



THE UNIVERSITY *of* EDINBURGH

Title	Dynamic syntax
Author	Tugwell, David
Qualification	PhD
Year	1999

Thesis scanned from best copy available: may contain faint or blurred text, and/or cropped or missing pages.

Digitisation Notes:

- Pagination errors in original: pages 1&2 are missing.

Dynamic Syntax


David Tugwell

PhD
The University of Edinburgh
1998



Declaration

I declare that this thesis has been composed by myself and that the research reported herein is my own. This thesis complies with all the regulations for the degree of PhD at the University of Edinburgh, and falls below the requisite word limit specified.



David Tugwell
October 1998

Abstract

In this thesis I argue that dynamic “left-to-right” grammars have been undeservedly neglected as models of natural language syntax. To support this I present such a model, which defines transitions between states representing the growing interpretation of the sentence. The syntactic rules specify in effect how the semantic content of each word may be added to the current interpretation.

In the first chapter I examine the question of the use of models in linguistics. Accepting the standard arguments for the modularity of the process of language comprehension, I argue nevertheless that a model of syntactic competence is only open to objective evaluation if it is embedded in an overall model of performance. I argue that a dynamic formulation of the competence grammar ensures a transparent relation to what is known about language comprehension, in particular its incremental nature. I argue that the proposed model does away with the need for a level of independent syntactic structure (either constituent or dependency-based), and is thus maximally parsimonious. Syntactic rules in the model make direct reference to the growing interpretation, thus distinguishing the model from other incremental approaches.

In the central chapters of the thesis, I examine a wide-range of syntactic constructions, predominantly from English, exploring the potential for analysis and explanation that is opened up by the change in syntactic viewpoint. Constructions considered include those involving unbounded dependencies, control, coordination, discontinuous constituents, clefts and others. I show that the dynamic model offers a simple and general account of a wide-range of coordination data. I show further that a number of constructions involving discontinuous, and scrambled, constituents do not pose particular problems since the model does not use any notion of syntactic constituency.

Acknowledgements

Thanks to my teachers, especially Ken Turner and Gerald Gazdar who started me off in linguistics, and Elisabet Engdahl who encouraged me greatly along the way.

Thanks to Betty Hughes and Dyane McGavin, plus all the wonderful computing support in the Centre for Cognitive Science.

Special thanks to everyone who has commented on my work or raised points when I have presented it in many locations in Edinburgh, Dublin, Manchester, Cambridge, Grenoble, Moscow and Pécs. Thanks to David Milward, Ruth Kempson, Wilfred Meyer-Viol for more detailed discussion.

Thanks to my supervisors Matt Crocker and Chris Brew, who supported me along the rather indirect path this thesis took.

I consider myself most fortunate to have had Dick Hudson and Mark Steedman as my examiners. They gave me much to think about and much encouragement and I thank them for it.

Thanks to Anna for much.

Thanks to the ESRC for giving me a grant to do this, and my mother and father for giving me the rest.

Thanks to my sons Magnus and Daniel for being such good boys (god dreng og jó fiú respectively).

Finally thanks to the busdrivers of Edinburgh for driving me backwards and forwards these last few months late at night and in the rain, pale and confused (me not them), clutching ever so slowly growing copies of this thesis.

Contents

Abstract	ii
Acknowledgements	iii
0 Introduction	1
0.1 Background	1
0.2 Organization of the thesis	3
1 Linguistic modelling	6
1.1 Input and output	7
1.2 Other knowledge	9
1.2.1 Modularity	9
1.2.2 Incrementality	12
1.2.3 The status of syntactic structure	13
1.3 Evaluation procedures	16
1.4 Linguistic models and linguistic theories	19
1.5 Summary	20
2 A dynamic model of syntax	23
2.1 General characterization	23
2.1.1 Markov processes as models of syntax?	25
2.1.2 Formal power	26
2.2 Interpretation states	27
2.2.1 An example	28
2.2.2 Limitations of the representation	30
2.3 State sequences	31
2.3.1 Assumption of monotonicity	32
2.3.2 A non-graphical representation	33
2.4 Transition rules	35
2.4.1 Accessibility ordering of constituents	36
2.4.2 Lexical entries	37

2.4.3	Structure of transition rules	38
2.4.4	A worked example	38
2.4.5	Restrictions on transition rules	41
2.4.6	Relation to phrase structure	42
2.5	The nature of complementation	43
2.5.1	Clausal complements	44
2.5.2	Governed prepositions	45
2.5.3	The argument–adjunct distinction	47
2.6	Relation to other approaches	47
2.6.1	Categorial grammar	47
2.6.2	Dependency grammars	51
2.6.3	Semantic parsing	53
2.6.4	Processing models: Parsifal & the ATN	54
2.6.5	Hausser’s Left-Associative Grammar	57
2.7	Summary	59
3	Unbounded dependencies	61
3.1	<i>Wh</i> -movement	61
3.1.1	Pied-piping	69
3.1.2	Inter-clausal movement	72
3.1.3	<i>That</i> -trace effect	76
3.1.4	Interrogative clauses as arguments	82
3.2	Relative clauses	84
3.2.1	Relative pronouns	86
3.2.2	Pied-piping	87
3.2.3	The <i>who</i> vs. <i>whom</i> controversy	92
3.2.4	Free relatives	93
3.3	<i>Wh</i> -exclamatives	97
3.4	Multiple- <i>wh</i>	99
3.5	Constraints on movement	101
3.6	Summary	104
4	Complement control	105
4.1	Raising	105
4.1.1	Extraposition	110
4.1.2	Raising adjectives and others	113
4.1.3	Auxiliaries as raising verbs	115
4.1.4	<i>Sounds like</i>	116
4.2	Equi	118
4.2.1	Subject control	118

4.2.2	Object control	120
4.2.3	Small clauses	125
4.3	Passive	128
4.3.1	<i>ing</i> -passives	131
4.3.2	Prepositional passive	132
4.4	<i>Tough</i> -movement	132
4.4.1	Attributive usage	136
4.4.2	Multiple extractions	138
4.5	Summary	140
5	Coordination	141
5.1	Basic coordination	141
5.1.1	Null conjuncts	148
5.1.2	Coordination of unlike constituents	148
5.1.3	Negative conjunctions	150
5.1.4	Correlative constructions	151
5.1.5	Embedded coordination	152
5.1.6	Non-binary coordination	153
5.1.7	Multiple copying	155
5.2	Right-node raising	156
5.3	Gapping	158
5.4	Summary	162
6	Discontinuous constituency	163
6.1	Extrapolation	163
6.2	Optional extraposition	163
6.2.1	Obligatory extraposition	166
6.3	Discontinuous word-orders	168
6.3.1	German	168
6.3.2	Dutch	174
6.3.3	Hungarian	176
6.4	Parenthetical interruptions	180
6.4.1	Slifting	181
6.4.2	Parenthetical coordination	184
6.5	Summary	186
7	Clefts	188
7.1	<i>It</i> -clefts	188
7.2	Pseudoclefts	194
7.2.1	Gapping and pseudoclefts	200

7.2.2	Propositional arguments	202
7.3	Existential <i>there</i>	203
7.4	Summary	206
8	Further topics	207
8.1	Some lexically-specific transitions	207
8.1.1	Subjunctive inversion	207
8.1.2	<i>Why go?</i>	208
8.2	Reduced relatives	210
8.3	Left-recursive structures	211
8.3.1	Clausal subjects	211
8.3.2	Possessives	213
8.3.3	Turkish nominalized clauses	216
8.4	Reduplication in Chinese	218
8.5	Japanese and other causatives	222
8.6	Parasitic gaps	225
8.6.1	Pre-parasitics	226
8.6.2	Post-parasitics	229
8.6.3	Summary	231
8.7	Accounting for processing limitations	231
8.7.1	Summary	236
9	Conclusion	238
9.1	Achievements	238
9.2	Prospects	239
	References	242

0.2 Organization of the thesis

In **chapter 1**, I examine the methodological issues involved in constructing models of language. I follow in the main the standard arguments of Chomsky, especially those establishing the modular nature of the languages process. The conclusions drawn from the modularity argument are quite different, however. I argue that objectively-verifiable models must be holistic models of linguistic behaviour and that therefore models of syntactic competence must be embeddable in models of language processing in order to be evaluable.

In **chapter 2**, I present the dynamic model of syntax being proposed. I assume that the semantic interpretations which the model incrementally constructs may be represented as a network of conceptual constituents (actions, entities, propositions, locations and so forth) which are connected by semantic relations. I argue that to formulate transition rules, unfinished constituents should be ordered on a stack. This has an obvious parallel with the parsing stack used in the processing of standard grammars, but I argue that the use made of it is quite distinct. I illustrate the operation of the model with the detailed derivation of a simple sentence of English.

The following six chapters, constituting the bulk of the thesis, are taken up with investigations into specific syntactic constructions and how they are to be modelled with the dynamic syntax by increasing and modifying the system of syntactic transition rules. Most of the examples are taken from English, although I also look briefly at languages of widely-differing types and argue that the dynamic approach minimizes their apparent syntactic differences.

Adopting a new tack on modelling syntax opens up whole new possibilities for approaching even standard constructions. These chapters are very much an exploration of what choices are available and what their consequences may be. The justification for taking such a broad view of syntactic structures, rather than concentrating on a single area, is that decisions made necessary in one area may have great effect on the conception of the grammar as a whole. My experience is that the analysis of any new construction, no matter how “peripheral”, is liable to bring to light phenomena that lead to revisions of previous assumptions. It therefore seems premature to focus attention on one particular construction or aspect of syntax.

The need for new theories to cover a wide-range of phenomena has been identified by Culicover when discussing the contribution of connectionist models to linguistics.

“...connectionist approaches have been notably deficient in failing to address, for whatever reason, the rich array of phenomena unearthed by linguistic theory in the past forty years. The descriptive successes of modern linguistic theory constitute a minimum criterion of success for any alternative (including, it should be noted, minimalist proposals).” Culicover (1998: 57).

In detail, the material in these chapters is divided in the following manner.

In **chapter 3**, I develop the use of the syntactic feature, *store*, holding any number of constituent values, and use it to tackle a range of unbounded dependency constructions. I argue that pied-piping in *wh*-questions and relative clauses is quite distinct, and that in the latter the relativizer acts in effect like a resumptive pronoun. I also argue that the *that*-trace effect may be given a very simple explanation if we approach it from the dynamic perspective.

In **chapter 4**, I look at a number of constructions which come under the heading of complement control. Much of the basic approach is indebted to recent work in HPSG, especially the insight that the syntax need not (and indeed cannot) impose control relations, but that these must arise out of the meanings of the constructions themselves. I also argue, contra a recent proposal by Grover (1995), that *tough*-movement is an example of an unbounded dependency construction, and attempt to explain the sensitivity it displays with regard to extraction from finite clauses.

In **chapter 5**, I follow the lead of Milward and argue that dynamic grammars provide simple and elegant solutions to coordination constructions, including phenomena that have been highly problematic for constituency-based approaches. In particular, I show that the dynamic grammar can capture the intuition that most coordination involves returning to a previous point in the derivation and continuing the second conjunct from that point. I also look at right-node raising and gapping constructions.

In **chapter 6**, I analyze a range of constructions that have been taken to require discontinuous syntactic constituents. In addition to the familiar cases of extraposition in English, I also take a look at phenomena in German and Dutch that have been problematic for accounts that assume that constituent structure has a direct connection to word order. I argue that the problems caused by the apparent scrambling of syntactic constituents should be solved not by adding to the levels of the grammar, as in some recent proposals, but rather by abandoning altogether the notion of syntactic constituents. I also consider some much-neglected parenthetical constructions which have proved problematic for constituency-based approaches and show how they can be modelled very simply with the dynamic grammar.

Chapter 7 deals with the syntax of cleft constructions. I argue that the *it*-cleft construction involves a relative clause restricting the reference of the dummy *it*, which has been equated via the copular construction with the element in focus. The pseudocleft has long caused problems for generative grammar in that it exhibits connectivity effects that appear to belie its surface structure. I argue that these may be predicted if we adopt a grammar where the only level of representation is that of the semantic interpretation. I finish with a brief proposal for the treatment of existential *there*.

Chapter 8 contains analyses of a range of constructions that illustrate various advantages of the approach being pursued. I claim firstly that certain highly idiomatic constructions may easily be modelled by associating a idiosyncratic syntactic transition directly with a particular word. I look at the analysis of left-recursive structures, which have been held to be problematic for incremental accounts. I argue that the systematic reduplication of verbs in Chinese can most naturally be captured in a model where the interpretation is available to the syntactic rules. I also show that the mismatch between the lexical nature and multiclausal meaning of causative verbs in Japanese and other languages is not in any way problematic if no constituent structure is assumed. I conclude the exploration of syntactic constructions with a tentative analysis of parasitic gaps. Finally in this chapter, I look at the relation between the complexity of states and the known limitations on processing centre-embedded constructions.

One important area, discussion of which unfortunately had to be excluded due to the constraints of time and space, is binding theory. However, a number of recent approaches⁴ have argued that the constraints on the binding of pronominals should be characterized not in terms of syntactic structure, but rather by a combination of position in semantic (or argument) structure, together with considerations of linear order. Consequently, as Kempson has demonstrated, questions of binding make an ideal arena to demonstrate the strengths of a dynamic approach which assumes a single semantic level of representation. This must be left to future research.

⁴For example the HPSG account in Pollard & Sag (1994), and the proposals of Bresnan (1994).

Chapter 1

Linguistic modelling

The central assumption of generative grammar is that it is possible to investigate the nature of language through the construction of formal models. In this chapter I will examine the nature of this modelling of language, asking the questions what we are building models of and what we can learn from them.

I shall adopt the *conceptualist*¹ position, according to which to model language is to model the human linguistic capacity, the make-up and operation of some aspect of the human brain. This is perhaps the dominant position in modern linguistics, established as such in Chomsky (1965), and places linguistics ultimately within the biological sciences. It contrasts with a *realist* or *formalist* approach which attempts to model language as an abstract mathematical construct.²

It is clear that in attempting to investigate the *language organ*, the system in the brain that enables us to use language, we are faced with quite a different task than with the investigation of a physically-discrete and easily-accessible organ such as, say, the kidney. The reasons for this range from the sheer complexity of language behaviour, our lack of knowledge of the faculty of thought with which the language faculty must in some way connect and the practical difficulties involved in investigating human brain function.

Given the inaccessibility of its operation, the language faculty may be seen as a *black box*. To investigate the workings of the box we may construct *linguistic models*, formal systems which share its input and output characteristics.

“Modelling becomes necessary whenever the object of a scientific discipline

¹Or alternatively *cognitive*, cf. Nuyts (1992: 3).

²Some rival generative theories, such as Generalized Phrase Structure Grammar (Gazdar *et al.*, 1985), are avowedly formalist, while Head-driven Phrase Structure Grammar declines to come down on either side of the question, noting that “a successful science does not have to have solved its foundational problems” (Pollard & Sag, 1994: 14). But this appears less a foundational *problem* than a question of foundational *intent*, and surely in building models it is best to be clear about what they are intended to be models of.

cannot be observed directly. The object is frequently compared then to a ‘black box’.., of which one knows only what materials it takes as ‘input’ and what it produces as ‘output’... As it is impossible to dismount the ‘black box’ without interfering with its operation, the only way one can learn about its contents is by constructing a model on the basis of its input and output. One has to form a certain hypothesis on the possible structure of the ‘black box’, the object, and then realize it in the form of a logical apparatus capable of transforming certain materials in the same manner. If the logical apparatus indeed functions like the object, then it is considered an approximation, or a model of it...” Apresjan (1973: 89).

It is as well to bear in mind, however, the limitations that any model will have, especially when we are modelling something as vastly complex as the language faculty. For example, limitations will arise from the fact that our models will not be using the same computational architecture as the brain. There will also necessarily be limitations from the fact that we cannot hope as yet to model realistically human thought and consciousness and neither, therefore, the way in which language interacts with it. We are therefore forced to acknowledge in advance that any model can only hope to be partial in nature.

“A theory, like any model, will have a *limit of verisimilitude*, the extent to which it approaches an exact copy of the real object it models... In human linguistics, our models will, of necessity, fall far short of being exact copies of the people we observe and wish to model. We may develop theories or models with varying limits of verisimilitude for different purposes and we will be interested in clarifying how these different models of the same real objects are related to each other and to the real objects they model.” Yngve (1996: 112).

Nevertheless, the generative enterprise in linguistics is founded on the belief that aspects of human language capability can be modelled with a sufficient closeness that insights may be gained into its nature.³

1.1 Input and output

A crucial decision to take in constructing a black box model is to specify the input and output. By necessity these have to be observable or measurable in some way. It has long been noted that the essential nature of language is to relate sound and meaning.

“*What is language?...* every sentence or word by which we express our

³And also on the belief, not forgetting one of the early driving forces behind the development of generative approaches, in terms both of funding and inspiration, that it could provide tools to perform useful tasks such as machine translation and automatic speech recognition.

ideas has a certain definite form of its own by virtue of the sounds of which it is made up, and has a more or less definite meaning.

The first thing in the study of language is to realize clearly this duality of form and meaning..." Sweet (1900)

It seems natural then to take sound and meaning as the elements that the model relates to each other, the input and output. In point of fact, this leaves us with two possible and distinct processes. Taking sound as input and meaning as output, we have a model of language understanding, while reversing these gives us a model of language production. There is no *a priori* reason to suppose that there is an inherent *reversibility* about these two processes, ie. that we can build a model with disregard for its direction of operation.⁴ In this thesis I shall concern myself with the problem of language understanding, primarily because it is the more accessible of the two processes to direct investigation.

In the broadest conception of the model the input will be a sound wave, but as it is standardly assumed that the *phonological* mapping from sound to strings of words and the *syntactic* mapping from strings of words to meaning are largely independent, it is this latter mapping that will be considered here. The notion *syntax* has been defined in many ways⁵, but the one that fits best with present approach is the functionalist definition given by Martinet.

"Syntax is often conceived as dealing with the combinations of significant units (usually conceived as 'words') in utterances. Functionally, it is best understood as the study of how the hearer can manage to derive a meaning from a succession of significant units, or, in other terms, to reconstruct, beyond the articuli of speech, the experience the speaker wants to communicate." Martinet (1996: 96)

The input, which can be idealized as being strings of words, is clearly observable and therefore unproblematic. There are obvious problems though in designating what the output *meaning* is to encompass, and how it is to be represented. However, it is clear that users of a language will generally agree on the meaning of texts in natural use (for otherwise the essential usefulness of language for communication would be compromised). We should also expect representations of meaning to provide a basis for derived notions such as equivalence, ambiguity, entailment and so forth, which also can be thought of as reasonably stable across language users.⁶

⁴As will become apparent, most linguists building models of language have taken the view that the model should be of a restricted portion of the present black box I am considering, the "linguistic competence", which would be process-independent. I will develop arguments for the necessity of taking a more global view of language modelling in this chapter.

⁵Perhaps the most unsatisfactory definition being the rather circular "the study of syntactic structure".

⁶The problems involved in evaluating the output of the model will be addressed in section 1.3.

It will be useful to compare this mapping from sound to meaning with Chomsky's conception of *I-language*.

“The system of knowledge attained—the I-language—assigns a status to every relevant physical event, say, every sound wave. Some are sentences with a definite meaning... Some are intelligible with, perhaps, a definite meaning, but are ill-formed in one way or another... Some are well-formed but unintelligible. Some are assigned a phonetic representation but no more... Some are mere noise.” Chomsky (1986: 26).

It is clear that Chomsky's I-language has output that goes beyond the meaning of the utterance, to include information about various kinds of ill-formedness of the input. It has been argued consistently by Chomsky that this is evidence that linguistic behaviour is not a homogenous process, but is composed of distinct modules. In the next section I shall examine the implications that such arguments have for the formulation of the model of language.

1.2 Other knowledge

In this section I look at what else we are able to deduce about the syntactic process, in addition to its mapping sound into meaning.

1.2.1 Modularity

The argument made by Chomsky (1957, 1965), and standardly followed by generative grammarians of whatever school, is that it is possible to identify a distinct component in the mapping from sound to meaning, corresponding to the the speaker's *knowledge of the language*. It is further argued that this linguistic *competence* should be seen as quite distinct from linguistic *performance*, defined as the use of the language on any particular occasion.

It is possible to identify two main strands in this distinction between competence and performance. One of these has an empirical content and a genuine significance for the nature of the model to be built and this will be discussed below. The other strand, however, is a statement of mere methodological choice, that when one is modelling a complex process it is sensible to idealize the problem and avoid having to deal with any inconsequential variation. So in modelling human language one assumes one has “an ideal speaker-hearer, in a completely homogenous speech community” (Chomsky 1965: 3). This idealization is purely for convenience, for there is nothing in general to stop us trying to model a particular individual on any particular occasion (for instance, there is nothing to

stop us modelling the language production of someone who is drunk).⁷

The idealization of the subject matter of investigation is familiar throughout the sciences. For example, returning to our earlier biological comparison, when one is interested in the general design and functioning of the kidney, one does not describe in an anatomical textbook any particular individual's kidney, but an idealized typical kidney. This does not stop interest in pathological conditions of the kidney, or variations of the organ among individuals. Similarly it would seem absurd to suggest that such investigations would belong to a different field of scientific enquiry than models of the typical organ.

Evidence for the contentful distinction, that is evidence for the inherent modularity of linguistic processes, can be summed up by observing, with apologies to Tolstoy, that good sentences are all good in the same way, but bad sentences are bad each in their own way. To see this, we may consider the following identity-parade of infamous sentences.

- (1.1) The man a bone to the dog gave.
- (1.2) The man the cat the dog the woman liked chased bit died.
- (1.3) The horse raced past the barn fell.
- (1.4) Colourless green ideas sleep furiously.

The argument goes that of these four sentences (1.1) is the only ungrammatical one. (1.2) is grammatical, in that it appears to follow the rules for constructing English sentences, but it is difficult to process as it exceeds for some reason the human processing capability.⁸ Sentence (1.3), meanwhile, typically defeats our processing strategy, but is understandable at the second attempt. Finally, (1.4) is grammatically sound but the interpretation arrived at conflicts with the way we imagine the world to be.

These are all standard arguments and the standard conclusion from them is that if we are to construct an accurate formal model of the entire process of language understanding it will be constrained to contain distinct components corresponding to syntactic rules or *competence*, processing strategy, processing limitations and the interface with conceptualization.⁹

However, knowing that these distinct modules are present in the mapping from sound to meaning does not mean that we are in a position to model them independently, or to

⁷The same point is made by Wilks in the discussion in Lyons & Wales (eds.) (1966: 94): "I do not see that one need not set out to make a model for ALL performances, for mistakes, squeaks and grunts."

⁸This question will be explored in section 8.7.

⁹A great deal of confusion has been created by using the same word *performance* to refer both to broad language behaviour including these extra-competence factors, as well as to non-ideal behaviour. There is no contradiction in building *idealized* models of performance in the former sense.

restrict our attention to modelling *competence*, as has by and large been the practise in generative linguistics. The problem is that to construct models we need the input and output to be directly observable or at least to be established in some independent manner. The four sentences considered above were extreme clear-cut examples whose badness could be confidently assigned to conflicts with particular modules, but it would not normally be the case that given a bad sentence we would be able to pinpoint exactly where its badness lies.

For example, looking at sentence (1.5) we know that it is uninterpretable, but do we have any intuitions as to why it should be so?

(1.5) He went that orange football.

An initial reaction may be that it is syntactically ill-formed, perhaps for the reason that the words '*that orange football*' form a noun phrase and the grammar stipulates that the verb *go* does not take noun phrase complements. However, in (1.6) the complement of *go* is also a noun phrase '*the other way*'.

(1.6) He went the other way.

This is a possible sentence since the noun phrase is interpreted as a direction, rather than a physical object, and directions can serve as arguments of motion verbs like *go*. So is the original sentence (1.5) conceptually bad or grammatically bad? It seems impossible to decide pre-theoretically either way. The evidence for decisions such as these are not open to inspection: they are buried in the black box. The input-output relation shows clear evidence of modularity, but does not define the architecture of the modules.

Similar arguments could be made for the difficulty of establishing the boundary between processing factors and syntactic ill-formedness. Although we know from the clear cases that there should be a distinction in the model, it will be a theory-internal matter exactly where that boundary is. So any model of competence will always be beset by problems of defining what it should cover.

It has nevertheless been argued that a sufficient source of data for the construction of models of competence is provided by grammaticality judgements. The problems with such judgements have often been noted in the literature¹⁰, for example that they vary greatly between linguists and even vary over time for the same linguist, and that they do not well reflect the essential *gradability* of our impressions.

An even more fundamental concern with such data, however, is that it can hardly be argued to be data purely about the component *linguistic competence*. It is impossi-

¹⁰See Schütze (1996) for discussion.

ble to analyse a sentence without using the other processing components, and without representing the meaning of the sentence internally, both of which will interfere with any judgements as to its soundness.¹¹ This point is made well by Schütze (1996: 26).

“It does not make any sense to speak of grammaticality judgments given Chomsky’s definitions, because people are incapable of judging grammaticality — it is not accessible to their intuitions... Linguists might construct arguments about the grammaticality of a sentence, but all that a linguistically naive subject can do is judge its acceptability.”

This is not to say that acceptability judgements, data about isolated sentences at the extremes of acceptability, even bearing in mind their variable and ephemeral nature, are not without value as diagnostic thought-experiments in constructing the model of competence, which is essentially a speculative enterprise not open to objective evaluation.

1.2.2 Incrementality

In addition to its modularity, the process of language understanding possesses a more obvious characteristic: we know, both introspectively and through experiment¹², that the process of language understanding is incremental in nature, that is interpretation takes place “left-to-right”. As we are trying to build a faithful model of human language ability, we should therefore insist that our linguistic model reflect this by also building its interpretation incrementally.

“... there is every reason for a theory of competence to be straightforwardly related to what is known about human linguistic performance. In particular, it is an advantage in a competence theory if it is clear how it could be turned into a processor which works ‘from left to right’ through the text, and appears to deal with semantics as nearly as possible in parallel with syntax...” Ades & Steedman (1982: 517).

Although it has been standardly assumed in generative grammar that models of competence can be constructed without thought to the constraints imposed by the necessity for incremental processing using that grammar, this seems to be at odds with the conceptualist position, that is that we are directly modelling the operations of the brain. If the brain is able to construct meanings incrementally, then clearly this puts the constraint on the competence module that it must be compatible with this.

¹¹The same point is made in Sells *et al.* (1991: 9) “Although it is standard for generative grammarians to take such judgements [ie. acceptability judgements, DT] as providing especially direct access to some internalized grammar, no justification for this practice has ever been offered.”

¹²The standard reference to support this is the speech shadowing studies of Marslen-Wilson (1973), though it has been confirmed in many studies since.

I will argue in the next chapter that the most straightforward way for the model of competence to comply with the dynamic nature of linguistic understanding, is to model the competence as a dynamic “left-to-right” system.

1.2.3 The status of syntactic structure

Since its inception, the paradigm of generative grammar has been coupled with the use of *syntactic* or *constituent structure*, a formalisation of the earlier immediate-constituent analysis. Over the years there appear to have been differing conceptions as to the status of syntactic structure within generative grammar: the essential question being is it a fact of language that all models will have to account for or merely part of the technical apparatus linguists have employed to account for the mapping between sound and meaning? In this section I shall review the claims.

The adoption of the constituent structure approach to generative grammar was justified by tradition: “Customarily, linguistic description on the syntactic level is formulated in terms of constituent analysis (parsing).”¹³ But it seems clear at this stage that syntactic structure is still considered to be a part of the linguist’s model and not part of the data to be accounted for.

“A linguistic level, such as ... phrase structure, is essentially a set of descriptive devices that are made available for the construction of grammars; it constitutes a certain method for representing utterances.” Chomsky (1957: 11)

By Chomsky (1965), however, constituent structure seems to have become part of the data that every model of language has to account for.

“It seems to me that... [“a traditional grammar’s subdivision of the string into continuous substrings, each of which is assigned to a certain category”] ...is, without question, substantially correct and is essential to any account of how the language is used or acquired.” Chomsky (1965: 64)

There is no evidence given for such a view, however, other than Chomsky’s own personal belief in it.¹⁴ Similarly, Postal (1972: 164) assumes without argument that there are

¹³Chomsky (1957: 26), though even the claim that Chomsky was following a generally accepted tradition is debatable as pointed out in Matthews (1993).

¹⁴With Chomsky’s adoption of the minimalist program, it appears that he no longer holds this belief, holding that S-structure, along with D-structure, have “only theory-internal motivation” (Chomsky, 1995: 191). The guiding ideas of the Minimalist Program are given as follows:

“We take [a particular language] L to be a generative procedure that constructs pairs (π, λ) that are interpreted at the articulatory-perceptual (A-P) and conceptual-intentional (C-I) interfaces... ...under minimalist assumptions... there are no levels of linguistic structure

“three unquestioned levels”, ie. semantic representations, phonetic representations and surface structures.

