, it is perhaps not surprising that they can appear inside non-decomposable idioms; it is no problem that *bucket* contributes no truth-conditional meaning of its own, because expressives do not need to operate on truth-conditional meaning (McClure 2011).

Compositionally, these pose no problems. Even in the non-MWE cases, such adjectives generally scope higher than the nouns they modify; for example, as Potts (2005: 7) points out, (87) can be loosely paraphrased by (88):

(87) I have to mow the damn lawn.

(88) I have to mow the lawn, which I hate doing.

In other words, they operate on type t (propositional) meanings (Potts 2005: Ch. 5), just like the domain delimiting adjectives. Potts's semantic system ensures that expressive meanings outscope everything else; on our analysis, an example like *kick the bloody bucket* is just the trivial case where the expressive meaning already scopes as high as it can.

Expressives are very much like evaluatives, discussed above, and just like them are often taken to contribute to some side-dimension of meaning. Nevertheless, for the sake of simplicity, we can do the same as we did for evaluatives, and treat expressives as simple $\langle tt \rangle$ predicates of propositions:

(89) $\llbracket \text{bloody} \rrbracket =$
 $\lambda p. \odot(p) : t_{(\text{ADJ} \in \uparrow)} \multimap t_{(\text{ADJ} \in \uparrow)}$

The predicate \odot stands for whatever negative evaluation we take this expressive to encode. When adjoined to idiomatic *bucket*, $(\text{ADJ} \in \uparrow)$ will refer to the clausal f -structure, just as above, and so we will obtain the correct interpretation whereby the expressive modifies the whole sentential meaning. For example:

- (90) $\llbracket \text{The computer's kicked the bloody bucket} \rrbracket =$
 $\odot(\exists e[\mathbf{kick-the-bucket}(e) \wedge \mathbf{pt}(e) = ix[\mathbf{computer}(x)]])$

One final class of adjectives which can appear inside non-decomposable idioms includes words like *proverbial* or *metaphorical*, which draw attention to the idiomaticity of the expression:

- (91) a. One day we all die, or as they say, kick the proverbial bucket. (iWeb)
 b. Margaery's position has changed quite a bit since Joff has kicked the metaphorical bucket. (WWW)

These have been described as meta-linguistic by Stathi (2007), on the basis that they say something about *bucket* as a linguistic object, rather than about buckets in the actual world. In this case, they lie outside of the scope of our analysis: 'words' do not exist in the ontology of our semantic system, only the meaning constructors which they contribute, and so meta-linguistic modification is not something it is equipped to describe. However, one thing we can say is that whatever the appropriate analysis of meta-linguistic modification turns out to be, it is not a problem if the words such adjectives adjoin to make no independent meaning contribution, since this is not what the adjective meanings operate on. All these adjectives need is a c-structure host, and that is what *bucket* provides.

6.2.3 Verbal MWEs: a summary

In this section, I have proposed that decomposable verbal MWEs are represented as 'pre-assembled' phrases: that is, VP idioms have the syntax of normal VPs, and contribute the same kinds of semantic units—verbal meaning constructors of type $\langle e_1, \dots, e_n, \langle vt, t \rangle \rangle$ and as many argument meaning constructors as are included in the idiomatic interpretation of the MWE. Owing to their regular syntax and semantics, they also offer up the same kinds of modification possibilities: for instance, adjectives can adjoin to nominal arguments and modify their meanings directly. Similarly, such decomposable verbal MWEs

are apt to match the left-hand side of various lexical rules, thereby participating in the normal syntactic alternations which the grammar makes available. In general, this is in keeping with Nunberg et al.'s (1994) contention that semantic decomposability correlates with syntactic flexibility. If there is a need for more fine-grained control, this can easily be achieved via lexically encoded restrictions. We also noted that the same kind of representation might be appropriate for 'institutionalised phrases', those expressions which are syntactically and semantically regular, but which exhibit a statistical idiomaticity: they represent the one and only (normal) way of expressing that particular meaning.

Non-decomposable verbal MWEs, on the other hand, have slightly less transparent representations. Their phrase structure is just what it seems, but their f-structure is not: a V+NP sequence, for instance, does not correspond to a verb and its object, but rather to a single, flat f-structure, which in turn corresponds to the simple, one-place predicate which it contributes to the semantics. This means that they are excluded from the scope of many lexical rules which make explicit reference to grammatical functions, and so are correspondingly less syntactically flexible. In terms of modification possibilities, such expressions allow adjectives to adjoin to their apparent object NPs because of their normal c-structure, but because of their flat f-structure, all but domain-delimiting, expressive, or meta-linguistic adjectives are rendered semantically inappropriate.

* * *

So far, we have seen the strengths of the Lexicalised LFG approach: the flexibility afforded by the potential for mismatches between c-structure and f-structure means that we can explain the differences between the surface syntax of MWEs and their semantic behaviour (e.g. the difference between decomposable and non-decomposable expressions); the grammatical regularity of decomposable MWEs explains why they are more likely to be syntactically flexible (since they look like normal objects in the grammar and so are subject to the same lexical rules); and by treating MWEs as richly detailed lexical entries we have enough fine-grained control to encode expression-by-expression

idiosyncrasies and constraints. Relatedly, we have also succeeded in localising any peculiarities associated with MWEs to the MWEs themselves, leaving the rest of the grammar untouched. However, despite these successes, there remain some challenges. In the next section, we look at one particularly perplexing phenomenon, that of lexical flexibility.

6.3 Lexical flexibility

One defining characteristic of most MWEs is that they must contain certain *words*—it is precisely the form of the words rather than their (literal) content which gives rise to the idiomatic meaning. Thus we cannot say *knock over the legumes* for *spill the beans* or *boot the pail* for *kick the bucket*. MWEs need not dictate the form of *all* the words inside them; we have seen above examples like *spill (the) beans* where the determiner is left open so that other determiners, or indeed no determiner at all, can appear there. There are also cases like *fly in (the) ointment*, where something must appear in the specifier of *ointment*, to make it definite, but it could be a possessive phrase instead of the determiner *the*. So we have MWEs which specify all of the words they contain (like *kick the bucket* or *by and large*), and ones which leave certain slots underspecified (either thereby making them optional or at most requiring that they contribute some particular feature). As we noted in Chapter 3.5.1.5, however, there is a third class, which permit variability in one of the words they contain, but neither in such an unconstrained fashion as *spill (the) beans*, nor in such a narrowly syntactically definable way as *fly in (the) ointment*. These are exemplified in (92):

- (92) a. It's time to put/place/lay/... our cards on the table.
 b. That gave me a kick up the backside/rear/bum/booty/....
 c. This adds/gives/brings/... grist to the mill.

In each of these examples, the variable words share certain semantic features (e.g. all of the nouns in (92b) refer to the backside), and we noted above that this is one area in which the semantic mapping approach shines, therefore.

Unfortunately, our construction-based theory does not do so well here: this is not the sort of constraint that is easily statable in the LFG description language—we cannot write equations which refer to the contents of meaning constructors, for example. What is more, unlike the optional/variable determiners discussed above, these words do not always contribute their normal meaning. The expressions in (92b) mean, roughly, ‘motivate someone (to stop being lazy)’, and this meaning does not mention backsides. This is unlike *spill the/some/a few beans*, where the different determiners contribute their usual meanings.

In this section, I discuss how we can handle this kind of lexical flexibility in the Lexicalised LFG framework. We begin with *grist to the mill*, since this is actually of a rather different kind to the others, where the variable verb *does* make its usual meaning contribution. We then turn to the other kinds of examples, and consider two different approaches to solving the problem that they pose.

6.3.1 *Grist to the mill*

The expression *grist to/for the mill* means, roughly, ‘support/fuel for the (relevant) enterprise/endeavour’, where *grist* corresponds to the fuel and *mill* to the endeavour. It can sometimes appear as a non-verbal MWE, a simple NP which can be used as a predicate itself or as an argument:

- (93) a. It would be grist to the mill of the tabloids if anything so bizarre happened next week. (BNC)
- b. [...] and we hear from that monster of objectivity inside every artist who sees everything as grist to the mill. (BNC)

This use would have an analysis along the lines of *fly in the ointment*, discussed in Section 6.1 above. Figure 6.32 shows how this might look, with the meaning templates spelled out below:²⁸

²⁸I am assuming that the meaning associated with *to/for* in the MWE is just the same as the meaning of regular *for*. In the text, I present an analysis of the version of the expression which has *to*, but the analysis of the *for* version will be identical except for a change of terminal node.

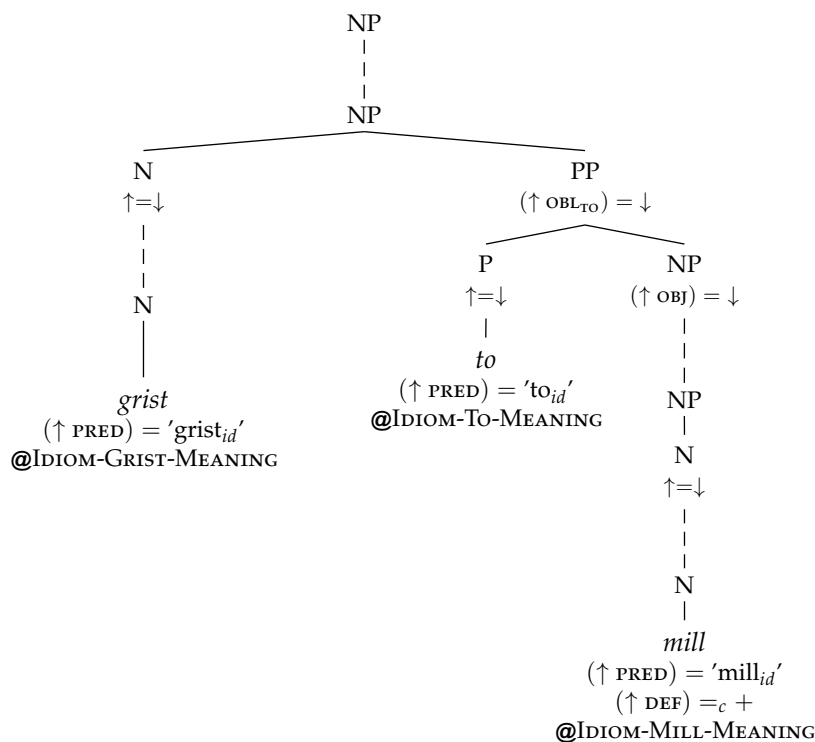


Figure 6.32: Lexical entry for *grist to the mill*

- (94) a. **IDIOM-GRIST-MEANING** :=
 $\lambda x.\mathbf{grist-id}(x) : e_{\uparrow} \multimap t_{\uparrow}$
- b. **IDIOM-TO-MEANING** :=
 $\lambda y \lambda P \lambda x.P(x) \wedge \mathbf{for}(x, y) :$
 $e_{(\uparrow \text{OBJ})} \multimap (e_{(\text{OBL}_{\text{TO}} \uparrow)} \multimap t_{(\text{OBL}_{\text{TO}} \uparrow)}) \multimap e_{(\text{OBL}_{\text{TO}} \uparrow)} \multimap t_{(\text{OBL}_{\text{TO}} \uparrow)}$
- c. **IDIOM-MILL-MEANING** :=
 $\lambda x.\mathbf{mill-id}(x) : e_{\uparrow} \multimap t_{\uparrow}$

However, this expression is also often found in a verbal configuration, where the NP *grist* and the PP *to the mill* are treated as two of the arguments of a ditransitive verb:

- (95) a. That the town has become New Zealand's best bicycle wine-touring destination is no surprise, while the Remutaka Cycle Trail has added grist to the mill. (NOW)
- b. This adds grist to the mill of those who claim that the industry has been operating in a far from satisfactory manner. (WWW)
- c. If this column has fixated a little too often on how sport it is only

- because every day brings fresh grist to the mill. (NOW)
- d. So aren't you afraid that your criticism of Islam brings grist to the mill of right-wing Islamophobia? (iWeb)
- e. The worst mass shooting in modern US history has given grist to the mill of Donald Trump. (WWW)

Notice firstly that this expression often appears with an *of*-phrase following *mill* which makes explicit precisely which endeavour or project is being furthered. I won't discuss this further here, but it presumably adheres to the same syntax and semantics as these *of*-complements do for any other noun (there must be a disjunction in the description of the MWE which allows for this argument slot to be available or not).

What is important for our purposes is that the verb, which can differ, nonetheless carries its usual meaning in the examples in (95). These sentences involve actually adding, bringing, or giving support to an enterprise. We therefore have a relatively straightforward analysis open to us: we give this version of *grist to the mill* the lexical entry in Figure 6.33, which has a sentential frame but with a missing verbal head. This will successfully unify with a verbal lexical entry which has both an NP OBJ and a PP OBL, and is compatible with a PP headed by *to*, like the lexical entry for *gave*, for instance, which is shown in Figure 6.34. In this instance, we do not provide an idiomatic meaning for *to*, since in verbs like *give*, it serves merely as a marker of thematic role—the semantics is all provided by the verb, which has the meaning in (96):