Surprisingly, given the fundamental nature of the question for generative grammar, the discussion of the status of syntactic structures has been largely confined to pedagogical treatments. Here many writers assume that constituent structure forms part of the data to be explained, rather than part of the explanation. For example, Pesetsky (1996) talks of the “discovery” of phrase structure, as though it were a physical phenomenon, rather than its “invention”:

A related discovery was the realization that the phrasal organization of sentences is hierarchical—generally characterizable by the sort of phrase-structure tree exemplified in ...” Pesetsky (1996: 137)

And in support of this claim he notes:

“The fact that so many phenomena refer to constituency—the grammar of sentence fragments, semantic interpretation, limitations on coreference, and much more besides—serves as evidence that constituent structure is a central property of syntax.” Pesetsky (1996: 137)

But does this amount to evidence that constituent structure is a central property of syntax, or rather that constituent structure is a central property of the model of syntax being proposed?¹⁵ Of course, the onus is placed on competing theories to model the range of data referred to by Pesetsky, but they do not have to account for constituent structure itself, which would be the case if it really were a central property of syntax.

And again Pinker (1994) concludes from a consideration of ambiguous sentences that:

“The two meanings in each sentence come from the different ways in which the words can be joined up in a tree... Phrase structure, clearly, is the kind of stuff language is made of.” Pinker (1994: 102)

Note that the claim is not that phrase structure is the kind of stuff that *models of language* are made of, but *language* itself, and once again the evidence for the claim is circumstantial: the argument runs that models employing constituent structure can relate

apart from the two interface levels PF and LF...” Chomsky (1995: 221)

This may be compared with Henry Sweet’s, slightly more elegant, definition of language quoted earlier in this chapter. The essential question that remains is how this new approach is going to characterize the relation between words and meaning, and it appears that this is very far from being answered in any concrete manner.

¹⁵Pesetsky’s paper comes with a warning that is pedagogical in nature, but that is precisely the problem with grappling with the basics of linguistic theory, that they are generally only spelled out in pedagogical texts.

word strings to meaning, it has not been demonstrated that this can be done in any other way, therefore constituent structures must exist.

In Borsley (1991), another textbook, this argument for constituent structure is set out more explicitly:

“What, then, is the motivation for the assumption that sentences have constituent structures? The most general answer is that without it is more or less impossible to specify what is and what is not possible in a language. This soon becomes clear if we try and specify what is and what is not possible without making this assumption.” Borsley (1991: 14)

This is followed by discussion of a simple model of language which does not make use of constituent structure and is clearly inadequate. Unfortunately for this method of argumentation, one could provide any number of such models without proving the contention that it is indispensable to a successful model and presumably therefore present in the mental syntax.¹⁶

As a welcome contrast, Pollard & Sag (1994) do take great pains to discuss the basic assumptions of the model and clearly point out that phrase structure is an assumed part of their model for which there is no direct evidence.

“Perhaps phrase structure itself ... is the nonobservable linguistic construct that enjoys the widest acceptance in current theoretical work. Surely the evidence for it is far less direct, robust, and compelling than that for phonological structure..., logical predicate-argument structure ... or underlying grammatical relations. But for all that a theory that successfully dispenses with a notion of surface constituent structure is to be preferred (other things being equal, of course), the explanatory power of such a notion is too great for many syntacticians to be willing to relinquish it.” Pollard & Sag (1994: 9-10)

And this is spelt out even more explicitly by Dowty.

“I suspect syntacticians today have almost come to think of the ‘primary empirical data’ of syntactic research as phrase structure trees, so firm are our convictions as to what the right S-structure tree for most any given sentence is. But speakers of natural languages do not speak trees, nor do they write trees when they communicate. The primary data for syntax are of course only STRINGS of words, and everything in syntactic description beyond that is part of a theory, invented by a linguist.” Dowty, (1996)

¹⁶It is strange that no mention is made in such arguments of dependency approaches to syntax, which also eschew constituent structures, or indeed the quite different conception of syntactic constituents assumed in categorial grammar. Both of these will be discussed at the end of the next chapter, section 2.6.

There is no doubting that the idea that words can be grouped together to form phrases has an inherent plausibility. There is a functional pressure on languages to arrange words relating to the same meaningful element close together in the discourse. This is a far step, however, from saying that hierarchical constituent structure descriptions play any part in the workings of the language faculty. Indeed, numerous phenomena in language such as non-constituent coordination and discontinuous constituency pose serious problems for such models. The purpose of this discussion, however, has not been to raise objections to constituent structure, but merely to make it clear that it should not be taken as part of the data that has to be modelled.

1.3 Evaluation procedures

Returning to the quotation at the beginning of this chapter, Apresjan stated that “If the logical apparatus indeed functions like the object, then it is considered an approximation, or a model of it...”. But how do we know if the model is “functioning like the object”? Or alternatively how do we know if one model is functioning more like the object than another? This is important not only if we want to compare two different models, but also if we want to improve the model we have. It is clear, therefore, that we need some procedure for evaluating models.

This is not a new concern, going back to the inception of generative grammar we find:

“One fundamental concern throughout this discussion of linguistic structure is the problem of justification of grammars.” Chomsky (1957: 49)

Chomsky (1957) argued against the possibility of having a *discovery procedure* to find the correct grammar for any given corpus, or a *decision procedure* to decide if a given grammar that correctly models a corpus is the best grammar possible. He did claim that it was possible to come up with an *evaluation procedure* that given two grammars correctly modelling the corpus could decide which one was superior, deciding with recourse to the notion of “economy of description”.

There are two major problems with applying such a proposal to the problem at hand. Firstly, it is abundantly clear, given the earlier discussion of the limits of modelling, that no model can hope to “correctly” model the mapping from words to meaning. So we are unlikely to be in the position of having to decide between competing “correct” models. Secondly, given that we are attempting to model a biological process, there is no reason to assume that *economy* of the model will be the only property to take into consideration. Indeed it is possible that models exhibiting obvious redundancy might have advantages in a biological system. The same point is made by Langacker:

“If claims of psychological reality are taken seriously, questions of economy assume the status of empirical issues, as opposed to methodological ones.” Langacker (1991: 262).

Given that we are modelling language as a black box, it is clear that the model, assuming that it conforms to the known facts of the language faculty such as modularity and incrementality discussed previously in this chapter, can only be judged on the closeness of its output to that of the human. It was stated above that the output of the model is meaning, which may be taken to be some idealized change in the state of mind of the individual directly associated with the string of words. But what do we know of the factors that such a meaning should be composed of? And how can we judge one representation of this meaning as somehow being more correct than another? It is true that to a certain extent we can insist that the representation reflect our own intuitions about synonymy, entailment of strings of words and so forth, but such concerns could still be satisfied by a vast number of competing and differing models.

The standard procedure in linguistics has been to let proponents of different models argue their case on subjective grounds, similar to the *economy* considerations of Chomsky, but it is doubtful if such an abandonment of objective evaluation criteria can lead to progress in science. This point is argued by Yngve.

“So the question... How can we know what to believe? has traditionally been answered in linguistics as follows: We know intuitively what is true and have rationalized those beliefs. If we don't know the answer intuitively ourselves, or are not sure, we take the word of someone who claims to know (a trendsetter). And if authorities differ among themselves, there is no possibility of appeal to an external physical reality to test their claims against. This leaves the matter at an impasse. No wonder linguistics dialog more often resembles verbal jousting for dominance by contending linguists than the discourse of scientists cooperatively trying to search out the truth about the world.” Yngve (1996: 68-69)

With no objective way to test conflicting opinions, we would continue to suffer from polemical warfare and general anarchy in the discipline.” Yngve (1996: 92)

When faced with the problem of making objective assessments of the language abilities of students, language teachers will typically use a *cloze* test. In these tests subjects are given real texts which have had words blanked out and their task is to guess what the original words were. Perhaps surprisingly, scores on such a simple test are held to be a very reliable indicator of the language ability of the subjects:

“...cloze testing is a good indicator of general linguistic ability, including the ability to use language appropriately according to particular linguistic and situational contexts.” Heaton, (1976: 17).

This technique for evaluating human performance has much in common with that for choosing the best probabilistic model of a language as being the one that makes the best prediction of the next word in a text. The best model, for a given corpus, can be objectively given as the one that can best predict the next word, that is the language model that can achieve the lowest cross-entropy estimate.¹⁷

Such an idea is not without precedent in linguistics, indeed it takes us back to the founder of the transformational analysis of language, Zellig Harris, who introduced transformations to capture the invariant collocational relationships between words and so predict whether a word would be more or less probable in a particular position.

“... a central problem for a theory of language is determining the departures from equiprobability in the successive parts of utterances.” Harris (1991: 54).

If we accept the idea that the best model is the one making the best predictions for missing words in text, however, we should be careful not to apply it indiscriminately. It might easily be the case that a linguistically “moronic” model¹⁸ trained on a lot of data might perform better than a sophisticated model trained on a little. However given any linguistically-credible model, any improvement in the syntax should result in an improvement at the level of word prediction.¹⁹

It also allows us a handle on the problem referred to above, that of deciding on an optimal representation for meaning, which given the current state of knowledge about the brain cannot avoid being speculative and partial. It should be true that improving the representation will allow better generalizations to be made and consequently the predictive power of the model to increase. So the specification of meaning is used as a tool to maximize the predictive power of the model, while at the same time measuring the predictive power of the model gives a way of making an objective assessment of the completeness and correctness of the meaning representation. The same point is made by Suppes:

Beginning with a probabilistic grammar, we want to improve the probabilistic predictions by taking into account the postulated semantic structure. The test of the correctness of the semantic structure is then in terms of the additional predictions we can make.” Suppes (1973: 392)

One constraint that such an evaluation metric places on a model is that it be holistic, that is that it model the entire interpretation process. Models of an abstracted linguistic

¹⁷See Brown *et al.* (1992) for discussion of this measure.

¹⁸To quote the description of n-gram models given in Brown *et al.* (1992).

¹⁹Of course this is not to dispute that improved models of knowledge of the world and the processes of reasoning will also have an important impact, perhaps far outweighing that of a good syntax. The point is that given two models with identical non-syntactic aspects, the one with the better modelling of syntax will be a better predictor.

competence, divorced from language use, are not open to objective evaluation and it is therefore unclear how they can be studied scientifically in isolation.

1.4 Linguistic models and linguistic theories

Few linguists would argue with the assertion that:

“A fundamental goal of linguistic theory is to explain why languages are the way they are...” Sag (1991: 70).

However, before one can know *why* something is in a particular state, it is first necessary to know *what* state it is in. Given the assumption that the only way to find out how language works is to build models of its operation, then explanation in linguistics cannot be divorced from the construction of precise models of linguistic behaviour.

Givón (1979), among others²⁰, has criticized the claim that models in linguistics can in any way be considered *explanatory*: “models cannot explain themselves”. He argues that if one wants to know why language is the way it is, one must take its function into account.

“To understand what grammar is, and how and why it comes to be this way, one must make reference to the natural parameters that shape language and grammar: cognition and communication, the brain and language processing, social interaction and culture, change and variation, acquisition and evolution.” Givón (1995: xv)

That is not to say that models of syntax should not be predictive, in the sense that given a novel string of words they should predict the meaning that it might have. And in as much as they reveal something about the object they model, this revelation might play an important part in the construction of explanations. But we should try to avoid the common trap of assuming because we have set up the rules of syntax in such a way that a particular set of linguistic data are modelled, then we have *explained* the data.

So far I have attempted to avoid the term *linguistic theory*. The precise distinction between model and theory in science has been the subject of much dispute and I will not add to it here. Thus far, I have used the term model in a pretheoretic way analogous to its use in a model aeroplane, a construction that has certain characteristics in common with the thing it models. One should certainly expect a linguistic theory to be more ambitious, encompassing restrictions on the class of possible models, as well as explanation of such restrictions.

²⁰Winograd (1977) also discusses “the fallacy of the simplest model”: that is the assumption that the placing of formal restrictions on models corresponds to some sort of *explanation*.

A comprehensive linguistic theory should also explain how language is acquired by children, but again it is a separate, and logically prior, task to know *what is acquired* before working out how this is done, though this is not to deny that it is important to take into account, when constructing a synchronous model, the constraints imposed by the process of language acquisition.

While acknowledging that a future goal must be to delineate the limits of a learnable human language, it can surely only come on the back of knowing how actual human languages work. One can have an accurate model of a particular language without a comprehensive linguistic theory to place it in, but not vice-versa.

It is also clear that functionalism and formal models are not incompatible in that functional explanations can be given for why formal models are as they are. To return to the biological analogy, it would be bizarre not to look for functional explanations to account for the structure of the kidney.²¹

Given the above, it seems that a reasonable way to proceed in linguistics is to construct explicit, objectively-evaluable models of particular languages and then seek to embed them in comprehensive theories providing explanation, which we can expect to be largely functional in nature. This thesis will be largely restricted to the former of these tasks.²²

1.5 Summary

Starting from the standard assumption of generative grammar that one can investigate the nature of human language ability by constructing models of it, I have attempted in this chapter to address the question of how best to go about this. I have identified an interesting process to model as the mapping from strings of words into meaning, and argued that, given that the process is largely inaccessible to investigation, it must be modelled as a *black box*.

It has been well-established through the brilliant foundational work of Chomsky that this process is not homogenous, but contains different facets, one of which is taken to be grammatical knowledge, or *competence*. Therefore, to aspire to verisimilitude a model must reflect these differences in itself. This notwithstanding, I have argued that the success of the model can only be evaluated as a whole, conjectured parts of it (such as linguistic competence) are not open to observation, their boundaries cannot be defined

²¹Commenting on Chomsky's statement that "language is not a task-oriented device", Wilks makes the same point, arguing that this is "all the more strange coming from one who has recently adopted the manner of speaking of the "language organ" and its similarity to other organs of the human body. For to speak of organs... is to speak of their function" Wilks (1987: 9).

²²Indeed, it is even narrower, for it will be largely restricted to considering the kind of competence model that might fit into an objectively-evaluable performance model.

pretheoretically, and any such model is not open to objective evaluation.

With regard to the evaluation of models, I have argued (following Zellig Harris) that a model should have to account for “the departures from equiprobability in language” and that this can only be achieved in an overall model of linguistic performance.

Setting out the task of linguistic modelling in this way would seem to run counter to the demarcation some have made between linguistics and psycholinguistics, namely that “whereas theoretical linguistics limits itself to studying knowledge of language, psycholinguistics aims to provide a complete computational model of the human language capacity and hence encompasses both knowledge and processing of language.”²³ But such a demarcation is argued against by Chomsky himself.

“The study of language is concerned with the system and its use. The linguist is thus concerned with the competence acquired and performance models that incorporate this competence and are concerned with its use. It is impossible to lay down *a priori* conditions as to the points in this complex system at which new insights will arise.” Chomsky, (1975: 197-198)²⁴

Rejecting such a demarcation also avoids the rather odd situation of linguistics being a sub-field of psycholinguistics, rather than vice-versa. Perhaps psycholinguistics would be better defined as a sub-field of linguistics, investigating language using experimental methods typical of psychology.

The idea that models of competence grammar should be shown to be embeddable in models of linguistic performance is one that has been expressed many times before.

“...if the study of language as conceived by the strong mentalist is to be ‘no different from the empirical investigation of other complex phenomena’ (Chomsky, 1965, 4) then a theory of competence is worth taking seriously as an explanatory theory *only if* it can be used in explaining and making predictions about linguistic behaviour. Indeed, the ability to do just that is the ultimate measure of whether the constructs in the linguistic theory are psychologically real. Often those who make this point are charged with misunderstanding the goals of mentalistic linguistics. I plead not guilty to this charge.” Ringen (1980: 106-107)

Any theory of competence grammar that cannot be incorporated into an incremental model of processing has a clear deficiency in terms of psychological plausibility. The only way of demonstrating to the contrary is to actually embed it in such a processing model,

²³Branigan (1995: 9).

²⁴This contrasts with the position adopted by Pollard & Sag (1994: 14): “...we have accepted the conventional wisdom that linguistic theory must account for linguistic knowledge ... but not necessarily for processes by which that knowledge is brought to bear...” Strangely enough, this does seem to be a more conventional position for Chomsky’s followers than the one espoused here by Chomsky himself.

which will then have the crucial scientific advantage of allowing objective evaluation.

Having set out the methodology to be pursued in building models of language, in the next chapter I will address the question of how to construct a model of competence that can be easily incorporated into an overall model of language performance. The following chapters will investigate the application of this model to the syntactic phenomena of English and other languages.

Chapter 2

A dynamic model of syntax

I argued in the previous chapter that a model of language comprehension should map incrementally from words to meaning. I also argued that it should be modular, isolating a system of syntactic rules, independently of strategies for employing these rules, processing limitations, or the interaction with knowledge of the world.

In this chapter I propose such a competence model. Given that it is to be employed in an overall system that operates in an incremental fashion, I have assumed that the competence itself should be modelled as a dynamic, “left-to-right” system.

2.1 General characterization

A dynamic model defines a set of states and a transition relation, specifying which states are accessible from which others.

In modelling syntax we are faced with the task of constructing a mapping between strings of words and their meanings. To cast this as a dynamic model I will take the initial state of each transition to consist of the current word in the input string, w_i , together with the current interpretation of that string, S_i , (assuming as argued in the previous chapter that interpretations are formed incrementally), to form the pair $\{w_i, S_i\}$. The states the transition must map to are the possible next states of the interpretation S_{i+1} .

We may alternatively say that the transition relation R is a three-place relation defining possible triplets $R(S_i, w_j, S_k)$ denoting present interpretation state, present word and next interpretation state.¹

¹I shall use the term *state* as shorthand for *interpretation state*, rather than for the more formal characterization in the model as word-interpretation pair.

So if we have an initial interpretation state S_0 (which for present purposes may be taken to be the null information state) and a string of words $w_1, w_2, w_3 \dots w_n$, the transition relation, applied successively, will define possible interpretations of the entire string S_n in the following manner.

$$S_0 \xrightarrow{w_1} S_1 \xrightarrow{w_2} S_2 \xrightarrow{w_3} S_3 \dots \dots \dots S_{n-1} \xrightarrow{w_n} S_n$$

Since the transition between states does not depend on the previous states in the sequence (this is clearly true from the definition of our transition relation), the model is a first-order Markov process.²

“Consider a stochastic process which moves through a countable set I of states. At stage n , the process decides where to go next by a random mechanism which depends only on the current state, and not on the previous states or even on the time n . These processes are called *Markov chains with stationary transitions and countable state space*.” (Freeman, 1982: 1).

In general we may take the word string to be an entire text, but since I shall be considering in this thesis only sentence internal syntactic processes, the string may be restricted to being a sentence and the state S_n , a possible interpretation of that sentence. As sentences may be indefinitely long, and assuming that all words add some information to the interpretation, it is clear that the interpretations may be indefinitely large, and therefore the set of interpretation states is countably infinite.³

Given these assumptions, it is clear that it is impossible to learn and store the possible triples $R(S_i, w_j, S_k)$ for any natural language, as these will also be countably infinite in number. The transition relation must therefore calculate the next state on the basis of the present state and word. If it can do this the fact that there are an infinite number of states is not problematic.

To illustrate this, consider the analogy of a Markov model with a countably-infinite number of states which performs an incremental arithmetic calculation, with each operation applying to the output of the last one.

$$10 \xrightarrow{+7} 17 \xrightarrow{\times 4} 68 \xrightarrow{-12} 56 \dots \dots \dots 1276 \xrightarrow{-20} 1256$$

Such a model is easy to use, without the need for storing possible transitions, as long as it is straightforward to calculate the next state in the sequence.

²This is assuming that the syntactic system is *static*, in the sense that it does not evolve over time. This is obviously an idealization: a snapshot of the actual evolving system of rules.

³Similarly, it may be argued that we should require the model to deal with unknown words, as humans do, and given that there is no natural limit to the length of words, we may assume that the possible words form a countably infinite set as well.

An arithmetic model such as this one will typically be deterministic, in that only one transition is possible at each step.⁴ This will not be the case for natural language, where global ambiguities are common and locally ambiguous continuations the norm rather than the exception. In general then, at each step the transition relation will provide a number of possible next states. As argued in the previous chapter, the strategy for which of the possible paths to pursue at any particular point (“the strategy of how best to play the game”) should be kept distinct from the “possible moves in the game”, which is what we are considering here.

This distinction between the underlying system of possible transitions and the control structure, the algorithm which decides how the paths are to be traversed, does not tally with the claim made for the dynamic model of Phillips (1996) that “The Parser is the Grammar”.⁵ However, the relation between parser and grammar in the present dynamic model is transparent, since the parser will be the grammar plus some control structure specifying which path to pursue at any particular step. Perhaps it is this desirably transparent relation between the two that Phillips wishes to emphasize, rather than complete identification.

2.1.1 Markov processes as models of syntax?

Modelling the syntactic process as a left-to-right mapping of words into meaning as outlined above has an obvious simplicity, avoiding the need for any level of syntactic structure and tying in with the established incrementality of language interpretation. Nevertheless, at first blush such a model appears to run counter to one of the founding tenets of generative grammar, namely the proof in Chomsky (1957) that left-to-right devices were fundamentally inadequate for the description of natural language syntax. This argument rested on their inability to describe structures involving an indefinite amount of nesting.

Crucially, however, the class of left-to-right devices ruled out by Chomsky had the restriction that the states should belong to a *finite* set (ie. finite-state grammars). As discussed above, the states in the present model are not restricted in this way and any such restriction would be completely arbitrary in such an idealized model of competence. The arguments about syntactic inadequacy, therefore, do not hold for the model being proposed.⁶

⁴If we were to take square roots, however, there would be two possible continuations.

⁵The same argument is made for another dynamic model, the LDS_{NL} of Kempson *et al.*, of which it is said that they are “modelling grammar as a parser” Davidson (1997: 35).

⁶That Chomsky’s arguments have led to the general disregard of left-to-right models, not just finite-state ones, can be seen from the following, far from untypical, quotation from a pedagogical work: “a probabilistic, left-to-right model cannot even generate all of the grammatical sentences of English because we have the possibility of embedding sentences inside of other sentences.” Slobin (1971: 10)

It is clear from context that the author intends the phrase *probabilistic, left-to-right model* to refer to a

Indeed, Chomsky did consider the possibility of linguistic description using such infinite-state left-to-right models. In a footnote, he writes:

“There are so many difficulties with the notion of linguistic level based on left to right generation, both in terms of complexity of description and lack of explanatory power..., that it seems pointless to pursue this approach any further. The grammars that we discuss below that do not generate from left to right [*ie. phrase structure and transformational grammars, DT*] also correspond to processes less elementary than finite state Markov processes. But they are perhaps less powerful than the kind of device that would be required for direct left-to-right generation of English.” Chomsky (1957: 24)

The argument against *non-finite* dynamic models of linguistic competence is therefore not that they are fundamentally inadequate, but that they will necessarily be overly complex and are perhaps over-powerful in language theoretic terms. This entire thesis is intended as a refutation of the charge that such an approach is pointless to pursue. The question of the formal power of the model will be discussed in the next section.

2.1.2 Formal power

The first thing to note about Chomsky’s criticism of the over-powerfulness of linguistically adequate left-to-right grammars is that it is doubtful that he this would recognised this argument as valid today. In more recent writings he explicitly rejects the applicability of formal language theory to modelling human linguistic capabilities. Indeed, from the conceptualist standpoint, questions of formal power in language theoretic terms seem of doubtful relevance. It has long been noted that the actual language processor is a finite-state model, in that there can only be a finite number of distinct states in the human brain. By placing processing restrictions on the infinite-state system modelling competence, one will arrive at a finite state model and the relation between grammar and processor will once again be transparent.⁷

Of course formal language theory is of interest in its own right. Hausser (1992) considers the power of his *Left-Associative Grammar*, a left-to-right grammar, sharing many of the characteristics of the one being proposed in this thesis.⁸ Hausser proves that an unrestricted Left-Associative Grammar belongs to the class of recursive languages. Furthermore, by placing various restrictions on the possible transitions, he reconstructs the

finite-state model (which in any case will allow the embedding of sentences within other sentences), but it is a common inaccuracy that may well have inhibited investigation of the type of model proposed in this thesis. The confusion between Markov models and finite-state models is perhaps not surprising, given the fact that even a defender of finite-state Markov models such as Damerau (1971) conflates the two.

⁷The same argument applies to the practical application of the model in language processing tasks. This is discussed in chapter 9.

⁸An explicit comparison between the two is made in section 2.6.5.

regular, context-free and context-sensitive languages, as well as delineating three separate classes of grammars corresponding to different degrees of computational complexity.⁹ It may well be the case that by making various restrictions on the class of transitions allowed, similar results could be obtained for the present model. This is an interesting question and is certainly worth further investigation.

2.2 Interpretation states

Having giving an outline of the model of syntax as characterizing transitions between states of the interpretation of a sentence, the questions to be addressed in this section are what these states are to represent, and how this information is to be represented. That is I wil start by setting out what we want to get out of the model, before discussing how it can be done.

Bearing in mind the conceptualist stance, we can assume that interpretations represent (aspects of) neurological events in the hearer's brain. The model is then aligned with the assumptions made in the cognitive grammar tradition.

“...we take a sentence (or other portion of discourse) to invoke in the listener a particular kind of experiential complex, here to be termed a cognitive representation...” Talmy (1988: 94).

This is described similarly by Langacker:

“There is nothing inherently mysterious about conceptualization: it is simply cognitive processing (neurological activity). Entertaining a particular conceptualization, or having a certain mental experience, resides in the occurrence of some complex “cognitive event” (reducing ultimately to the coordinated firing of neurons). An established concept is simply a “cognitive routine”, i.e. a cognitive event (or event type) sufficiently well “entrenched” to be elicited as an integral whole. Cognitive grammar embraces a “subjectivist” view of meaning. The semantic value of an expression does not reside solely in the inherent properties of the entity or situation it describes, but crucially involves as well the way we choose to think about this entity or situation and mentally portray it.” Langacker (1988: 6-7).

This approach contrasts with the truth-theoretical approach of formal semantics, the approach to meaning adopted by most formal generative approaches. Interestingly, although categorial grammar has generally been associated with a truth-theoretical seman-

⁹It is interesting to note that Hausser (p.c.) considers one of the main advantages of Left-Associative Grammar its “far superior properties in terms of mathematical complexity”.

tics, Steedman (1996) does concede the necessity for some form of conceptual representation.

”...it is worth noting that an explicitly representationalist theory, involving autonomous psychologically real predicate-argument structure, remains distinctly plausible. The only obvious source for the innate knowledge that the child must bring to language learning has always been some kind of prelinguistic conceptual representation, which may well be identical to some form of predicate-argument structure.” (Steedman, 1996: 107).

It should be noted that this level of conceptual representation is conceived of as being linguistically determined level, rather than being some language-independent “language of thought”.¹⁰ It is thus on a par with the “semantic structure” of cognitive grammar.

“Semantic structure is viewed in cognitive grammar as “conventionalized conceptual structure”, i.e. the form our conceptualizations must assume for ready linguistic expression given the symbolic conventions and resources of the language.” Langacker (1988: 94).

Establishing what it is that we wish to represent is an important initial step, but the question of how to represent such “cognitive routines” or “cognitive events” remains largely conjectural, although certain aspects of it have been explored in the cognitive grammar tradition initiated by Langacker (eg. Langacker 1987, 1991), the work on conceptual structure by Jackendoff (Jackendoff 1990, 1997) and Fauconnier’s theory of “mental spaces” (Fauconnier, 1994).

The paradigm of cognitive grammar has, however, been little concerned with the formal modelling of syntax. Instead of hypothesizing about the nature of such conceptual representations, I shall assume a minimal representation and devote attention instead to the *syntactic process*, that is how such representations relate to the words which invoke them in the language hearer.

2.2.1 An example

Let us take a simple sentence of English, (2.1), and consider what we might minimally expect to be represented in the final interpretation (S_n) of the sentence.

(2.1) He gave the dog a tasty bone today.

It is once more appropriate to take a cue from work done in the conceptualist approach to language.

¹⁰Hudson (1990) takes the same position. See the discussion of Word Grammar in section 2.6.2.

“we will assume that *clauses* are the syntactic unit that codes *mental propositions* and that propositions in turn code *cognized events/states*.” Givón (1990: 515)

Fleshing this out, we could further specify that the proposition refers to some past event and following conventional arguments we could further identify the event as an *action*. Specifically this is an action of *giving* which involves an *agent*, a *beneficiary* and a *theme*. We know further that the three participants identified are singular entities, with the restrictions that the *agent* is a male human, the *beneficiary* a definitely identifiable dog, and that the *theme* is some indefinite bone, which is specified as having the quality of being tasty. Finally, there is a restriction on the action having taken place during the time specified by the period *today*.

I will assume that this information can be represented as a network of conceptual entities, or constituents, connected by semantic relations. This network is represented graphically in figure 2.1.

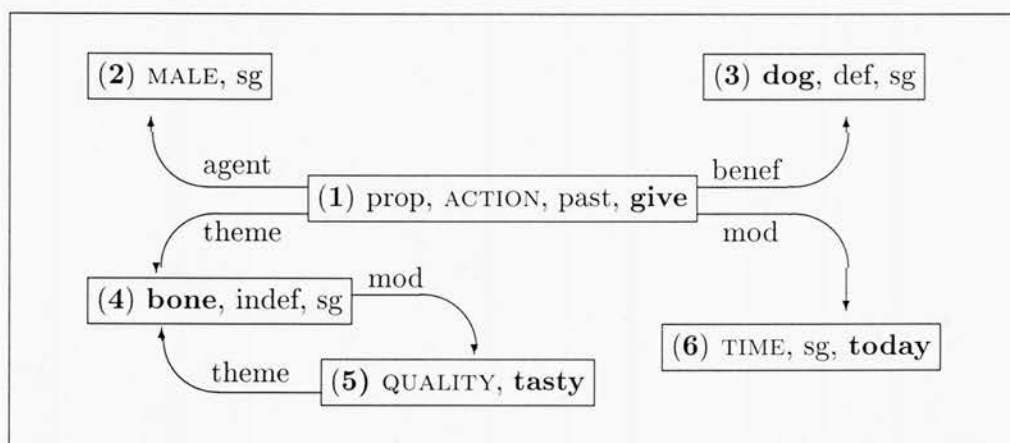


Figure 2.1: **He gave the dog a tasty bone today.**

Such conceptual representations are directly inspired by the work of Jackendoff and others (Jackendoff, 1990). However, no attempt will be made to rival Jackendoff’s work in terms of depth of analysis, since the sole focus of this thesis is on the syntactic construction of such representations, and I will assume that this can be tackled independently of that of the detailed content of the representation.

Six conceptual entities are identified, which I shall refer to as constituents. They are composed of the following features.

- An arbitrary identification number.
- **lexeme value** (in bold).

- **BROAD SEMANTIC TYPE** (in small capitals). It is assumed that the conceptual entities identified in the interpretation are located in some sort of semantic ontology. The broad semantic type given here is a shorthand, indicating the place occupied in the hierarchy, where this is relevant to the syntactic process.
- A set of thematic relations, which take as their values other constituents in the network. I shall assume a fairly standard inventory of roles: theme, benef(iciary), agent, patient, exper(iencer), mod(ifier), possess(or).¹¹
- Plurality, with possible values of **sg** and **plur**.
- Definiteness, with possible values of **def** and **indef**.¹²
- Tense, with possible values of **past**, **pres**, **inf**.
- Whether the constituent is propositional in nature, abbreviated as **prop**.

Instead of representing these feature-value pairs explicitly (eg. type: male, number: sg), I consider it sufficient to give only the values that these features take. More possible features will be introduced as further examples are examined later. It may be possible to place co-occurrence restrictions on features, but this has no direct relevance to the syntactic analysis in this thesis. The information given above could, of course, be expressed in other equivalent representation systems, for instance the PATR-II formalism which has been extensively used in computational linguistics. The main concern in this thesis is to set out the linguistic examples in a way which is as easy to follow as possible.

The ordering of features within a constituent is without significance, except in the case of thematic relations, where the ordering represents an ordering on a list (corresponding to an obliqueness ordering of arguments). Similarly, the layout of constituents in the conceptual graph is not significant.

2.2.2 Limitations of the representation

As noted in the previous chapter any model will have a limit of verisimilitude, and the representation of meaning adopted here certainly has many obvious restrictions, some of which will be discussed here.

¹¹As stated above, I will not be interested in questions of how these roles are established or distinguished, as this has little direct relevance to the syntactic processes being considered.

¹²See below for discussion of reference.

2.2.2.1 Relation to information structure

For the purposes of explicating the syntactic model I shall take no account of the informational structure of the discourse, that is how the conceptual entities identified in the sentence relate to previous information. Similarly no consideration will be made of the question of reference, and the identification of participants.

There can be no doubt however that an embedding within a theory of information structure would be of great value in increasing the predictive power of the model. One only has to consider the improvement in human performance in cloze tests in running text compared to the case when only a single sentence is given. However, such an extension is not central to the aim of this thesis which is to construct a model of sentence-internal syntax.¹³

2.2.2.2 Semantic decomposition

Another self-imposed limitation on the representation of conceptual structure is that I will not attempt to decompose the meaning of lexemes into more elemental constituents of meaning. Much work in cognitive grammar and conceptual structure analysis investigates precisely this issue and, if such decomposition can be shown to give the chance of capturing further generalizations and improving the predictive performance of the model as a whole, it would be useful to incorporate such insights into the model. Once again, however, consideration of such matters should not have too direct a bearing on the analysis of syntactic constructions and will therefore not be pursued here.¹⁴

2.3 State sequences

If we take figure 2.1 to be an acceptable, if primitive, representation of the final interpretation of our example sentence (2.1), the task still remains to determine the sequence of states that were involved in its derivation.

Clearly, for much of the information contained in the final state it is quite straightforward to unambiguously identify the state at which it must have been added. For example the lexical value of words could only reasonably be added at the transitions involving these words themselves: it is not possible to predict that the lexeme **bone** will be part of the interpretation before the word *bone* appears, and if not added then it will be impossible

¹³Similarly, incorporating aspects of *event structure* into the representation would surely be of value, but remains outside the scope of this thesis.

¹⁴It may be that if we introduce into the model a deeper decomposition of the conceptual structure then thematic roles will not be primary relations, but rather derived notions as argued in Jackendoff (1990).

to recover it subsequently, given that each new state is formed only on the basis of the current word and the previous state.

Similar arguments apply for the features *def* and *indef* being associated with the transitions for determiners, and the tense feature *past* with the finite verb.¹⁵

2.3.1 Assumption of monotonicity

One simplifying assumption that can be made is to stipulate that at each transition the new interpretation state is formed only by adding information to the previous one, and that no information is removed. As far as the syntactic constructions considered in this thesis are concerned, I have had no occasion to postulate destructive operations in transitions, and the monotonicity assumption will prove to be very convenient. Not only does it place a restriction on the kinds of operations that transitions may perform, but it also allows us to conclude that there was no information present at an intermediate state that is not present in the final one.

If we explicitly mark each piece of information in the final interpretation of our example sentence with a numerical superscript indicating the number of the state at which it was added, it will give us a final interpretation as in figure 2.2.

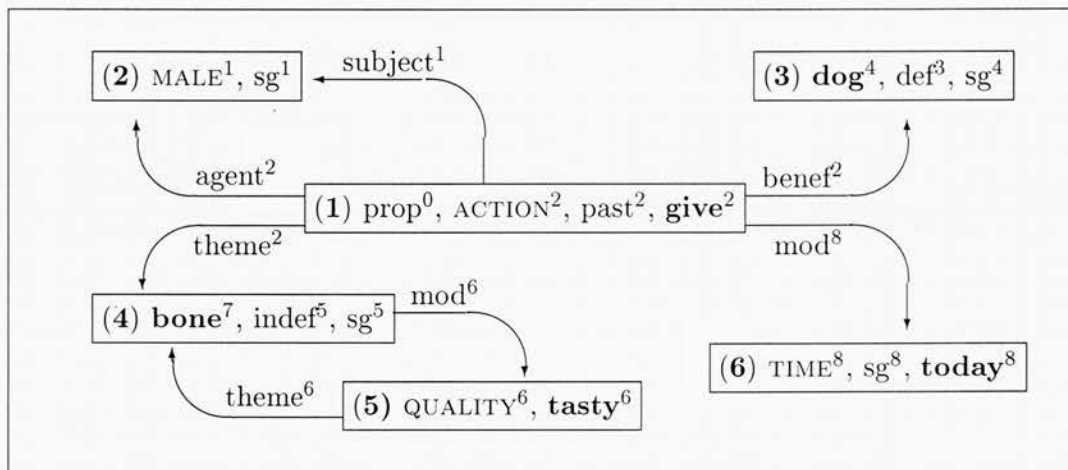


Figure 2.2: **He gave the dog a tasty bone today.**

Each of the intermediate states in the derivation is now directly recoverable from this final state by removing any information added at subsequent states. The superscript “date-stamping” borne by a constituent is used here merely as a device to allow the final

¹⁵There do remain pieces of information, however, for which it is not immediately obvious at which state they were added. The only way we can decide in such cases is to look at the system of transitions as a whole and determine how the various possibilities will affect the formulation of transition rules.

interpretation state to represent unambiguously the entire derivation—these superscripts will play no part in the syntactic transition rules to be introduced in this chapter.¹⁶

Aside from the dating of the information, another innovation in figure 2.2 is the relation **subject** which specifies the surface subject of a constituent. This relation makes no contribution to the meaning of the sentence and is necessary for purely *syntactic* reasons: the constituent in subject position cannot be given its thematic role until the main verb is reached, but it must nevertheless be connected as each state must be a connected graph.¹⁷ Another such syntactic feature, *store*, will be introduced for the analysis of long-distance movement in chapter 3.

Interpretation states thus contain a mix of syntactic and semantic features, the former making no contribution in the final interpretation of the sentence, but playing a role in the construction of the interpretation, ie. playing a role in the syntactic process.¹⁸ For other languages we will have to extend the stock of syntactic features to include grammatical gender, syntactic case and possibly others.

The alternative to allowing purely syntactic features in the representation of meaning would be to separate the two types of feature out into different levels. However, it is clear that semantic features do play a part in the syntactic process and thus would have to be duplicated if we had a separate syntactic level of representation. It has often been noted that syntactic constituents and conceptual constituents have a close correspondence.¹⁹ Instead of having such duplication of constituents, the model proposed here opts for a single level composed of conceptual constituents which may carry semantically empty features which are present merely as an aid in the construction of the representation. Syntactic features could be thought of as the scaffold which allows the meaning to be constructed, but is not itself part of the interpretation.

2.3.2 A non-graphical representation

The graphical network diagrams were used above to give a clear picture of the approach, and avoid confusion with syntactic structures. However, as such diagrams become more difficult to reproduce and read as the examples get more complex, I shall switch to a non-graphical representation of the same information, which can be seen in the following

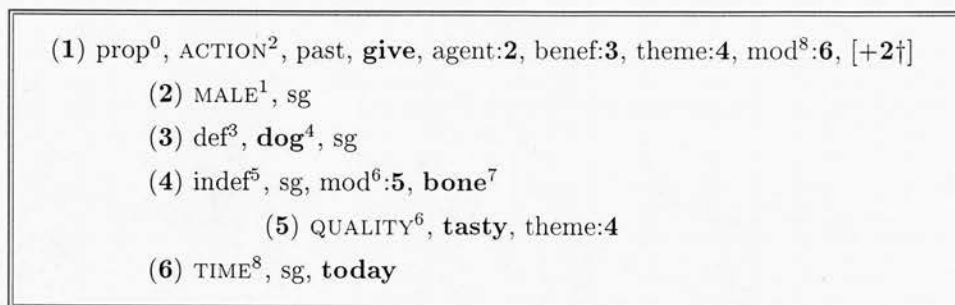
¹⁶However, it might be speculated that this date-stamping may be used to weight the probabilities of modification and complementation with respect to the distance in time between the modifier and the modified. This will then allow the modelling of such processing effects as heavy-element shift.

¹⁷Similarly for Word Grammar, Hudson (1984: 162) argues the need for the relation *subject* to be present at semantic, as well as syntactic, structure.

¹⁸While *syntactic* features make no semantic contribution, *semantic* features will in general make syntactic contributions, for example the feature **number** contributes both to the meaning and is involved in the syntactic operation of agreement.

¹⁹See the discussion of the arguments for constituent structure in Zwicky (1978): “constituents tend to have semantic unity”.

derivation diagram.²⁰



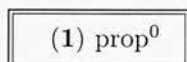
(2.1): **He gave the dog a tasty bone today.**

The main innovation here is in the representation of constituent-valued features, such as thematic relations. Instead of arrows pointing to the constituent itself, the *address*, ie. the identification number, of the constituent is indicated after the feature which takes it as its value.²¹ The syntactic relation of *subject* is represented by + and the feature is placed on another data-object, *store*, belong to the constituent, a stack which holds the addresses of semantically unintegrated constituents. We shall have much greater use of this stack in the following chapters, but for present purposes it merely holds the subj (+) feature. We will explicitly show that such constituents have been semantically interpreted by adding the sign †.

The constituents are arranged as far as possible to reflect the semantic dependency relations between them, ie. complements and adjuncts of a constituent are listed indented below it. It should be emphasized that this arrangement of the constituents is made purely to ease viewing (and will in any case be impossible where a constituent depends on more than one constituent, as will be the case, for example, with relative clauses). Where a constituent is not semantically dependent, it will be positioned according to its syntactic dependency.

From the above derivation it follows that the initial sequence of states was as follows:

S_0



²⁰This representation has the additional advantage of being essentially the same as that used in the computational implementation, which is outlined in the appendix.

²¹Superscripts indicate date-stamping as before. As an abbreviatory convention, all information which shares the date-stamping of a constituent to its left is unmarked. Similarly, a constituent address which is added as soon as the constituent itself is formed is also left unmarked. So the first constituent above represents the following fully-dated constituent:

(1) prop⁰, ACTION², past², **give**², agent²: 2², benef²: 3³, theme²: 4⁵, mod⁸: 6⁸, [2¹†²]

S_1 : He...	<div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>(1) prop⁰, [+2] (2) MALE¹, sg</p> </div>
S_2 : He gave...	<div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>(1) prop⁰, ACTION², past, give, agent:2, benef: , theme: , [+2†] (2) MALE¹, sg</p> </div>
S_3 : He gave the...	<div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>(1) prop⁰, ACTION², past, give, agent :2, benef:3, theme: , [+2†] (2) MALE¹, sg (3) def³</p> </div>

It may be argued that the sequence of states involved in the derivation of an interpretation has an independent validity in that it is potentially open to direct investigation. This does not appear to be the case for the transition rules which allow the derivations—these rules could be expressed in various ways and yet if they allowed equivalent derivations, then there is no way of choosing one formulation over another without appealing to vague notions such as simplicity and elegance.

There will be a concentration in this thesis, therefore, on establishing the sequence of states, given a final interpretation, and less on the questions involved in establishing one particular formulation of the transition rules which will generate such sequences. It may be argued that the sequence of states in the derivation of the meaning of a sentence essentially characterizes the syntax. However, to have a predictive model providing analyses of unseen texts it will be necessary to specify the transition relation between states and this will be the subject of the next section.

2.4 Transition rules

Transition rules in the model are essentially syntactic rules: routinized operations on the growing conceptual structure. Typically they will add the semantic content associated with the current word onto this structure. It has been argued that meaning-constructing operations exist on a continuum of automaticity, with the type of operations in the model of syntax being towards the automatic end of this continuum, what Givón (1989: 248) terms the “automatic language processing mode”. Or similarly:

“We assume that language comprehension is mediated by a stable set of highly skilled, automatized processes that apply obligatorily to their characteristic inputs... The function of this set of core processes is to project the speech input onto a representation of the world.” Marslen-Wilson & Tyler (1987: 58)

There are clearly situations where conceptual meaning is constructed by processes that are less automatic, as in working out the meaning of cryptic newspaper hoardings or interpreting non-standard varieties of English, such as Elizabethan poetry. This is not to deny that the end result of such a process, the interpretation, is essentially the same as that of a normal sentence, merely that the process of interpretation may have used non-syntactic, that is non-routinized means, to arrive at this interpretation. For the purposes of this thesis I shall assume that we are interested in modelling the routinized operations, that is productive syntactic rules.

From a developmental perspective, it seems plausible to take conceptual representations as primary and the entrenchment of rules involving manipulations of these representations as developing subsequently. The task of learning language then is to pair representations with sequences of words that describe them, establishing syntactic rules as repeated typical manipulations.

2.4.1 Accessibility ordering of constituents

A central problem facing any attempt to specify rules for how information may be added to states is that, as detailed previously, states belong to a countably infinite set, for they may have any number of constituents. To be able to formulate transition rules it is clear that states must be partitioned into a manageably small number of equivalence classes.

One obvious way to do this is to add a structure to the state, dividing the constituents into those that are still potentially active, in the sense that information may still be added to them, and those that are inactive (or inaccessible). The active constituents will consist of those that are unsaturated, ie. those with argument slots still to be filled, plus those which were altered or added at the previous state. In addition, those constituents which have dependencies to such constituents (that is, their *parents*) will in general also be active.²²

Furthermore, these active constituents can be arranged in order of accessibility and a natural data structure to hold these active constituents is a stack. The top element of the stack may be termed the *active constituent* and it is this which transition rules will in general add information to.²³

In many respects, this corresponds to the *parse stack* in a processing account. However, the identification of the two is far from being complete. In the present dynamic model of competence the constituents are conceptual, and they are at all times linked together in

²²Although there are exceptions to this generalization, for example the parenthetical constructions discussed in 6.4.2.

²³Although this is not to rule out the possibility that information in other constituents may not be accessed in constructing the new state, but in this case it will be accessed *through* the active constituent.

a network. The constituents on a parse stack, assuming a standard constituent structure grammar, are unlinked syntactic constituents and crucially they *combine*, according to the grammar rules, to form new constituents. No such combination takes place with the conceptual constituents in the present model.

Referring back to the derivation of our example sentence, we may trace the following states of the active stack (with the leftmost element on the stack being the topmost, ie. the active constituent). It is assumed that all new constituents are initially placed on the stack, but if they are saturated they may be immediately removed from it. The constituents removed at each transition are underlined in the diagram below.

[1]	<u>[2,1]</u>	[1]	[3,1]	<u>[3,1]</u>	[4,1]	<u>[5,4,1]</u>	<u>[4,1]</u>	[6,1]
He	gave	the	dog	a	tasty	bone	today	

Before we look at how rules are formulated to add information to the active constituent, we must consider the types of information to be added, which will be found in the lexical entry for each word.

2.4.2 Lexical entries

An essential resource for the transitions must be information about particular words, which is accessed when they are found, or hypothesized, in the input string. The nature of lexical entries in the present model differs little from that in other generative approaches, although there is a difference in the way this information is used.

Entries contain information on how the word may allow the information to be added to the present information state (ie. *syntactic* information), as well as the informational content itself (ie. *semantic* information). Both will be represented by sets of features, with all lexical entries being of the following form.²⁴

wordform {Syntactic type} [SEMANTIC CONTENT]

The features comprising the semantic content have already been seen in the interpretation of our example sentence. Syntactic features include a basic **wordclass** taking the values *verb*, *noun*, *adjective* and *determiner*. Other features are pertinent only to certain wordclasses, such as tense, countability, case and so forth.

As an illustration, we may assume the following lexical entries for the words occurring in the example sentence.

²⁴Questions of morphology will be ignored here as having no direct bearing on the formulation of the syntactic process.

<i>he</i>	{pronoun, nominative}	[MALE, sg]
<i>gave</i>	{verb, past}	[ACTION, give , ⟨agent, benef, theme⟩]
<i>the</i>	{determiner}	[def]
<i>dog</i>	{noun, count}	[dog , sg]
<i>a</i>	{determiner}	[indef, sg]
<i>tasty</i>	{adjective}	[QUALITY, tasty , ⟨theme⟩]
<i>bone</i>	{noun, count}	[bone , sg]
<i>today</i>	{noun, noncount}	[TIME, sg, today]

It should be made clear that wordclasses such as noun and verb are not included as part of the interpretation. They are only present in the lexicon, where they may be referred to by transition rules.²⁵

2.4.3 Structure of transition rules

As outlined above transition rules determine how the semantic content of the present word may be added to the current interpretation. This content will typically be added or attached to the *active constituent*. I will assume that these transitions may be expressed as prototypical transition schemas, which refer to aspects of the active constituent and the lexical entry of the current word.

I shall also assume that transitions may take place without reference to the current word, so that an actual word-to-word transition will consist of multiple sub-transitions. In this way an indefinite number of new constituents may be added to the the active stack in a single transition.²⁶

After each transition is performed any number of active constituents may be removed from the top of the stack, a process which I shall refer to as *rationalizing* the stack. Generally, unsaturated constituents will not be removed from the active stack.

2.4.4 A worked example

Returning to our example sentence (2.1), let us consider the steps in its derivation in detail, state by state.

((2.1) He gave the dog a tasty bone today.

²⁵Although that is not to say that they could not be added as syntactic features if this was necessary for the characterization of the syntactic process.

²⁶The number of new constituents at any transition will correspond to the increase in depth or *rank* (to use Jespersen's term) of the new construction, and in practice is unlikely to exceed four. For discussion of situations where it may be argued that the number of new constituents may be non finite, see the discussion of left-recursion in chapter 8.

At each transition, the present state will be given, showing the division of constituents into the active stack (on the left, with the active constituent uppermost), with the inactive constituents on the right.

S_0 : [1, prop]

he Content added as subject.

S_1 : [1, prop, [+2]] [2, male, sg]

gave Content added to proposition. Subject interpreted as first argument.

S_2 : [1, prop, action, gave, past, ag:2, benef: , theme: ,[+2†]] [2, male, sg]

the Content added as argument.

S_3 : [3, def]
[1, prop, action, gave, past, ag:2, benef:3, theme:] [2, male, sg]

dog Content added to active constituent.

S_4 : [1, prop, action, gave, past, ag:2, benef:3, theme:] [3, def, sg, dog]
[2, male, sg]

a Content added as argument.

S_5 : [4, indef, sg] [3, def, sg, dog]
[1, prop, action, gave, past, ag:2, benef:3, theme:4] [2, male, sg]

tasty Content added as modifier. Active constituent interpreted as theme.

S_6 : [4, indef, sg, mod:5] [5, quality, tasty, theme:4]
[1, prop, action, gave, past, ag:2, benef:3, theme:4] [3, def, sg, dog]
[2, male, sg]

bone Content added to active constituent.

S_7 : [1, prop, action, gave, past, ag:2, benef:3, theme:4] [4, indef, sg, mod:5, bone]
[5, quality, tasty, theme:4]
[3, def, sg, dog]
[2, male, sg]

today Content added as modifier.

S_8 :	[1, prop, action, gave, past, ag:2, benef:3, theme:4, mod:6]	[6, time, today] [4, indef, sg, mod:5, bone] [5, quality, tasty, theme:4] [3, def, sg, dog] [2, male, sg]
---------	--	---

It will be seen that the formulation of transition rules in the present model corresponds closely to the supposition made in Crain & Steedman (1985) that rules of syntax “describe what a processor does in assembling a semantic interpretation”.

Of the eight transition rules employed in the derivation, six of them may be thought of as distinct. These are represented and named in the following table. In the schemas **X** stands for an eXisting constituent (which will generally be the active constituent), and **N** stands for a New constituent (the semantic content of the lexical entry of the current word). $X \cup N$ represents the unification of the two sets of features.

The first two rules in the following table are those that do not add a new constituent, but only add info to the active constituent.

transition rule	transition	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
add-fin-verb	<i>gave</i>	{verb, fin}	(X, prop, [$+\alpha$])	($X \cup N$, prop, arg1: α , [$+\alpha^\dagger$])
add-head	<i>dog, bone</i>	{noun}	(X)	($X \cup N$)

The following rules are those that add new constituents in various relationships with the active constituent.²⁷

transition rule	transition	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
add-subj	<i>he</i>	{nominative}	(X, prop)	(N) (X, prop, [$+N$])
add-arg	<i>the, a</i>	{det}	(X, arg:)	(N) (X, arg:N)
premodify	<i>tasty</i>	{adj}	(X)	(N, arg1:X) (X, mod:N)
modify	<i>today</i>	{adv}	(X)	(N) (X, mod:N)

There will be various further conditions on the rules applying. For example, subjects must not be accusative, arguments not nominative, and modifiers must be of an appropriate semantic type. Similarly, the rule for adding finite verbs must check the number

²⁷The question of how many distinct rules there are, for example should *add-subj* and *add-arg* be merged into one rule, or *modify* and *premodify*, is a matter for the implementation, but not of direct relevance to the operation of the syntax. Indeed an argument could also be made for merging *modify* and *add-arg*, leaving in the end, perhaps, one rule for adding new constituents and one rule for adding information to a constituent.

and person of the subject to ensure agreement, although I shall not consider questions of agreement in this thesis.

Although I will not spell out all the conditions on the rules in detail, it may be instructive to consider a negative example, that is why an ungrammatical sentence is not given an interpretation. Consider the following partial sentences.

- (2.2) a. *Gave he...
 b. *The tasty gave...

There is no way that **add-fin-verb** can apply for the verb *gave* in (2.2)a as there is no subject in position to be interpreted as demanded by the rule. In (2.2)b the initial constituent started by *the* and further modified by *tasty* cannot be removed from the stack as it is incomplete, missing its head. Consequently, the proposition is not exposed and again the rule **add-fin-verb** cannot apply. In this way we can see that many of the constraints imposed by syntactic constituent structure are paralleled.

It should be stressed again that although the sentence is unambiguous, the syntactic rules will not derive it a deterministic fashion. They will allow local ambiguities at many of the transitions.

2.4.5 Restrictions on transition rules

Although I have given examples of schematic transition rules in the previous section, I have made no attempt to restrict the form that these rules may take.²⁸ It would be expected of a theory of syntax that it specifies what form these rules may take for languages in general.

One response to this is to argue that the formulation of such restrictions is premature. I am presenting a particular *model* of (some aspects of) syntax in this thesis and am not presenting a true *theory*.

Another response is to argue that as the model is meant to be taken as psychologically real then the only limits that should be placed on possible rules are that they must be psychologically plausible. But the whole question of syntactic rules is so much a theory internal matter and so far from psycholinguistic investigation that judging what may or may not be plausible is pure speculation.

Be that as it may, the rules I present in this thesis generally add information to the active constituent and sometimes dependents of that constituent. They may also be

²⁸Apart from the earlier assumption that transitions must be monotonic, that is that they can only add information.

sensitive to whether a particular constituent is or is not on the active stack, or has or has not been semantically interpreted. It may perhaps be useful to place restrictions, dividing rules up into those that only refer to the active constituent, and those that may take account of more global information and this is an area that would benefit from further investigation.

2.4.6 Relation to phrase structure

It will be clear from the above that in the derivation of the interpretation of the sentence no use is made of constituent structure, which may be defined as a complete bracketing of the words in the sentence. It is possible, however, to define via the syntactic derivation a derivative notion of syntactic phrase as *the span of words being processed while a particular conceptual constituent is on the active stack*.

Using this definition the phrase corresponding to each of the six constituents in our example can be calculated as follows, adding in the third column the traditional syntactic type given to such phrases.

constituent of interpretation	associated phrase	syntactic type
(1) ACTION, prop, past, give	<i>he gave the dog a tasty bone today</i>	S
(2) MALE, sg	<i>he</i>	NP
(3) dog , def, sg	<i>the dog</i>	NP
(4) bone , indef, sg,	<i>a tasty bone</i>	NP
(5) QUALITY, tasty	<i>tasty</i>	AP
(6) TIME, sg, today	<i>today</i>	NP

It should be noted that none of the constituents corresponds to a verb phrase and also that there are in general no phrases lower than maximal projections, eg. in the phrase *a tasty bone*, there are no sub-phrases such as *a tasty* or *tasty bone*.

The definition of phrase given here does not stipulate that the span of words constituting a phrase must be contiguous. As we shall see in chapter 6, the analysis of discontinuous constituents involves replacing deactivated constituents back onto the active stack. Therefore in a sentence such as *a dog came in which had no nose* the phrase corresponding to the constituent [indef, sg, **dog**] will be the words *a dog which had no nose*.

It has been argued that one of the grounds for phrase-structure based analyses is that language-users do have some concept of the division of a sentence into phrases. This would follow just as well from a grammar where phrase boundaries were marked by moving conceptual constituents on and off the stack, as well as in a system with explicitly represented constituent structures. Similarly the assumed link between constituent and intonational

structure²⁹ might also prove to have a reflection in the present model, though again this requires further investigation.

2.5 The nature of complementation

The only words taking complements in the example derivation above were *give* and *tasty*, whose lexical entries contain the ordered lists of complements ⟨**agent**, **benef**, **theme**⟩ and ⟨**theme**⟩ respectively. The ordering relates to the canonical order in which the arguments are filled.³⁰

It is also necessary to place some restrictions on what kind of constituents can fill these argument roles. In general, such restrictions would seem to come under the interface with world knowledge, and so outside of the syntactic rules that we are considering here (although as argued in the previous section the border is not easy to define pretheoretically). It is important though to consider how this selectional information can be combined with the syntactic model. We may assume that each role is restricted to a certain sub-type of constituent identified in the overall semantic ontology. So for the argument frame of *give* we might have a further specification on the thematic roles as follows.

give ⟨ **agent**: ANIM, **benef**: ANIM, **theme**: THING ⟩

All of these arguments will typically be filled by constituents headed by a noun so that the restriction on complementation given here would be almost equivalent to a syntactically-defined argument frame such as [NP, NP, NP]. But consider the complements of a verb like *put*, which we might assume to be:

put ⟨ **agent**: ANIM, **theme**: THING, **goal**: LOCATION ⟩

In a model using syntactic restrictions on arguments, *put* might typically be assigned a subcategorization frame of [NP, NP, PP], perhaps with additional semantic restrictions on what these constituents may represent. But such syntactic restrictions appear to be of doubtful validity, for although the prepositional phrase is perhaps most typical for the **goal** argument, this is only due to the fact that a prepositional phrase is the most typical realisation of LOCATION.³¹ As the following examples show, LOCATION could also be realised by adverbial, nominal or even adjectival phrases.

²⁹For example, Steedman (1991).

³⁰For a discussion of argument filling in non-canonical order see the analysis of *heavy-constituent shift*, which is presented in chapter 4.

³¹The third argument of *put* could also be realised as a PATH as in:

(2.3) He put the plates onto the table / away / further toward the window.

This could be captured either by a separate lexical entry, or preferably by stipulating that the argument should be superset of LOCATION and PATH, but I will not go into this question here.

- (2.4)
- a. I put the plates on the table
 - b. I put the plates right there.
 - c. I put the plates the same place I always put them.
 - d. I put the plates too high for him to reach.