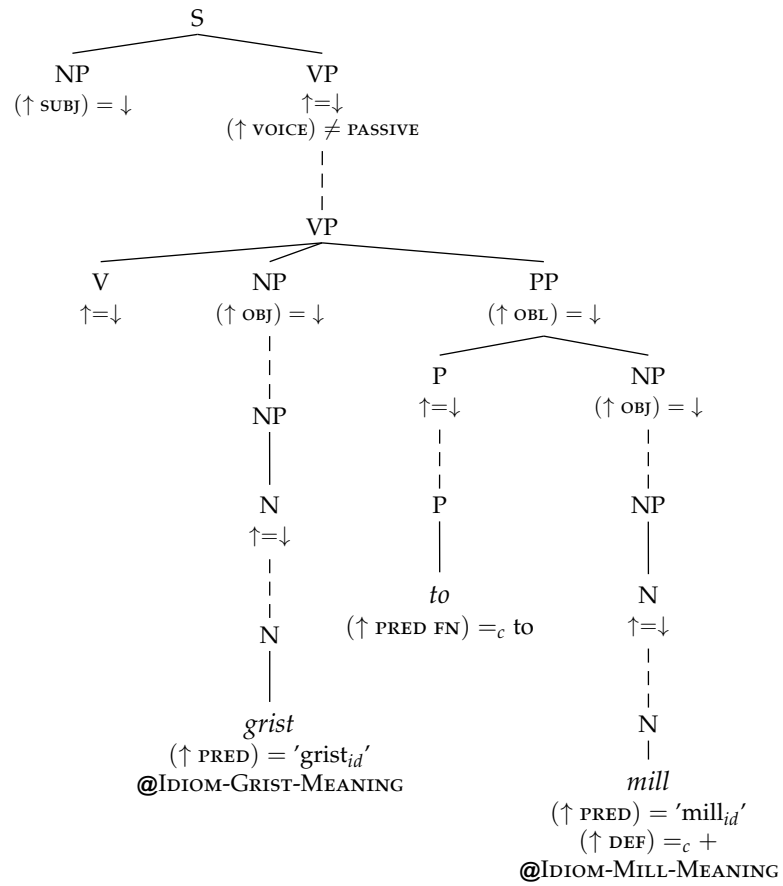


Figure 6.33: Lexical entry for verbal *grist to the mill*

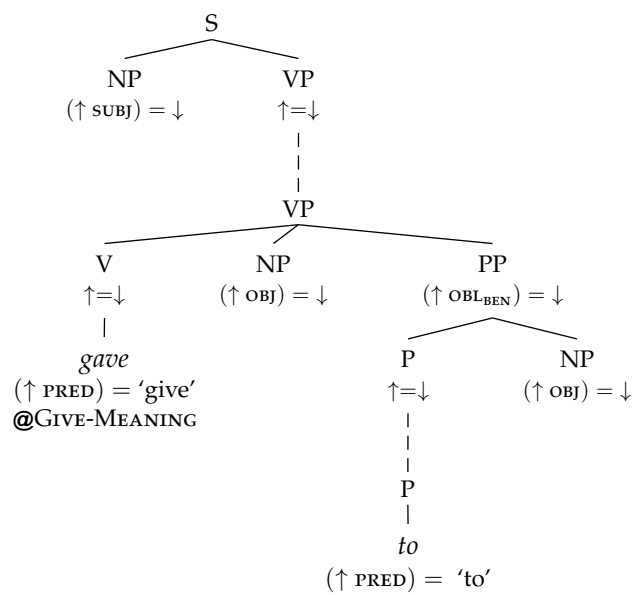


Figure 6.34: Lexical entry for *gave*

$$(96) \quad \lambda x \lambda y \lambda z \lambda F. \exists e [\mathbf{give}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{th}(e) = y \wedge \mathbf{ben}(e) = z \wedge F(e)] : \\ e_{(\uparrow_{\text{SUBJ}})} \multimap e_{(\uparrow_{\text{OBJ}})} \multimap e_{(\uparrow_{\text{OBL}_{\text{BEN}} \text{ OBJ}})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$$

This consumes the meaning of the prepositional object directly, and assigns it to the beneficiary role of the verbal event.

Note that the OBL function is underspecified in the lexical entry for verbal *grist to the mill*; we can treat bare OBL as a disjunction over all possible OBL functions, thematic and idiosyncratic:

$$(97) \quad \text{OBL} \equiv \{ \text{OBL}_{\text{GOAL}} | \text{OBL}_{\text{LOC}} | \dots | \text{OBL}_{\text{TO}} | \text{OBL}_{\text{ON}} | \dots \}$$

We need this underspecification because other compatible verbs bring a slightly different meaning. For instance, the oblique argument of *add (to)* might be thought to be a goal rather than a beneficiary. This is entirely compatible with our analysis, provided *add (to)* has a lexical meaning constructor like *give*, where the $\text{OBL}_{\text{GOAL}} \text{ OBJ}$ is consumed directly and assigned the goal role:

$$(98) \quad \lambda x \lambda y \lambda z \lambda F. \exists e [\mathbf{add}(e) \wedge \mathbf{ag}(e) = x \wedge \mathbf{th}(e) = y \wedge \mathbf{goal}(e) = z \wedge F(e)] : \\ e_{(\uparrow_{\text{SUBJ}})} \multimap e_{(\uparrow_{\text{OBJ}})} \multimap e_{(\uparrow_{\text{OBL}_{\text{GOAL}} \text{ OBJ}})} \multimap (v_{\uparrow} \multimap t_{\uparrow}) \multimap t_{\uparrow}$$

Verbs like *put* are not compatible with *grist to the mill*, because they require a semantic contribution from the preposition. *Put* takes two complements, an NP theme and a PP location.²⁹ Whereas *give* really expresses a relation between three entities—the agent, theme, and beneficiary, represented by the object of *to*—*put* expresses a relation between two entities and a property—the agent, theme, and location, which is not an entity like *the table* but rather a property like *on the table*. Since the verbal lexical entry for *grist to the mill* does not include a meaning for the preposition, it is therefore incompatible with the meaning of *put*. The only way it could be made compatible is if we also included the semantically contentful version of *to*. This would successfully unify, but does not provide the right kind of meaning for *put* to consume, as we can see from a non-MWE example like (99):

²⁹*Put* can also take an adverbial location argument like *there*, but this version of the verb is not relevant here.

(99) *Kira put the reports to the desk.

And if we were to use any other preposition, the constraining equation on the preposition in *grist to the mill*'s description would not be satisfied. Thus, we correctly rule out the use of any verbs other than those which take two complements of the right type, and which contribute the correct preposition.³⁰

6.3.2 Other kinds of lexical flexibility

The *grist to the mill* example is relatively straightforward because the variable word contributes its usual meaning. In many respects, then, it is just like the cases with optional determiners we discussed earlier, except that it is rarer to specify multiple arguments of a verb without also specifying the verb itself.³¹ Far more troubling for the present approach, however, are cases where the lexical variation appears to once again be constrained to a particular semantic family, but where the word in question does *not* contribute its usual meaning. Examples include (100), mentioned above:

- (100) a. It's time to put/place/lay/... our cards on the table.
 b. That gave me a kick up the backside/rear/bum/booty/....

³⁰Some verbs, like *provide*, are compatible with both versions of the MWE, with *to* and with *for*:

- (i) a. However, ruling the supernatural out of science by fiat actually provides grist to the mill of anti-evolutionism. (WWW)
 b. Material production provides grist for the mill of economic analysis. (WWW)

I take it that there are simply two versions of the MWE, and two versions of *provide* (one which selects a PP headed by *to* and one which selects a PP headed by *for*), with whatever subtle distinction of meaning exists between the latter carrying over to expressions which contain the MWE. Note that the other version of *provide*, *provide X with Y* is incompatible on the assumption that there is no version of the MWE which has the PP headed by *with*. This configuration would give a nonsense meaning with the MWE anyway, since the semantic roles of the two arguments are switched: X is the goal/beneficiary and Y is the theme, so the meaning of *provide grist with the mill* would be 'provide fuel with the endeavour'. (But even with the arguments in the 'correct' order, *provide the mill with grist* also lacks the idiomatic interpretation—this once again highlights that this is a question of incompatible form more than incompatible meaning.)

³¹Other MWEs which do this include *out of the frying pan into the fire*, and *cat... out of the bag*, as mentioned earlier.

We could perhaps make a case that (100a) does involve the normal meaning of *put*, albeit in some extended sense, if we take it to mean something like ‘put our information in the open’. However, this does seem less obviously acceptable for *place* or *lay*, which are much more concrete. More problematic is (100b), where it is clear that the idiomatic meaning, approximately ‘motivated me (to stop being lazy)’, does not involve anything resembling a bottom.

Let us consider the latter class in more detail, therefore. How can we handle such data? Firstly, we could abandon the ‘direct encoding’ approach which we have developed in the form of Lexicalised LFG and opt for a semantic mapping approach to idiomatic meaning instead. Examples like (100b) are very good *prima facie* evidence for such an approach, after all: it is the *meaning* which is important, and not the form of words used. Certainly, even granting that the criticisms I levelled against this approach in Chapter 3 are justified, we cannot rule out the possibility that it might nonetheless be the correct way to treat certain kinds of expression: perhaps there are some MWEs which do operate this way, and we simply have to have different analyses for different MWEs. It does not seem intuitively as if the kinds of expressions in (100) are so radically different from the other kinds of MWEs we have been discussing, though, and so this strategy should be a last resort. Here I would like to explore two other alternatives that do not abandon the direct encoding approach.

One option is to simply assume there are multiple versions of the MWE listed in the lexicon, as would seem prudent for cases of minor lexical variation like the preposition in *grist to/for the mill*, for example. These need not even reside in different lexical entries: if the only variation is in the form of terminal node, we can simply have a disjunctive constraint in the description. However, this makes the semantic relationship between the various possible words entirely coincidental, which is less than satisfying: what the different words which can fill the gap in *kick up the* — have in common is precisely that they all mean ‘backside’.

At the same time, there are cases where such semantic links can be misleading. Consider (101):

(101) They threw me to the wolves/dogs/lions.

It seems at first glance that this is just like the *kick up the backside* example, and any dangerous animal will do for the object of *to*. However, this is not the case:

(102) #They threw me to the coyotes/foxes/tigers/leopards.

Even other wild canines or big cats aren't acceptable. If there is a semantic criterion at work here, it is a peculiarly specific one. It seems instead that the potential nouns just have to be listed in the lexicon.

The question, then, is whether *kick up the backside* is like this as well: is the apparent semantic link merely illusory? If it is, and there is a closed class of words which can appear here, then we can safely list them. If not, any word can appear here provided it means 'backside'. Call these the extensional and intensional approaches, respectively, since in the former case the relevant constraint takes the form "word w can appear here if it is in the set $\{w_1, w_2, w_3, \dots\}$ ", whereas in the latter it takes the form "word w can appear here if it has property P ". How could we tease these approaches apart? For the synchronic grammar, it would be impossible: there are a finite number of words in any speaker's lexicon, and so there are certainly a finite number of words meaning 'backside'; we could therefore quite readily list them.³² One place in which the two approaches make different predictions, though, is with regard to new coinages. According to the extensional approach, any new words for backside should *not* automatically be able to appear in this expression, since they do not appear on the pre-specified

³²Presumably Gricean strictures against excessive prolixity would save us having to worry about the potentially infinite class of expressions like (i):

(i) They gave me a kick up the portion of my body directly above my legs but below my waist.

These might, though, be acceptable as witticisms or other kinds of 'extended' use, as discussed in Chapter 2.2.9.

list. On the other hand, if all that matters is meaning, then the intensional approach predicts that they could do.

I have not been able to determine the age of *kick up the backside*, so I cannot come to any firm conclusions, but as I pointed out in Chapter 3.5.1.5, even relatively new coinages like *booty* and *tushy*, for which the OED gives a first citation of 1959 and 1962 respectively, are attested:

- (103) a. Even instructors sometimes need a kick up the booty to reassess, re-motivate and make things fresh again. (WWW)
 b. That provided the kick up the bootie I needed and I found an extra gear. (WWW)
- (104) a. Introducing the emo-tinged Aussies giving indie rock a feminist kick up the tushie. (WWW)
 b. Thank you so much for being the kick up the tushy that I needed. (WWW)

This strongly suggests that what we need is precisely a constraint which requires a particular meaning. Unfortunately, such a constraint is not directly statable in LFG: we cannot refer to the contents of the meaning language side of meaning constructors in our description language. If we want to emulate something like this, we would need to make reference to a feature or a collection of features at some level of representation—perhaps s-structure. Some such features are generally taken to be part of the theory: linguistic phenomena which pay attention to whether something is animate, or human, for instance, are relatively common, and so these motivate corresponding binary-valued ANIMATE and HUMAN s-structure features (see e.g. Dalrymple & Nikolaeva 2011). But it seems much less plausible that there will be a cross-linguistically relevant feature \pm BACKSIDE or similar.

However, all is not lost. Ash Asudeh (pers. comm.) has suggested to me a way in which we can simulate the effect of semantic selection by using the flexibility of Glue composition. First of all, we give *kick up the backside* the lexical

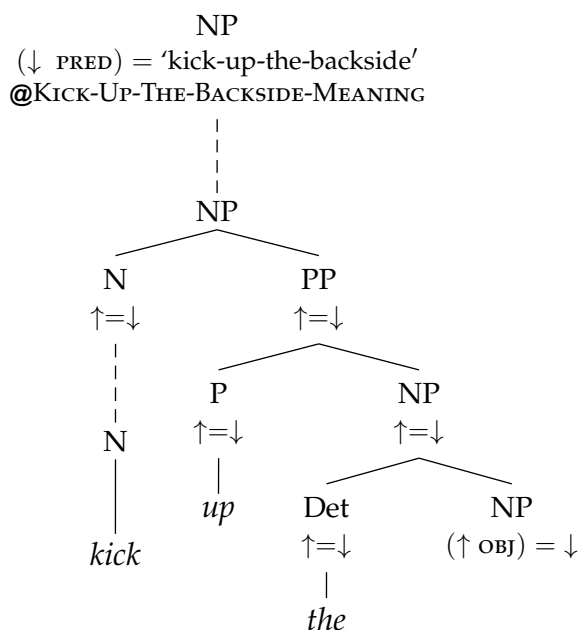


Figure 6.35: Lexical entry for *kick up the backside* under the intensional approach

entry in Figure 6.35; this includes an open OBJ position for the ‘backside’ word. The contents of **KICK-UP-THE-BACKSIDE-MEANING** is the following:

$$(105) \quad \lambda P \lambda x. \mathbf{kick-up-the-backside}(x) \wedge P = \mathbf{backside} : \\ (e_{(\downarrow \text{OBJ})} \multimap t_{(\downarrow \text{OBJ})}) \multimap e_{\downarrow} \multimap t_{\downarrow}$$

This provides the MWE meaning as well as consuming the object, the flexible slot in the MWE, and equating its meaning with the meaning **backside**. Assuming that **tushy**, **booty**, **arse**, etc. are coextensive with **backside**, this will enable any of them to fill the OBJ slot. Anything that does not mean ‘backside’ will not satisfy the equality statement.

However, there are two reasons we might not like this approach. Firstly, it suggests that e.g. *#kick up the tortoise* is infelicitous because it is *false* rather than because it is ill-formed—after all, it can compose perfectly well; it’s just that the conjunct **tortoise** = **backside** in its meaning will come out false. I do not find this terribly satisfying, since it seems to me that what we want to say about *kick up the tortoise* is that *it is not an instance of the MWE in question*, not that it is an instance of it which always contributes a false meaning.

For one thing, under the present approach sentences like (106) will be given the wrong interpretation:

(106) Kim didn't give me a kick up the tortoise.

The meaning of such a sentence will include (107):

(107) $\neg[\dots \mathbf{kick-up-the-backside}(x) \wedge \mathbf{tortoise} = \mathbf{backside} \dots]$

But (107) entails (108):

(108) $\dots \neg \mathbf{kick-up-the-backside}(x) \vee \neg(\mathbf{tortoise} = \mathbf{backside}) \dots$

That is, the negation targets one or more sub-parts of what it asserted. But if **tortoise = backside** is part of what is asserted, then one way for the sentence to be true is for all the other conjuncts in (107) to be true and for this one to be false. But of course in the actual world **tortoise = backside** is always false, and so (106) should be able to mean that Kim gave me a kick up the backside. In fact, it should have the same meaning as (109):

(109) Kim gave me a kick up the backside and tortoises are not backsides.