Similarly, despite the prototypical restriction of surface subjects being headed by nouns (expressed in the syntactic rule $S \rightarrow NP VP$), subjects seem not to be restricted by syntactic category.

- (2.5)
- a. That he left surprised me.
 - b. Sometimes is too often for her mother.
 - c. On the wall seems a good place to hang it.

And conversely, adjuncts are not restricted to adverbial or prepositional phrases, but are often realised nominally as here.

- (2.6) I did it the same way this summer.

The conclusion is that, in general, restrictions on arguments cannot be captured well by placing restrictions on the syntactic category of the constituents that may fill them. They are part of the interaction between syntax and world-knowledge and are best modelled with reference to (soft, ie. probabilistic) restrictions on semantic types.

In recognition of these facts, Chomsky (1986: 87-90) assumes that lexical heads *s-select* arguments of a particular semantic category C, and that the syntactic category of that argument is independently stipulated as one that is the “canonical structural realization of C”. The result is the same as is assumed here: that lexical entries only have to specify the semantic category of their arguments. But in the present model there is no choice in the matter, since syntactic constituents do not exist, and so it can be claimed to be more parsimonious.

2.5.1 Clausal complements

In this section I will address the question of whether there are lexical heads which stipulate that their arguments should be clausal. From the above arguments it should be sufficient for the lexical entry of the verb *think*, for example, to indicate the required semantic type of its theme, that is that it should be a proposition.

think ⟨ **agent:** HUMAN, **theme:** PROPOSITION ⟩

This allows the following distinction to be made.

- (2.7) a. He thought he was funny.
 b. *He thought the problem.

Any purely syntactic restriction on the complement *think* would have problems with the fact that it also allows non-clausal arguments, such as the nominal phrase headed by *thing* in 2.8, which appears to act as a “pro-proposition”.

- (2.8) He was thinking the same thing I was.

So once again it seems problematic to restrict subcategorization syntactically.

2.5.2 Governed prepositions

There are, however, instances of seemingly arbitrary *syntactic* (ie. non-meaningful) restrictions on complements. For example, it is often the case that complements are not just selected according to their conceptual meaning, but in addition they must be preceded by a particular preposition, which retains very little or none of its meaning, as in the non-literal reading (ie. \approx *supported*) of (2.9).

- (2.9) He always stood by me.

The question then arises how such a restriction on complementation is to be expressed in a system based on relations between conceptual constituents.

Let us suppose that there is a lexical entry for the word *stand*, in addition to its literal meaning, of the following type.

stand [ACTION, **stand-by**, < **agent**: ANIM, **benef**: ANIM [lex: by] >]

Here the benefactive argument includes the extra restriction that it is introduced by the lexical item *by*. I shall then assume that the *by*-phrase forms a distinct constituent, whose semantic content is identical with that of its own argument.³² These constituents, contributing only the value of their lexical head, may be thought of as being bleached of conceptual content, allowing the following analysis.

³²This is much the same as the HPSG analysis of these prepositions in Pollard & Sag (1994), where the CONTENT value of the preposition is structure-shared with that of its complement.

- (1) prop⁰, mod²:3, EVENT³, past, **stand-by**, theme:2, benef:4, [+2†]
 (2) MALE¹, sg
 (4) **by**⁴, content:5
 (5) 1ST-PERSON⁵, sg
 (3) TIME², **always**

(2.9): **He always stood by me.**

Given the assumptions of the present system, an alternative might have been to treat the prepositions as features on the conceptual argument itself. However, one would then be forced to make arbitrary decisions about at which point prepositions become non-meaningful and stop forming separate conceptual constituents. The analysis given here allows for a continuum of semantic bleaching.³³

A similar approach is taken with the analysis of phrasal verbs taking particles as in the idiomatic reading (ie. \approx *rebuked*) of (2.10):

(2.10) She told him off.

Here we may analyse the lexical entry as involving three arguments with the last one being a relation having no *semantic* content, but simply specifying that its lexical head must be *off*.

tell [ACTION, **tell-off**, < **agent**: ANIM, **patient**: ANIM, \emptyset : [lex: **off**]]

Thus the sentence will receive the following analysis, where the constituent headed by *off* is entirely devoid of semantic content.

- (1) prop⁰, ACTION², past, **tell-off**, agent: 2, patient:3, \emptyset :4, [+2†]
 (2) FEMALE¹, sg
 (3) MALE³, sg
 (4) **off**⁴

(2.10): **She told him off.**

Once again the arguments for retaining such a constituent is both for uniformity and to avoid having to choose an arbitrary cut-off point in the continuum of semantic bleaching.

³³As will be seen in the later treatment of unbounded dependencies, the analysis also allows a uniform treatment of such constructions as “Who did he always stand by?”. Note incidentally that in the sentence “By whom did he always stand?”, the literal meaning of *by* comes to the fore to such an extent that the idiomatic meaning of the construction is lost.

2.5.3 The argument–adjunct distinction

It will be noted from the examples given so far, that the difference between arguments and adjuncts in the model is not based on structural differences, or indeed semantic ones, but is rather one of expectation. Arguments are essential to the meaning of a constituent and must be filled.³⁴ Adjuncts are not anticipated in the same way, even though they may still comprise the same relation. For example, the same *duration* constituent is found in the following two sentences, the first being an argument, however, and the second an adjunct.

- (2.11) a. The play lasted for two hours.
 b. He read the book for two hours.

In the first sentence it is essential to the meaning of the verb and is anticipated as an argument, while in the second it is not anticipated even though the constituent and its relation to the event is identical.³⁵

2.6 Relation to other approaches

In this final section, I will attempt to relate the model of syntax that I have presented in this chapter with other approaches to syntax, particularly those which also reject the orthodoxy of constituent structure approaches.

2.6.1 Categorial grammar

The syntactic system of categorial grammars includes a small number of basic syntactic types, rules for deriving more complex types from these (type-raising), and a small number of rules specifying the possible combination of these types. These syntactic types are assigned to words in the lexicon, the theory consequently being highly lexicalist, and the syntactic rules determine how the semantics of each word may combine to give an interpretation for a complete sentence.

The specific categorial system that I shall consider is Combinatory Categorial Grammar (CCG), a variant developed by Steedman and others in a number of papers going back

³⁴Unless they can be filled by contextual information, consideration of which is outside the scope of this thesis.

³⁵It has often be noted that arguments and adjuncts in practice are difficult to distinguish and appear to be best described as lying at the extremes of a continuum. It would be interesting to attempt to model this as a degree of expectation, but in the present model only a binary division between anticipated and not anticipated is employed.

to Ades & Steedman (1982). This approach shares at least one common factor with the present dynamic model, in that it is expressly motivated by the need for the grammar to be compatible with the incremental and integrative nature of language processing.

“...it is an advantage in a competence theory if it is clear how it could be turned into a processor which works ‘from left to right’ through the text, and appears to deal with semantics as nearly as possible in parallel with syntax...”
Ades & Steedman (1982: 517).

The conception of syntax in CCG also has obvious similarities with the present dynamic model.

“...it does not require that the processor build autonomous syntactic representations, however temporary, for subsequent semantic interpretations. According to this model, syntax is something that a speaker or a hearer *does* in getting from strings to meanings, or vice versa, not something that is built.”
(Ades & Steedman 1982: 550).

This alternative view of syntactic constituents, viewing them as mere steps to building a meaning representation, is the springboard to tackling a range of problematic data, such as non-constituent coordination, by allowing the use of what other theories would see as non-standard constituents. This approach is also used in an attempt to model the largely incremental construction of meaning.

This latter point may be illustrated by considering the derivation of the simple example sentence (2.12).³⁶

(2.12) Keats eats apples.

With a standard assignation of syntactic types to the words, the combinatory rules would allow an interpretation of the sentence to be derived in the following way. The transitive verb *eats* (of type $(S \setminus NP)/NP$) combines with the object *apples* (NP) to form a constituent of type $S \setminus NP$. This then combines with the subject NP *Keats* and the interpretation is complete.³⁷

Keats	eats	apples
NP	$(S \setminus NP)/NP$	NP
	$S \setminus NP$	
	S	

³⁶Taken from Steedman (1996: 10).

³⁷The derivation thus corresponds to the application of the phrase structure rules $S \rightarrow NP \quad VP$ and $VP \rightarrow V \quad NP$.

It will be seen that in this derivation the verb *eats* cannot immediately interpret *Keats* as its subject (and thereby fill its *agent* role), for it must wait for the object to combine with the verb first. However, applying type-raising to the subject *NP* will create the raised type $T/(T \setminus NP)$ (stating that the constituent will combine to the right with any constituent which will combine with an *NP* to the left). Using this raised type we can then achieve an incremental derivation in which the subject can be interpreted as soon as the verb is reached.

$$\begin{array}{c}
 \begin{array}{ccc}
 \text{Keats} & \text{eats} & \text{apples} \\
 \hline
 NP & (S \setminus NP) / NP & NP \\
 \hline
 \text{---} >T & & \\
 T / (T \setminus NP) & & \\
 \hline
 & S / NP & \\
 \hline
 & S &
 \end{array}
 \end{array}$$

The type-raising of subjects also allows the derivation of examples of non-constituent coordination, such as (2.13), where the strings *Keats eats* and *Mary steals* both compose to categories of type S/NP and may be conjoined.³⁸

(2.13) Keats eats and Mary steals apples.

It will be noted that allowing such flexibility of categories means that the simple sentence (2.12) may now be derived in two distinct ways, although the interpretation arrived at will be the same in each case. Clearly as sentences become more complex, the number of possible derivations will also increase. Although this feature of categorial grammar has frequently been cited as a serious problem, such a view is explicitly denied by Steedman himself.

“This property, which is sometimes misleadingly referred to as ‘spurious ambiguity,’ increases the problem of nondeterminism for the parser. This fact is quite irrelevant to the present concerns with competence grammar. *Any* grammar that covers the range of phenomena considered here engenders exactly the same nondeterminism in the parser.” Steedman (1996: 93)

In this thesis, I challenge this claim by showing that the dynamic model does cover such examples and does not allow multiple derivations of the same interpretation. That such non-determinism cannot be removed from the categorial grammar, for example by obligatory type-raising all subjects, may be seen by considering an example of *VP* coordination.

³⁸The rule for coordination assuming that conjuncts of like types may coordinate.

(2.14) Keats sang and danced.

Here the subject NP, even if it is type-raised, cannot combine immediately and be interpreted as the **theme** of *sang* as then a constituent of type *S* would be formed and the coordination would not go through. Instead the *NP Keats* must wait to be interpreted as the subject of the coordinated verb phrases, as shown in the following derivation.

Keats	sang	and	danced
<i>NP</i>	<i>S\NP</i>	<i>CONJ</i>	<i>S\NP</i>
	<i>S\NP</i>		
<i>S</i>			

However, there is no reason to assume that *Keats* is not interpreted as subject immediately the verb is reached in both (2.13) and (2.14) above, although as we have seen in the latter case this is impossible. As shall be shown in chapter 5, this problem does not appear in the dynamic model, in both sentences the subject will be interpreted as the subject of the verb as soon as it is reached.

If we take a step back from the details of derivations, a fundamental reason for the extra non-determinism, and the reduction in the potential for incremental interpretation, must be that in the categorial grammar syntactic rules have no access to the semantic representation being constructed, they are expressed purely in terms of syntactic types.

“Syntactic rules and phonological processes have no access to this level of representation [ie. predicate-argument structure]” (Steedman, 1996: x).

So if the string *Keats sang* were interpreted as a proposition there is no way for the syntactic rules to apply to some sub-part of it. In the dynamic model rules have access in general to all of the information in the semantic interpretation being constructed (although it is usually sufficient to limit this to the active constituent).

It should also be noted that the constituents on the parsing stack in the processing of categorial grammars are not related to each other, whereas the semantic constituents on the active stack in the dynamic model are at all times connected by some relation either semantically or syntactically into a unified network.

In conclusion it may be said that the restriction of the syntactic process (ie. the process of the construction of an interpretation) in categorial grammar to the combination of syntactic types, although making for a conceptually clean and elegant system leads inevitably to limitations in its application to incremental interpretation and the admission of multiple derivations of the same interpretation. Such a restriction also raises questions

as to how categorial systems scale up to produce a wide-coverage grammar for practical use.³⁹

2.6.2 Dependency grammars

Dependency approaches, which can be traced back to Tesnière (1959), characterize syntax as the possible dependency relations between words, eschewing representation of constituent structure. Of the many dependency-based theories proposed over the years, I will focus on a comparison with Hudson's Word Grammar (Hudson, 1984 & 1990), as being perhaps the most fully developed and widely known.

Firstly, we should note that Word Grammar is explicitly conceptualist in that "language is studied as a mental phenomenon" (Hudson, 1990: 53) and includes a level of (language-dependent) semantic structures which are comparable in conception to those assumed for the dynamic model in this thesis.⁴⁰

"...semantic structures are ordinary cognitive representations of the events, etc., that the sentence describes (though some of the categories might not have existed without the influence of language)..." Hudson (1990: 123)

Hudson assumes that there is "no boundary between linguistic and non-linguistic concepts" (1990: 75), and that linguistic theories must therefore make use of a detailed conceptual hierarchy, similar to that deemed to be necessary for the specification of semantic types earlier in this chapter.

The major difference between the two models lies in the way these semantic structures are arrived at, that is, the syntax. Word Grammar has a separate level of syntactic structure, representing dependency relations between words: "dependency structures are simply sets of pairwise relations between single words, without implications for word order" (1990: 117). In the present dynamic model the only relevant factor is the way a word adds to the current interpretation. No relations exist between words, apart from that of serial order.

The closeness of semantic dependency relations to syntactic ones is taken to be a strong point of Word Grammar.

"One of the advantages of the WG approach to semantic analysis is the close integration of the syntactic and semantic dependencies: every syntactic

³⁹The generally restricted coverage of grammatical constructions is regretted by Wood (1993: 148) in her survey of work on categorial grammar.

⁴⁰Although in many aspects, particularly in regard to such questions of sense and reference, it must be admitted that the Word Grammar structures have been worked out with greater depth and thoroughness. It remains a task to incorporate many of these insights into the present model.



dependency corresponds either to a semantic dependency or to a semantic identity.” (Hudson, 1997: 32)

According to the assumptions of the present model, however, this correspondence is regarded as redundancy and syntactic dependencies are not represented, except where this is unavoidable as in the case of subjects.

One problematic area for pure dependency grammars is coordination, which has proved difficult to handle purely in terms of dependencies between individual words. Word Grammar tackles this by the exceptional use of constituent structures, but it appears to be a point of weakness of the theory that constituent structure is required for this one construction and nowhere else. I will show in chapter 5, that the dynamic model does allow a simple analysis of coordination constructions, requiring no major innovations in the syntax.

Although dependency models have not in general been developed with an explicit view to their integration into a model of processing, Fraser (1989) discusses an attempt to construct such a model for Word Grammar. In this model words are successively placed onto a processing stack and reduced when a dependency can be found between them.

The main point of commonality between Word Grammar and the present model appears to lie then in the recognition and representation of a psychologically-real level of semantic or conceptual structure. As such, both models also are part of a conceptually-based approach to linguistics whose origins may be traced back to Chafe (1970), and among whose present adherents may be counted the school of Cognitive Grammar centred around Langacker⁴¹, as well Jackendoff’s research into conceptual structures.⁴² One common assumption of these latter approaches is that lexical meanings are decomposable into a small number of basic meaning constituents.

Where all of the models discussed here differ is in their assumptions about syntax. We have seen that dependency approaches take it to be characterized by word-to-word dependencies, while Jackendoff assumes that constituent structures are constructed in parallel with conceptual structures and that the two are linked by correspondence rules. The Cognitive Grammar tradition is suspicious of formal models of syntax and the method by which the meanings of words are combined into the meanings of expressions has, as far as I am aware, not been formalized to any degree. The defining characteristic of the present model may then be taken to be the assumption that syntax is formalizable as a set of transition rules that perform operations on interpretation states.

⁴¹Langacker (1987).

⁴²Jackendoff (1990), (1997).

2.6.3 Semantic parsing

A separate strand of research into conceptual structures and semantic decomposition may be classified under the term *semantic parsing*, taken from Wilks (1983), where it is used to characterize an approach to automatic language analysis practised by researchers in Artificial Intelligence, whose central claim is the following.

“... parsing is fundamentally a matter of mapping texts onto semantic structures, and doing so without the need for a separate, identifiable, syntactic component.” Wilks (1983: 182)

The question of whether the present model could be classified under this heading depends on the definition one gives to the expression *syntactic component*. For example, one of the leading exponents of the approach claims that: “We have never been convinced of the need for grammars at all” (Shank, 1975: 12) and that “processes should not really depend on syntax, but on semantics”. Such views may be taken to be reaction to the syntactocentric position of the dominant linguistic paradigm⁴³, but nevertheless an over-reaction. Citing work by Schank and himself, Wilks makes the same point.

“Early AI work on the analysis of general natural language... was much preoccupied by what its practitioners saw as the wrongheadedness of Chomsky’s preoccupation with syntax and the need for semantics-based methods to understand language. In doing so they may well have overlooked much that was of value in Chomsky’s system...” Wilks (1987: 8)

And the same point is made in Sag (1991: 73) with reference to knowledge-based approaches: “Grammar—detailed grammar of the sort that card-carrying linguists have spent decades investigating—is an essential constraint in NLP [*Natural Language Processing*], both for parsing and for generation.”

Although the basic assumptions of semantic parsing may not be far removed from those of the present model, there is clearly a big divergence in which parts of the model are thought to be of interest. Practitioners of semantic parsing have been generally concerned with the details of the semantic analysis of expressions, the focus in the present thesis, however, is on the syntactic process.

⁴³Typified by the claim in Chomsky (1987: 191) that “most of the theory of meaning is called syntax.”

2.6.4 Processing models: Parsifal & the ATN

The model that I am arguing for in this thesis amounts to a redrawing of the line between competence and performance with respect to the norm in generative grammar. The competence component now contains the dynamic element of language interpretation rather than leaving this to the processor. It thus turns out that the resulting competence grammar has many points of similarity with previous models which attempted to address the problem of language processing. In this section I will explore these similarities and attempt to clarify the fundamental points of divergence.

I shall begin by comparing the present approach to syntactic modelling with that of Marcus' influential model Parsifal (Marcus, 1980).⁴⁴ The central foundation of Marcus' approach is his *Determinism Hypothesis*, according to which "the syntax of any natural language can be parsed by a mechanism which operates strictly deterministically". To this end Marcus constructs a *grammar-interpreter* such that "the grammar must specify exactly what action the interpreter should take at each and every point in the analysis process."⁴⁵

Parsifal's grammar-interpreter is able to manipulate two data objects: a stack of constituents that could still be given daughters and a buffer holding items that are yet to be integrated into the analysis. The former stack has clear parallels with the "active constituent stack" introduced above for the present model, although it should be noted that Marcus is assuming that these constituents are syntactic constituents. There is no object in the present dynamic grammar corresponding to Parsifal's buffer of unintegrated items: all constituents are integrated immediately into the interpretation.⁴⁶ Marcus notes that this constituent buffer "is the primary source of the power of the parser" in that it enables the determinism hypothesis to be upheld.

"The ability to parse a constituent before its higher level grammatical role can be determined and then drop that constituent back into the buffer is exactly the mechanism needed to cut through the seeming nondeterminism." Marcus (1980: 53).

Once we limit the objective to constructing a grammar rather than a complete parser, then this extra data object can be dispensed with.⁴⁷ Marcus himself notes that restricting the mechanisms available to a model often leads to simplification of the rules that it requires: the argument of this thesis is that concentrating on a dynamic model of competence allows simplicity both in the grammar and in the processor that employs it.

⁴⁴In broad outline much of the following comparison applies equally to the processing model proposed in Church (1982).

⁴⁵Marcus (1980: 14).

⁴⁶That is, to the extent that they can be. Constituents appearing before the item that will give them a semantic interpretation, such as subjects and topics, are given a syntactic link to the active constituent and their addresses placed on its store. The constituent buffer in Parsifal is used also for post-head arguments.

⁴⁷And so, by Occam's razor, should be dispensed with.

The grammar rules in Parsifal are formalized as pattern-action rules, again with clear similarities to those of the dynamic model presented here. However, Parsifal's rules are divided into "*rule packets*, clumps of grammar rules which can be turned on and off as a group", and they are explicitly ordered, unlike the declarative formulation of the dynamic syntax. Once again the question of which rule to apply first is not addressed in the dynamic competence model presented here but left to other performance principles to be worked out independently.

There are obvious similarities in design to be found between Parsifal and the dynamic model of syntax presented in this thesis. It indeed may be speculated that any attempt to model the incremental interpretation of language is bound to require comparable constructs⁴⁸, such as the stack of active constituents, which would indeed tend to support the assumption that this has some underlying psychological reality. The important difference between these previous models and the present one is rather one of conception and scope: the approach in this thesis respects the traditional distinction of modularity of the process of language interpretation. To be sure the system of grammatical rules works incrementally, but it is a non-deterministic abstraction and says nothing about how and when these rules are to be applied in processing.

This distinction is perhaps brought into even sharper relief by comparing the approach to dynamic syntax adopted here with the work done on processing language using augmented transition networks (ATN's), which are described by Woods thus:

"a transition network grammar is a finite-state transition diagram which has been generalized to a pushdown store automaton by adding a recursion mechanism and then further generalized (up to the power of a Turing machine) by the addition... of a set of registers that can hold arbitrary pieces of tree structure and... arbitrary conditions and actions that can set and test these registers on the arcs of the network." Woods (1973: 112)

ATN's were typically conceived as processing models incorporating a declarative grammar base (the pushdown store automaton) with extra general operations on the states. They were also viewed as possible psychological models of comprehension, for example in Kaplan (1972).

It will be worthwhile to examine why ATN's fell into disfavour in computational linguistics. It seems that this was not due to any practical problems resulting from their greater power in language theoretic terms. The reasons for the decline in use are discussed lucidly by Kaplan, himself involved in their development, as follows:

"Richer formalisms basically allow outputs to be determined by intermedi-

⁴⁸This is illustrated in the following chapter by consideration of Hausser's incremental model of syntax, which again shares clear similarities of architecture.

ate process steps, intermediate operations and intermediate data that are not theoretically committed. In any linguistic system there are certain kinds of structures that really have some theoretical significance... But in these richer procedural formalisms you get the possibility of determining inputs and outputs by things that you don't really care about, that have to do with particular details, either of your specification or your implementation. You have various combinations of indexing, caches, strategies, extra registers, and the ability to look at the internal implementation state of the parser... By letting process-control specifications into the grammar, you thought you were going to get something that ran more efficiently... In hindsight it seems that whenever the sequential, procedural facilities of the ATN formalism were used, some sort of linguistic error was being committed, in the sense that generalizations were not being expressed." Kaplan (1987)

In contrast, we are committed to the objects in the present dynamic model. The intermediate interpretation states are held to be psychologically real and necessary for the operation of the grammar, rather than arbitrary constructs of the processor. Although the scope of the grammar is extended to include the dynamic flow of interpretation, *process-control specifications* are outside of its scope. This is not the case with the models of Marcus and Church.

To give a idea of how the present approach differs from the a typical ATN in practice, consider the characterisation given in Woods (1973: 126-127):

"... a configuration consists of a list:

(STRING WEIGHT STATE STACK REGS HOLD PATH)

where STRING remembers where in the string the configuration is, WEIGHT is a measure of the estimated "likelihood" ..., STATE is the state of the network..., STACK remembers the higher-level computations in which the current one is embedded ..., REGS is the current set of register contents, HOLD is the current hold list for constituents that have been found out of place, and PATH is a complete record of how this configuration was reached..."

Some of these objects have parallels in the states of the present model⁴⁹ We shall also see in chapter 5 that there are parallels in the treatment of coordination. And there are no underlying string substitution grammar rules, which are assumed in the pushdown store automaton.⁵⁰

⁴⁹In particular the stack of non-integrated constituent addresses, or *store*, has some similarities with the HOLD register and in chapter 3 will be similarly used in the analysis of unbounded dependency constructions. There is an important difference though that the store is a feature of constituents themselves rather than of the state. In this way it is more akin to the SLASH category mechanism of GPSG and HPSG.

⁵⁰This is of course a point about the actual grammar used as the basis of the ATN, rather than a characteristic of the ATN itself.

To sum up, the present dynamic model is a competence model, following the Chomskyan assumption of modularity of the interpretation task set out in the previous chapter. In adopting a dynamic competence, however, we are cutting the interpretation cake differently, in particular proposing a declarative grammar which pays respect to the incremental nature of interpretation. There is a clear separation between the parser and the grammar, with transition possibilities being specified by the grammar and search strategy left to the parser. This is a different approach to Parsifal or the ATN, where the two components are conflated.

2.6.5 Hausser's Left-Associative Grammar

Finally, it will be useful to compare the present approach with one which shares most of its basic premisses, Hausser's *left-associative grammar*.⁵¹ This is the most directly comparable of the dynamic approaches mentioned in the introduction. Those of Kempson *et al.* and Milward are based on systems of logical deduction, while Philips assumes a more standard role for constituent structure. Such a comparison is particularly instructive for it shows how an alternative architecture can be given to the syntax, while arriving at much the same result.

The essential insight is that the syntactic model itself should take account of the time-flow of language.

“a time-linear syntax does away with semantically motivated constituent structures in the syntax, making room for the systematic derivation of semantic hierarchies”. (Hausser, 1989)

Syntactic rules are formulated declaratively as operations of the following kind.

$$\text{rule}_i: [\text{CAT1 CAT2}] \Rightarrow [\text{rule-package}_i \text{ CAT3}]$$

Here CAT1, CAT2 and CAT3 are the categories of the present *sentence start*, next word and next *sentence start* respectively, and the rule package contains a list of the rules that can apply at the next word.

The operation of the syntax may be explained with respect to the following example. The rule NOM+FVERB, one of those that may apply sentence-initially⁵², applies to the sentence start “(NH) Peter” and the next word “(N D A V) gave” to give a new sentence start “(D A V) Peter gave” and a new rule package. One of the rules in this package is FVERB+MAIN which applies to the sentence start and the next word “(N-H) Fido” to

⁵¹Presented in Hausser (1989), from where all the quotations and examples in this section are taken.

⁵²That is, it is in the rule package of the special rule START.

give a new sentence start and a new rule package and so forth throughout the sentence.⁵³

*START		
	(NH) Peter	<i>Name Human</i>
	(N D A V) gave	<i>Nominative, Dative, Accusative, Verb</i>
*NOM+FVERB		
	(D A V) Peter gave	
	(N-H) Fido	<i>Name non-Human</i>
*FVERB+MAIN		
	(A V) Peter gave Fido	
	(GQ) the	<i>General Quantifier</i>
*FVERB+MAIN		
	(GQ V) Peter gave Fido the	
	(S-H) bone	<i>Singular noun non-Human</i>
*DET+NOUN		
	(V) Peter gave Fido the bone	
	(V DECL) .	<i>Declarative</i>
*CMPLT		
	(DECL) Peter gave Fido the bone .	

A representation of the meaning may be built up simultaneously as a semantic hierarchy to give the following result. Numbers indicate the word at which the information is added.⁵⁴ Again Hausser adopts the conceptualist position, according to which: “meanings are in the head” (p. 15).

```
(SENT-2
  (SUBJ ((NP-1 (NAME (Peter-1))))))
  (VERB (give-2))
  (INDIR-OBJ ((NP-3 (NAME (Fido-3))))))
  (DIR-OBJ ((NP-4 (REF (the-4 SG-5)) (NOUN ((bone-5))))))
```

The major difference in the architecture of the two models appears to be that in the one presented in this thesis, syntactic rules operate directly on the semantic representation itself, there is no separate syntactic representation for them to apply to. In Hausser’s system the construction of the semantic representation is entirely dependent on the syntax, and the syntax can run independently without any semantic representation being built. It may be possible to see a parallel between Hausser’s use of sentence start categories, with the necessity of introducing the active stack into the semantic representation of the

⁵³The abbreviations used are explained on the right.

⁵⁴Other numbers indicating the number of the derivation and number of the sentence in the text are not shown here in the interests of simplicity.

present model, in order to pick out a constituent to which the rule should apply. It is not immediately apparent if either of these approaches is advantageous in terms of its simplicity of use or in writing grammars.⁵⁵

One point that Hausser makes clear is that a dynamic grammar of the type he proposes is quite distinct from the Augmented Transition Network (ATN), discussed in the previous section:

“Conventional transition networks are not left-associative because (i) they represent rules which cover several transitions in the network, and (ii) they permit the embedding of sub-networks into higher networks. LA-grammar, on the other hand, is left-associative, because its rules always cover a single transition.” Hausser (1989: 174)

To sum up, although the models share many properties, their biggest divergence is probably in the fundamental difference in emphasis of the uses to which they have been put. Although Hausser demonstrates that the model covers a range of constructions, he does not seek to investigate questions which have been of central importance to syntactic theory over the past decades. Instead his emphasis has been on questions of formal power (as discussed at the beginning of this chapter), on issues of language generation, as well as on more philosophical areas of semantic theory.

2.7 Summary

In this chapter I have presented a model of the rules (linguistic competence) to be used in an overall model of linguistic behaviour. The model is a dynamic system in that it defines a set of states and the possible transitions between these states. I have made clear that although “left-to-right” it is not a finite-state grammar, and furthermore argued that issues of formal power in language theoretic terms are of doubtful relevance, both in terms of effective processing and psychological validity.

The states of the dynamic model represent the growing interpretation of the sentence, which I take to be idealized cognitive events. I have assumed that important aspects of these interpretations may be represented as a network of conceptual constituents connected by semantic relations. I further assume that such constituents may contain *syntactic* (that is, semantically empty) features, which help in the construction of the interpretation but do not themselves contribute to the meaning.

In order to draw up rules which specifying the possible transitions between these states, I have argued that a stack of active constituents must be specified for each state, with

⁵⁵ An argument could be made that the model in this thesis is slightly more economical in that there is no separate syntactic representation, however minimal.

rules referring to the topmost constituent on this stack, which I refer to as the *active constituent*. I argue that it is premature to place formal restrictions on what these rules can do, although I assume that derivations are monotonic, in that at each transition information is only added and not removed.

I argued that the lack of syntactic constituents in the model was in keeping with the recognized fact that subcategorization had to be expressed largely in terms of the semantic requirements of lexical heads.

Finally, I compared the approach with other non-standard approaches to syntactic modelling. In contrast to categorial grammar, the model assumes that syntactic rules do have access to the growing semantic interpretation. It has much in common with conceptually-based dependency approaches, such as Word Grammar, but I argue that the explicitly left-to-right nature of the grammar obviates the need for the representation of syntactic dependency relations, just as it does constituent structure.

The approach may be summarized as a “recutting of the cake”, modularity is respected, but the dynamic characterization of competence means that it has much in common with previous models of processing such as that of Marcus and the ATN. Nevertheless restricting the model to exclude processing considerations is argued not only to be justified given the Chomskyan arguments for modularity, but also to allow a simpler characterization of the job the processor has to do in using the model.

In the following chapters, I shall examine how rules have to be added to this basic model to account for a wide range of the syntactic phenomena that have formed the backbone of generative investigations into the syntactic process.

Chapter 3

Unbounded dependencies

Constructions involving phrases appearing an indefinite distance in time before their point of interpretation are among the most studied and revealing sets of data in the syntax of English. The data presents many challenges and it behoves any aspiring model of syntax to deal with its complexities. In this chapter I introduce an additional feature **store** (which owes a clear debt to the **SLASH** feature of GPSG and its descendants) and discuss how it can be used to model, and even on occasion help explain, the data.

Although the analysis given here of basic operations such as *wh*-movement and relative clauses is reasonably well established and has been successfully implemented computationally, the analysis for other more marginal constructions is still somewhat tentative and should be taken as suggestive of the possibilities of analysis in the model, rather than an established optimal one.

3.1 *Wh*-movement

There is a general constraint in English that phrases containing (*wh*) question words are positioned sentence initially.¹ The problem this poses for the syntax is how these questioned constituents are correctly interpreted when there may be an indefinite lapse in time until the position in the sentence where they would normally receive their interpretation.

For example, the questioned constituent in (3.1) is interpreted as the theme of *put*, in (3.2) as the goal of *put*, and in (3.3) as the *loc*-role of *in*, interpretations which according to the syntactic rules sketched so far would require them to occupy the marked positions.²

¹It is possible to construct a functional explanation for this, ie. to communicate the fact that the sentence is not an assertion, but typically a request to the hearer to supply a specific piece of information. This is made clear as early as possible to avoid the need for subsequent reanalysis.

²Note that here and throughout the rest of this thesis, the marking of these “gap sites” is purely in the interests of exposition, and is in no way part of the theory being presented.

- (3.1) What did he put _ in the tea?
 (3.2) Where did he put the tea _ ?
 (3.3) What did he put the tea in _ ?

The Chomskyan tradition in syntax has employed the metaphor of movement to describe this phenomenon, supposing that the *wh*-constituent appears in the canonical position at some more abstract level of syntactic structure, and is then moved to a position in the structure such that it appears initially in the sentence. There is no possibility of providing such an analysis here for we lack any such abstract levels of structure. There is therefore no alternative but to say that the *wh*-constituent is able to wait an indefinite amount of time in the conceptual structure that is being built up, before it can be semantically integrated into the meaning of the sentence.³

To accommodate these displaced constituents we shall use the feature **store** already introduced in the previous chapter to hold the subject feature marked +. To mark non-subjects we shall use the feature -. To put initial *wh*-constituents onto the store of the proposition we must introduce a new transition rule *add-to-store*. Given an initial *wh*-word, the rule *add-to-store* will mark the proposition as interrogative **Q** and add the address of the new constituent to its store. Employing the notation introduced in the previous chapter this may be expressed as follows.

transition rule	Lexical entry	top of stack (S_n)	top of stack (S_{n+1})
<i>add-to-store</i>	{ wh }	(X, prop, [])	(N) (X, prop, Q, [-N])

Let us assume the lexical entry for the word *what* is:

what {pronoun, wh} [NON-HUMAN, ref:-]

The fact that it is the identity of this constituent that is questioned is represented informally by placing a dash after its referent feature *ref*.⁴

The first transition of (3.1) will then result, following application of *add-to-store* to the initial empty proposition, in the following state.

³This also provides a ready explanation for the often-noted anti-symmetry of movement: elements can wait for an interpretation but the interpretation cannot go ahead without such elements appearing, or at least this movement where it does occur, as in heavy-element shift is highly restricted. Ross (1967) gives a detailed list of such antisymmetries.

⁴Although, I shall make use of such a feature at various times during the following chapters, its precise nature still has to be worked out.

- (1) prop⁰, Q¹, [-2]
 (2) NON-HUMAN¹, ref:-

(3.1): S_1 **What** || did he put in the tea?

A clear distinction should be made between the *store*, a feature on constituents holding as yet uninterpreted addresses, and the *active stack*, the partial ordering of the constituents making up an entire state. For example, the *wh*-constituent formed by *what* is on the store of the proposition, but will leave the active stack after the first transition as it is saturated and not to be modified at the next transition.

The rule for adding surface subjects could of course also apply here to give a subject *wh*-phrase as in (3.4).⁵

(3.4) What did all this damage?

Indeed it appears problematic to assume that when a non-pronominal (ie. non case-marked) constituent appears initially in the sentence it must be marked as either subject or non-subject. To see this we can consider the two alternative resolutions of (3.5).

(3.5) Which of those gorgeous Geordie women that you met last week at the football
 liked you best?
 game did you like _ best?

If we assume that suboptimal analyses are lost after a certain amount of time then there is no way of explaining why one of these continuations does not present difficulties in interpretation. Instead we shall suppose that such constituents are on the stack unmarked until there status is disambiguated, either by the addition of a lexical verb which requires a subject or by an auxiliary which introduces the two possibilities of subject or non-subject. We shall largely gloss over this issue in the following derivations however. We must slightly amend the **add-fin-verb** transition rule, however, for instead of requiring a subject in place, it will now require a stored constituent that is not explicitly *non-subject* (ie. marked -).

Returning to sentence (3.1), the transition after *what* introduces an inverted auxiliary verb with the following transition.⁶

⁵This rule must then be modified with a similar *wh*-condition to ensure the proposition is marked as interrogative.

⁶As the rule indicates this will also apply with pre-subject negative phrases (the proposition in this case marked with the feature *neg*).

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{aux, finite, inverted}	(X, prop)	($X \cup N$, prop)

where X contains **Q** or **neg**

Again the features added by the auxiliary verb, represented **N**, will include its lexical head, semantic type, tense and subcategorization requirements. A similar rule for auxiliary verbs will produce yes-no questions.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{aux, finite, inverted}	(X, prop, [])	($X \cup N$, prop, Q, [])

The idea of marking a certain subset of auxiliary verbs in the lexicon as being **inverted** is adopted from GPSG (thus enabling the distinction *Aren't I* vs. **I aren't* to be modelled), although here the feature **inverted** does not propagate outside the lexicon into the structures themselves.

The information added by the auxiliary varies⁷, but I shall assume that the support verb *do* is semantically vacuous, its effect restricted to adding tense and an argument slot **content** which takes a tenseless EVENTUALITY, ie. a STATE or an EVENT.

- | |
|--|
| (1) prop ⁰ , Q ¹ , past ² , subj: , content: , [-2] |
| (2) NON-HUMAN ¹ , ref:- |

(3.1): S_2 **What did** || he put in the tea?

The word *he* adds the surface subject to the clause by way of the previously given transition rule **add-subj**.

- | |
|--|
| (1) prop ⁰ , Q ¹ , past ² , content: , [+3, -2] |
| (3) MALE ³ , sg |
| (2) NON-HUMAN ¹ , ref:- |

(3.1): S_3 **What did he** || put in the tea?

When the head verb *put* is reached the action is added as the content of the proposition. Both of the constituents in store are inherited by this new constituent. The subject value is then interpreted as the first argument, *agent*, of the action *put*, and the non-subject is interpreted as the second argument.

⁷Auxiliaries are discussed at greater length in the section on raising verbs, 4.1.3.

The rule for adding arguments will have to be extended to allow the store values to be inherited, with the previous values marked with an asterisk to show that they have been removed from the constituent but not yet interpreted. We shall also have to introduce the following transition schema **interpret-store** to interpret these elements and cancel them on the store.

transition rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
interpret-store		(X, [-N])	(X, arg: N, [-N†])

As we shall see, *arg* here can be any of the unfilled argument roles, if it is semantically compatible, or even a modifier. The fact that the stored constituent has been interpreted in this constituent is explicitly marked as before.

Thus, the modified rules outlined above allows the derivation to move into the following next state with after the addition of the main verb *put*, with *he* being interpreted as its **agent** and the *wh*-constituent being interpreted as its **theme**.

- | |
|---|
| <p>(1) prop⁰, Q¹, past², content:4, [+3*, -2*]
 (4) ACTION⁴, put, agent:3, theme:2, goal: , [+3†, -2†]
 (3) MALE³, sg
 (2) NON-HUMAN¹, ref:-</p> |
|---|

(3.1): S_4 **What did he put** || in the tea?

The final locational *goal* argument, *in the tea*, is added routinely to complete the derivation.

- | |
|--|
| <p>(1) prop⁰, Q¹, past², content:4, [+3*, -2*]
 (4) ACTION⁴, put, agent:3, theme:2, goal:5, [+3†, -2†]
 (3) MALE³, sg
 (2) NON-HUMAN¹, ref:-
 (5) LOCATION⁵, in, loc:6
 (6) def⁶, tea⁷, sg</p> |
|--|

(3.1): **What did he put in the tea?**

Let us now return to our second example of *wh*-movement (3.2), repeated here, which involves the movement of the verb's second argument.

(3.2) Where did he put the tea _ ?

The rule *add-to-store* will apply as above with the initial word *where*, to create a first state entirely parallel with that of (3.1).

(1) prop ⁰ , Q ¹ , [2] (2) LOCATION ¹ , ref:–

(3.2): S_1 **Where** || did he put the tea?

The derivation proceeds as before, but in this sentence when the head verb *put* is reached, the non-subject constituent in store is of a type to fill the third **goal** argument of the verb, rather than the **theme** role. We will assume that the argument is once again filled immediately the head verb is added. The constituent, *the tea*, is then added as **theme** of *put* to complete the derivation.

(1) prop ⁰ , Q ¹ , past ² , content:4, [+3*, –2*] (4) ACTION ⁴ , put , agent:3, theme:5, goal:2, [+3†, –2†] (3) MALE ³ , sg (5) def ⁵ , tea ⁶ , sg (2) LOCATION ¹ , ref:–

(3.2): **Where did he put the tea?**

Stipulating that arguments may only be filled by constituents in store immediately, and without regard for the canonical order of argument filling, is equivalent to a strategy of early gap-filling. Importantly, it also avoids the pitfalls that some trace-based accounts may run into, pointed out by Pickering & Barry (1991). These centre on the fact that extractions such as (3.6) do not get worse no matter how far the supposed extraction site is from the head verb.

(3.6) Where did she put the old jar that she had borrowed from his uncle __ ?

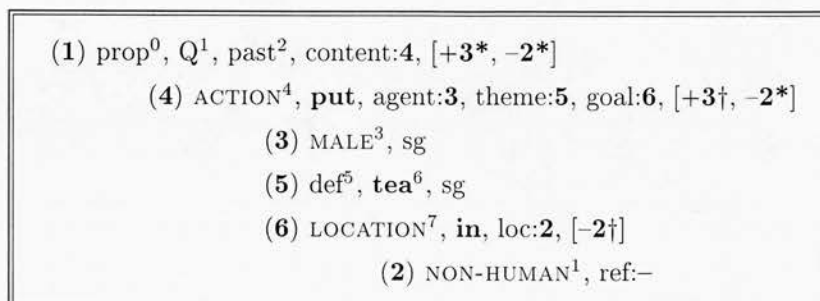
We shall instead assume that interpretation of stored constituents is made as early as possible.

Our third example of *wh*-movement, involves extraction from a position inside a subphrase of the sentence.

((3.3) What did he put the tea in __ ?

For the first three words the derivation follows step for step that of (3.1). At the head verb *put*, however, the *wh*-constituent inherited from store is not interpreted immediately.

The phrase *the tea* fills the theme role of *put*. Finally, the word *in* creates a constituent of type LOCATION to fill the remaining goal argument of the verb, inheriting the stored *wh*-constituent as it does so. This is then immediately interpreted to fill its own *loc* argument role, as shown in the following diagram.



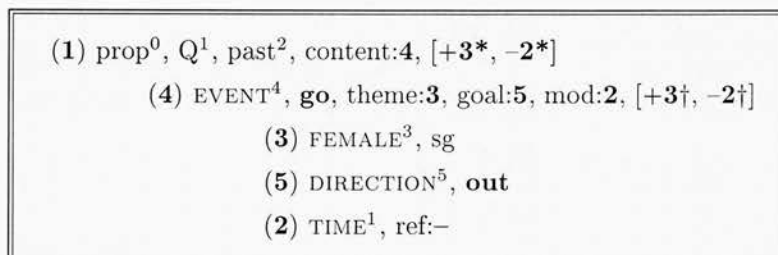
(3.3): **What did he put the tea in?**

3.1.0.1 Wh-adjuncts

The interpretation of *wh*-constituents extracted from non-argument positions, as in (3.7), is allowed by exactly the same rules as extraction from argument positions.

(3.7) When did she go out?

Once again the *wh*-constituent is interpreted as soon as the head *go* is added.



(3.2): **When did she go out?**

Addresses of *wh*-adjuncts can also be passed down into dependent clauses, resulting in the ambiguity of the string in (3.8), with one reading resulting from the adjunct being interpreted at *say* and the other from interpretation at *went*.⁸

- (3.8) a. When did you say she went out?
 b. When did you say she went out ?

⁸Questions of the relative opacity of matrix verbs, ie. their tendency to allow transmission of the *wh*-constituents will not be addressed in this thesis.

3.1.0.2 The coordination diagnostic

This is a good opportunity to introduce the diagnostic of possible coordination. Roughly speaking, the analysis of coordination in the model (to be presented in detail in chapter 5) assumes that when a conjunction is reached the derivation in effect returns to a previous state and continues from there with the second conjunct. It is thus a useful tool for probing at which state in the derivation information is added.

The first four words of (3.1) and (3.3) are identical (ie. *What did he put...*), but their analyses diverge as the fourth word is added: in (3.1) the stored *wh*-constituent is interpreted by *put*, while in (3.3) it is left in store. If we attempt to coordinate the two sentences, we see that we can return after the first three words of the first sentence and continue with the second, (3.9)a, but this is not possible after the first four words.

- (3.9) a. What did he put the tea in __, and
then put __ in the cupboard?
- b. *What did he put the tea in __, and
then __ in the cupboard?

The test therefore supports the analysis given above.

3.1.0.3 A restriction on inheriting store values

The rule for the inheritance of store values from the active constituent given above assumed that inheritance could take place, optionally, into any argument.⁹ This must be modified however in light of a known constraint on extraction originally formulated as Kuno's Clause Nonfinal Incomplete Constituent Constraint (Kuno, 1973), which states that movement is impossible out of non-final arguments, as illustrated by the following examples.

- (3.11) a. Who did he give a picture of Harry to __ ?
- b. *Who did he give a picture of __ to Harry?
- c. Who did he give Harry a picture of __ ?

To capture this constraint in the present model it is necessary to stipulate that there can be no inheritance of store values from the active constituent while it has an unfilled

⁹Indeed adjuncts may also inherit store values, as long as the constituent is saturated, although the goodness of the result typically seems to demand some sort of semantic coherence.

- (3.10) a. Which cloth should I clean the windows with __ ?
- b. Which room should I read the book in __ ?

argument role. For example, sentence (3.11)b will be blocked as the **goal** argument of *put* is unfilled at the point when the **theme** constituent *a picture* is added and therefore the stored *wh*-element cannot be inherited by this constituent.

It is possible to construct a functional explanation for such a restriction. Without it there would be an increase in the potential for local ambiguity in the grammar. For example, (3.11)a would allow two competing derivations for the phrase *a picture of ...* one with an inherited store value and one without. With the restriction in place this ambiguity is removed for the phrase is not allowed to inherit a value for store.¹⁰

3.1.1 Pied-piping

The transition rules set out above will now allow all the basic examples of *wh*-movement, either of arguments or adjuncts, where the *wh*-word occurs clause initially.¹¹ However, *wh*-words do not always occur phrase-initially, as evidenced by (3.15), which has an identical interpretation to the earlier (3.3). The word *what* appears phrase-internally as a complement of *in*. This phenomenon, which since Ross (1967) has gone under name of “pied-piping”, is not allowed by the rules given so far.

(3.15) {In what} did he put the tea _ ?

To interpret such a construction using the dynamic model and arrive at the same interpretation as (3.3), two possible options present themselves. These are that:

1. The interrogative nature of the construction is already predicted from the initial word *in* and later confirmed by *what*.
2. The interrogative is not predicted at *in* but is only added by the explicit *wh*-word *what*.

¹⁰It appears, however, that it is possible to construct examples which are exceptions to this rule, for example (3.12) from Fodor (1992) (quoted along with other examples in Pollard & Sag, 1994: 193).

(3.12) ?Which cousin did you put a picture of _ in the family album.

It will be seen that the acceptability of (3.12) is significantly reduced by making the *wh*-constituent of a type more likely to fill the second argument:

(3.13) ??What did you put a picture of _ in the family album.

The restriction should perhaps be viewed then as a strong negative weighting, rather than an absolute one, and conditional on the stored constituent being of a type unlikely to fill the remaining argument.

¹¹More precisely, where it occurs phrase-initially in a constituent occurring before the subject and any verb. Sentence adjuncts will be allowed to precede the *wh*-constituent as in (3.14).

(3.14) As for John, today where did he put the tea?

In the analysis that follows I shall assume the latter of these approaches, it being perhaps the minimal assumption. The alternative choice would involve introducing the feature **Q** at *in* and would make all initial prepositional phrases locally ambiguous. There would also need to be some mechanism for ensuring that a *wh*-constituent really does appear later in the phrase.

So following option 2 above, the first word *in* will create a constituent of type LOCATION, with an unfilled **loc** role. This will be placed on the store of the main proposition by the rule *add-to-store* as a standard (non-*wh*) topicalisation.¹²

<p>(1) prop⁰, [2] (2) LOCATION¹, in, loc:</p>

(3.15): S_1 **In** || what did he put the tea?

The next word *what* will create a *wh*-constituent, which fills the *loc* role of the active constituent *in*.

<p>(1) prop⁰, Q², [2] (2) LOCATION¹, in, loc:3 (3) NON-HUMAN², ref:–</p>
--

(3.15): S_2 **In what** || did he put the tea?

Crucially, this addition is accompanied by marking the proposition (1) as interrogative by adding the feature **Q**. It should be noted that this was not the active category, for the LOCATION (2) was. To allow for pied-piping, the rule *add-arg* must then be then be extended as follows.¹³

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ wh }	(X1, arg:) (X, prop)	(N) (X1, arg: N) (X, prop, Q)

¹²That is it will be the same as the first transition of a sentence with a topicalized prepositional phrase as in (3.16).

(3.16) In the cupboard, he put the tea _ .

¹³In addition to interrogatives, negatives also pied-pipe, as shown in (3.17).

(3.17) In no way was he surprised at her suspicion.

The rest of the derivation follows the course of (3.2) and the final interpretation is identical with the synonymous (3.3), modulo the difference in the order in which the information was added.

<p>(1) prop⁰, Q², past³, content:5, [+4*, -2*] (5) ACTION⁵, put, agent:4, theme:6, goal:2, [+4†, -2†] (4) MALE⁴, sg (6) def⁶, tea⁷, sg (2) LOCATION¹, in, loc:3² (3) NON-HUMAN², ref:-</p>
--

(3.15): **In what did he put the tea?**

There are of course severe restrictions on the occurrence of pied-piping in English *wh*-movement.¹⁴ Acceptability of the construction decreases rapidly with the depth of embedding of the pied-piped *wh*-constituent.

- (3.18) a. ??In the shadow of who(m) did he live?
 b. Who(m) did he live in the shadow of?
- (3.19) a. *The brother of who(m) did he like?
 b. Who(m) did he like the brother of?

As the proposition being marked interrogative must be directly below the active constituent on the stack, the above schema will allow only the complements of pre-subject prepositions to pied-pipe. Pied-piping out of noun phrases, as in (3.18)a and (3.19)a will be ruled out as nouns do not take direct complements.¹⁵

It is certainly true that deeper embeddings of *wh*-constituents are possible as in (3.21).

- (3.21) In the pockets of which of his trousers did he eventually find the letter?

It may be argued, however that these are approaching the status “quiz”-questions, a construction where *wh*-constituents are allowed in any part of the sentence, even in positions from which *wh*-extraction would be impossible.

¹⁴It is much more restricted than pied-piping in relative clauses, which is analyzed in quite a different fashion later in this chapter.

¹⁵Initial manner verbal phrases do not allow pied-piping either, as shown in (3.20).

- (3.20) *Seeing which man did he shout in alarm?

An analysis of such constructions is beyond the scope of this thesis, but this might be evidence for restricting the class of constituents that are “transparent” to the *wh*-constituent appearing within them.

(3.22) He eventually found the letter in the pocket of which of his trousers?

An analysis of this construction is outside the scope of the thesis.

3.1.2 Inter-clausal movement

I have so far only considered *wh*-movement within a clause, but the interpretation site of the fronted constituent may be found in an indefinitely deeply embedded clause, and aspects of this inter-clausal movement will be examined in this section. For example, in (3.23) and (3.24) the *wh*-constituent is to be interpreted respectively as the object or subject of the embedded clause. In this section I shall investigate the transition rules that allow this to take place and the consequences that they have.

(3.23) Which bone did the man think the dog wanted _ ?

(3.24) Which dog did the man think _ wanted a bone?

The first six transitions of the object-extraction sentence (3.23) follow those we have seen previously in this chapter.

- (1) prop⁰, Q¹, past³, content:4, [+3*, -2*]
 (4) STATE⁶, **think**, exper:3, theme: , [+3†, -2]
 (3) def⁴, MALE⁵, sg, **man**
 (2) def¹, ref:-, **bone**², sg

(3.23): S_6 **Which bone did the man think** || the dog wanted?

The *wh*-constituent in store is of the wrong type to serve as the theme of *think* and so remains uninterpreted.¹⁶

The argument clause is created by a rule **add-proposition**. This rule is not related to any lexical item, but may apply whenever the active constituent requires a proposition as argument.¹⁷

transition rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
add-proposition		(X, arg:)	(N, prop) (X, arg: N)

¹⁶Note that a *wh*-constituent can serve as the **theme** argument of *think* only if its semantic type is suitably wide, as in (3.25).

(3.25) What did the man think _ ?

¹⁷An account of the fact that in English finite propositional arguments do not cooccur with prepositions is given in the discussion on raising in 4.1.0.3, and will require a slight modification of this rule.

At this transition the store will also be inherited, and the determiner *the* will create a new constituent which is installed as the surface subject of the newly-created proposition via the rule **add-subj**. This combined action will result in the following state.

- | |
|---|
| <p>(1) prop⁰, Q¹, past³, content:4, [+3*, -2*]
 (4) STATE⁶, think, exper:3, theme:5, [+3†, -2*]
 (3) def⁴, MALE⁵, sg, man
 (5) prop⁷, [+6, -2]
 (6) def⁷
 (2) def¹, ref:-, bone², sg</p> |
|---|

(3.23): *S*₇ **Which bone did the man think the** || dog wanted?

The derivation now proceeds in the familiar manner, ending in the interpretation of the subject and the stored *wh*-constituent as the **exper** and **theme** of *want* respectively.

- | |
|---|
| <p>(1) prop⁰, Q¹, past³, content:4, [+3*, -2*]
 (4) STATE⁶, think, exper:3, theme:5, [+3†, -2*]
 (3) def⁴, MALE⁵, sg, man
 (5) prop⁷, STATE⁹, past, want, exper:6, theme:2, [+6†, -2†]
 (6) def⁷, dog⁸, sg
 (2) def¹, ref:-, bone², sg</p> |
|---|

(3.23): **Which bone did the man think the dog** wanted?

So we have seen that object-extraction from an embedded clause follows from the rules of complementation and store inheritance that we have already established, providing that we introduce a rule **add-proposition** to introduce finite complement clauses. It will be evident, however, that these rules will not allow a stored constituent to be interpreted as the subject of an embedded clause, as in (3.24). This follows from the fact that the inherited *wh*-constituent will have been already explicitly marked as non-subject. We will now consider the addition that must be made to the grammar to allow such a subject extraction.

Given that the first six words of (3.24) are the same as those of the object-extraction sentence (3.23), it might be supposed that the derivation was identical up to that point. If so it would merely be necessary for the finite verb of the embedded clause *wanted* to interpret the inherited store value as its subject and the derivation would go through. However, if allowed, such a transition would create severe problems for the grammar as a whole, for if main verbs could interpret non-subject store values as their subjects, it would

in effect eliminate the distinctiveness of the subject position. Assuming the previously-mentioned coordination test it would then be impossible to rule out a sentence such as (3.26), where *who* is subject of the first clause and an object in the second.

(3.26) *Who did the man see _ and _ saw him?

The only way to get around this, and to allow the transition for *wanted* to be identical to other transitions for finite verbs, is to introduce a new rule which creates a proposition and switches the status of its inherited store constituent from non-subject to subject, which I shall refer to as a *switching transition*. It may be represented as follows.

transition rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
add-proposition (subj-switching)		(X, arg: , [-A])	(N, prop, [+A]) (X, arg: N, [-A*])

Applying this rule immediately the matrix verb *think* is added will result in the following state.

<p>(1) prop⁰, Q¹, past³, content:4, [+3*, -2*] (4) STATE⁶, think, exper:3, theme:5, [+3†, -2*] (3) def⁴, MALE⁵, sg, man (5) prop⁶, [+2] (2) def¹, ref:-, dog², sg</p>

(3.24): S_6 **Which dog did the man think** || wanted a bone?

The next transition, for the head verb *wanted*, would then be completely standard, with it interpreting its surface subject as its first argument. The derivation thereafter proceeds without further incident.

<p>(1) prop⁰, Q¹, past³, subj:3, content:4, [+3*, -2*] (4) STATE⁶, think, exper:3, theme:5, [+3†, -2*] (3) def⁴, MALE⁵, sg, man (5) prop⁶, STATE⁷, want, exper: 2, theme:6, [+2†] (2) def¹, ref:-, dog², sg (6) indef⁸, sg, bone⁹</p>

(3.24): **Which dog did the man think wanted a bone?**

To sum up, I have argued that the derivation of subject-extraction sentences necessarily requires a further transition rule, which I have termed a *switching rule* and that this rule

allows the embedded propositional argument to be created as soon as the matrix verb requiring one is added.

In addition to the theory-internal considerations discussed above, we can test if the sixth states of (3.23) and (3.24) are identical or not by employing the diagnostic of coordination.

- (3.27) a. Which dog did the man think _ was hungry and _ wanted a bone?
 b. Which dog did the man think Anna liked _ and Peter had already fed _ ?
 c. *Which dog did the man think _ wanted a bone and Anna liked _ ?
 d. *Which dog did the man think Anna liked _ and _ wanted a bone?

If both derivations were identical at state 6, we should expect (3.27)c and (3.27)d to be as grammatical as the other two coordinations, which is clearly not the case. This therefore provides further evidence that the two derivations diverge already at the word *think*.

The switching rule introduced above is equivalent in effect to the Slash Termination Metarule 2 of GPSG.¹⁸

$$X \rightarrow W, V^2[+\text{SUBJ}, \text{FIN}] \quad \Rightarrow \quad X/\text{NP} \rightarrow W, V^2[-\text{SUBJ}]$$

This states that for any rule introducing a lexical item that subcategorizes for a finite sentential argument, there will be a corresponding rule which terminates a slashed NP and allows the lexical item to take a VP as its argument.¹⁹ This insight of the need for a separate rule to terminate subject extractions has been incorporated into HPSG (Pollard & Sag, 1994) and Combinatory Categorical Grammar (Steedman, 1996). Having separate rules for terminating subject and object transitions does seem to correspond with the fact that subject extraction has a marked status, being disallowed in many languages.²⁰

¹⁸Gazdar *et al.* (1985: 161).

¹⁹It should be noted that the switching transition is not dependent on the lexical category of the word, merely on it having a finite clausal complement, as shown by the sentences in (3.28).

- (3.28) a. Who was it most likely (*that) would want to invite him?
 b. Who was it a certainty (?*that) would want to invite him?