This is not what we want.

One solution to help us avoid this outcome would be to make the equality between the meaning of the object and the meaning of **backside** a presupposition, rather than an assertion (e.g. by making use of Beaver's (2001) presupposition operator). Then an expression containing *kick up the tortoise* is infelicitous in the same way an expression containing a non-referring definite description is. And since presuppositions project past negation, (106) would correctly come out as infelicitous, at least under the intended reading.

This is an improvement, but I must admit I still have qualms. It seems to me that *the present King of France*, for instance, is well formed in a way that *kick up the tortoise* is not; that is, the latter *does not count as an instance of the MWE in question*, whereas the former is purely compositional and so is of impeccable formal standing. I concede, however, that this might not amount to much in practical

terms: an expression which makes any statement it is used in infelicitous is, to all intents and purposes, ill-formed.

The second objection to this kind of approach is that it requires an articulated f-structure, as in (110):

$$(110) \quad \left[\begin{array}{l} \text{PRED} \quad \text{'kick-up-the-backside'} \\ \text{OBJ} \quad \left[\text{PRED} \quad \text{'backside'} \right] \end{array} \right]$$

But *kick up the backside* is non-decomposable; the only modifiers which appear on *backside* are meta-linguistic, expressive, and domain-delimiting, as for *kick the bucket*:

- (111) a. He proceeded to give Hamburg a kick up the proverbial backside, and the city loved him for it. (WWW)
- b. 'Vault' is the biggest kick up the metaphorical backside drum & bass has ever had. (WWW)
- (112) a. Any fit healthy man that's OK with sitting on his arse and living of another bloke's taxes needs a kick up the bloody backside. (WWW)
- b. Well if that wasn't a kick up the fucking backside then I don't know what else will be. (WWW)
- (113) a. This has been the biggest enabler to give me a good kick up the spiritual backside. (WWW)
[= to encourage me in the spiritual domain]
- b. Six years ago Guy Ritchie gave Conan Doyle's Sherlock Holmes a comprehensive kick up the creative backside. (WWW)
[= creatively speaking, gave it a kick up the backside]

If we follow the same logic as before, then, we should prefer to give *kick up the backside* a flat representation at f-structure:

$$(114) \quad \left[\text{PRED} \quad \text{'kick-up-the-backside'} \right]$$

An articulated f-structure is inevitable in any version of the intensional approach, however: since individuable meanings require separate f-structures, if we want to refer to the meaning of the ‘backside’ word separately, it had better have its own f-structure.

To sum up, then, of the three ways we might handle true lexical flexibility, one, resorting to the semantic mapping approach, radically complicates both our theory and our typology of MWEs (we have to decide where the idiom mappings required by that approach are stored, and we have to find a way of working out which MWEs are analysed via lexical encoding and which are computed via a mapping—how lexically flexible do they have to be before they are put in the latter camp?); another, the extensional approach, is empirically deficient (it fails to predict that new coinages will be acceptable in the relevant MWEs); and the third, the intensional approach, gives the best coverage, but may not be wholly theoretically satisfying. The first option is clearly the more theoretically fraught, but neither of the extensional or intensional options is really very appealing either: one does not describe the facts correctly, and the other does so in a theoretically suspect manner.

6.3.3 Summary

Lexical flexibility is not something naturally expected under the kind of theory proposed in this thesis. MWEs, just like every other lexical item, are pairings of form and meaning, with both parts fully specified, and so we do not expect the form of an expression to vary any more than its meaning. However, it is clear that such lexically flexible expressions do exist, and so we face a challenge. Some kinds of variability are readily incorporated into the framework, namely where there is an empty slot in the MWE which is filled by a normal word which brings with it its normal meaning—examples include specifiers for nominal MWE components like *spill* ___ *beans* or *fly in* ___ *ointment*, and the verbal part of *grist to the mill* (and perhaps *V one’s cards on the table*). So, too, when the variation is arbitrarily restricted, as in *throw someone to the wolves/dogs/lions*:

arbitrary variation must be lexically encoded on anyone's theory, and the present framework can quite readily describe a disjunction of possible terminal nodes.

What is problematic, though, is when the variation appears to be restricted by meaning alone. In cases like *kick up the backside*, it seems that any noun will do as long as it means the same thing as *backside*, and this is not easy to model. We can treat this case as equivalent to the previous one, simply listing a disjunction of possibilities, but this seems like the wrong generalisation, assuming that newly coined words with the correct meaning are also permitted in the MWE. If we want to truly capture the fact that it is the meaning which is held constant, and we do not want to abandon the direct encoding approach, then we are forced into a theoretically awkward position. Because LFG lacks any *grammatical* means of imposing extreme semantic selectional restrictions, we are forced to do so in the semantics itself. But this means that the illicit examples are claimed to be cases of presupposition failure, or simply untrue, rather than ill-formed, which strikes me as the wrong conclusion to reach. Another alternative is that some expressions really can't be explained along direct encoding lines, and must instead be analysed using a theory whereby literal meanings are computed and then mapped to idiomatic ones. We saw good reasons to reject such a theory as a general approach to idiomatic meaning in Chapter 3, but that is not to say that it has no role to play. Perhaps in the typology of MWEs there are some which are more metaphorical than others, and therefore more amenable to a mapping analysis. However, this requires a dramatic complexification of our theory, and also poses a host of new questions about how we divide MWEs between the two approaches.

6.4 Chapter summary

In this chapter, we have seen how Lexicalised LFG can handle a variety of different MWEs. The framework has a number of attractive features that make it well suited to this task: most prominently, the separation between c-structure and f-structure. This enables us to capture the fact that non-verbal MWEs

like *all of a sudden* might have apparently quite complex surface syntax, even though functionally they correspond to simple, atomic units. This possibility for mismatches between c- and f-structure also enables us to neatly capture the fact that decomposable and non-decomposable MWEs may appear similar on the surface, but behave quite differently in terms of semantics and syntactic flexibility. For example, *kick the bucket* and *spill the beans* both have the same phrase structure, but look quite different at f-structure, accounting for their different modification and extraction possibilities. The separation of levels also allows us to explain many of the apparently unusual modification possibilities of some non-verbal MWEs, since the apparent mismatch between form and meaning only applies at c-structure: when we interpret the semantics at f-structure, many of the meanings follow quite naturally. The same goes for non-decomposable verbal MWEs, with domain-delimiting adjectives which take the whole clause in their scope nonetheless appearing inside the VP at c-structure.

Our analysis also gives some formal explanation of the claim made by Nunberg et al. (1994) that syntactic flexibility is a reflex of decomposability (with the caveats mentioned in Chapter 2.2.7 that the full picture is more complicated than this would suggest). Syntactically decomposable expressions like *spill the beans* are more flexible because they have a more 'normal' syntax, and so are able to trigger the various lexical rules in the grammar. But this syntactic normality is in part due to their decomposability: the 'object' NP *the beans* is also a genuine f-structure OBJ, as it would be in the case of a normal transitive verb, but this is possible precisely because it has its own meaning, motivating a more articulated f-structure for the MWE. This is unlike a non-decomposable idiom such as *kick the bucket*, which, because of its unitary semantics, has a flat f-structure, and so treats *the bucket* as a c-structure complement but not an f-structure OBJ. It is this disconnection between c-structure and f-structure which prevents *kick the bucket* from triggering the relevant lexical rules, and this mismatch is motivated by its non-decomposability. Thus, decomposability leads to an articulated f-structure, which leads to a more normal syntax and

so a greater degree of syntactic flexibility. Non-decomposability leads to a flat f-structure, which means a greater separation between c- and f-structure and so a reduced syntactic flexibility.

On the other hand, our theory, like all theories which do not fall under the semantic mapping umbrella, fails to accommodate lexical flexibility very naturally. With no easy way to refer to meanings in lexical entries, we are left with either an extensional approach which simply lists possible lexical variants, or an intensional approach which claims that the constraints on lexical variation are to do with felicity/truthhood rather than well-formedness. An alternative solution would be to separate MWEs into those which are understood via the approach outlined in this chapter, and those which are understood via some variety of the semantic mapping approach. This raises as many questions as it answers, since we now must decide how to divide up the space of MWEs, and exactly how to formally implement a semantic mapping theory. It is to be hoped that future research might determine whether there are independent differences between lexically flexible and lexically inflexible MWEs (perhaps in terms of psycholinguistic properties) which might provide independent motivation for such a move.

7

Conclusion

Contents

7.1 Summary of the thesis	339
7.2 Assessing the proposal	345
7.3 Future research	347

This chapter brings the thesis to a close. I begin in Section 7.1 by reviewing the main claims which have been made in the previous six chapters. Then, in Section 7.2, I return to the properties of MWEs we sought to explain, and show how the approach developed in the thesis addresses them. Finally, in Section 7.3, I suggest directions for future research and further expansions of the proposals made here.

7.1 Summary of the thesis

This thesis has explored a number of topics relating to MWEs, their proper analysis, and what a theory of grammar should look like more generally. The central claims are as follows:

- MWEs are of interest to linguistic theory because they do not fit neatly into the word-phrase dichotomy: they appear to be made up of multiple

words, like phrases, but their form and meaning can be unpredictable, in which case they must be learned, like words. (*Chapter 1*)

- MWEs possess a number of properties which make them analytically challenging:
 1. **semantic idiomatcity**, the aforementioned unpredictable meaning;
 2. **syntactic idiomatcity**, the fact that their syntax may not match that of other, normally composed expressions;
 3. **lexical idiomatcity**, the fact that they may contain 'words' which have no independent existence outside of the MWE;
 4. the fact that they are, however, still overwhelmingly made up of words that *do* exist outside of the expression (the **real words problem**);
 5. **morphosyntactic flexibility**, the fact that their sub-parts can inflect individually;
 6. **syntactic flexibility**, the fact that their sub-parts can be reordered and separated by syntactic processes;
 7. **decomposability**, the fact that their meanings can sometimes be distributed across their sub-parts;
 8. the fact that they seem to be **psycholinguistically** different from compositional expressions, being processed faster; and
 9. the possibility of there being two different ways to interpret some idiomatic expressions, including an **extended** sense which treats them more like metaphors.

The first three of these make MWEs look word-like (i.e. stored), since the lexicon is the repository of idiosyncrasies. The next four make MWEs look phrase-like, since they must be structurally complex and flexible. The psycholinguistic findings show that idiomatic expressions are at least *different* from compositional ones, and the extended uses add a whole other dimension of complexity. (*Chapter 2*)

- Any theory which hopes to explain these properties must take seriously the tension which MWEs exhibit between their word-like and phrase-like properties. Unfortunately, many existing theories fail to do this. The space of analytical possibilities is relatively limited, and in fact there are essentially four different kinds of theory we might entertain:
 1. A **monolexic** approach, which treats MWEs as big words. This does well at describing the first three properties, but fails to account for the ways in which MWEs behave like phrases.
 2. A **lexical ambiguity** approach, which treats MWEs as conspiracies of individual words which carry (parts of) the overall meaning separately. This explains MWEs' flexibility, but fails to account for idiosyncrasy at the phrasal level (i.e. syntactic idiomaticity), and does not predict any difference in psycholinguistic behaviour between MWEs and regular compositional phrases.
 3. A **semantic mapping** approach, which first composes the words making up the MWE literally, and then computes the idiomatic meaning from the literal one. This holds some promise for explaining the extended uses of idioms, as it is in many respects a theory of metaphor. It also strongly predicts that MWEs will be made up of real words. However, this approach requires there to be a literal parse available for the MWE string, and so fails to account for lexical and syntactic idiomaticity. It also does not offer any way to control the syntactic flexibility of MWEs, and predicts they should be just as flexible as their literal counterparts, contrary to fact. It does predict a psycholinguistic difference between MWEs and non-idiomatic compositional expressions, but in the wrong direction: MWEs should be slower, not faster.
 4. A **construction-based** approach, which treats MWEs as stored phrases with internal structure. This achieves a good coverage, although it

fails to find a solution to the real words problem or offer any account of the extended uses.

The first approach treats MWEs as wholly word-like (i.e. as atomic units), and the next two treat them as wholly phrase-like (i.e. as composed of smaller independent units). It is only the fourth which takes the tension inherent in MWEs seriously, recognising that they are units in their own right, but also have internal structure. This is why it achieves the best coverage of the facts. (*Chapter 3*)

- The theory that I propose incorporates two syntactic frameworks, Lexical Functional Grammar (LFG) and Tree-Adjoining Grammar (TAG). LFG is a constraint-based, declarative theory which takes a modular view of the grammar. Syntax is divided into two levels, c-structure and f-structure, which represent, respectively, surface phenomena like constituency and more abstract, functional phenomena like grammatical relations. It is computationally and psychologically appealing, and its modularity makes it very flexible. However, it is based on a context-free grammar (CFG), and so does not have a natural means of describing MWEs, since CFGs lack the necessary extended domain of locality. TAGs, by contrast, do have this property, and this, coupled with the operation of adjunction, makes TAG a natural framework in which to implement the construction-based approach mentioned previously. (*Chapter 4*)
- It is possible to combine the two theories once we view TAG as a constraint-based formalism operating on tree descriptions rather than as a derivational formalism operating on trees. To achieve this, we introduce the concept of quasi-trees, which have their topology underspecified by making use of dominance instead of immediate dominance relations in their descriptions. With this in place, it is possible to include such descriptions in LFG lexical entries. This means we can reduce such lexical entries to nothing but descriptions, containing information about all of the different

levels of linguistic representation, including c-structure. This affords LFG lexical entries the extended domain of locality needed to describe MWEs, and also make for a more elegant theory overall. (*Chapter 5*)

- Such a theory also enables us to lexicalise LFG. Lexicalised grammars are those where each elementary structure in the grammar is associated with a lexical item. Such grammars strongly instantiate the lexicalist position that the lexicon holds a central place in the grammar, and their analyses are also straightforward to translate into dependency structures, enabling connections to other theories. In standard LFG, the elementary structures are CFG phrase-structure rules, not all of which are associated with a lexical item, and so the grammar is not lexicalised. Now that lexical entries can contain extended c-structure information, however, we can replace the phrase-structure rules wholesale, and take lexical entries (descriptions) to be the elementary structures instead. This lexicalises the grammar, and so I call this version of the theory Lexicalised LFG. (*Chapter 5*)
- With these tools in hand, we can give a satisfying account of many kinds of MWEs. Non-verbal MWEs like *by and large* or *all of a sudden* are associated with an articulated c-structure but a flat f-structure. This explains the possibility of adjoining adjectives to their sub-parts, e.g. *all of a complete sudden*, which nonetheless semantically scope over the whole expression, or even over some component of the clause containing it. The contrast between decomposable and non-decomposable MWEs corresponds to the presence or absence of multiple meaning constructors in the lexical entry for the MWE: if multiple meanings are present, they can be modified separately; if they are not, they cannot, and any modification will affect the whole expression. This also often corresponds to a mismatch between c-structure and f-structure, especially in verbal MWEs: non-decomposable expressions like *kick the bucket* and decomposable ones like *spill the beans* have identical c-structures, but very different f-structures, since the complex meaning

corresponds to an articulated f-structure, while the unitary meaning corresponds to a flat f-structure. This also explains the differing modification possibilities: modifiers which adjoin to *beans* can modify its $\langle et \rangle$ meaning directly, and so regular intersective adjectives are permitted here; modifiers which attach to *bucket*, on the other hand, can only modify the expression as a whole, since *bucket* corresponds to the same f-structure as the other parts of the MWE—this means that regular intersective adjectives are impossible here, but domain-delimiters, expressives, and meta-linguistic adjectives are all permitted. (*Chapter 6*)

- Finally, we saw that this kind of theory is not so well placed to handle lexical flexibility, which is the property whereby certain MWEs permit some of their words to vary, often according to semantic criteria. There are three responses to such a phenomenon: firstly, we might decide that we should abandon any ‘direct encoding’ theory for these expressions, and instead opt for the semantic mapping approach despite its flaws as a general theory of MWEs; secondly, we might still encode the meaning in a lexical entry, but simply list the various possible options for the flexible position; thirdly, we might require the presence of a particular class of word by specifying a particular meaning. The first of these poses a number of unappealing theoretical questions. For example, where will the semantic mappings live in the grammar? (They don’t seem like lexical entries, but what are they?) How do we decide which MWEs should be analysed according to which theory? The second option is easy to implement, but theoretically inelegant, and empirically deficient, if it is the case that new coinages which have the correct meaning can appear in such expressions (cf. *kick up the tushy*). The third option is not easily implementable in LFG, since there is no direct way of imposing constraints on meaning. We can use Glue semantics to match the meaning of the flexible word against a pre-defined standard, but this requires a more articulated f-structure than we might like, and predicts that the problem with the incorrect forms

is that they cause presupposition failures or are false rather than being ill-formed. None of these solutions is wholly satisfying. (*Chapter 6*)

7.2 Assessing the proposal

How well does the theory developed in this thesis fare against the criteria first introduced in Chapter 2? It is an instance of the construction-based approach, and so for the most part inherits that approach's strengths and flaws. Semantic idiomaticity, for example, is easily described because MWEs are stored in the lexicon, which is precisely the repository of arbitrary form-meaning pairings.

Syntactic idiomaticity can be modelled straightforwardly because MWEs are also *units* in the lexicon; since they are stored as a whole, along with their syntactic frame, it is no challenge to encode any syntactic idiosyncrasies directly. We have the added advantage in LFG of being able to represent the fact that some expressions may have unusual surface syntax but still remain functionally similar to single-word expressions: for instance, *all of a sudden* has a complex and idiosyncratic c-structure representation, but corresponds to a single, flat f-structure, just like an adverb such as *suddenly*.

Morphosyntactic flexibility has not really been addressed directly in this thesis. However, each of the terminal nodes in a MWE has exactly the same formal status as a terminal node in a single-word expression's lexical entry in Lexicalised LFG, and so whatever morphological theory we adopt should be able to cover the MWE cases without the need for any undue modification.

Decomposability is simply modelled by the inclusion of multiple meaning constructors in a lexical entry, and these sub-parts can be the target of modification, quantification, etc., as usual.

Syntactic flexibility is handled by lexical rules: MWEs which match the left-hand side of a lexical rule are predicted to have the corresponding derived syntactic structure as well. What is more, this flexibility follows in large part from decomposability, in keeping with Nunberg et al.'s (1994) proposal. When verbal idioms are decomposable, their c-structure annotations look more like

those of non-MWE verbs in the grammar, because they have a more articulated f-structure, just like compositional expressions; this means that they match more lexical rules, and so participate in more syntactic alternations. Non-decomposable idioms, on the other hand, have more 'defective' c- to f-structure mappings, resulting in forms which do not match the left-hand side of so many lexical rules, explaining their relative inflexibility. Of course, further idiosyncratic constraints on flexibility may also be required, and these can be achieved by adding constraints to the lexical entries of MWEs which make them incompatible with the relevant lexical rules.

The psycholinguistic findings receive the same explanation as they do in general for any construction-based approach: since MWEs are stored as units, they can be retrieved in a single step, rather than having to be combined online by the parser. This makes MWEs faster to process than the parallel compositional expressions, accounting for the decreased reading/reaction times reported in the experimental literature.

Because MWEs are stored in the lexicon as units and not composed of other parts, the theory proposed here has no real explanation for the real words problem. There is no reason why the terminal nodes in MWEs should be words which have an independent existence: their form is entirely arbitrarily encoded in each lexical entry. However, as I argued in Chapter 3.7.4, it in fact seems likely that the real words problem is not a phenomenon for which we should seek a synchronic explanation. Rather, the reason MWEs have the forms they do is because of the historical processes which led to their creation. In this case, failure to explain the real words problem synchronically may be more of a feature than a bug—semantic mapping theories which take the real words problem seriously, for example, are singularly unable to explain the existence of lexical idiomaticity.

The one area that the theory proposed here has nothing to say about is the extended uses of idioms. But as discussed in Chapter 3.5.3, only a semantic mapping approach has any chance of explaining this phenomenon. This thesis has presented a theory of the core uses, which are directly encoded in the

lexicon. It is to be hoped that future research will elaborate the connection between the two, but it does not seem likely that a single kind of theory will explain both. Rather, we need a theory of how the two very different interpretation techniques are related.

Thus, the Lexicalised LFG approach described in this thesis accurately accounts for all of the properties of MWEs which we identified in Chapter 2. The only areas where it is deficient are in handling the real words problem and the extended uses of idioms. However, the real words problem probably does not merit a synchronic explanation, and we need a theory of the core uses before we can worry about a theory of the extended uses—that is precisely what this thesis provides.

7.3 Future research

Nonetheless, there remain a number of important topics which I have left unaddressed. Constraints of time and space have prevented me from giving an account of many other kinds of MWE, such as light verb constructions or prepositional verbs, which bring with them their own unique challenges. Light verb constructions (LVCs), for example, are semi-productive in a way which idioms are not; their meanings are often inferrable from the meaning of their nominal component:

- (1) a. to take a bath → to bathe
 b. to take a walk → to walk
 c. to have a conversation → to converse

This makes it seem unappealing to store them in the lexicon, since their meanings are not idiosyncratic. We might prefer instead to have alternative versions of the light verbs, equipped with meaning constructors which turn the nouns into verbal predicates. But there are two problems with treating LVCs as derived compositionally rather than stored. Firstly, sometimes the meaning is not compositional in any straightforward sense:

- (2) to take a breather ↗ to breathe

To take a breather does not exactly mean *to breathe*, even if there is some metaphorical sense in which they may be related. Secondly, even when the meaning is predictable, the choice of light verb is not:

- (3) a. to have a bath/walk/conversation
 b. to take a bath/walk/*conversation

It is hard to see what the semantic motivation for the exclusion of *conversation* in the second series might be here, and if this is really just a matter of lexical idiosyncrasy then it must be listed somewhere—so we are back to storing light verb constructions in the lexicon, in spite of their apparent compositionality.

In addition to expanding the coverage of MWEs, it would be fruitful to do more cross-linguistic research. We already noted in Chapter 2.2.7 that non-decomposable verbs in German seem to be more flexible than their equivalents in English. However, my analysis in Chapter 6.2.2 claimed that the inflexibility of non-decomposable MWEs was a reflex of their non-decomposability. If this flexibility can vary cross-linguistically, then the relationship between flexibility and decomposability cannot be so direct. I suggested there that the passive rule in German may be more permissive than in English. This clearly bears investigating, and data from additional languages could also be used to determine in what ways languages can vary in this respect.

Certainly, more needs to be said about the extended uses of idioms. How exactly are the mappings between literal and figurative situations to be encoded in the grammar, and how do we know when to interpret idioms via this method and not the direct access strategy? These are the sorts of questions that my colleagues and I hope to soon provide answers to (Findlay et al. in preparation).

Finally, there are a number of implications of the Lexicalised LFG framework itself which should be explored. For example, by adopting the TAG approach to long-distance dependencies, we do away with one requirement for functional uncertainty in LFG. The question then arises: is it possible to find alternative

theories of e.g. binding, and thereby do away with it all together? Secondly, what is the role of mapping theory (in the sense of Lexical Mapping Theory: e.g. Bresnan & Kanerva 1989) in Lexicalised LFG? Since verbs encode their argument structures directly in their lexical entries and voice alternations are handled by lexical rules, it might not be necessary in the same way. However, here we have only considered the very simple case of the English passive; lexical rules may not be enough to capture argument alternations perspicuously in languages with more complex voice systems than English. [Removed references to Stefan Müller's work, since the kinds of 'phrasal' argument structures he argues against are not actually like the proposals in this thesis.] Lastly, one of the parade examples of the theoretical utility of LFG is its ability to handle non-configurational languages like Warlpiri with relative ease (Bresnan et al. 2016 devote their first chapter to the issue of non-configurationality and its role in motivating the LFG architecture, for example). However, TAG, being resolutely configurational, is not very well suited to this challenge. It would be an added strength of Lexicalised LFG if it could be shown that incorporating a TAG into LFG does not disrupt its ability to describe non-configurational languages. I believe that the unification-based approach to TAG adopted in this thesis means that this should be possible, but the demonstration must await future work.

Appendices



Sources for attested examples

Corpora

iWeb: The Intelligent Web-based corpus (Davies 2018–).

GloWbE: Corpus of Global Web-based English (Davies 2013a).

NOW: News on the Web (Davies 2013b)

COCA: The Corpus of Contemporary American English (Davies 2008–).

BNC: British National Corpus (Aston & Burnard 1998; Davies 2004–).

Chapter 2

(10a): <http://daldale-rhod.blogspot.com/2012/11/>

(10b): song lyrics, from 'Gimme the prize', by Queen (<https://youtu.be/vf4UhPuwoGE>)

(19b): *Lust: under the billionaire's control*, an e-book by Megan Harold; direct link to relevant page: <http://bit.ly/2y8gIqF>

(20): Stavans (2001: 54)

(21): adapted from a review highlighted on the front page of <http://www.crockettlawfirm.com/practice-areas/>

(30): <https://www.facebook.com/SciFiontheRock/photos/looks-like-the-beans-have/987038601478227/>

(33a): <https://www.colleenhartleycreative.com/blog/2018/1/12/word-of-the-year-2018>

(35a): <https://www.telegraph.co.uk/news/6179497/Clarissa-Dickson-Wright-They-dont-call-me-Krakatoa-for-nothing.html>

(35b): <https://kentucky.forums.rivals.com/threads/ul-players-suing-ncaa.268835/>

(35c): adapted from <https://nJORogejustus.wordpress.com/2012/07/03/before-you-say-i-love-you/>

(36a): <http://watchersonthewall.com/major-kings-landing-season-6-spoilers-jeremy-podeswa-and-pilou-asbaek-interviewed/>

(36b): adapted from <https://www.chronofhorse.com/forum/forum/discussion-forums/off-course/338005-animal-communicator>

(36c): <https://uproxx.com/hitfix/vin-diesel-implies-marvel-inhumans-connection-on-facebook/>

(39a): <http://nerdbastards.com/2012/12/07/nerd-art-dump-terminator-the-sith-batman-does-the-wash-galactus-the-smoker-moar/>

(39b): <http://www.refinery29.com/2015/06/88395/chad-michael-murray-baby-announcement-twitter>

(39c): <https://www.indiebound.org/book/9780544176676>

(40a): <http://tt.tickettannoy.com/tt-news/qa-slamboree-soundsystem/>

(40b): <http://emotivalounge.proboards.com/thread/13916/xda-1-add-on-product?page=1>

(40c): *Shearer wonderland: Duncan Shearer: The Autobiography*, an e-book by Duncan Shearer; direct link to relevant page: <http://bit.ly/2y56ibi>

(41b): adapted from <https://variety.com/2010/dirt/real-estalker/hilary-duff-lists-the-duff-family-digs-1201230736/>

(41c): adapted from https://www.huffingtonpost.com/em-and-lo/7-tips-for-getting-over-a-sexual-hang-up_b_9481224.html?guccounter=1

(43a): <https://www.veooz.com/news/jHUG7jp.html>

(44a): <http://www.huffingtonpost.com/steve-robertson/the-bell-of-liberty->

part-1_b_3661083.html

(52b): www.sevenstoreys.com/ChristmasLetter2012.docx

(52c): <https://thetruthaboutplas.com/2013/04/02/told-you-so-project-labor-agreement-on-connecticut-high-school-increased-costs-and-reduced-competition/>

(52d): song lyrics, from 'Miss missing you' by Fall Out Boy (<https://youtu.be/0ecLVSWGcCM>).

Chapter 3

(4a): adapted from <https://www.truckersforum.net/threads/best-big-company.78417/>

(4b): <https://www.omorashi.org/forums/topic/21534-the-case-of-the-litter-box/>

(4c): <http://www.peaktestosterone.com/forum/index.php?topic=3716.0>

(5a): <http://www.thepictame.com/share/BayDCRBHc2F>

(5b): <https://www.fanfiction.net/s/10160894/2/YGOAS-Drabbles>

(15): TV script, *Person of Interest*, s2e19 (https://www.springfieldspringfield.co.uk/view_episode_scripts.php?tv-show=person-of-interest&episode=s02e19)

(30a): adapted from <https://www.linkedin.com/pulse/who-your-inspiration-andrea-springthorpe>

(30b): <http://www.racheljoyce.org/kansas-70-3-2/>

(31a): <https://www.timeout.com/london/music/whats-the-deal-with-camp-cope>

(31b): https://www.reddit.com/r/C25K/comments/2jt09v/my_success_story_how_i_went_from_couch_to_half/

(32a): *The Evening Star*, Washington, D.C., Tues. 9 July, 1907, 15/3; accessed at <https://chroniclingamerica.loc.gov/lccn/sn83045462/1907-07-09/ed-1/seq-15.pdf>

(42b): a Brexit-related headline from satirical website *The Daily Mash* (<https://www.thedailymash.co.uk/politics/politics-headlines/shit-finally-completes-29-month-journey-towards-fan-20181114179403>)

(42c): <https://www.olbg.com/forum/viewtopic.php?f=6&t=29262&start=36>

- (42d): <https://www.stuff.co.nz/national/politics/80309478/Usual-scaremongers-get-on-Maori-rights-bandwagon>
- (43b): www.sevenstoreys.com/ChristmasLetter2012.docx
- (43c): <https://thetruthaboutplas.com/2013/04/02/told-you-so-project-labor-agreement-on-connecticut-high-school-increased-costs-and-reduced-competition/>
- (43d): song lyrics, from ‘Miss missing you’ by Fall Out Boy (<https://youtu.be/0ecLVSWGcCM>).