This is predictable given the switching rule above.

²⁰For example, I have found that native informants confirm the following asymmetric pattern in Mandarin Chinese, showing that extraction from the pre-verbal subject position was strongly dispreferred.

- (3.29) a. zhè-běn shū John rěnwéi nǎ-yī-ge jiàoshou dú-guò _ ?
 this book John think which professor read
 b. *?nǎ-yī-ge jiàoshou John rěnwéi _ dú-guò zhè-běn shū?
 which professor John think read this book

There is also independent support for the supposition made above that the embedded clause with a subject gap can be created at the matrix verb. It is established in the psycholinguistic literature that there is a measurable delay in reading *think* in (3.23) (repeated here) compared to *thought* in (3.30).

(3.23) Which bone did the man think the dog wanted __ ?

(3.30) The man thought the dog wanted a bone.

This is explainable if the sentence is locally ambiguous at *think*, that is the hearer does not wait until the next word *the* which would resolve the ambiguity, but posits a possible subject gap at *think*, which is then disconfirmed by the next word. This observation is referred to as the *filled-gap effect*.

Although I am not presenting a model of processing in this part of the thesis, but a model of the underlying syntactic rules, this psycholinguist evidence is nevertheless relevant due to the transparent relation between the two. In actual fact, the psycholinguist evidence has only a *positive* relevance to the grammar: if it shows that something must be happening at the *n*th word in the sentence, then the derivation allowed by the syntax must allow for this to happen at least by the *n*th word, as here. However, if the evidence were negative, for instance if there were no filled-gap effect, then one could not conclude definitely that a syntax allowing early subject-gap filling were wrong, as it might be the case that there is that possibility available in the grammar, but it is not the one immediately chosen, the chosen processing strategy leading to a delay in its use.

It is clearly an advantage that experimental evidence is applicable, at least partly, to the model of syntax itself, which is not the case for syntactic models bearing a less transparent relation to a processing model.

3.1.3 *That-trace effect*

Let us consider the effect of inserting the usually optional finite-clause complementizer *that* [ðət] at the start of the embedded clauses of the previous examples (3.23) and (3.24). Clearly the complementizer can introduce clauses with object gaps, (3.31), but not those with subject gaps, (3.32).

c. *?zhè-ge jiàoshou John rènwei __ dú-guò nǎ-yī-bēn shū?
 this professor John think read which book

d. nǎ-yī-bēn shū John rènwei zhè-ge jiàoshou dú-guò __ ?
 which book John think this professor read

The native speaker comment was that sentences (3.29)b and (3.29)c were in some way comprehensible, but clearly substandard. However, the data seem to contradict the conclusion of Xu and Langendoen (1985: 26) that “there is no subject-object asymmetry in the binding of empty categories in Chinese” and is therefore in need of further verification.

- (3.31) Which bone did the man think that the dog wanted _ ?
 (3.32) *Which dog did the man think that _ wanted a bone?

This phenomenon, known as the *that*-trace effect, has been the subject of intense linguistic examination and speculation over the years. The various proposals to account for the effect, in the main couched as configurational restrictions on the syntactic structures involved, are too numerous and rely too much on theory-internal motivations for it to be profitable to go into them here. Indeed, it is generally accepted that the numerous attempts at explanation have ended up being essentially stipulative in character.

Under the present analysis, the badness of (3.32) falls out from the rules we have already adopted. It is a consequence of the early switching rule introduced to allow subject extraction, together with the rules associated with the finite complementizer [ðət]. Having no semantic content, the sole function of this word is to create a new proposition, whether as a complement clause, subject clause²¹ or relative clause. In sentence (3.31), therefore, *that* is in an appropriate position to initiate a complement clause using the previously given rule **add-proposition**. However, in the subject extraction sentence (3.32) the embedded finite clause is already formed at the matrix verb *think* as discussed above. The complementizer *that* is not needed to start a complement clause, and adding an additional proposition will cause the derivation to fail. Thus the effect is explained in a simple fashion and without any need for stipulation.

The *that*-trace effect is an important diagnostic for syntactic models as it seems clearly to be a side-effect of the syntactic process, rather than something that has a functional explanation, or is learned directly from the data. Indeed, it is even been claimed in a linguistics textbook that the unlearnability of this phenomenon is direct evidence for the innateness of grammar: “It is hard to see how the child can infer this [ie. the ungrammaticality of a *that*-trace sentence] from evidence to which he is exposed.”²² The account given here does not require children that have any innate knowledge to reject the *that*-trace construction, as long as they learn from the available data the following points.²³

1. To process the subject extraction sentences they hear, they must develop a transition for *early switching*. This is consequential on them appreciating the distinguished role of the sentence subject.
2. That the function of the complementizer *that* is to initiate a finite proposition.

²¹This will be discussed in 4.1.0.3.

²²Haegemann (1994: 11).

²³In addition, they need to know that interpretations are constructed incrementally, but this is surely not so much learned as a biological constraint on processing.

If learners have learnt these two facts from the data, then they will have also acquired the *that*-trace effect.

3.1.3.1 Other strategies for subject extraction

It is one thing to explain how children can learn from the data a syntax that makes complementizers and subject extraction mutually incompatible. It is quite a separate task to explain why subject extraction should be performed in this way in English, when there are other possible strategies open to languages. In this section I will examine some of these strategies and speculate as to the constraints on their development.

Pro-drop languages It has long been noted that in languages exhibiting *subject pro-drop*, such as Italian, *that*-trace sentences are possible. This is quite predictable given that the arguments for the switching rule in English rested on the observation that English verbs require a surface subject in place, unlike, by definition, those in subject pro-drop languages. It may be assumed that main verbs in these languages can interpret constituents in store as their subjects.

French In French, which does not exhibit subject pro-drop, there is a literary construction distinguishing the complementizer *qui*, which is used with subject extraction, from *que*, which is used with sentences involving either object extraction or no extraction. The following illustrative data is taken from Radford (1988: 592-593).

(3.33) Laquelle as-tu dit que tu crois que je préfère __?
Which have you said that you think that I prefer?

(3.34) Laquelle as-tu dit que tu crois qui __ va gagner?
Which have you said that you think that will win

It appears that here the switching transition has been lexicalized as the specialized function of the complementizer *qui*. It should be noted that in literary French the early-switching option is not grammatical.

(3.35) *Laquelle as-tu dit que tu crois __ va gagner?
Which have you said that you think will win?

Danish The situation in literary French is paralleled to a certain extent in Danish where a similarly specialized complementizer *der* (with a non-standard variant *a(t) der*) can perform the switching transition, but not the standard complementizer *at*. The data are from Engdahl (1985: 123).

- (3.36) Hvem troede du han sagde $\{\emptyset / *at / der\}$ havde malet huset.
 Who thought you he said had painted the-house

Here though the early-switching option does seem to be possible, since subject extraction is grammatical with no complementizer as in English. It should be noted that the specialized subject-extraction complementizer *der* is also the specialized subject relative pronoun, *som* being the relative pronoun unmarked for case.

Norwegian The evidence from Norwegian has often been presented as a challenge to explanations of the *that*-trace effect. The reason is that in this non pro-drop language it appears possible to initiate a subject-gapped embedded clause with the standard finite complementizer, as in the following topicalized sentence from Engdahl (1985).

- (3.37) Denne boken er jeg sikker på at kom ut i Russland.
 This book am I sure on that came out in Russia

If we assume that, as in Danish and English, subject extraction without *at* is also possible, it must be conjectured that two strategies are operating in Norwegian: early switching and also a separate switching transition available for the complementizer *at*, in addition to its standard transition adding a finite complement.

Conclusions The employment of the different strategies for subject extraction outlined above is presented in the following table. I will call hypothetical varieties of Norwegian that allow the complementizer-gap sequence, but differ as to whether they disallow or allow early switching, Norwegian A & B respectively.

Lang/Strategy	Early switching	Special switching complementizer	Switching transition for standard complementizer
<i>pro</i> -drop	—	—	—
English	✓	—	—
Literary French	—	✓	—
Danish	✓	✓	—
Norwegian A	—	—	✓
Norwegian B	✓	—	✓

It is interesting to speculate about the effect that these strategies will have on two functional demands of the language system: minimizing local ambiguity in processing and employing the minimum number of function words. It appears that English is the most economic in the latter respect, having no switching complementizers. This is at the price, however, of an addition to local ambiguity inherent in early-switching (the filled-gap

effect).²⁴ The French solution presented here involves no extra local ambiguity and a single extra function word. Danish seems at first sight to be non-optimal in having two strategies available: however, this apparent superfluity may make practical sense as, Danish being a verb second language, the main verb will often be found some distance from the extraction site. If this happens, as in (3.38), the explicit switching complementizer *der* (or *a' der*) becomes much preferred²⁵, suggesting that it may not be so redundant after all.

- (3.38) Hvem sagde den gamle mand {??/ *at / der} havde malet huset.
 Who said the old man had painted the-house
 Who did the old man say had painted the house?

The two hypothetical varieties of Norwegian both seem to cause extra local ambiguity. Norwegian A, the variety without early switching, has the same problem of local ambiguity as English but delays its resolution even longer, for in English it is resolved as soon as the complementizer or embedded main verb is reached, while in Norwegian A, the local ambiguity starts at the complementizer and must be resolved later. In Norwegian B, with early switching, this problem is compounded. A local ambiguity starts on the matrix verb, as in English, which is then potentially resolved by a following embedded main verb, but if the complementizer follows as in (3.37), a new local ambiguity arises similar to the one in Norwegian A.

These considerations are sufficient to show that there is no need in English for the complementizer to take on this extra switching role, and if it did so it would only compound the problems of local ambiguity. It may be said then that English and literary French are in near optimal states, as may Danish also be, given that the special switching complementizer appears to be necessary. The actual situation in genuine dialects of Norwegian needs investigation, before we can pronounce that it is as non-optimal as it appears from the data here.

3.1.3.2 Subject extraction from second argument

So far we have assumed that switching in subject-extracted English sentences takes place as the word taking a finite clausal complement is added (that is *think* in the above examples). This assumption has to be modified, however, to deal with cases where the clause is the second argument as in “tell somebody that...”. If this simultaneity restriction is dropped, and the switching rule can apply any time the constituent with an unsaturated clausal

²⁴It should be borne in mind, however, that in English the local ambiguity is nearly always resolved extremely quickly, that is at the next word. The exception is where a subject gap site occurs in the verb's second argument causing extreme processing difficulty. This is discussed in the next section.

²⁵My thanks to Line Mikkelsen for the judgement.

argument is the active constituent then extraction from subject position of the second clause will be allowed as in (3.39)a. This also predicts that it is again incompatible with a complementizer as shown in (3.39)b.

- (3.39) a. Which boy did they tell you _ had lied?
 b. *Which boy did they tell you that _ had lied?

However as the first *beneficiary* argument gets longer, the sentences get markedly worse.

- (3.40) ??Which boy did they tell the local police _ had done it?

I shall assume that (3.40) is grammatical, but leads to considerable processing difficulties. One obvious reason for difficulty is that the *wh*-constituent may be taken as the beneficiary of *tell* rather than *the local police*. However, this cannot be the sole reason for difficulty, as we should then expect an object extraction from the second argument to be similarly problematic, which does not appear to be the case.

- (3.41) Which boy did they tell the local police they had caught red-handed _ ?

It appears then that the possibility for subject-extraction decreases rapidly with the time-distance from the matrix verb. To model this sort of phenomenon appears to require a weighted model of syntax of the kind to be discussed in chapter 9.

3.1.3.3 Intervening adverbials

One problem which remains with this account of the *that*-trace effect is the construction, reported in Culicover (1993), in which adverbials following the complementizer *that* in an embedded clause appear to permit a subsequent subject gap.

- (3.42) Which man does he think (that) under no circumstances _ would surrender?

Steedman (1996: 71) deals with this phenomenon by allowing adverbial phrases to have the same switching category as verbs that take clausal complements. A similar solution is available in the present model, saying that adverbials may also have the store-to-subject switching transition, but this seems a poorly motivated solution, for it seems unlikely that such a transition is learnable from the available data.

There are a number of strange things going on here. Firstly, that the phenomenon improves if the adverbial is longer, compare (3.42) and (3.43)a. Secondly, that it appears to work best if the adverbial is negative, compare (3.43)a and b. And lastly, that it is only

really good if the finite verb in the embedded clause is a modal, ie. a verb that does not enforce subject-verb agreement, compare (3.43)b and c.

- (3.43) a. ?Which man does he think that never _ would surrender?
 b. ??Which man does he think that always _ would surrender?
 c. *Which man does he think that always _ surrenders?

I can suggest no solution to this conundrum, but it appears that the data rule out any easy syntactic solution. Perhaps we have to accept that we are approaching the limits of verisimilitude of the model here.

3.1.4 Interrogative clauses as arguments

To conclude the section on *wh*-movement I will take a brief look at interrogative non-matrix clauses. In particular I will speculate on how to model the difference between these and main-clause interrogatives, the most obvious of which is the lack of subject-auxiliary inversion, as shown in (3.44).

- (3.44) He knew which bone the dog wanted _ .

Such a construction will in fact be allowed by the rules we have established already, provided that we modify the previous condition on subject-auxiliary inversion and stipulate that propositions marked interrogative **Q** only trigger inversion if they are matrix propositions.²⁶ This is clearly non-explanatory, but perhaps the only explanation for this is functionality, that is that clearly distinguishing true interrogatives and embedded ones leads to greater communicative effectiveness. Certainly, non-inversion in embedded interrogatives seems to be a point that learners of English as a second language often have problems with, suggesting that it has to be explicitly learnt.

It must be somehow specified which verbs take interrogative arguments, although to state this directly in the lexicon seems to pose problems, given patterns such as the following.

- (3.46) a. *I think why he likes her.
 b. I can't think why he likes her.

²⁶Note that non-matrix negative propositions do show inversion, as in (3.45), so it appears difficult to deduce the non-inversion in non-matrix interrogatives from some general rule.

- (3.45) He knew that under no circumstances would she be found guilty.

It seems therefore that complementation potentialities need to be responsive to more general semantic principles of the growing conceptual structure.

The transmission and eventual interpretation of the *wh*-element will be governed by the rules already given to allow the following derivation of (3.44).

- (1) prop⁰, STATE², past, **know**, exper:2, theme:3, [+2†]
 (2) MALE¹, sg
 (3) prop³, Q, STATE⁷, past, **want**, exper:5, theme:4, [+5†, -4†]
 (5) def⁵, **dog**⁶, sg
 (4) def³, ref:-, **bone**⁴, sg

(3.44): **He knew which bone the dog wanted.**

Note that this analysis predicts the same acceptability of pied-piping in these *wh*-constituents as was found in those appearing initially in matrix clauses.

- (3.47) a. ??He wondered in which town she lived __ .
 b. In which town does she live __ ?

There are also *wh*-complementizers, *whether* and *if*, which correspond to the non-interrogative *that*. They occur in the same positions, that is introducing clauses in subject or argument position, but never matrix sentences.

- (3.48) He wondered whether she danced.

The transition for *whether* will initiate an explicitly interrogative proposition and allow the following derivation of (3.48).

- (1) prop⁰, STATE², past, **wonder**, exper:2, theme:3, [+2†]
 (2) MALE¹, sg
 (3) prop³, Q, ACTION⁵, past, **dance**, agent:4, [+4†]
 (4) FEMALE⁴, sg

(3.48): **He wondered whether she danced.**

It must then be the case that a proposition initiated by the complementizer *that* is not neutral as to whether it is interrogative or not, as the initial matrix proposition must be assumed to be. It must be marked as non-interrogative and thus disallow any *wh*-constituents to be added to it to prevent sentences such as (3.49).

(3.49) *He didn't know that [ðət] what she wanted _ .

This introduces a difference between propositions initiated without a complementizer and those initiated with *that*.

3.2 Relative clauses

We turn now to another class of English constructions involving unbounded movement, finite relative clauses. The essential problem faced in analysing such constructions is similar to that for *wh*-movement, in that the grammar has to interpret *the cat* as the object of the modifying relative clause in (3.50) and as the subject of the modifying clause (3.51), even though neither of these phrases is in the canonical position to receive such an interpretation. They are distinguished from the constructions we have considered so far, however, in that these constituents are also to receive other semantic interpretations: in both sentences *the cat* also needs to be interpreted as the patient of *feed*.

(3.50) He fed the cat that the dog bit _ .

(3.51) He fed the cat that _ caught mice.

Given that we are representing interpretations as a network, rather than a tree, there is no inherent difficulty in the same conceptual constituent playing a role in more than one clause at the same time. What has to be added to the grammar is a rule for introducing a modifying proposition which contains the address of the modified constituent either as its subject or store value. This is achieved by adding one rule, which will be possible given the same environment as for *add-proposition*, that is either for no lexical item or the complementizer *that*.²⁷

transition rule	top of stack (S_n)	top of stack (S_{n+1})
add-finite-relative	(X)	(N, prop, [-X]) (X, mod: N)

So in the derivation of (3.50), this rule applies at the complementizer *that*, creating a proposition with the address of the constituent *the cat* in its store. Subsequently, the derivation continues completely in line with principles we have already established, with the constituent *the cat* receiving its second interpretation as the patient of *chase*.

²⁷In standard English, subject-gapped relatives must be initiated by *that* or an explicit relative pronoun, which will be discussed below.

- | |
|---|
| <p>(1) prop⁰, ACTION², past, feed, agent:2, theme:3, [+2†]</p> <p style="padding-left: 2em;">(2) MALE¹, sg</p> <p style="padding-left: 2em;">(3) def³, cat⁴, sg, mod⁵:4</p> <p style="padding-left: 4em;">(4) prop⁵, ACTION⁸, past, bite, agent:5, patient:3, [+5†, -3†]</p> <p style="padding-left: 6em;">(5) def⁶, dog⁷, sg</p> |
|---|

(3.50): **He fed the cat that the dog bit.**

The derivation of sentence (3.51), with the subject-gapped relative clause, is entirely parallel except that the address of the modified constituent is installed as the subject of the relative clause, resulting in it receiving its second interpretation as the agent of *catch*.

- | |
|---|
| <p>(1) prop⁰, ACTION², past, feed, agent:2, theme:3, [+2†]</p> <p style="padding-left: 2em;">(2) MALE¹, sg</p> <p style="padding-left: 2em;">(3) def³, cat⁴, sg, mod⁵:4</p> <p style="padding-left: 4em;">(4) prop⁵, ACTION⁶, past, catch, agent:3, patient:5, [+3†]</p> <p style="padding-left: 6em;">(5) mouse⁷, indef, plur</p> |
|---|

(3.51): **He fed the cat that caught mice.**

Such an approach treats the function of the complementizer *that* in relative pronouns as essentially the same as it is in initiating complement clauses. It thus runs counter to arguments that *that* in relative clauses should be seen as a relative pronoun, and that *that*-less relatives and *that*-relatives should be treated separately.²⁸ This approach agrees with the intuitions of Jespersen.

“We have thus brought together a great many phenomena, which traditional grammar puts into various separate pigeonholes, though they are in reality identical means of connecting a clause with the rest of the sentence, either without any form-word or with the empty and therefore in many cases superfluous particle *that*.” Jespersen (1927: 167)

Of course, finite complement clauses and finite relative clauses *are* quite different, and one could therefore argue that there is an *a priori* distinction to be drawn between *that* initiating one or the other. But it should be noted that this would also apply to a noun such as *he*, beginning a relative clause as in “the man he saw” or a complement clause as in “She thought he had left”. It would be distinctly odd to argue that *he* is of a fundamentally different nature in the two sentences, yet the same situation holds for the complementizer *that*.²⁹

²⁸As is the case, for example, in the HPSG treatment of relative clauses given in Sag (1997).

²⁹One of the arguments made in Sag (1997) to support this position is as follows: if relative *that* were a complementizer then it would appear to allow a sequence “complementizer + subject gap” which should

3.2.0.1 Adjunct relative clauses

The possibility of relative clauses in which the gap is realised as an adjunct in the modified clause, as in 3.52, follows directly from the transition schema given above.

(3.52) He liked the way (that) she whistled.

The derivation is given here, with the constituent *the way* being simultaneously the *theme* of the main clause and a manner modifier in the relative clause.

- (1) prop⁰, STATE², past, **like**, exper:2, theme:3, [+2†]
 (2) MALE¹, sg
 (3) def³, MANNER⁴, sg, **way**, mod⁵:4
 (4) prop⁵, ACTION⁶, past, **whistle**, agent:5, mod:3, [+5†, -3†]
 (5) FEMALE⁵, sg

(3.52): He liked the way she whistled.

3.2.1 Relative pronouns

We now have to consider relative clauses introduced by *which*, *who(m)*, *whose*, *where*, *when* and *why*. The simplest hypothesis is to suppose that these words initiate relative clauses using the same transitions given above, with the additional property that they place semantic restrictions on the identity of the constituent modified. This entails that relative pronouns do not form independent constituents in the conceptual representation, unlike *wh*-pronouns, so I shall refer to them instead as *relativizers*.³⁰

As an illustration, the derivation of the following subject relative example (3.53) introduces no constituent representing *whose* directly, the word merely serves to introduce the relative clause and start a constituent whose possessor is the constituent being modified by the clause.

(3.53) He pitied the woman whose cat had died.

The full derivation is as follows.

be ruled out by the *that*-trace effect. As we have seen, however, the present account of this effect only holds if a non-subject stored constituent must become a subject when inherited by the embedded clause, it would not stop a complementizer initiating a proposition with a subject gap.

³⁰That relativizers must be listed distinctly from *wh*-pronouns can be seen from the non-occurrence of *what* as a relative pronoun (in standard English), the non-occurrence of *how* and the fact that *whose* may have a non-human referent as a relative but not as a *wh*-pronoun.

- (1) prop⁰, STATE², past, **pity**, exper :2, theme:3, [+2†]
 (2) MALE¹, sg
 (3) def³, FEMALE⁴, sg, **woman**, mod⁵:4
 (4) prop⁵, STATE⁷, perf, past, content:6, [+5†]
 (6) EVENT⁸, ppart, **die**, theme:5
 (5) possess⁵:3, **cat**⁶, sg

(3.53): **He pitied the woman whose cat had died.**

Some evidence in support of this position, that relativizers do not create independent constituents, comes from the data in (3.54), which shows that it is possible to modify pronouns and *wh*-constituents with relative clauses, but impossible to do the same with relative pronouns.

- (3.54) a. He who likes eggs will pay the price.
 b. Who, that you like, did you see?
 c. *I know a girl who, who/that I know you would like, knows Bill.
 d. I know a girl whose father, who I know you would like, knows Bill.

This evidence is by no means conclusive, for there may be other reasons why sentences such as (3.54)c are uninterpretable. However, unless clear evidence can be found that relative pronouns do give rise to an independent constituent, the simplest assumption to make is that they do not.

3.2.2 Pied-piping

As was the case for *wh*-movement, we have to account also for constructions where the relative pronouns are non-initial in constituents. It should be noted that the distribution of pied-piped relative pronouns is much wider than that of pied-piped *wh*-pronouns, which leads us to expect a different type of analysis. Not only are indefinitely deep embeddings allowable, but also pied-piping within nominal phrases as in (3.55)b, and even clauses as in (3.55)d.

- (3.55) a. The committee objected to all books the size of the writing on the covers of which the government had prescribed.
 b. He pitied Mary, the cat of whom had died.
 c. *The cat of whom had died?
 d. The elegant parties, to be admitted to which was a privilege, ...
 e. ?*To be admitted to which parties was a privilege?

I will assume that these pied-piped relative clauses are initiated by the same rule that initiates *that*-less relative clauses, set out above. In these sentences, however, the address of the modifier placed in store is further inherited immediately by the subject or topic of the relative clause.³¹ This combination of transitions does not constitute a separate transition rule, but it may be helpful to show the combined effect of the two rules applying here in a single transition schema, as follows.

transition rules	top of stack (S_n)	top of stack (S_{n+1})
add-finite-relative & add-store	(X)	(N1, [-X]) (N, prop, [N1, -X*]) (X, mod:N)

This is in many respects a minimal assumption that makes use of the existing apparatus rather than introducing additional features. The derivation up to the first word of the relative clause in (3.55) will then be as follows.³²

(1) prop ⁰ , STATE ² , past, pity , exper:2, theme:3, [+2†]
(2) MALE ¹ , sg
(3) FEMALE ³ , sg, Mary
(4) prop ⁴ , asserted, [5, -3*]
(5) def ⁴ , [-3]

(3.55): S_4 **He pitied Mary, the** || cat of whom had died.

Having put an address on the store in this way, it is then necessary to ensure that this can only be interpreted by a transition involving an explicit relative pronoun. I will assume the following transition where the relativizer in effect performs the fixing of the stored constituent into an argument position. An added condition is that the type of the relative must match that of the stored constituent (Y).

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ relativizer, TYPE }	(X, arg: , [-Y])	(X, arg: Y, [-Y†])

Relative pronouns in this position then function rather like resumptive pronouns.³³

³¹Note that this is the only way the new constituents can be topicalised.

³²As this is a non-restrictive relative clause, as most pied-piped relatives are, I will assume that the modified constituent is not marked as restricted, and that the relative clause itself is marked as being asserted, in contrast with non-restrictive relatives, which are presupposed to be true. Clearly, these are just suggestive stopgap measures and a serious representation of these differences awaits further analysis.

³³Note, the set of relative pronouns that can be pied-piped appears to be slightly more restricted: *which*, *whom*, *whose*, but not *where*, *when* or *why*.

With this rule the address of *Mary* in the store of *the cat* can be interpreted with the aid of the relative pronoun and the derivation is completed as follows.

- (1) prop⁰, STATE², past, **pity**, exper:2, theme:3, [+2†]
 (2) MALE¹, sg
 (3) FEMALE³, sg, **Mary**
 (4) prop⁴, asserted, STATE⁸, perf, past, content:7, [+5*, -3*]
 (7) EVENT⁹, ppart, **die**, theme: 5, [+5†]
 (5) def⁴, **cat**⁵, sg, possess⁶:6, [-3*]
 (6) POSS⁶, **of**, content:3⁷, [-3†]

(3.55): **He pitied Mary, the cat of whom had died.**

The essential question to address with this approach to pied-piping is to rule out the element on the stack being interpreted as a “gap” as in (3.57)a, or alternatively to rule out gaps which are not inside constituents in the subject or topic position of relative clauses being interpreted with the help of relative pronouns as in (3.57)c.

- (3.57) a. *He pitied Mary, the cat of _ had died.
 b. Who did you say you saw a picture of _ ?
 c. *Who did you say you saw a picture of whom?

It seems the only property that distinguishes the positions of pied-piped relative pronouns is that the constituent the store addresses are in is itself yet to receive its semantic interpretation. For example, in (3.55) the constituent *the cat* has not yet been interpreted as the theme of *die*.³⁴ Therefore, it is necessary to add to the transition rule for pied-piped relativizers given above the condition that the active constituent (X) must not be interpreted. Similarly, we must add to the previous rule for the standard “non-lexical” interpretation of store values (that is interpretation as the head is added) the condition that the active category must be interpreted, in order to ensure a complementary distribution.

transition rule	lexical entry	top of stack (S _n)	top of stack (S _{n+1})
relative pied-piping	{ relativizer, TYPE }	(X, arg: , [-Y])	(X, arg:Y, [-Y†])

Condition: X is in store.

- (3.56) *He liked Liverpool, at the meeting where he had met Susie.

There is an interesting parallel here with the fact that obligatory parasitic gaps must also be referential and are thus restricted to the nominal types. As will be shown in the analysis of parasitic gaps in chapter 8, there are many similarities between the two constructions.

³⁴Similar appeal to the accessibility of constituents on the active stack is made in the analysis of parasitic gaps in chapter 8.

transition rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
interpret-store		(X, arg: , [-Y])	(X, arg: Y, [-Y†])

Condition: X is not in store.

Note that this approach places no limit on the depth of embedding of the relative pronoun, which agrees with the data.

Nanni (1978) notes that the constituent containing the pied-piped relative is not restricted to an NP or PP, giving the following examples.

- (3.58) a. That woman, compared to whom Attila the Hun was an angel, ...
 b. The elegant parties, to be admitted to which was a privilege, ...

I will not go into the proper analyses of these clauses at this time, but no matter if they are treated as subjects or topics, they will be uninterpreted and so the relative pronoun will be allowed to fix the store value.

It should also be noted that the pied-piped relativizer does not have to be final in the constituent. Steedman (1996: 51) notes that to allow such examples as (3.59), due to Morrill, would require extra generalization of the pied-piping categories, along with the assumption that the modifier *by Matisse* is an argument rather than an adjunct.³⁵

- (3.60) ...the woman a painting of whom by Matisse fetched a fortune.

Such examples would create no particular problem for the present approach.

3.2.2.1 Modification of non-nominals

The rules as we have formulated them do not refer to the syntactic category of the modified element, indeed they cannot do as syntactic categories are not present in the interpretation. They will therefore allow relative clauses to modify propositions, prepositional phrases, adverbs and adjectives, examples of which are shown in (3.61).

³⁵The present approach would treat the *by*-phrase as an adjunct, though others may not. However it is not difficult to construct a similar sentence with a constituent, or two, after the relativizer which cannot be taken to be anything else than an adjunct, as in (3.59).

- (3.59) The president, the shaming of whom on television yesterday caused a market crash today, has decided....