Chapter 6

- (2b): from s8e12 of *The Goon Show*; script available here: http://www.thegoonshow.net/scripts_show.asp?title=s08e12_the_great_british_revolution
- fn. 5 (ia): Godoy et al. (1998: 57)
- fn. 5 (ib): song lyrics, from ‘Minotaur’, by Clutch (<https://youtu.be/qYTXNyZrqYo>)
- (10b): Bogarde (1992: 41)
- (10c): <http://seasidedonkey.co.uk/give-a-gift-twice/>
- (56b): <http://www.ibtimes.com.au/xiaomi-mi-5-pricing-leaked-video-surface-render-reveal-pink-gold-variants-1496076>
- (73a): <https://www.export.gov/apex/article2?id=Norway-shipping-and-maritime-equipment-and-services>
- (86b): <https://www.beatzo.net/blog/2005/04/>
- (91b): <https://www.bustle.com/articles/22113-game-of-thrones-natalie-dormer-is-as-polite-as-margaery-so-go-ahead-touch-her-hair>
- (95b): <http://www.acas.org.uk/index.aspx?articleid=4893>
- (95e): <https://economictimes.indiatimes.com/blogs/et-commentary/the-worst-mass-shooting-in-modern-us-history-has-given-grist-to-the-mill-of-donald-trump/>
- fn. 30 (ia): adapted from *Issues in education by subject, profession, and vocation: 2013 edition*, an e-book edited by Q. Ashton Acton; direct link to relevant page: <https://books.google.co.uk/books?id=MoI6Pfz5MG4C&pg=PA1021&lpg=PA1021>
- fn. 30 (ib): adapted from *Corporate pharaohs: a vicious circle of globalization*, a

self-published book by Richard and June Brinkman; direct link to relevant page:

https://books.google.co.uk/books?id=HM2_jd7MbnEC&pg=PA67&lpg=PA67

(103a): adapted from <https://www.linkedin.com/pulse/who-your-inspiration-andrea-springthorpe>

(103b): <http://www.racheljoyce.org/kansas-70-3-2/>

(104a): <https://www.timeout.com/london/music/whats-the-deal-with-camp-cope>

(104b): https://www.reddit.com/r/C25K/comments/2jt09v/my_success_story_how_i_went_from_couch_to_half/

(111a): https://www.theguardian.com/friday_review/story/0,,408027,00.html

(111b): <https://www.bassrush.com/5-tunes-that-changed-our-lives-mind-vortex/>

(112a): adapted from <https://forums.spacebattles.com/threads/a-serious-question-on-depression-anxiety-etc-and-the-workforce.293218/>

(112b): <http://therepublikofmancunia.com/confirmed-team-live-blog-wolves-vs-united/comment-page-13/>

(113a): <https://www.out-of-body-encounters.co.uk/about/>

(113b): <https://www.dailymail.co.uk/home/event/article-3344855/Agyness-Deyn-Sunset-Song-Stunning-Sunset-shame-script-s-not-song.html>

B

Semantic types

(1) **Definition of types:**

1. e, t, v, s , and ε are types.
2. If σ and τ are types, then $\langle \sigma, \tau \rangle$ is a type.
3. Nothing else is a type.

(2) **Domains of types:**

1. The domain D_e of e is the set of entities.
2. The domain D_t of t is the set of propositions, i.e. $\mathcal{P}(W)$, the power set of the set of worlds.
3. The domain D_v of v is the set of events.
4. The domain D_s of s is the set of states.
5. The domain D_ε of ε is the set of eventualities, $D_v \cup D_s$.
6. The domain $D_{\langle \sigma, \tau \rangle}$ of a functional type $\langle \sigma, \tau \rangle$ is the set of all functions from D_σ into D_τ .

For consistency, I will use the same variable labels for variables of particular types, according to the list below:

(3) **Variable labels:**

1. x, y, z are variables of type e .
2. e, e', e'', \dots are variables of type v .
3. s, s', s'', \dots are variables of type s .
4. $\varepsilon, \varepsilon', \varepsilon'', \dots$ are variables of type ε .
5. p, q are variables of type t .
6. X, Y, Z are variables of derived simple types.
7. P, Q, R are variables of type $\langle e_1, \dots, e_n, t \rangle$.
8. F, G are variables of type $\langle vt \rangle$.
9. \mathcal{V} is a variable of type $\langle vt, t \rangle$.
10. \mathcal{Q} is a variable of type $\langle et, t \rangle$.
11. A, B, C are variables of other constructed types.

Bibliography

- Abeillé, Anne. 1988. Parsing French with Tree Adjoining Grammar: some linguistic accounts. In *Proceedings of the 12th conference on computational linguistics*, 7–12. Budapest, HU.
- Abeillé, Anne. 1990. Lexical and syntactic rules in a Tree Adjoining Grammar. In *Proceedings of the 28th annual meeting of the Association for Computational Linguistics*, 292–298. Stroudsburg, PA: Association for Computational Linguistics.
- Abeillé, Anne. 1995. The flexibility of French idioms: a representation with Lexicalized Tree Adjoining Grammar. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: structural and psychological perspectives*, Hove, UK: Lawrence Erlbaum.
- Abeillé, Anne & Owen Rambow. 2000a. Tree Adjoining Grammar: an overview. In Anne Abeillé & Owen Rambow (eds.), *Tree Adjoining Grammars: formalisms, linguistic analysis and processing*, 1–68. Stanford, CA: CSLI Publications.
- Abeillé, Anne & Owen Rambow (eds.). 2000b. *Tree Adjoining Grammars: formalisms, linguistic analysis and processing*. Stanford, CA: CSLI Publications.
- Ackerman, Farrell, Gregory T. Stump & Gert Webelhuth. 2011. Lexicalism, periphrasis, and implicative morphology. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar*, 325–358. Oxford, UK: Wiley-Blackwell.
- Andrews, Avery. 2010. ‘Grammatical’ vs. ‘lexical’ meaning constructors for Glue Semantics. In Yvonne Treis & Rik De Busser (eds.), *Conference of the Australian Linguistic Society 2009*, Melbourne, AU: Australian Linguistic Society. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.602.7907&rep=rep1&type=pdf>.
- Andrews, Avery D. 2008. The role of PRED in LFG+Glue. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, 46–67. Stanford, CA: CSLI Publications. <http://www.stanford.edu/group/cslipublications/cslipublications/LFG/13/papers/lfg08andrews.pdf>.
- Arnold, Doug. 2015. A Glue Semantics for structurally regular MWEs. Poster presented at the PARSEME 5th general meeting, 23–24th September 2015, Iași, Romania.
- Arnold, Doug & Louisa Sadler. 2010. Pottsian LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG10 Conference*, 43–63. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/15/papers/lfg10arnoldsadler.pdf>.

- Aston, Guy & Lou Burnard. 1998. *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.
- Asudeh, Ash. 2004. Resumption as resource management. Doctoral dissertation, Stanford University.
- Asudeh, Ash. 2005. Control and semantic resource sensitivity. *Journal of Linguistics* 41(3). 465–511. <https://doi.org/10.1017/S0022226705003427>.
- Asudeh, Ash. 2012. *The logic of pronominal resumption*. Oxford, UK: Oxford University Press.
- Asudeh, Ash & Richard Crouch. 2001. Glue semantics: a general theory of meaning composition. Talk given at Stanford Semantics Fest 2, March 16 2001.
- Asudeh, Ash & Richard Crouch. 2002. Glue semantics for HPSG. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 1–19. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2001/Ash-Crouch-pn.pdf>.
- Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1(1). 1–54. <http://jlm.ipipan.waw.pl/index.php/JLM/article/view/56/49>.
- Asudeh, Ash & Gianluca Giorgolo. 2012. Flexible composition for optional and derived arguments. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG12 Conference*, 64–84. Stanford, CA: CSLI Publications. <http://www.stanford.edu/group/cslipublications/cslipublications/LFG/17/papers/lfg12asudehgiorgolo.pdf>.
- Asudeh, Ash & Gianluca Giorgolo. 2016. Perspectives. *Semantics & Pragmatics* 9. Article 21. <http://semprag.org/article/view/sp.9.21>.
- Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG14 Conference*, 68–88. CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/papers/lfg14asudehetal.pdf>.
- Asudeh, Ash & Ida Toivonen. 2006. Symptomatic imperfections. *Journal of Linguistics* 42(2). 395–422. <https://doi.org/10.1017/S0022226706003963>.
- Asudeh, Ash & Ida Toivonen. 2012. Copy raising and perception. *Natural Language and Linguistic Theory* 30(2). 321–380. <https://doi.org/10.1007/s11049-012-9168-2>.
- Austin, Peter & Simon Musgrave (eds.). 2008. *Voice and grammatical relations in Austronesian languages*. Stanford, CA: CSLI Publications.
- Bach, Emmon. 1981. On time, tense and aspect: an essay on English metaphysics. In Peter Cole (ed.), *Radical pragmatics*, 62–81. New York, NY: Academic Press.
- Bach, Emmon. 1989. *Informal lectures on formal semantics*. Albany, NY: State University of New York Press.

- Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–374. Amsterdam, NL: John Benjamins.
- Baldwin, Timothy & Su Nam Kim. 2010. Multiword Expressions. In Nitin Indurkha & Fred J. Damerau (eds.), *Handbook of Natural Language Processing* (2nd edn.), 267–292. Boca Raton, FL: CRC Press.
- Bargmann, Sascha. 2017. How frozen are frozen idioms? Talk given at the Recent Trends in Linguistic Research Colloquium, University of Frankfurt, 14 November 2017.
- Bargmann, Sascha, Berit Gehrke & Frank Richter. 2018. How to modify idioms. Talk given at the Workshop on One-to-Many Relations in Morphology, Syntax and Semantics at the 40th Annual Meeting of the Deutsche Gesellschaft für Sprachwissenschaft (DGfS 2018), Universität Stuttgart, 8 March 2017. [link to slides].
- Bargmann, Sascha & Manfred Sailer. 2018. The syntactic flexibility of semantically non-decomposable idioms. In Manfred Sailer & Stella Markantonatou (eds.), *Multiword expressions: insights from a multi-lingual perspective*, 1–29. Berlin, DE: Language Science Press. <http://langsci-press.org/catalog/book/184>.
- Bauer, Laurie. 1983. *English word-formation*. Cambridge, UK: Cambridge University Press.
- Beaver, David I. 2001. *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.
- Beesley, Kenneth R. 1998. Arabic morphology using only finite-state operations. In *Proceedings of the workshop on computational approaches to Semitic languages*, Association for Computational Linguistics. <http://www.aclweb.org/anthology/W98-1007>.
- van Benthem, Johan. 1988. Strategies of intensionalization. In István Bodnár, András Máté & László Pólos (eds.), *Intensional logic, history of philosophy, and methodology: to Imre Ruzsa on the occasion of his 65th birthday*, 41–59. Budapest, HU: Department of Symbolic Logic, Eötvös University.
- van Benthem, Johan. 1991. *Language in action: categories, lambdas, and dynamic logic*. Amsterdam, NL: Elsevier Science.
- Berwick, Robert C. 1982. Computational complexity and Lexical-Functional Grammar. *American Journal of Computational Linguistics* 8((3–4)). 97–109. <https://dl.acm.org/citation.cfm?id=981926>.
- Bhatia, Archna, Claire Bonial, Marie Candito, Fabienne Cap, Silvio Cordeiro, Vassiliki Foufi, Polona Gantar, Voula Giouli, Carlos Herrero, Uxoá Iñurrieta, Mihaela Ionescu, Alfredo Maldonado, Verginica Mititelu, Johanna Monti, Joakim Nivre, Mihaela Onofrei, Viola Ow, Carla Parra Escartín, Manfred Sailer, Carlos Ramisch, Renata Ramisch, Monica-Mihaela Rizea, Agata Savary, Nathan Schneider, Ivelina Stonayova, Sara Stymne, Ashwini Vaidya, Veronika Vincze & Abigail Walsh. 2018. Annotation guidelines for the PARSEME shared task on automatic identification of verbal MWEs (edition 1.1). <http://parsemefr.lif.univ-mrs.fr/parseme-st-guidelines/1.1/>.

- Biswas, Alexander A. 2017. Inflectional periphrasis in LFG. MPhil thesis, University of Oxford. https://www.academia.edu/33800486/Inflectional_Periphrasis_in_LFG.
- Blackburn, Patrick & Edith Spaan. 1993. A modal perspective on the computational complexity of Attribute Value grammar. *Journal of Logic, Language and Information* 2(2). 129–169. <http://doi.org/10.1007/BF01050635>.
- Bleam, Tonia, Chung-hye Han, Rashmi Prasad, Carlos Prolo & Anoop Sarkar. 2001. A Lexicalized Tree Adjoining Grammar for English. Tech. rep. from the Institute for Research in Cognitive Science, University of Pennsylvania. <http://www.cis.upenn.edu/~xtag/>.
- Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-Based Construction Grammar*. Stanford, CA: CSLI Publications.
- Bobrow, Samuel A. & Susan M. Bell. 1973. On catching on to idiomatic expressions. *Memory & Cognition* 1(3). 343–346. <https://doi.org/10.3758/BF03198118>.
- Bock, J. Kathryn & Willem J. M. Levelt. 1994. Language production: grammatical encoding. In Morton A. Gernsbacher (ed.), *Handbook of psycholinguistics*, 945–984. San Diego, CA: Academic Press.
- Bogarde, Dirk. 1992. *Great meadow: a memoir*. London, UK: Bloomsbury.
- Bolinger, Dwight. 1967. Adjectives in English: attribution and predication. *Lingua* 18. 1–34. [https://doi.org/10.1016/0024-3841\(67\)90018-6](https://doi.org/10.1016/0024-3841(67)90018-6).
- Bolinger, Dwight. 1975. On the passive in English. In Adam Makkai & Valerie Becker Makkai (eds.), *The first LACUS Forum 1974*, 57–77. Columbia, SC: Hornbeam Press.
- Bonami, Olivier. 2015. Periphrasis as collocation. *Morphology* 25(1). 63–110. <https://doi.org/10.1007/s11525-015-9254-3>.
- Bonami, Olivier & Danièle Godard. 2008. Lexical semantics and pragmatics of evaluative adverbs. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: syntax, semantics, and discourse*, 274–302. Oxford, UK: Oxford University Press.
- Borer, Hagit. 1984. *Parametric syntax: case studies in Semitic and Romance languages*. Dordrecht, NL: Foris.
- Bowdle, Brian F. & Derdre Gentner. 2005. The career of metaphor. *Psychological Review* 112(1). 193–216. <http://doi.org/10.1037/0033-295X.112.1.193>.
- Bresnan, Joan. 1978. A realistic transformational grammar. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, Cambridge, MA: MIT Press.
- Bresnan, Joan. 1982a. Control and complementation. *Linguistic Inquiry* 13(3). 343–434. <https://www.jstor.org/stable/4178286>.
- Bresnan, Joan. 1982b. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 3–86. Cambridge, MA: MIT Press.