- (3.61) a. *He disliked Mary*, which _ surprised her friends.
 b. He lived *under a table*, which he found _ a warm and comfortable place.
 c. He lied *rarely*, which _ was too often for his mother.
 d. He was *happy*, which he had every right to be _ .

The derivation of (3.61)a follows the rules we have already established.³⁶

- (1) prop⁰, STATE², past, **dislike**, exper :2, theme:3, mod⁴:4, [+2†]
 (2) MALE¹, sg
 (3) FEMALE³, sg, **Mary**
 (4) prop⁴, asserted, EVENT⁵, past, **surprise**, theme:1, exper:5, [+1†]
 (5) def⁶, possess:6, **friend**⁷, plur
 (6) FEMALE⁶, sg

(3.61)a: **He disliked Mary, which surprised her friends.**

Hudson (1997) notes that a relative clause may also pick out a part of the proposition, seemingly corresponding to a verb phrase, as in (3.62)³⁷, where the referent appears to be '*eat(x, muesli)*'.

- (3.62) Peter ate muesli, which Joe never does.

He uses this example to motivate a Stepwise Composition Principle, according to which the meaning of verbs combines first with the meaning of its object, and then this predicate combines with the meaning of the subject to form the meaning of the whole proposition.

However, as Hudson himself points out, it also seems possible, though admittedly more difficult, to construct examples where any elements of the proposition are left out in forming the entity referred to. For example, in (3.63)a the meaning picked up is '*kiss(Peter, x)*' and in (3.63)b it is '*kiss(x, y, on-cheek)*'.

- (3.63) a. Peter kissed Mary, which is something that never happened to Jane.
 b. Peter certainly kissed Mary on the cheek, which John never did to Jane.

Although I provide no analyses for these sentences, it appears that they must require some more general principle of abstraction, rather than one that only picks out the predicate.

³⁶The fact that the modification of non-nominals is limited to non-restrictive relative clauses is surely connected to the fact that restrictive relative clauses restrict the reference of properly-referring objects, which are generally taken to be nominal. Again, work needs to be done on clarifying precisely what the notion of reference involves.

³⁷Hudson's ex. 67a.

3.2.3 The *who* vs. *whom* controversy

We have so far not considered the use of *wh*-words which are explicitly marked for case. In terms of the present model these are *nominative* which create surface-subjects in store, and *non-nominative* or *accusative* which create non-subjects. In modern colloquial English *who* is generally used in both of these roles, ie. it is unmarked for case.³⁸ However the distinction of the nominative *who* and accusative *whom* was a feature of earlier English and is preserved in formal registers.

There has been some controversy about the correct usage of *who/whom*, in those registers where the distinction is preserved, when the pronoun is the filler for a subject gap in a lower clause.³⁹

- (3.64) a. I dislike the man who you say _ has already left.
 b. I dislike the man whom you say _ has already left.

The typical view of traditional and prescriptivist grammarians has been that the pronoun should accord with the form it would have *in situ*, and therefore that (3.64)a is the correct form of the two above. Although sentences such as (3.64)b are often met with they are considered by this tradition to be a careless or even hypercorrect mistake. This is stated most vehemently by Fowler.⁴⁰

“the *whom* form, though probably no grammarian would have a word to say for it, is now so prevalent in the newspapers that there is real danger of its becoming one of those STURDY INDEFENSIBLES of which the fewer we have the better...”

That every *whom* in those quotations ought to be *who* is beyond question, & to prove it is a waste of time since the offenders themselves would admit the offence...” Fowler (1926: 724-725).

The opposing view, that the accusative form is essentially correct in such cases, is persuasively championed by Jespersen. He supports his argument by noting the overwhelming preponderance of *whom* in such positions in earlier writings, from times when the distinction of *who* and *whom* was still maintained in the colloquial language.⁴¹ These include the following quotations from Shakespeare and the King James Bible.⁴²

³⁸Although, as discussed above, generally not in pied-piped positions in those registers which allow pied-piping, again suggesting that this is a functionally separate role.

³⁹For an extended discussion of this point see Stuurman (1990: 261-266).

⁴⁰Similarly in Quirk *et al.* (1985: 368): “The hypercorrect use of *whom* is common in examples of pushdown relative clauses such as this: *The Ambassador, *whom* we hope will arrive at 10 a.m.,...”.

⁴¹Jespersen (1927: 197-201).

⁴²See Dot Wordsworth’s column on language matters in *The Spectator*, 16 September 1995, where a reader claims that example (3.65)b is one of only two “grammatical errors” in the Bible.

- (3.65) a. Young Ferdinand, whom they suppose is drown'd. *Tempest*, III, iii, 92.
 b. But whom say ye that I am? *Matt.xvi.15*.⁴³

The present approach to unbounded movement is only compatible with the latter analysis (fortunately, for one does not wish to argue against the combined authority of Jespersen, Shakespeare and the King James Bible). As the grammar stands there is no way to distinguish those constituents in store that are destined to become subject of an embedded clause from those destined to become objects. It is the switching transition that creates subjects out of these stored constituents. This lack of distinction is backed up by the evidence of coordination, as the following examples, varying somewhat in literary quality, both attest.

- (3.66) a. Which was the girl that you liked __ , and thought __ might marry Peter?
 b. There is a vice that most I doe abhorre,
 and most desire should meet the blow of iustice.⁴⁴

In addition to his historical evidence, Jespersen rationalizes the argument in the following manner:

“The form *whom* is used because... the speech-instinct would be bewildered by the contiguity of two nominatives, as it were two subjects in the same clause.” (ibid: 199).

This argument has a direct parallel in the present model in that each clause can only contain one surface subject.

It appears that these data have not been widely appreciated in many current models of syntax,⁴⁵ most of which appear to assume that the case of the filler and the gap must match, although it is difficult to see how this position can be reconciled with the coordination evidence given above.⁴⁶

3.2.4 Free relatives

These constructions are characterized by a phrase-initial *wh*-word and an obligatory relative clause restricting the (often almost vacuous) meaning of the matrix phrase. Cann &

⁴³Interestingly, recent editions of the Bible, as well as “correcting” the *who* to *whom*, have also removed the *that*. “Who do you think that I am?” feels suspiciously like a *that*-trace violation, although my intuitions about this are unclear, and I am not aware of any discussion of such instances in the literature.

⁴⁴Shakespeare, *Measure for Measure* II 2.31, quoted in Jespersen (1927: 200) in order to make precisely this point.

⁴⁵An exception being Steedman (1987).

⁴⁶Radford (1988: 591-592) agrees with the judgement that either *who* or *whom* is possible in the initial position of such deeply subject-gapped relative clauses, but provides no account of it, setting it instead as an exercise for the reader.

Tait (1989) make the distinction between *headed* free relatives, as in (3.67)a, and *incorporated* free relatives, as in (3.67)b.

- (3.67) a. He bought her whatever she wanted _ .
 b. He bought her whichever cake she wanted _ .

Unlike the relative clauses we have seen so far the relative clause modifier in such constructions is not optional.

- (3.68) a. *He bought her what.
 b. *He bought her whichever cake.

As modifiers are by definition optional elements, the only way such constructions can be modelled in the grammar is if the relative clause is initiated by the initial *wh*-word itself. So in the two example sentences the relative clauses are initiated at *what* and *whichever* respectively. These are specified directly by the lexicon and are attached in argument positions using the rules already given. For example, the transition for *whatever* in argument position, as in (3.67)a, will be the following.

rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
free relative	{rel, emphatic}	(X)	(N1, prop, [N]) (N, mod: N1) (X, arg: N)

This transition combined with the ones already established will allow the following derivation of (3.67)a.

(1) prop ⁰ , ACTION ² , past, buy , agent:2, benef:3, theme:4, [+2†]
(2) MALE ¹ , sg
(3) FEMALE ³ , sg
(4) NON-HUMAN ⁵ , mod ⁴ :5
(5) prop ⁴ , STATE ⁶ , past, want , exper:6, theme:4, [+6†, -4†]
(6) FEMALE ⁵ , sg

(3.67)a: **He bought her whatever she wanted.**

Again class membership of this group of *wh*-items is different to both standard *wh* and relative pronouns, indicating the need for listing in the lexicon.⁴⁷

⁴⁷Note that *why* is not a member of this group, leaving a gap in the paradigm:

It should also be clear from the previous discussion that the creation of the relative clause by the free relative head itself will preclude the possibility of the complementizer *that* appearing in such constructions, since its sole function is to initiate finite clauses.

- (3.70) a. *He bought her what that she wanted.
 b. *He bought her whichever cake that she wanted.

Thus the dynamic grammar provides a unified explanation for this phenomenon and the *that*-trace phenomenon discussed previously, although in this instance there is not necessarily a subject gap.⁴⁸

3.2.4.1 Case mismatches

The relativizer in free relative constructions is in some sense playing two roles at the same time, one external role of argument or modifier in the external clause and another internal role in the restrictive clause. This often results in conflict as to which case the relativizer itself should take. Given the archaic nature of the accusative *whomever* and *whomsoever* in English, linguistic intuitions are lacking. However, free-relative case mismatches in other languages do provide interesting insights into the syntactic process.

For example, Cann & Tait (1989) discuss the following data in German. When free-relatives appear in their canonical position before the verb, which must check their case to fulfill subcategorization requirements, the case required by the matrix and the modifying clause must match. For example, the verb *nehmen* requires an object with accusative case and will not admit free relatives whose case differs to accord with the requirements of the restricting clause.

- (3.72) a. Ich muss wen du mir empfiehlst nehmen.
 I must who(acc) you me(dat) recommend take
 I must take whoever you recommend to me.
 b. *Ich muss wer einen guten Eindruck macht nehmen.
 I must who(nom) a good impression makes take

- (3.69) a. He will do it when he always does it.
 b. He did it how he always does it.
 c. *He did it why he always does it.

⁴⁸Grimshaw (1977) notes that a long incorporated free relative phrase may allow a complementizer: "In general, increasing the distance between *that* and the *wh* phrase in free relatives leads to greater acceptability." Grimshaw (1977: 99). And the following example is given.

- (3.71) I'll read whatever long boring books with red covers that you give me.

I must take whoever makes a good impression.

- c. *Ich muss wem du vertraust nehmen.
 I must who(dat) you trust take
 I must take whoever you trust.

However, if the free relatives are extraposed (a possibility which they share with other finite clauses) then it is reported that such mismatches are either acceptable (in some Southern German dialects) or at least their badness is to some extent ameliorated.

- (3.73) a. Ich muss nehmen, wen du mir empfiehlst.
 b. %Ich muss nehmen, wer einen guten Eindruck macht.
 c. %Ich muss nehmen, wem du vertraust.

Here the case of the free-relative argument cannot be checked immediately the verb *nehmen* is added, since it has been extraposed. It will be assumed that such extraposition is analyzed in a similar manner to heavy-element shift in English and that some place holder is temporarily employed to fill the missing object position and that this is then substituted by the extraposed constituent. It appears that at least in Southern German dialects this indirect relation between the verb and its argument allows the case value of the extraposed free relative to escape being checked.⁴⁹

It might be seen as a drawback of the dynamic model model that checking the case of pre-head and extraposed arguments must be carried out by two separate mechanisms, but here such a division seems to find some support from the data. Indeed it may be the case that modelling such a distinction might prove to be problematic for unification-based models of agreement.

Another instance of case mismatching in German free relatives is noted by Ingria (1989). As shown above in (3.72)b & c, case mismatches where the free relative is in canonical argument position are normally disallowed. However, entirely parallel sentences are admissible just in case the pronoun used has the same form for both cases required, as shown in the following examples.

- (3.75) a. Was du mir gegeben hast, ist prächtig.
 what_{nom/acc} you me given have is wonderful

⁴⁹It is assumed that finite verbs in V2 position can leave their external arguments temporarily unfilled, as in English, checking the required case when they are filled, so the following example will be ruled out.

(3.74) *Ich nehme wer einen guten Eindruck macht.

What you have given to me is wonderful.

- b. Ich habe gegessen was noch übrig var.
 I have eaten what_{nom/acc} still left was
 I ate what was left.

As Ingria points out this is incompatible with an account of case constraints on arguments that is based on unification. It suggests rather that constituents may carry disjunctive values for case, and that these features are then checked by the verb rather than unified with the argument requirements.

3.3 Wh-exclamatives

The *wh*-exclamative construction is similar to the *wh*-interrogative in that a *wh*-word is placed sentence initially. The chief differences are that there is no subject-auxiliary inversion and the *wh*-words participating in the construction are severely limited to those linked to degree: the degree modifier *how* and a degree usage of *what* peculiar to this construction.

Let us look at a simple example (3.76)a, which contrasts with the corresponding interrogative (3.76)b, and work out what the transition rule introducing it must be.⁵⁰

- (3.76) a. How nice she was _ !
 b. How nice was _ she?

As the exclamative construction is limited to being initiated by one of two words we can link the transition directly to their lexical entry.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ how , wh, degree }	(X, prop, [])	(N1, DEGREE, ref:-) (N, mod: N1) (X, prop, exclam, [N])

Before we can consider the derivation of the exclamative, we must consider that of a simple copular construction with a predicate adjective as in (3.77).

- (3.77) She was nice.

In the derivation, it will be assumed that the adjective introduces a constituent of type QUALITY which inherits the subject from the copular and interprets it as its theme.

⁵⁰It will also be necessary to introduce here for the first time an analysis of predicate constructions.

- (1) prop⁰, STATE², past, **be**, content:3, [+2*]
 (3) QUALITY³, **nice**, theme:2, [+2†]
 (2) FEMALE¹, sg

(3.77): **She was nice.**

In the exclamative version, where by necessity the identification of the theme is delayed, I shall have to assume that a placeholder α is available to mark the address of the future subject, and which is therefore interpreted as the theme of *nice*. When the *nice* constituent is taken off the store to fill the content role of the copular, the two subjects must be equated with each other, with the concrete address of the *she* constituent, (4), substituting for the placeholder α . This ensures that *she* is eventually interpreted as the theme of *nice*, as shown in the following derivation diagram.⁵¹

- (1) prop⁰, exclam¹, STATE⁴, past, **be**, content:2, [+4*, -2†]
 (2) degree:3, QUALITY², **nice**, theme:4, [+4†]
 (3) DEGREE¹, ref:-
 (4) FEMALE³, sg

(3.76): **How nice she was!**

Turning now to the exclamative construction with exclamative *what*, it is first necessary to note that there is no parallel *wh* construction.

- (3.78) a. What a writer she is!
 b. *What a writer is she?

The position of *what* preceding the indefinite determiner is typical of degree phrases, as is its non-occurrence with other determiners.

- (3.79) a. How good a writer is she?
 b. She was too good a writer to write this kind of thing.
 c. *What the writer she is!

Again the exclamative transition can be associated directly with the lexical entry:

⁵¹The derivation of the corresponding interrogative, (3.78)b above, is similar, again requiring a temporary placeholder to mark the theme of *nice*.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ what , exclamative }	(X, prop, [])	(N1, DEGREE, ref:-) (N, indef, degree: N1) (X, prop, exclam, [N])

This transition stipulates that the constituent on the stack is indefinite⁵², and will allow an analysis of (3.78)a as follows.

(1) prop ⁰ , exclamative ¹ , STATE ⁵ , pres, be , content:2, [+4*, -2†]
(4) FEMALE ⁴ , sg
(2) degree ¹ :3, indef, sg ² , writer ³ , ref ⁵ :4, [+4†]
(3) DEGREE ¹ , ref:-

(3.78)a: **What a writer she is!**

The fact that exclamative *what* marks its dependent constituent as indefinite rules out (3.79)c, but will allow an indefinite plural as in (3.81).⁵³

(3.81) What writers they are!

3.3.0.2 Embedded exclamatives

Exclamative constructions may also occur in complement clauses, which will be allowed by the rules given above. However, as there is no inversion in embedded interrogative constructions, the same string may often be interpreted as either an interrogative or an exclamative clause, as is the case for (3.82)a. Sentence (3.82)b can only be an embedded exclamative, however.

- (3.82) a. He told me how good she was __ .
b. He told me what a great place Leeds was __ to live in.

3.4 Multiple-*wh*

The rules for introducing clause-initial *wh*-constituents given above already rule out two distinct *wh*-constituents appearing one after the other in clause initial position, assuming

⁵²It would of course be preferable if this was derivable from other, semantic factors.

⁵³Interestingly other degree constructions appear to be at best extremely awkward with indefinite plurals.

- (3.80) a. ??How good writers are they?
b. ??They were too good writers to write this kind of thing.

Why this should be the case and is not the case for the exclamative is something of a mystery.

that the interrogative feature Q may only be assigned once to a clause. So a question such as (3.84) is impossible.⁵⁴

(3.84) *Which dog_{*i*} which bone_{*j*} did you give ___{*j*} to ___{*i*} ?

However, it makes perfect sense to form questions asking for multiple pieces of information and the language has to permit a way for this to be done. In English this is achieved by placing any additional *wh* constituents in their canonical positions later in the clause. The restrictions on the placement of these constituents have been the subject of intense speculation in theoretical syntax over the last decade, much of it attempting to determine the hierarchical restrictions imposed on the appearance of these secondary *wh*-constituents by the hierarchical constituent structures.⁵⁵

Consider the following set of data.

(3.85) a. Who said what?
b. *What did who say __ ?

Ignoring echo questions, which will not be dealt with here, the first thing to note is that acceptability of a *wh*-constituent appearing in situ is dependent on there being an initial *wh*-constituent in the clause. We can rule out the unacceptable (3.85)b by adding the condition that this *wh*-constituent must have received a semantic interpretation before the *wh* in-situ constituent is added, as here when *who* is added the initial *what* is yet to be interpreted and is still in store. Again in employing the dynamic model we are replacing hierarchical considerations by linear ones.

If this assumption is correct we would expect the pattern of data found in the two following examples. In (3.86)a, the initial *wh*-constituent *which dog* is interpreted at *give* as the beneficiary and so the *wh* in-situ theme should be possible, while in (3.86)b it does not receive an interpretation until *to*, *after* the second *wh*-constituent is added, and the *wh* in-situ theme is predicted to be impossible.⁵⁶

(3.86) a. Which dog did you give __ which bone?
b. ??Which dog did you give which bone to __ ?

⁵⁴It might be thought that this might be related to a general information packaging restriction on having more than one item as sentence topic, but interestingly double topicalization, although undoubtedly difficult, does not appear ungrammatical.

(3.83) ?Fido_{*i*}, that type of bone_{*j*} I would never give ___{*j*} to ___{*i*} .

⁵⁵Kuno & Takami (1993) offer alternative functional explanations of the data, some aspects of which have been incorporated into the present brief analysis.

⁵⁶Sentence (3.86)a will only be acceptable to those English speakers (predominately British?) who allow extraction of the benefactive argument in such constructions.

Acceptability judgements on multiple *wh* questions are more fickle and unreliable than most, but there does seem to be a difference between the two. Again although judgements are less than clear it does seem worse to add two *wh*-constituents violating this condition, as in (3.87)c, than an otherwise identical sentence with just one violation, (3.87)b.

- (3.87) a. I can't remember who said what to whom.
 b. ??I can't remember what who said _ to whom.
 c. *I can't remember who who said what to _ .

The situation is rather more complex, however, as the appearance of *wh* in situ improves markedly with increased specificity of the *wh*-constituent, resulting in examples parallel to the unacceptable (3.85)b and (3.85)d sounding much better.

- (3.88) a. ?Which of the boys did which of the girls like _ ?
 b. ?I can't remember which of the girls which of the boys danced with _ .

The correlation between level of specificity of *wh*-phrases and their amelioration of syntactic awkwardness is known in other constructions, but modelling such phenomena is outside the scope of the system of syntactic rules as given here.

The idea that we can replace hierarchical relations in syntactic structure with questions of whether constituents have or have not received an interpretation will be a recurrent theme, one involved in many of the constructions considered in this thesis that involve binding and coreference.

3.5 Constraints on movement

Ross (1967) introduced the notion of island constraints, detailing many situations in which extracting phrases out of their canonical position is impossible. In this section, I will briefly consider how these constraints relate to the account of unbounded movement presented in this chapter.

Earlier in the chapter a constraint was placed on the inheritance of stored constituents, based on Kuno's Clause Nonfinal Incomplete Constituent Constraint, prohibiting an argument inheriting a store value from a constituent which still had argument slots unfilled. It was argued that such a restriction has a functional explanation in terms of minimizing local ambiguity.

Extraction from subjects It so happens that the architecture of the model is such that this restriction simultaneously rules out extraction from subjects. This follows from the fact that *wh*-movement entails that an auxiliary comes before the subject and places its subcategorization requirements, typically a *content* role, on the proposition. When the subject is reached then, there is always an unfilled role, and thus Kuno's constraint blocks store inheritance into the subject without further stipulation.

So, for example in (3.89)b, inheritance of the store into the subject *a picture* is disallowed as there is an unfilled *content* role introduced by the inverted auxiliary *did*.

- (3.89) a. Who did he want to see a picture of _ ?
 b. *Who did a picture of _ appear in the paper?

Note that extraction from subjects is possible in relative clauses, but apparently only seems to be better where the relativizer is pied-piped, as the following contrast shows.

- (3.90) a. These people, of whom the majority _ appear to be women, wear no socks.
 b. ?I liked that band who the majority of _ appeared to be women?
 c. *Which group do the majority of _ appear to be women?

As the model stands it predicts that extraction from all subjects in relative clauses should be admissible as there is no inversion, but this question is clearly in need of further investigation.

Extraction from complex noun phrases The Complex NP Constraint of Ross (1967) is exemplified by the following data, where (3.91b) shows that extraction of a *wh*-constituent from a position inside a subject relative clause is disallowed, while (3.92b) shows that the same is true of an object relative clause.

- (3.91) a. He met the man that wanted to buy the dogs.
 b. *What did he meet the man that wanted to buy _ ?
- (3.92) a. He met the dogs that Peter had given _ to Jimmy.
 b. *Who_i did he meet the dogs_j that Peter had given __j to __i ?

As mentioned previously postmodifiers do on occasion allow extraction out of them⁵⁷ although the results are variable depending on semantic factors. To rule out extraction

⁵⁷Although the difficulty of such extraction has been formalized as the Adjunct Island Constraint.

from relative clauses for syntactic reasons one would have to introduce a specific condition on relative clauses that unlike other modifiers they may never inherit store values. This constraint could certainly be added to the transition schemas that introduced them, given in the previous section of this chapter.

However, there do seem occasions where extraction is possible, helped by increased specificity of the extracted constituent and “semantic lightness” of the noun phrase being extracted out of.⁵⁸ For example, it is quite easy to construct passable sentences involving extraction from a position inside subject relatives as in (3.93)a. This is much more difficult for extraction from object relatives, as witnessed by (3.93)b, although I would hesitate to call it ungrammatical.

- (3.93) a. ?Which of these dogs did you say you knew someone that wanted to buy _
?
b. ??Which of these guys_i did you say you had seen something_j that Peter had sent __j to __i ?

If this is the case, then it seems that the grammar must allow extraction from, at least subject-gapped, relative clauses as grammatical. If we construct a weighted grammar, which will be discussed in chapter 9, then the relative badness of such sentences must arise from an interaction of various factors, including the information structure, relative specificity of constituents and the relative low weighting of extraction from modifiers.

Extraction from interrogative clauses Similarly a constraint on extraction from embedded interrogative clauses has also been suggested, the *Wh*-island Condition.

- (3.94) a. He knew who liked fish.
b. *What did he know who liked _ ?

Again the felicity of such sentences improves with a more specific extracted constituent, and even seems marginally possible with extraction from an object interrogative, which will have for a time two stored constituents, as in (3.95)b.

- (3.95) a. Which of the girls can you not remember who was in love with _ ?
b. ?Which of the girls can you not remember what you gave _ _ ?

It seems that again we are forced to model this effect as the combination of a complex of factors in a weighted model, but that a blanket ban is unsupportable.

⁵⁸A number of examples are given in Pollard & Sag (1994: 225).

3.6 Summary

The main innovation of this chapter has been an extension of use of the syntactic feature *store*. Previously used to hold the subject in declarative sentences until it is in a position to be interpreted, it is here also used to hold fronted *wh*-constituents until they can be interpreted. This usage has obvious parallels with the SLASH feature of GPSG and its descendants, holding the addresses of constituents that still need to be interpreted, as well as the HOLD feature of ATN's. I introduced rules, generally associated with *wh*-words, for introducing constituents onto the store, modified the rule of complementation to allow store addresses to be inherited by complements, and a single rule for their eventual interpretation by lexical heads. Together these rules allow the grammar to model the basic patterns of *wh*-movement, interrogative argument clauses, finite relative clauses, free-relatives and *wh*-exclamatives.

I argued the need for the adoption of a constraint, originally due to Kuno, that blocks inheritance of store values into non-final arguments. It later transpired that without further stipulation this constraint also disallows extraction from subject position, but only where the subject is inverted, which again appears to agree with the data.

In the analysis of extraction from embedded finite clauses, I showed that the most economic analysis of subject extraction in English requires “early switching” of the extracted constituent from a non-subject to the subject of the embedded clause. This is not only consistent with the psycholinguistic evidence, but also gives rise without stipulation to the *that*-trace effect.

I addressed the question of how to capture the phenomenon of pied-piping, speculating that two dissimilar processes were involved in *wh*-movement and relative clauses, with the relative pronoun in the latter behaving more like a resumptive pronoun. Adopting this analysis it transpires that the position of constituents on the active stack is of crucial importance here. The choice of analysis has been motivated by a concern with keeping the extra syntactic machinery to a minimum.

There are many other constructions involving unbounded movement, all of which it is assumed will make similar use of the store feature. Some of these constructions will receive an analysis later in the thesis as they depend on the prior consideration of other phenomena: *tough* movement is considered at the end of the next chapter on complement control, cleft and pseudoclefts are tackled in chapter 7, and parasitic gaps in the final syntax chapter, chapter 8.

Chapter 4

Complement control

In the previous chapter I examined how unbounded movement constructions allow constituents to be placed an indefinite distance from their canonical place of interpretation. There are a large family of other constructions where the positions of elements can be thought of as moved from their canonical positions for interpretation, but where this dislocation is bounded in distance and thus cannot be analyzed with the store mechanism. These will be examined in this chapter, with a great deal of the analysis following or reinterpreting the HPSG analysis of Pollard & Sag (1994).

4.1 Raising

The sentences (4.1)a and (4.1)b have, modulo any slight differences in point of view, identical interpretations. Logically speaking, *appear* is a one-place predicate which takes a proposition as an argument. The “surface structure” of the two sentences is quite different, however, as *he* is the surface subject of *appear* in (4.1)a and the surface subject of *likes* in (4.1)b. Traditionally, the subject of the matrix clause in (4.1)a is viewed as having been raised up from the complement clause. The challenge for the model of syntax is to achieve identical interpretations for the two sentences.

- (4.1) a. He appears to like dogs.
 b. It appears that he likes dogs.

It might be expected, given the fact that it has one propositional argument, that the lexical entry for the verb *appear* would have the following form.

sem: ⟨ STATE, **appear**, [theme: prop] ⟩

But we also have to encode the fact that from a syntactic perspective *appear* does not interpret its surface subject as its argument (otherwise the derivation of (4.1)a would fail at “*he appears*” since the surface subject does not satisfy the selectional restrictions). In fact, the surface subject in (4.1)a bears no direct semantic relation to *appear* at all. A simple solution to this problem would be to suppose that the lexical entry contains another purely syntactic argument, which I shall term the empty argument.

sem: $\langle \text{STATE, appear, } [\emptyset : \text{NULL, theme: prop}] \rangle$

Naturally, the empty argument places no semantic restrictions on the constituent filling it.

So with this entry for *appear* and the standard syntactic rules which have been used so far, the interpretation state after two words of (4.1)a will be as follows.

<p>(1) prop⁰, STATE², pres, appear, \emptyset:2, theme: , [+2] (2) MALE¹, sg</p>
--

(4.1)a: S_3 **He appears** || to like dogs.

We know that the constituent *he* has to be interpreted as the first argument of the following verb. We also know from the previous chapter that verbs in English have to have a surface subject in place before they can be added to the sentence. From this we can infer that the function of the infinitival marker *to* must be to create a new proposition, with its subject that of the empty argument.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{to, inf-complementizer }	(X, arg:Y)	(N, prop, inf, [+Y]) (X, arg:Y, arg:N)

The proposition is marked *inf* which will restrict the verb being added to it to being in base form.

This transition then takes the derivation of (4.1)a to the following state.

<p>(1) prop⁰, STATE², pres, appear, \emptyset:2, theme:3, [+2*] (3) prop³, inf, [+2] (2) MALE¹, sg</p>
--

(4.1)a: S_3 **He appears to** || like dogs.

The rest of the derivation now proceeds in the customary fashion.

- (1) prop⁰, STATE², pres, **appear**, \emptyset :2, theme:3, [+2*]
 (3) prop³, inf, STATE⁴, **like**, exper:2, theme:4, [+2†]
 (2) MALE¹, sg
 (4) **dog**⁵, plur

(4.1)a: **He appears to like dogs.**

It will be seen that the whole analysis of the raising sentence (4.1)a has been achieved by two innovations. Firstly by ensuring that raising verbs subcategorize for a null first argument, and secondly by assigning the infinitival marker *to* a special clause-initiating function giving it a status not unlike that of a complementizer.¹

Returning to the alternative formulation of the sentence, (4.1)b, it transpires that we only have to add a new lexical entry for the “dummy” pronoun *it* and the derivation follows without any addition to the rules we have established. The entry for dummy *it* contains the syntactic information that it is nominal and singular² but is devoid of semantic content. It has no problems functioning as the surface subject of a raising verb as this only checks the syntactic features of its subject in any case.³

At the third transition a finite clause is created by the finite complementizer *that* and this clause is interpreted as the final complement of the verb *appear* in entirely the usual manner.