- Bresnan, Joan. 1993. Interaction between grammar and discourse in Chicheŵa (Bantu). In William A. Foley (ed.), *The role of theory in language description* (Trends in Linguistics: Studies and Monographs 69), 45–60. Berlin, DE: Mouton de Gruyter.
- Bresnan, Joan. 1995. Linear order, syntactic rank, and empty categories: on weak crossover. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, III & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar*, 241–274. Stanford, CA: CSLI Publications.
- Bresnan, Joan. 1998. Morphology competes with syntax: explaining typological variation in weak crossover effects. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, Cambridge, MA: MIT Press and MIT Working Papers in Linguistics.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax* (2nd edn.). Oxford, UK: Wiley-Blackwell.
- Bresnan, Joan & Jonni M. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20(1). 1–50. <https://www.jstor.org/stable/4178612>.
- Bresnan, Joan, Ronald M. Kaplan, Stanley Peters & Annie Zaenen. 1982. Cross-serial dependencies in Dutch. *Linguistic Inquiry* 13(4). 613–635. <https://www.jstor.org/stable/4178298>.
- Bresnan, Joan & Sam A. Mchombo. 1995. The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13(2). 181–254. <https://doi.org/10.1007/BF00992782>.
- Bruening, Benjamin. 2018. The lexicalist hypothesis: both wrong and superfluous. *Language* 94(1). 1–42. <https://doi.org/10.1353/lan.2018.0000>.
- Butt, Miriam, Mary Dalrymple & Anette Frank. 1997. An architecture for linking theory in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*, Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG2-1997/lfg97butt-dalrymple-frank.pdf>.
- Butt, Miriam, María-Eugenia Niño & Frédérique Segond. 1996a. Multilingual processing of auxiliaries within LFG. In Dafydd Gibbon (ed.), *Natural language processing and speech technology: Results of the 3rd KONVENS Conference, Universität Bielefeld, October 1996*, 111–122. Berlin: Mouton de Gruyter.
- Butt, Miriam, María-Eugenia Niño & Frédérique Segond. 1996b. Multilingual processing of auxiliaries in LFG. In Dafydd Gibbon (ed.), *Natural language processing and speech technology: results of the 3rd KONVENS conference, Bielefeld, October 1996*, 111–122.
- Cacciari, Cristina & Patrizia Tabossi. 1988. The comprehension of idioms. *Journal of Memory and Language* 27(6). 668–683. [https://doi.org/10.1016/0749-596X\(88\)90014-9](https://doi.org/10.1016/0749-596X(88)90014-9).

- Candito, Marie-Hélène. 1996. A principle-based hierarchical representation of LTAGs. In *Proceedings of the 16th conference on Computational Linguistics (COLING)*, 194–199. Association for Computational Linguistics. <http://dx.doi.org/10.3115/992628.992664>.
- Carrol, Gareth & Kathy Conklin. 2017. Cross language lexical priming extends to formulaic units: evidence from eye-tracking suggests that this idea 'has legs'. *Bilingualism: Language and Cognition* 20(2). 299–317. <https://doi.org/10.1017/S1366728915000103>.
- Champollion, Lucas. 2010. Cumulative readings of *every* do not provide evidence for events and thematic roles. In Maria Aloni, Harald Bastiaanse, Tikitou de Jager & Katrin Schulz (eds.), *Logic, language and meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 2009* (Lecture Notes in Computer Science 6042), 213–222. Berlin, DE: Springer. https://doi.org/10.1007/978-3-642-14287-1_22.
- Champollion, Lucas. 2015. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38(1). 31–66. <https://doi.org/10.1007/s10988-014-9162-8>.
- Chierchia, Gennaro. 1982. Bare plurals, mass nouns, and nominalizations. In *Proceedings of the First West Coast Conference on Formal Linguistics*, 243–255. Stanford, CA: Stanford Linguistics Association.
- Chierchia, Gennaro & Sally McConnell-Ginet. 2000. *Meaning and grammar: an introduction to semantics* (2nd edn.). Cambridge, MA: MIT Press.
- Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2(3). 113–124.
- Chomsky, Noam. 1957. *Syntactic structures*. London, UK: Mouton.
- Chomsky, Noam. 1959. On certain formal properties of grammars. *Information and control* 2. 137–167.
- Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague, NL: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, UK: Cambridge University Press.
- Church, Alonzo. 1951. The need for abstract entities in semantic analysis. *Proceedings of the American Academy of Arts and Sciences* 80(1). 100–112.
- Collins, Chris. 1988. Part 1: conjunction adverbs. Unpublished ms., MIT.

- Crabbé, Benoît, Denys Duchier, Claire Gardent, Joseph Le Roux & Yannick Parmentier. 2013. XMG: eXtensible MetaGrammar. *Computational Linguistics* 39(3). 591–629. https://doi.org/10.1162/COLI_a_00144.
- Cronk, Brian C., Susan D. Lima & Wendy A. Schweigert. 1993. Idioms in sentences: effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research* 22(1). 59–82. <https://doi.org/10.1007/BF01068157>.
- Cronk, Brian C. & Wendy A. Schweigert. 1992. The comprehension of idioms: the effects of familiarity, literalness, and usage. *Applied Psycholinguistics* 13(2). 131–146. <https://doi.org/10.1017/S0142716400005531>.
- Crouch, Dick, Mary Dalrymple, Ron Kaplan, Tracy King, John Maxwell & Paula Newman. 2012. *XLE documentation*. Palo Alto Research Center (PARC), Palo Alto, CA. <http://www2.parc.com/isl/groups/nlitt/xle/doc/xle.html>.
- Crouch, Dick & Josef van Genabith. 2000. Linear logic for linguists: ESSLI-2000 course notes. Unpublished ms. available at <http://www.ling.ohio-state.edu/~pollard/681/crouch.pdf>.
- Cruse, D. A. 1986. *Lexical semantics*. Cambridge, UK: Cambridge University Press.
- Culy, Chris. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8(3). 345–351. https://doi.org/10.1007/978-94-009-3401-6_14.
- Culy, Christopher D. 1983. An extension of phrase structure rules and its application to natural language. Master's thesis, Stanford University.
- Dahlstrom, Amy. 1991. *Plains Cree morphosyntax*. New York, NY: Garland.
- Dahlstrom, Amy. 2009. OBJ_θ without OBJ: a typology of Meskwaki objects. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG09 Conference*, 222–239. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/14/papers/lfg09dahlstrom.pdf>.
- Dalrymple, Mary. 1993. *The syntax of anaphoric binding* (CSLI Lecture Notes 36). Stanford, CA: CSLI Publications.
- Dalrymple, Mary (ed.). 1999. *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*. Cambridge, MA: MIT Press.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). Stanford, CA: Academic Press.
- Dalrymple, Mary. 2015a. Morphology in the LFG architecture. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG15 Conference*, 64–83. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/20/papers/lfg15dalrymple.pdf>.
- Dalrymple, Mary. 2015b. Periphrasis and morphology in LFG. Paper presented at the Workshop on Morphology, Waseda University, 21 July 2015.

- Dalrymple, Mary & Jamie Y. Findlay. 2019. Lexical functional grammar. In András Kertész, Edith Moravcsik & Csilla Rákosi (eds.), *Current approaches to syntax – a comparative handbook*, Berlin, DE: De Gruyter Mouton.
- Dalrymple, Mary, Dag T. T. Haug & John J. Lowe. 2018. Integrating LFG's binding theory with PCDRT. *Journal of Language Modelling* 6(1). 87–129. <https://doi.org/10.15398/jlm.v6i1.204>.
- Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG04 Conference, 199–208*. Stanford, CA: CSLI Publications. <http://www.stanford.edu/group/cslipublications/cslipublications/LFG/9/lfg04dkk.pdf>.
- Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2015. Economy of expression as a principle of syntax. *Journal of Language Modelling* 2(3). 377–412. <http://jlm.ipipan.waw.pl/index.php/JLM/article/view/82>.
- Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell, III & Annie Zaenen. 1995a. Nonlocal dependencies. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, III & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar*, 131–135. Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John Lamping, Fernando Pereira & Vijay Saraswat. 1995b. Linear logic for meaning assembly. In Suresh Manandhar, G. P. Lopes & W. Nutt (eds.), *Proceedings of Computational Logic for Natural Language Processing*, Edinburgh, UK.
- Dalrymple, Mary, John Lamping, Fernando Pereira & Vijay Saraswat. 1999a. Overview and introduction. In Mary Dalrymple (ed.), *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*, Cambridge, MA: MIT Press.
- Dalrymple, Mary, John Lamping & Fernando Pereira Vijay Saraswat. 1999b. Quantification, anaphora, and intensionality. In Mary Dalrymple (ed.), *Semantics and syntax in Lexical Functional Grammar: the resource logic approach*, 39–90. Cambridge, MA: MIT Press.
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge, UK: Cambridge University Press.
- Dalrymple, Mary, Stuart M. Shieber & Fernando C. N. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14(4). 399–452. <https://doi.org/10.1007/BF00630923>.
- Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81–120. Pittsburgh, PA: University of Pittsburgh Press.
- Davies, Mark. 2004–. *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). <http://corpus.byu.edu/bnc>.
- Davies, Mark. 2008–. *The Corpus of Contemporary American English: 450 million words, 1990–present*. <http://corpus.byu.edu/coca>.
- Davies, Mark. 2013a. *Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries*. <http://corpus.byu.edu/glowbe>.

- Davies, Mark. 2013b. *Corpus of News on the Web (NOW): 3+ billion words from 20 countries, updated every day*. <https://corpus.byu.edu/now/>.
- Davies, Mark. 2018-. *iWeb: the intelligent web-based corpus*. <http://corpus.byu.edu/iweb>.
- Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the definition of word* (Linguistic Inquiry monographs 14). Cambridge, MA: MIT Press.
- Dummett, Michael. 1973. *Frege: philosophy of language*. London, UK: Duckworth.
- Egan, Andy. 2008. Pretense for the complete idiom. *Noûs* 42(3). 381–409. <https://www.jstor.org/stable/25177172>.
- Égré, Paul. 2018. Philosophy of linguistics. In Anouk Baberousse, Denis Bonnay & Mikael Cozic (eds.), *Philosophy of science: a companion*, 654–726. Oxford, UK: Oxford University Press.
- Eisner, Jason & Giorgio Satta. 2000. A faster parsing algorithm for Lexicalized Tree-Adjoining Grammars. In *Proceedings of the Fifth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+5)*, 79–84. <http://aclweb.org/anthology/W00-2011>.
- Ernst, Thomas. 1981. Grist for the linguistic mill: idioms and ‘extra’ adjectives. *Journal of Linguistic Research* 1(3). 51–68.
- Ernst, Thomas. 1983. Towards an integrated theory of adverb position in English. Doctoral dissertation, Indiana University.
- Ernst, Thomas. 2001. *The syntax of adjuncts* (Cambridge Studies in Linguistics 96). Cambridge, UK: Cambridge University Press.
- Estill, Robert B. & Susan Kemper. 1982. Interpreting idioms. *Journal of Psycholinguistic Research* 11(6). 559–568. <https://doi.org/10.1007/BF01067612>.
- Falk, Yehuda N. 2001. *Lexical-Functional Grammar: an introduction to parallel constraint-based syntax*. Stanford, CA: CSLI Publications.
- Falk, Yehuda N. 2003. The English auxiliary system revisited. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG03 Conference*, 184–204. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/8/pdfs/lfg03falk.pdf>.
- Falk, Yehuda N. 2010. An unmediated analysis of relative clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG10 Conference*, 207–227. CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/15/papers/lfg10falk.pdf>.
- Fernando, Chitra & Roger H. Flavell. 1981. *On idiom: critical views and perspectives*. Exeter, UK: University of Exeter Press.

- Ferreira, Fernanda. 2000. Syntax in language production: an approach using Tree-Adjoining Grammars. In Linda Wheeldon (ed.), *Aspects of language production*, 291–330. Hove, UK: Psychology Press.
- Ferreira, Fernanda, Ellen F. Lau & Karl G. D. Bailey. 2004. Disfluencies, language comprehension, and Tree Adjoining Grammars. *Cognitive Science* 28(5). 721–749. https://doi.org/10.1207/s15516709cog2805_5.
- Fillmore, Charles J. 1987. Varieties of conditional sentences. In Fred Marshall (ed.), *Proceedings of the Third Eastern States Conference on Linguistics (ESCOL 86): October 10–11, 1986*, 163–182. Pittsburgh, PA: University of Pittsburgh.
- Fillmore, Charles J. 1988. The mechanisms of ‘Construction Grammar’. In Shelley Axmaker, Annie Jaissner & Helen Singmaster (eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35–55. Berkeley, CA: Berkeley Linguistics Society, University of California, Berkeley.
- Fillmore, Charles J. & Paul Kay. 1995. Construction Grammar coursebook. Unpublished ms., Department of Linguistics, University of California, Berkeley.
- Fillmore, Charles J., Paul Kay & Mary Catherine O’Connor. 1988. Regularity and idiomatity in grammatical constructions: the case of *let alone*. *Language* 64(3). 501–538. <https://doi.org/10.2307/414531>.
- Findlay, Jamie Y. 2016a. Mapping theory without argument structure. *Journal of Language Modelling* 4(2). 293–338. <http://jlm.ipipan.waw.pl/index.php/JLM/article/view/171/148>.
- Findlay, Jamie Y. 2016b. The prepositional passive in Lexical Functional Grammar. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 255–275. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2016/headlex2016-findlay.pdf>.
- Findlay, Jamie Y. 2017a. Multiword expressions and lexicalism. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG17 Conference*, 209–229. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2017/lfg2017-findlay.pdf>.
- Findlay, Jamie Y. 2017b. Multiword expressions and lexicalism: the view from LFG. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, 73–79. Association for Computational Linguistics. <http://aclweb.org/anthology/W17-1709>.
- Findlay, Jamie Y. 2018. When the cat you let out of the bag has claws: the role of metaphor in understanding idioms. Talk given at the Linguistics and English Language Research Seminar, University of Manchester, 27 November 2018 [link to slides].
- Findlay, Jamie Y., Sascha Bargmann & Manfred Sailer. 2018. Pulling a pretence rabbit out of the hat. Talk given at the Workshop on Form-Meaning Mismatches in Natural Language (FMM2018), Georg-August-Universität Göttingen, 4 August 2018 [link to slides].

- Findlay, Jamie Y., Sascha Bargmann & Manfred Sailer. 2019. Why the butterflies in your stomach can have big wings: combining formal and cognitive theories to explain productive extensions of idioms. Talk given at the EUROPHRAS 2019 Productive Patterns in Phraseology Conference, 24 January 2019 [link to slides].
- Findlay, Jamie Y., Sascha Bargmann & Manfred Sailer. in preparation. The double life of idioms: lexical encoding, metaphor, and analogy in the analysis of VP idioms.
- Fodor, Jerry A. 1975. *The language of thought*. New York, NY: Thomas Y. Crowell Company.
- Fodor, Jerry A. 1983. *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: MIT Press.
- Francis, Dick. 1978. *In the frame*. London, UK: Pan Books.
- Frank, Anette & Josef van Genabith. 2001. GlueTag: linear logic based semantics for LTAG—and what it teaches us about LFG and LTAG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG01 Conference*, 104–126. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/6/pdfs/lfg01frankgenabith.pdf>.
- Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6(1). 22–42. <https://www.jstor.org/stable/25000426>.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.
- Gehrke, Berit. To appear. Multiple event readings and *occasional*-type adjectives. In Patricia Cabredo Hofherr & Jenny Doetjes (eds.), *The Oxford handbook of grammatical number*, Oxford, UK: Oxford University Press.
- Gentner, Dedre. 1983. Structure-mapping: a theoretical framework for analogy. *Cognitive Science* 7(2). 155–170. [https://doi.org/10.1016/S0364-0213\(83\)80009-3](https://doi.org/10.1016/S0364-0213(83)80009-3).
- Gentner, Dedre & Francisco Maravilla. 2018. Analogical reasoning. In Linden J. Ball & Valerie A. Thompson (eds.), *International handbook of thinking and reasoning*, 186–203. Oxford, UK: Routledge.
- Gibbs, Raymond W., Jr. 1980. Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition* 8(2). 149–156. <https://doi.org/10.3758/BF03213418>.
- Gibbs, Raymond W., Jr. 1986. Skating on thin ice: Literal meaning and understanding idioms in context. *Discourse Processes* 9(1). 17–30. <https://doi.org/10.1080/01638538609544629>.
- Giorgolo, Gianluca & Ash Asudeh. 2011. Multidimensional semantics with unidimensional glue logic. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG11 Conference*, 236–256. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/16/papers/lfg11giorgoloasudeh1.pdf>.

- Girard, Jean-Yves. 1987. Linear logic. *Theoretical Computer Science* 50(1). 1–102. [https://doi.org/10.1016/0304-3975\(87\)90045-4](https://doi.org/10.1016/0304-3975(87)90045-4).
- Glucksberg, Sam. 1993. Idiom meanings and allusional content. In Cristina Cacciari & Patrizia Tabossi (eds.), *Idioms: Processing, structure, and interpretation*, 3–26. London, UK: Lawrence Erlbaum.
- Glucksberg, Sam. 2001. *Understanding figurative language: from metaphor to idioms* (Oxford Psychology Series 36). Oxford, UK: Oxford University Press.
- Godoy, Ricardo, Marc Jacobson & David Wilkie. 1998. Strategies of rain-forest dwellers against misfortunes: the Tsimane' Indians of Bolivia. *Ethnology* 37(1). 55–69. <https://www.jstor.org/stable/3773848>.
- Goldberg, Adele E. 1995. *Constructions: a Construction Grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford, UK: Oxford University Press.
- Goldberg, Adele E. & Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80(3). 532–568. <https://www.jstor.org/stable/4489722>.
- Greibach, Sheila A. 1965. A new normal-form theorem for context-free phrase structure grammars. *Journal of the Association for Computing Machinery (JACM)* 12(1). 42–52. <http://doi.acm.org/10.1145/321250.321254>.
- Grice, H. Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Speech acts*, vol. 3 in the Syntax and Semantics Series, 41–58. New York, NY: Academic Press.
- Grice, H. Paul. 1978. Further notes on logic and conversation. In Peter Cole (ed.), *Pragmatics* (Syntax and Semantics 9), 113–128. New York, NY: Academic Press.
- Grimshaw, Jane. 1992. *Argument structure*. Cambridge, MA: MIT Press.
- Grimshaw, Jane. 2005. Extended projection. In *Words and structure*, 1–74. Stanford, CA: CSLI Publications. [Published version of 1991 manuscript by the same name].
- Hall, Robert A., Jr. 1973. The transferred epithet in P. G. Wodehouse. *Linguistic Inquiry* 4(1). 92–94. <https://www.jstor.org/stable/4177754>.
- Halle, Morris. 1973. Prolegomena to a theory of word formation. *Linguistic Inquiry* 4(1). <https://www.jstor.org/stable/4177749>.
- Halvorsen, Per-Kristian & Ronald M. Kaplan. 1988. Projections and semantic descriptions in Lexical-Functional Grammar. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, 1116–1122. Institute for New Generation Systems.
- Harkema, Henk. 2001. A characterization of minimalist languages. In Philippe de Groot, Glyn Morrill & Christian Retoré (eds.), *Logical Aspects of Computational Linguistics (LACL 2001)* (Lecture Notes in Computer Science 2099), 193–211. Berlin, DE: Springer. https://doi.org/10.1007/3-540-48199-0_12.

- Haug, Dag T. T. 2008. Tense and aspect for glue semantics: the case of participial XADJ's. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, 291–311. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/13/papers/lfg08haug.pdf>.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar* (Blackwell Textbooks in Linguistics 13). Oxford, UK: Blackwell.
- Hemmings, Charlotte. 2016. The Kelabit language, Austronesian voice and syntactic typology. Doctoral dissertation, SOAS, University of London.
- Horn, George M. 2003. Idioms, metaphors and syntactic mobility. *Journal of Linguistics* 39(2). 245–273. <https://doi.org/10.1017/S0022226703002020>.
- Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51(3). 639–671. <https://www.jstor.org/stable/412891>.
- Jackendoff, Ray. 1977. *X' syntax: a study of phrase structure*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 2007. *Language, consciousness, culture: essays on mental structure*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 2009. Compounding in the Parallel Architecture and conceptual semantics. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford handbook of compounding*, 105–128. Oxford, UK: Oxford University Press.
- Janssen, Theo M. V. 1997. Compositionality. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic and language*, 417–473. Amsterdam, NL: Elsevier Science.
- Jayez, Jacques & Corinne Rossari. 2004. Parentheticals as conventional implicatures. In Francis Corbin & Henriëtte de Swart (eds.), *Handbook of French semantics*, 211–229. Stanford, CA: CSLI Publications.
- Joshi, Aravind K. 1985. Tree adjoining grammars: how much context-sensitivity is required to provide reasonable structural descriptions? In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives*, 206–250. Cambridge, UK: Cambridge University Press.
- Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10(1). 136–163. [https://doi.org/10.1016/S0022-0000\(75\)80019-5](https://doi.org/10.1016/S0022-0000(75)80019-5).
- Joshi, Aravind K. & Yves Schabes. 1997. Tree-Adjoining Grammars. In Grzegorz Rozenberg & Arto Salomaa (eds.), *Handbook of formal languages*, vol. 3, 69–123. Berlin, DE: Springer.
- Kallmeyer, Laura. 2010. *Parsing beyond context-free grammars*. Heidelberg, DE: Springer.
- Kallmeyer, Laura & Maribel Romero. 2008. Scope and situation binding in LTAG using semantic unification. *Research on language and computation* 6(1). 3–52. <https://doi.org/10.1007/s11168-008-9046-6>.

- Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson & Paul Bennett (eds.), *Linguistic theory and computer applications*, London: Academic Press. Reprinted in Dalrymple et al. (1995: 339–367).
- Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell, III, & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar*, 7–28. Stanford, CA: CSLI Publications.
- Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-Functional Grammar: a formal system for grammatical representation. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 173–281. Cambridge, MA: MIT Press.
- Kaplan, Ronald M., John T. Maxwell, III & Annie Zaenen. 1987. Functional uncertainty. In *CSLI Publications Monthly Newsletter*, Stanford, CA: Stanford University.
- Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin & Antony Kroch (eds.), *Alternative conceptions of phrase structure*, 17–42. Chicago, IL: Chicago University Press.
- Kasai, Jungo, Bob Frank, Tom McCoy, Owen Rambow & Alexis Nasr. 2017. TAG parsing with neural networks and vector representations of supertags. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1712–1722. Association for Computational Linguistics. <http://aclweb.org/anthology/D17-1180>.
- Kastovsky, Dieter. 1982. *Wortbildung und Semantik*. Düsseldorf, DE: Bagel/Francke.
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language* 75(1). 1–33. <http://doi.org/10.2307/417472>.
- Kay, Paul & Ivan A. Sag. 2012. A lexical theory of phrasal idioms. Unpublished ms., Stanford University.
- Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. A lexical theory of phrasal idioms. Unpublished ms., CSLI, Stanford. <http://www1.icsi.berkeley.edu/~kay/idiom-pdflatex.11-13-15.pdf>.
- Kibort, Anna. 2001. The Polish passive and impersonal in Lexical Mapping Theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG01 Conference*, 163–183. Stanford, CA: CSLI Publications. <http://www.stanford.edu/group/cslipublications/cslipublications/LFG/6/lfg01kibort.pdf>.
- Kibort, Anna. 2007. Extending the applicability of Lexical Mapping Theory. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*, 250–270. Stanford, CA: CSLI Publications. <http://www.stanford.edu/group/cslipublications/cslipublications/LFG/12/papers/lfg07kibort.pdf>.
- Kokkonidis, Miltiadis. 2005. Why glue a donkey to an f-structure when you can constrain and bind it instead. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, 238–252. Stanford, CA: CSLI Publications. <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/10/pdfs/lfg05kokkonidis.pdf>.

- Kokkonidis, Miltiadis. 2008. First-order Glue. *Journal of Logic, Language and Information* 17(1). 43–68. <https://doi.org/10.1007/s10849-006-9031-0>.
- Krenn, Brigitte & Gregor Erbach. 1994. Idioms and support verb constructions. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, Stanford, CA: CSLI Publications.
- Kripke, Saul A. 1977. Speaker's reference and semantic reference. *Midwest Studies in Philosophy* 2. 255–276.
- Kroch, Anthony. 1987. Unbounded dependencies and ubjacency in a Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *Mathematics of language: proceedings of a conference held at the University of Michigan, Ann Arbor, October 1984*, 143–172. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.35.09kro>.
- Kroch, Anthony & Aravind K. Joshi. 1985. The linguistic relevance of Tree Adjoining Grammar. Tech. Rep. MC-CS-85-16 Department of Computer and Information Sciences, University of Pennsylvania.
- Kroeger, Paul. 2007. LFG as a framework for descriptive grammar. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*, 293–301. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/12/papers/lfg07kroeger.pdf>.
- Kuhlmann, Marco. 2007. Dependency structures and lexicalized grammars. Doctoral dissertation, Saarland University.
- Kuhn, Jonas. 2001. Resource sensitivity in the syntax-semantics interface and the German split NP construction. In W. Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax*, Stanford, CA: CSLI Publications.
- Kuno, Susumu & Ken-ichi Takami. 2004. *Functional constraints in grammar: on the unergative-unaccusative distinction*. Amsterdam, NL and Philadelphia, PA: John Benjamins.
- Kyriacou, Marianna, Kathy Conklin & Dominic Thompson. 2018. Fathoming the depths of the idiomatic advantage: an eye-tracking investigation. Talk given at the 4th International Conference on Figurative Thought and Language (FTL4), 25 October 2018.
- Lakoff, George. 1984. *There*-constructions: a case study in grammatical construction theory and prototype theory. Berkeley Cognitive Science Program Working Paper No. 18.
- Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.
- Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Larson, Richard K. 1998. Events and modification in nominals. In Devon Strolovitch & Aaron Lawson (eds.), *Proceedings of the 8th Semantics and Linguistic Theory Conference (SALT 8)*, 145–168. <https://doi.org/10.3765/salt.v8i0.2803>.

- Lee-Schoenfeld, Vera & Gabriele Diewald. 2017. Passivization possibilities in double-accusative constructions. *Proceedings of the Linguistic Society of America* 2. 9:1–14. <http://dx.doi.org/10.3765/plsa.v2i0.4050>.
- Levin, Beth & Malka Rappaport Hovav. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: lexical and compositional factors*, 97–134. Stanford, CA: CSLI Publications.
- Lichte, Timm & Laura Kallmeyer. 2016. Same syntax, different semantics: a compositional approach to idiomaticity in multi-word expressions. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics* 11, 111–140. Paris, FR: Colloque de Syntaxe et Sémantique à Paris (CSSP). http://www.cssp.cnrs.fr/eiss11/eiss11_lichte-and-kallmeyer.pdf.
- Link, Godehard. 1983. The logical analysis of plurals and mass terms: a lattice-theoretic approach. In Rainer Bäuerle, Christoph Schwartze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 302–323. Berlin, DE: Walter de Gruyter.
- Lipps, Jonathan. 2011. A lexical-functional analysis of Swahili relative clauses. MPhil thesis, University of Oxford.
- Lowe, John J. 2014. Gluing meanings and semantic structures. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG14 Conference*, 387–407. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/papers/lfg14lowe.pdf>.
- Lowe, John J. 2015a. Complex predicates: an LFG + glue analysis. *Journal of Language Modelling* 3(2). 413–462. <http://jlm.ipipan.waw.pl/index.php/JLM/article/view/125/116>.
- Lowe, John J. 2015b. *Participles in Rigvedic Sanskrit: the syntax and semantics of adjectival verb forms*. Oxford, UK: Oxford University Press.
- Maher, Zachary. 2013. Opening a can of worms: idiom flexibility, decomposability, and the mental lexicon. Unpublished ms., Yale University. <https://ling.yale.edu/sites/default/files/files/alumni%20senior%20essays/Maher%20Zach%20-%20Senior%20Essay.pdf>.
- Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *University of Pennsylvania working papers in linguistics, volume 4.2: Proceedings of the 21st Annual Penn Linguistics Colloquium*, 201–225. Philadelphia, PA: University of Pennsylvania Department of Linguistics.
- Marcus, Mitchell P., Donald Hindle & Margaret M. Fleck. 1983. D-Theory: talking about talking about trees. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, 129–136. Association for Computational Linguistics. <https://doi.org/10.3115/981311.981337>.
- Markantonatou, Stella & Niki Samaridi. 2018. Revisiting the grammatical function “object” (OBJ and OBJ_θ). In Manfred Sailer & Stella Markantonatou (eds.), *Multiword expressions: insights from a multi-lingual perspective*, 187–213. Berlin, DE: Language Science Press. <http://langsci-press.org/catalog/book/184>.

- McCarthy, John J. 1982. Prosodic structure and expletive infixation. *Language* 58(3). 574–590. <https://www.jstor.org/stable/413849>.
- McCawley, James D. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4(3). 243–269. <https://www.jstor.org/stable/25000330>.
- McCawley, James D. 1988. The comparative conditional construction in English, German, and Chinese. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society, 176–187*. Berkeley, CA: Berkeley Linguistics Society, University of California, Berkeley.
- McClure, Scott. 2011. Modification in non-combining idioms. *Semantics & Pragmatics* 4. 1–7. <https://doi.org/10.3765/sp.4.7>.
- McGinnis, Martha. 2002. On the systematic aspect of idioms. *Linguistic Inquiry* 33(4). 665–672. <https://doi.org/10.1162/ling.2002.33.4.665>.
- McGlone, Matthew S., Sam Glucksberg & Cristina Cacciari. 1994. Semantic productivity and idiom comprehension. *Discourse Processes* 17(2). 167–190. <https://doi.org/10.1080/01638539409544865>.
- McMillan, James B. 1980. Infixing and interposing in English. *American Speech* 55(3). 163–183. <https://doi.org/10.2307/455082>.
- Montague, Richard. 1970. Universal grammar. *Theoria* 36(3). 373–398.
- Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In K. Jaakko J. Hintikka, Julius M. E. Moravcsik & Patrick Suppes (eds.), *Approaches to natural language: proceedings of the 1970 Stanford workshop on grammar and semantics* (Synthese Library 49), 221–243. Dordrecht, NL: Reidel.
- Müller, Stefan. 2018. The end of lexicalism as we know it? *Language* 94(1). e54–e66. <https://doi.org/10.1353/lan.2018.0014>.
- Mycock, Louise & John Lowe. 2013. The prosodic marking of discourse functions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG13 Conference*, 440–460. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/LFG/18/papers/lfg13mycocklowe.pdf>.
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty & Daniel Zeman. 2016. Universal Dependencies v1: A multilingual treebank collection. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, 1659–1666. Portorož, SI: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2016/pdf/348_Paper.pdf.
- Nunberg, Geoffrey. 1977. The pragmatics of reference. Doctoral dissertation, City University of New York Graduate Centre.

- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538. <https://doi.org/10.1353/lan.1994.0007>.
- Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge, MA: MIT Press.
- Parsons, Terence. 1995. Thematic relations and arguments. *Linguistic Inquiry* 26(4). 635–662. <https://www.jstor.org/stable/4178917>.
- Partee, Barbara, Alice ter Meulen & Robert E. Wall. 1990. *Mathematical methods in linguistics*. Dordrecht, NL: Kluwer.
- Partee, Barbara Hall & Vladimir Borschev. 1998. Integrating lexical and formal semantics: genitives, relational nouns, and type-shifting. In Robin Cooper & Thomas Gamkrelidze (eds.), *Proceedings of the Second Tbilisi Symposium on Language, Logic, and Computation*, 229–241. Tbilisi, GE: Center on Language, Logic, Speech, Tbilisi State University. <http://www-unix.oit.umass.edu/~partee/docs/tbilisi17.pdf>.
- Perrin, Laurent. 2003. Citation, lexicalisation et interprétation des expressions idiomatiques. In Jacqueline Authier-Revuz, Marianne Doury & Sandrine Reboul-Touré (eds.), *Parler des mots: le fait autonymique en discours*, 281–292. Paris: Presses Sorbonne Nouvelle.
- Pietroski, Paul M. 2002. Function and concatenation. In Gerhard Preyer & Georg Peter (eds.), *Logical form and language*, 91–117. Oxford, UK: Clarendon Press.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.
- Potts, Chris. 2005. *The logic of conventional implicatures* (Oxford Studies in Theoretical Linguistics 7). Oxford, UK: Oxford University Press.
- Potts, Chris. 2007. The expressive dimension. *Theoretical Linguistics* 33(2). 165–198.
- Potts, Christopher, Luis Alonso-Ovalle, Ash Asudeh, Rajesh Bhatt, Seth Cable, Christopher Davis, Yurie Hara, Angelika Kratzer, Eric McCready, Tom Roeper & Martin Walkow. 2009. Expressives and identity conditions. *Linguistic Inquiry* 40(2). 356–366. <https://doi.org/10.1162/ling.2009.40.2.356>.
- Przepiórkowski, Adam. 1999. Case assignment and the complement-adjunct dichotomy: a non-configurational constraint-based approach. Doctoral dissertation, Universität Tübingen.
- Przepiórkowski, Adam. 2016. How *not* to distinguish arguments from adjuncts in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 539–559. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2016/headlex2016-przepiorkowski.pdf>.
- Przepiórkowski, Adam. 2017. A full-fledged hierarchical lexicon in LFG: the FrameNet approach. In Victoria Rosén & Koenraad De Smedt (eds.), *The very model of a modern linguist*, vol. 8 Bergen Language and Linguistics Studies, 202–219. Bergen: University of Bergen Library. <https://doi.org/10.15845/bells.v8i1.1336>.

- Pullum, Geoffrey K. 2013. The central question in comparative syntactic metatheory. *Mind and Language* 28(4). 492–521. <https://doi.org/10.1111/mila.12029>.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retoré (eds.), *Logical Aspects of Computational Linguistics (LACL 2001)* (Lecture Notes in Computer Science 2099), 17–43. Berlin, DE: Springer. https://doi.org/10.1007/3-540-48199-0_2.
- Pulman, Stephen G. 1993. The recognition and interpretation of idioms. In Cristina Cacciari & Patrizia Tabossi (eds.), *Idioms: processing, structure, and interpretation*, 249–270. London, UK: Lawrence Erlbaum.
- Putnam, Hilary. 1967. Psychological predicates. In William H. Capitan & Daniel D. Merrill (eds.), *Art, mind, and religion*, 37–48. Pittsburgh, PA: University of Pittsburgh Press.
- Pylkkänen, Liina, Rodolfo Llinás & Gregory Murphy. 2006. The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience* 18(1). 97–109. <https://doi.org/10.1162/089892906775250003>.
- Restall, Greg. 2000. *An introduction to substructural logics*. London, UK: Routledge.
- Richter, Frank & Manfred Sailer. 2003. Cranberry words in formal grammar. In Claire Beyssade, Olivier Bonami, Patricia Cabredo Hofherr & Francis Corblin (eds.), *Empirical issues in syntax and semantics 4*, 155–172. Paris, FR: Presses Universitaires de Paris Sorbonne.
- Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In Arne Beckmann & Norbert Preining (eds.), *ESSLLI 2003: course material I* (Collegium Logicum 5), 87–143. Vienna, AT: Kurt Gödel Society Wien.
- Riehemann, Susanne. 2001. A constructional approach to idioms and word formation. Doctoral dissertation, Stanford University.
- Ross, John R. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT. Published as Ross, John R. 1986 *Infinite syntax!* Norwood, NJ: ALEX.
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson & Jan Scheffczyk. 2010. *FrameNet II: extended theory and practice*. Berkeley, CA: International Computer Science Institute. Distributed with the FrameNet data.
- Sadler, Louisa & Douglas J. Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30(1). 187–226. <https://www.jstor.org/stable/4176260>.
- Sag, Ivan. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2007/sag.pdf>.

- Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake & Dan Flickinger. 2002. Multiword expressions: a pain in the neck for NLP. In Alexander Gelbukh (ed.), *Computational Linguistics and Intelligent Text Processing: Third International Conference (CICLing 2002), Mexico City, Mexico, 17–23 February 2002* (Lecture Notes in Computer Science 2276), 1–15. Berlin, DE: Springer. https://doi.org/10.1007/3-540-45715-1_1.
- Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 1–30. Stanford, CA: CSLI Publications.
- Sailer, Manfred. 2000. Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar. Doctoral dissertation, Eberhard-Karls-Universität Tübingen.
- Sailer, Manfred. To appear. Idioms: *pull strings, spill beans, and kick the bucket*. In Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.), *The Blackwell companion to semantics*, Oxford, UK: Blackwell.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.
- Schabes, Yves, Anne Abeillé & Aravind K. Joshi. 1988. Parsing strategies with 'lexicalized' grammars: application to Tree Adjoining Grammars. In *Proceedings of the 12th Conference on Computational Linguistics (COLING)*, 578–583. Association for Computational Linguistics. <http://dx.doi.org/10.3115/991719.991757>.
- Schenk, André. 1995. The syntactic behavior of idioms. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: structural and psychological perspectives*, 253–271. Hove, UK: Lawrence Erlbaum.
- Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343. https://doi.org/10.1007/978-94-009-3401-6_12.
- Shieber, Stuart M. & Yves Schabes. 1990. Synchronous Tree Adjoining Grammars. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING '90)*, 253–258. Helsinki, FI.
- Simpson, Jane. 1983. Aspects of Warlpiri morphology and syntax. Doctoral dissertation, MIT.
- Simpson, Jane. 1991. *Warlpiri morpho-syntax: a lexicalist approach*. Dordrecht: Kluwer Academic Publishers.
- Soehn, Jan-Philipp & Manfred Sailer. 2008 [2003]. At first blush on tenterhooks: about selectional restrictions imposed by nonheads. In Gerald Penn (ed.), *Proceedings of FGVienna: the 8th Conference on Formal Grammar*, 123–138. Stanford, CA: CSLI Publications. <https://web.stanford.edu/group/cslipublications/cslipublications/FG/2003/soehn.pdf>.
- Spencer, Andrew. 1988. Bracketing paradoxes and the English lexicon. *Language* 64(4). 663–682. <https://doi.org/10.2307/414563>.

- Spencer, Andrew. 1991. *Morphological theory*. Oxford, UK: Basil Blackwell.
- Stabler, Edward P. 1997. Derivational minimalism. In Christian Retoré (ed.), *Logical Aspects of Computational Linguistics (LACL 1996)* (Lecture Notes in Computer Science 1328), 68–95. Berlin, DE: Springer. <https://doi.org/10.1007/BFb0052152>.
- Stabler, Edward P. 2011. Computational perspectives on minimalism. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 617–642. Oxford, UK: Oxford University Press.
- Stalnaker, Robert C. 1976. Possible worlds. *Noûs* 10(1). 65–75. <https://www.jstor.org/stable/2214477>.
- Stathi, Katerina. 2007. A corpus-based analysis of adjectival modification in German idioms. In Christiane Fellbaum (ed.), *Idioms and collocations: corpus-based linguistic and lexicographic studies*, 81–108. New York, NY: Continuum.
- Stavans, Ilan. 2001. *The inveterate dreamer: essays and conversations on Jewish culture*. Lincoln & London: University of Nebraska Press.
- Steedman, Mark. 1987. Combinatory grammars and parasitic gaps. *Natural Language & Linguistic Theory* 5(3). 403–439. <https://doi.org/10.1007/BF00134555>.
- Steen, Gerard. 2008. The paradox of metaphor: why we need a three-dimensional model of metaphor. *Metaphor & Symbol* 23(4). 213–241. <https://doi.org/10.1080/10926480802426753>.
- Steen, Gerard. 2017. Deliberate Metaphor Theory: basic assumptions, main tenets, urgent issues. *Intercultural Pragmatics* 14(1). 1–24. <https://doi.org/10.1515/ip-2017-0001>.
- Stump, Gregory T. 2001. *Inflectional morphology*. Cambridge, UK: Cambridge University Press.
- Swinney, David A. & Anne Cutler. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior* 18(5). 523–534. [https://doi.org/10.1016/S0022-5371\(79\)90284-6](https://doi.org/10.1016/S0022-5371(79)90284-6).
- Szabó, Zoltán Gendler. 2017. Compositionality. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Stanford, CA: The Metaphysics Research Lab. <https://plato.stanford.edu/entries/compositionality/>.
- Tabossi, Patrizia, Rachele Fanar & Kinou Wolf. 2009. Why are idioms recognized fast? *Memory and Cognition* 37(4). 529–540. <https://doi.org/10.3758/MC.37.4.529>.
- Tseng, Jesse. 2007. English prepositional passive constructions. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 271–286. Stanford, CA: CSLI Publications. <http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2007/tseng.pdf>.
- Vande Wiele, Héloïse. 2016. The loss of poetic effects: from indeterminate to conventionalized meaning. *Language and Literature* 25(1). 54–71. <https://doi.org/10.1177/0963947015623419>.

- Vijay-Shanker, K. 1992. Using descriptions of trees in a Tree Adjoining Grammar. *Computational Linguistics* 18(4). 481–517. <https://dl.acm.org/citation.cfm?id=176317>.
- Vijay-Shanker, K. & Aravind K. Joshi. 1988. Feature structure based Tree Adjoining Grammars. In *Proceedings of the 12th Conference on Computational Linguistics (COLING '88)*, 714–719. Association for Computational Linguistics. <http://dx.doi.org/10.3115/991719.991783>.
- Walther, Markus. 2000. Finite-state reduplication in One-Level Prosodic Morphology. In *Proceedings of the First Conference of the North American Chapter of the Association for Computational Linguistics*, 296–302. Association for Computational Linguistics. <http://www.aclweb.org/anthology/A00-2039>.
- Waszczuk, Jakub & Agata Savary. 2015. Modeling syntactic properties of MWEs in LFG. Poster presented at the PARSEME 4th general meeting, 19–20 March 2015, Valetta, Malta.
- Wedekind, Jürgen & Ronald M. Kaplan. 2018. Tractable Lexical-Functional Grammar. Submitted to *Computational Linguistics*.
- Whitehead, Alfred North & Bertrand Russell. 1910. *Principia mathematica*, vol. I. Cambridge, UK: Cambridge University Press.
- Wiese, Richard. 1996. Phrasal compounds and the theory of word syntax. *Linguistic Inquiry* 27(1). 183–193. <https://www.jstor.org/stable/4178931>.
- Wigner, Eugene. 1960. The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13(1). 1–14. <https://doi.org/10.1002/cpa.3160130102>.
- Wittenberg, Eva, Ray S. Jackendoff, Gina Kuperberg, Jesse Paczynski, Martinand Snedeker & Heike Wiese. 2014. The processing and representation of light verb constructions. In Asaf Bachrach, Isabelle Roy & Linnaea Stockall (eds.), *Structuring the argument*, 61–82. Amsterdam, NL: John Benjamins.
- Wittenberg, Eva & Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3). 393–413. <https://doi.org/10.1075/ml.6.3.03wit>.
- Zwicky, Arnold. 2007. Extramarital toes. *Language Log* (December 4, 2007). <http://itre.cis.upenn.edu/~myl/language-log/archives/005186.html>.