- (1) prop⁰, STATE², pres, **appear**, \emptyset :2, theme:3, [+2†]
 (2) sg¹
 (3) prop³, STATE⁵, pres, **like**, exper:4, theme:5, [+4†]
 (4) MALE⁴, sg
 (5) **dog**⁶, plur

(4.1)b: **It appears that he likes dogs.**

Comparing the final interpretations of (4.1)a and (4.1)b it will be seen that they are semantically identical, the only difference being in the surface subject of the proposition,

¹In GPSG and HPSG *to* is taken to be a type of auxiliary verb, following Pullum (1982a). Lencho (1992) makes a number of arguments for regarding *to* as a complementizer however. Of course, it makes little difference how it is classified as long as its function is described correctly.

²Or at least so I will assume for the present. In chapter 7, I argue that it has no number but exhibits singular agreement by default as do propositions.

³I will assume that this “dummy *it*” is distinct from “weather *it*”, which does appear to have some minimal ambient referent, and therefore will be assigned a role by weather verbs and even on occasion from other verbs, as in (4.2).

- (4.2) a. It's trying to rain.
 b. *It's trying to appear that he likes dogs.

with (4.1)b containing the semantically empty constituent formed by the dummy pronoun *it*. Only such semantically vacuous constituents will be allowed to remain uninterpreted in a successful derivation. Furthermore, it should be noted that the two sentences are possible with a single lexical entry for the raising verb *appear*.

It should be clear from the approach taken above, and the discussion in the previous chapter, why the derivation of the following sentence (4.3) fails.

(4.3) *He appears likes dogs.

Although the constituent *he* will be uninterpreted when the verb *likes* is reached, there is no proposition for this verb to be placed in and the derivation fails.

Similarly, the derivation of (4.4) also fails without further stipulation. Although the sentential surface subject would appear to be a suitable argument for the verb *appear*, its lexical entry stipulates that the subject is not to receive an interpretation and so the derivation cannot go through.

(4.4) *That he likes dogs appears.

Raising verbs can of course appear with sentential surface subjects, as in (4.5), but then these subjects must be interpreted as the logical subject of the complement clause.

(4.5) That he likes dogs appears to surprise you.

The next section will be a brief interlude on the syntax of propositional subjects.

4.1.0.3 Propositional subjects

The derivation of (4.5) demands a transition to introduce a propositional subject, which must be linked directly to the complementizer *that*.⁴

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ that , complementizer}	(X, prop)	(N, prop) (X, prop, [N])

With this transition and supposing a lexical entry for *surprise* allowing a propositional subject, the derivation is unproblematic.

⁴Of course, this transition would also allow the proposition to turn out to be a topic as discussed at the beginning of the previous chapter.

- (1) prop⁰, STATE⁵, pres, **appear**, \emptyset :2, theme:5, [+2*]
 (5) prop⁶, inf, STATE⁷, **surprise**, theme:2, exper:6, [+2†]
 (2) prop¹, STATE³, pres, **like**, exper:3, theme:4, [+3†]
 (3) MALE², sg
 (4) dog⁴, plur
 (6) 2ND-PERSON⁸

(4.5): **That he likes dogs appears to surprise you.**

Although they have often been cited as examples of grammatical sentences which are unacceptable for processing reasons (ie. having a status akin to that of the garden-path sentences discussed in chapter 1), Fodor (1992) argues convincingly that sentences with propositional subjects lacking a complementizer, such as (4.6), are ungrammatical, and not just unprocessable. That is the presence of a complementizer in the subject clause is a necessity specified in the grammar, although the *reason* for the syntax being so may safely be ascribed to the constraints of processing.

(4.6) *He likes dogs appears to surprise you.

Given that the rule for introducing propositional subjects and topics is associated directly with the complementizer *that*, there is no way of deriving (4.6) with the present grammar.

There are other restrictions on sentential subjects, many of them first described in Iwakura (1976). Firstly, they may not appear in inverted subject position. The transition rule given above will not generalize to inverted subjects and so the examples in (4.7) will not be allowed.

- (4.7) a. *Does that he likes dogs surprise you?
 b. *How likely is that John will win the election?

Judging by the unacceptability of the exclamative example (4.8)a, which does not involve subject inversion, it may appear necessary to place a further restriction on sentential subjects, that the store of the proposition to which they are added be empty.

- (4.8) a. *How strange that he doesn't like pickled onions is __ !
 b. How strange that man is __ !

However there do seem to be occasions where topicalized constituents do appear before the sentential subject.⁵

⁵Iwakura rates this example as ??, but to me it seems impeccable.

(4.9) To me, that the world is round is obvious _ .

Such a restriction on the store is therefore unwarranted.⁶

The rule given above for introducing propositional subjects places no restriction on their being the subject of a matrix clause, and thus also allows them to appear as the subjects of complement clauses. This seems to be correct, for although they are generally very difficult to process, they can be improved in various ways, which suggests that they are grammatical. For example, provided the right intonation is supplied, (4.12) seems perfectly acceptable.⁷

(4.12) I don't think that he likes pickled onions is at all strange.

4.1.1 Extraposition

There is a clear correlation between sentences with propositional subjects and those allowing an apparent dummy expletive subject and an extraposed sentential argument. For example, (4.13)a & b are equivalent in meaning.

- (4.13) a. That he liked pickled onions shocked her.
 b. It shocked her that he liked pickled onions.

From (4.13)a it follows that, in distinction to raising verbs, the verb *shock* can assign a thematic role to its surface subject and will therefore have the following lexical entry.

sem: ⟨ STATE, **shock**, [theme: prop, exper: ANIM] ⟩

A common approach to dealing with the fact that the same sentence may also appear with initial *it*, as in (4.13)b, is to associate them, via some general lexical rule applying to all propositional subject verbs, with another lexical entry explicitly subcategorizing for an expletive subject and moving the propositional argument to final place on the list of arguments, giving the following entry.

⁶Perhaps the reason for the badness of the exclamative construction (4.10)b is related rather to the existence of a reduced exclamative construction, consisting solely of a exclamative adjectival phrase.

- (4.10) a. How lovely!
 b. How strange that he doesn't like pickled onions!

It appears that (4.10)b is to be analyzed as an example of extraposition, which is discussed in the following section.

⁷It is perhaps even clearer following a subordinating conjunction where the complementizer *that* must unambiguously begin a subject clause.

- (4.11) It sounds like that he ate all the pickled onions has really shocked her.

This construction is discussed at the end of this section.

sem: ⟨ STATE, **shock**, [∅ : NULL, exper: ANIM, theme: prop] ⟩

Such a solution is also available in the present model, but it has a number of problems with it and so will not be adopted. Instead I shall argue that the initial pronoun in (4.13)b is not the semantically-vacuous expletive which occurred previously in the raising examples. Instead it will be analyzed as a genuine pronoun of type PROPOSITION standing in a cataphoric relation to the extraposed proposition. Such an analysis makes it akin to nominal cataphoric construction such as (4.14), with the difference that the propositional construction, being much more common and useful for reasons of placing long constituents last, has become more routinized.

(4.14) He won the election, that funny guy with the beard.

The first advantage to note of this approach is that it entails that only one lexical entry is necessary for verbs (and other parts of speech) that take propositional subjects. Encoding in the lexicon does not seem a good option, since all constructions involving propositional subjects have a variant with extraposition, even in those instances where it appears difficult to analyze the extraposed clause as being an external argument of some lexical head, as in (4.15)b.

(4.15) a. That he lost is a shame.
b. It is a shame that he lost.

Furthermore, it explains the fact that extraposed constituents (both clausal and non-clausal) in general appear clause finally, rather than in the final argument position.⁸

(4.17) a. It shocked him to the very core that she liked pickled eggs.
b. ??It shocked him that she liked pickled eggs to the very core.

It also predicts the difference between raising verbs and propositional subject verbs, in that the former never and the latter always give semantic roles to their subjects, and therefore the former can never appear without an external argument while the latter can.

⁸There is of course a tendency to position clausal arguments as late as possible, but a difference can still be seen between extraposed clauses and those in second argument position.

(4.16) a. ?*It was accepted that the argument was inconclusive by everybody there.
b. She was finally persuaded that John was no good by a man she met on the bus.

These examples are from Tugwell (1994), which sets out at greater length reasons for not treating these extraposed clauses as arguments.

- (4.18) a. *It appears.
b. It sucks.

Additionally, it sheds new light on a very common, but little-studied, construction, in which the apparent extraposed subject is not a proposition, but apparently a modifying *when*-clause, which cannot appear as subject of the same verb.⁹

- (4.20) a. It surprised her when he swore.
b. *When he swore surprised her.
- (4.21) a. It was obvious if he liked someone.
b. *If he liked someone was obvious.
c. If he liked someone it was obvious.

This construction does not present a problem to the cataphoric approach to extraposition. The pronoun *it* in (4.20)c refers not to the time constituent formed by the phrase *when he swore*, but rather to the presupposed proposition that is the complement of *when* inside that constituent, ie. *that he swore*. Therefore, (4.20)a is equivalent not to (4.20)b, but rather to the following sentence (4.22).

- (4.22) When he swore, that he swore surprised her.

The same meaning can be expressed in (4.23)a by an initial positioning of the *when*-clause and an anaphoric reference to the proposition contained in it. The corresponding sentence with an initial proposition is not possible as it is not of the semantic type to have a role of modifier.

- (4.23) a. When he swore it surprised her.
b. *That he swore it surprised her.

Another piece of evidence in favour of distinguishing extraposition from raising is that extraposed clauses are more likely than argument clauses to be initiated by the complementizer *that*, resulting in the following pattern.

⁹Of course, *when*-clauses can appear in subject position if is compatible with the meaning of the sentence, as in (4.19).

(4.19) When he was playing football seemed to be when he was happiest.

- (4.24) a. ?It shocked him she swore.
 b. It seemed to him she was going to swear.

Finally, the analysis predicts, unlike the raising analysis of these constructions, that this cataphoric proposition construction should also appear in object position, which is indeed the case.¹⁰

- (4.25) a. He likes it when she swears.
 b. ??He likes when she swears.

As was noted in the introduction to the model in Chapter 2, the mechanisms involved in coreferencing and the representation of reference have not been worked out, which presents difficulties for representing the derivation of sentences involving extraposition, as I have argued that it is essentially a cataphoric construction. Nevertheless, a tentative analysis of the extraposed sentence (4.13)b is given here.¹¹

- | |
|--|
| <p>(1) prop⁰, EVENT², past, shock, theme:2, exper:3, [+2†]
 (2) prop¹, sg, ref⁴:4
 (4) prop⁴, STATE⁶, past, like, exper:5, theme:6, [+5†]
 (5) MALE⁵, sg
 (6) onion⁷, plur
 (3) FEMALE³, sg</p> |
|--|

(4.13)b: **It shocked her that he liked onions.**

4.1.2 Raising adjectives and others

The same raising construction found with verbs such as *appear* can also be found with adjectives.

- (4.26) a. He is likely to surprise you.
 b. It is likely that he will surprise you.

Assuming a lexical entry for *likely* that has a similar argument list as that for *appear* above and the analysis of predicate constructions introduced in the section on exclamatives in the previous chapter, the derivation of (4.26) is straightforward.

¹⁰See Postal & Pullum (1988) for a systematic review of the data. These should be kept distinct from genuine dummy semantically-bleached objects which appear in various idioms (for example, *to wag it = play truant*).

¹¹Extraposition will be examined in more detail in chapter 6, which deals with discontinuous constituents.

- (1) prop⁰, STATE², pres, **be**, content:3, [+2*]
 (3) QUALITY³, **likely**, Ø:2, theme:4, [+2*]
 (4) prop⁴, inf, STATE⁵, **surprise**, theme:2, exper:5, [+2†]
 (2) MALE¹, sg
 (5) 2ND-PERSON⁶

(4.26)a: **He is likely to surprise you.**

Such lexical entries must also be given to certain nouns expressing a degree of difficulty.¹²

- (4.28) a. He's a certainty to win the cup.
 b. It's a certainty that he will win the cup.

One difference from raising verbs, however, is that raising adjectives and nouns *can* give a semantic role to their subjects.

- (4.29) That he will surprise you is highly likely.

The conclusion to be drawn is that *likely*, along with other non-verbal raisers, has two distinct lexical entries, the raising one with an empty subject position and a standard propositional subject entry that allows (4.29). One consequence of this is that it allows sentence (4.26)b above to be derived twice, once by the raising entry and once by the combination of a sentential subject entry and extraposition, although the actual semantic content of these derivations will be identical.

There does, however, appear to be an adjective that is an obligatory raiser, which is the word *worth* as demonstrated in the following examples.¹³

- (4.31) a. It is worth reading this book.
 b. This book is worth reading __ .
 c. *Reading this book is worth.

¹²Though this is not entirely predictable from the type and therefore it seems must be individually listed.

- (4.27) a. *He's probable to win the cup.
 b. *He's a probability to win the cup.

¹³Although *worth* is far from being a prototypical adjective, not being gradable for example, and taking nominal arguments.

- (4.30) a. *How worth is it reading?
 b. This book is worth a lot of money.

Whether its status as an obligatory raiser is linked to this is unclear.

As (4.31)b suggests, *worth* has many similarities to *tough*-adjectives, which will be discussed later in this chapter.

To recap, the lexical entries for raising and sentential subject verbs are summed up in the following table.

verb	adjective	raising [\emptyset : NULL, theme: prop]	sent-subj [theme: PROP]
appear	worth	✓	—
—	likely	✓	✓
suck	probable	—	✓

Unfortunately this analysis offers no clues as to why there are no normal gradable adjectives with purely raising characteristics or why there appear to be no verbs with both lexical entries.

4.1.3 Auxiliaries as raising verbs

The analysis of auxiliaries as being in essence raising verbs goes back to Ross (1969) and McCawley (1971). The basic insight is that none of the three auxiliary verbs in (4.32) give a semantic role to the surface subject *he*, but instead subcategorize for a single propositional argument, just like a raising verb such as *appear*. The auxiliaries, however, place characteristic conditions on the tense of these arguments.

(4.32) He could have been swimming.

The derivation of (4.32) is as follows.

- | |
|--|
| <p>(1) prop⁰, STATE², past, can, content:3, [+2*]
 (3) STATE³, perf, inf, content:4, [+2*]
 (4) STATE⁴, be, ppart, content:5, [+2*]
 (5) ACTION⁵, swim, ing, agent: 2, [+2+]
 (2) MALE¹, sg</p> |
|--|

(4.32): **He could have been swimming.**

Although I will not go deeply into how the tense requirements should be enforced, it should be borne in mind that the checking of these requirements appears to be subject to some sort of immediacy requirement.¹⁴ For example, if the complement of the auxiliary is

¹⁴This can be compared with the case of extraposed German relative clauses with mismatched case previously examined in 3.2.4.1.

topicalized then the tense requirements can, on occasion at least, be disregarded as shown in (4.33).

- (4.33) He said he would eat them up and eat them up he certainly has _ .

It appears then that this tense must be checked immediately the past participle complement is formed, or else this tense information is lost. It is once again the kind of behaviour that may be easier to model in a dynamic model of syntax than in a static one.

4.1.4 *Sounds like*

The final construction to be examined in the section on raising is exemplified by the following sentences, whose interpretations appear to be practically identical.¹⁵

- (4.35) a. He sounds like he's rich.
b. It sounds like he's rich.

This construction is definitely a raising construction as there is only one logical argument (a proposition) and no role can be given to the subject.

- (4.36) a. *A: He's very nice. B: It sounds like.
b. A: He's very nice. B: It sounds like it.

It is natural to assume, therefore, that there is a lexical entry *sound-like* (along with the parallel *seem-like* and *look-like*) which has an empty first argument and a propositional second argument which must be introduced by the lexical item *like*.¹⁶

sem: ⟨ STATE, **sound-like**, [∅ : NULL, theme: prop [lex: like]] ⟩

This entry gets us the raising derivation of (4.35)b. It also follows from this why the sentence (4.37)a, corresponding to the normal raising construction (4.37)b, is impossible for *sound like*.

- (4.37) a. *He sounds like to be rich.
b. He seems/appears to be rich.

¹⁵ Actually (4.35)a is ambiguous, but I shall ignore in this discussion the interpretation which is connected with the noise made by the subject, an interpretation that comes to the fore in a sentence such as (4.34).
(4.34) He sounds like he's got a sore throat.

¹⁶ Such lexically-specified arguments were introduced in the section on complementation in chapter 2.

The word *like* is a manner subordinator, although here semantically bleached, and like other subordinators initiates its own complement clause (ruling out **because that*, **before that* and so forth). With a complement clause already initiated the finite complementizer *to* is unable to perform its function.

The question that remains is how is (4.35)a possible, when it appears that the surface subject *he* is given no semantic role at all? I shall adopt the line that the subject pronoun in the embedded clause in this construction is a type of resumptive pronoun, similar to the one used to fix an illegal extraction in the rather non-standard construction illustrated in (4.38)a.

- (4.38) a. ?Which girl were you waiting to hear whether she had passed?
 b. *Which girl were you waiting to hear whether _ had passed?

This analysis is supported by the fact that there must be a coreferential pronoun in the complement clause and that it must be in subject position.¹⁷

- (4.41) a. He sounds like he's hated by everyone.
 b. ?*He sounds like everyone hates him.
 c. It sounds like everyone hates him.

The second *he* in the following derivation of (4.35)a will act as a resumptive pronoun fixing the uninterpreted subject of the matrix clause as the subject of the embedded clause.¹⁸ The derivation of (4.35)a can then be made as follows to give a final interpretation identical to that of (4.35)b.

¹⁷Although Heycock (1995: 288) presents a list of counterexamples to this claim all of them sound quite jarring to me, although of course they are quite interpretable.

- (4.39) a. That book sounds like everyone will want to buy it.
 b. That book sounds like its publication should cause quite a stir.
 c. That book sounds like everyone thinks it should be banned.

Perhaps here the initial constituent is more of a general sentence topic, as in (4.40).

- (4.40) That woman, I don't know what she think she's doing!

¹⁸If we adopt the tentative proposal made in chapter 9, that subjects are held in store as well, then the parallel with resumptive pronouns becomes even closer.

- (1) prop⁰, STATE², pres, **sound-like**, \emptyset :2, theme:3, [+2*]
 (3) **like**³, content:4
 (4) prop³, STATE⁵, pres, **be**, content:5, [+2*]
 (5) QUALITY⁶, **rich**, theme:2, [+2†]
 (2) MALE¹, sg

(4.35)a: **He sounds like he 's rich.**

Finally, how do we then rule out the same process of resumptive subject pronouns with a standard raising verb as in (4.43). The only explanation is to suppose that the fact there is an alternative construction available in these cases, means that it must be taken, ie. a familiar winner-takes-all situation.¹⁹

(4.43) *He appears he's rich.

This parallels the fact that resumptive pronouns can only be used when there is no grammatical alternative.

(4.44) *Which girl_i do you think she_i will pass?

4.2 Equi

Another term derived from the early transformationalists, *equi* constructions are similar to raising, but as we shall see involve the assignation of an extra thematic role.

4.2.1 Subject control

The problem presented by subject control constructions such as (4.45) is similar to that of the raising construction in that the surface subject of the main clause must be interpreted as the subject of the embedded clause. It differs however in that the matrix verb *expect* does assign a semantic role to its surface subject, ie. there is someone “doing the expecting”, whereas there was no-one “doing the seeming”.

(4.45) He expects to win.

To address the problem let us suppose the following lexical entry for *expect*.

¹⁹Although it must be said that on occasion these do not seem totally bad to me, especially with *seem*.
 (4.42) ??He seems that he's in trouble.

(4.46) sem: ⟨ STATE, **expect**, [exper: ANIM, theme: prop] ⟩

This does not seem sufficient in that there is nothing stating that the experiencer of *expect* must also be the subject of its theme. However, if we recall the transition given above for the infinitival complementizer *to* it turns out that this is sufficient, in combination with the lexical entry above, to ensure the correct derivation.

Consider the state sequence in the derivation of (4.45). The verb *expects* will mark its subject with the first of its argument roles, in accordance with its lexical entry. At the next transition the infinitive complementizer will do exactly what it did in the previous raising examples: create a new proposition, make the value of the subject the same as that of the main clause and attach the clause as the final argument. This leaves the following interpretation after *to*.²⁰

<p>(1) prop⁰, STATE², pres, expect, exper:2, theme:3, [+2†] (2) MALE¹, sg (3) prop³, inf, [+2]</p>

(4.45): S_3 **He expects to** || win.

The completion of the derivation is completely standard once more: the verb *win* interprets the subject as its first argument.

<p>(1) prop⁰, STATE², pres, expect, exper:2, theme:3, [+2†] (2) MALE¹, sg (3) prop³, inf, EVENT⁴, win, exper:2, [+2†]</p>
--

(4.45): **He expects to win.**

Adjectives also exhibit subject control and so must be provided with lexical entries with the same argument pattern of subject control verbs given above.

(4.47) He is eager to leave.

It will be noted that the lexical entries which allow equi constructions will also allow constructions where the second argument is a finite proposition.

(4.48) He expects that Lee will be there.

²⁰Note that by convention in the diagram I shall order constituents which are semantically dependent upon more than than other constituent under the first of these.

If the same lexical entries are used in the subject control and finite argument constructions then the state of the interpretation after “he expects” in (4.45) and (4.48) will be the same. Our coordination diagnostic would then predict that it should be possible to coordinate these arguments of different types and the data, adapted from Borsley (1996), bears this out.

- (4.49) a. Kim expects to be there and that Lee will also be there.
 b. Kim expects that Lee will be there and also to be there himself.

This coordination data presents a challenge for theories which assume two separate entries for *expect* here.²¹

4.2.2 Object control

In object control constructions the first external argument of a verb is interpreted as the subject of its second clausal complement as in (4.50).

- (4.50) He persuaded her to leave quietly.

Again adopting a minimally specific approach to the lexical entry, I will suppose that it is sufficient to state that *persuade* takes three arguments and that the control relation between the two objects will be decided by separate principles.

sem: ⟨ ACTION, **persuade**, [agent: ANIM, patient: ANIM, theme: prop] ⟩

However, returning to the rule given above for the infinitival complementizer *to*, it was stated that the first argument (ie. subject) of the verb is to be the subject of the new complement clause. This would result in the subject of *leave* in (4.50) being *he*, which is obviously not what is required.²² The rule for *to* must therefore be generalized to picking out the most recent argument of the verb, rather than the first one.

This constraint on the interpretation of the subject of the infinitive is in effect a statement of the Minimal Distance Principle of Rosenbaum (1967), which states that the subject of the infinitive clause is the nearest NP to its left. It will of course still allow subject control in single object verbs as in the previous section, and with this modification in place the derivation of the object control (4.50) proceeds in an equivalent manner.

²¹Even if they were able to handle the coordination of unlike categories: see the next chapter for details.

²²The status of double object verbs which still exhibit subject control, eg. *promise* will be discussed below.

- (1) prop⁰, ACTION², past, **persuade**, agent:2, patient:3, theme:4, [+2⁺]
 (2) MALE¹, sg
 (3) FEMALE³, sg
 (4) prop⁴, inf, EVENT⁵, **leave**, theme:3, mod⁶:5, [+3⁺]
 (5) MANNER⁶, **quiet**, theme:4

(4.50): **He persuaded her to leave quietly.**

Once again the lexical entry allows for the propositional argument to be realised as a standard finite clause, as in (4.51).

(4.51) We persuaded her that he was no good.

This analysis can again be tested by attempting to coordinate arguments of differing types and the data appear to support it.

- (4.52) a. ?We persuaded her that he was no good and to leave him straightaway.
 b. We persuaded her to leave him and that California was a good place to go.

4.2.2.1 Heavy element shift

A final point to note about object control is that, as with all double object constructions, the first object if sufficiently long and tortuous may be placed second. This is illustrated in (4.53), which is slightly awkward, but acceptable.

(4.53) He finally persuaded _ to leave all of the people who had been involved in the prolonged disturbances.

The problem this poses is how the heavy-shifted patient, *all of the people...*, gets to be interpreted as the subject of the embedded clause, since when this clause is initiated it has not even appeared in the sentence.

However, this problem disappears with the analysis of heavy-element shift that I adopt, and which will have independent motivation in a number of places in this thesis. It assumes that the address of the heavy-shifted argument is temporarily held by a placeholder (seen also in the analysis of fronted adjectival phrases in the previous chapter) which is later replaced with the address of the actual argument. This allows the following (partial) derivation of (4.53).

- (1) prop⁰, mod²:**3**, ACTION³, past, **persuade**, agent:**2**, patient: $\alpha^4/5$, theme:**4**, [+**2**†]
 (2) MALE¹, sg
 (5) ref⁶: \forall
 (4) prop⁴, inf, EVENT⁵, **leave**, theme: $\alpha/5$, [+ α †]
 (3) TIME², **finally**

(4.53): **He finally persuaded to leave all** || of the people who ...

The placeholder α which ends up being interpreted as the theme of *leave* is substituted with the address of the constituent *all of the people...* as soon as this heavy-shifted constituent is added.

4.2.2.2 Object raising

At first glance sentences such as (4.54) appear to be instances of the object control constructions discussed above. However, although they may be outwardly similar, a little reflection makes it clear that the verb *expect* has only two logical arguments: the expecter and the thing expected. In this section, I shall show how adopting the above analysis of object control, and combining it with the an insight from the earlier analysis of raising sentences, we can arrive at a satisfactory analysis of such sentences.

(4.54) He expects her to win.

The fact that *expect* takes only one external semantic argument, a proposition, can be demonstrated by the equivalence of (4.55)a and (4.55)b.

- (4.55) a. He expects the doctor to examine the boy.
 b. He expects the boy to be examined by the doctor.

Despite this, however, there is good evidence²³ that this is a double object construction. For example, a modifier of the verb may be placed between the two objects in (4.56)a, something not possible with a genuine clausal complement (4.56)b.

- (4.56) a. He expected her yesterday to change her mind.
 b. *He expected she yesterday would change her mind.

Similarly the subject can be shifted to the right if it is sufficiently heavy, another characteristic of double object constructions.

²³Following the arguments of Pollard & Sag (1994).

- (4.57) He expected to be there all the people he had ever met or would meet or wished to meet.

We can account for the double object behaviour, as well as the fact that the first object bears no semantic relation to *expect*, by adopting the empty-argument solution used in raising. We then state in the lexical entry that one of the arguments is semantically vacuous (in this case the second rather than the first).

sem: \langle STATE, **expect**, [exper: ANIM, \emptyset : NULL, theme: prop] \rangle

It will be seen that the advantage of this approach is that it allows the infinitive *to* to have exactly the same function in these constructions as in all others we have examined so far. It creates an infinitival proposition which takes the address of the neighbouring argument as its subject. The derivation, therefore, is entirely parallel to the object-control construction above, with the caveat that the first external argument is not interpreted by the verb, but only by the complement clause.

<p>(1) prop⁰, STATE², pres, expect, exper:2, \emptyset:3, theme:4, [+2†]</p> <p style="padding-left: 2em;">(2) MALE¹, sg</p> <p style="padding-left: 2em;">(4) prop⁴, inf, EVENT⁵, win, theme:4, [+3†]</p> <p style="padding-left: 4em;">(3) FEMALE³, sg</p>

(4.54): **He expects her to win.**

Although this approach does a good job at modelling the data using the existing machinery, it appears to be not entirely unproblematic. I have assumed that there are two lexical entries for *expect*, one with a single propositional object which allows *he expects that...* and *he expects to...*, as well as the double object entry given above which allows *he expects her to...*. The need for two separate entries seems to be placed in doubt by the fact that there are no verbs in just one of these classes. Also the coordination data, although variable, does seem to suggest just one entry would be preferable.

- (4.58) a. ?He expects Peter to win, or perhaps to win himself.
 b. ?He expects Peter to go away and that Mary will follow him.
 c. ?*He expects that Mary will leave and Peter to follow her.
 d. ?*He expects to go away and Mary to follow him.

Clearly, further work needs to be done to capture this generalization, if it is indeed desirable to do so.

4.2.2.3 *Promise*

As set out above, the rules predict that all constructions involving two objects, the second of which is a proposition, should show object control. However, there are verbs with two objects that appear to buck the trend and insist on subject control. For example, the verb *promise* in (4.59), where the subject of the complement clause is identical with that of the matrix clause.

(4.59) She promised me to win.

Given the approach taken here, there are no mechanisms that could be used to force subject agreement, and the verb *promise* must have a lexical entry entirely the same as that of the object control *persuade* with the slight difference that the second argument is a beneficiary rather than a patient.

sem: ⟨ ACTION, **promise**, [agent: ANIM, benef: ANIM, theme: prop] ⟩

It seems that the beneficiary in (4.59) is not acceptable as a subject due to the semantics of the verb, only the one promising can be the one doing the thing promised, which overrides (though not for all speakers)²⁴ the structural constraints.

Similarly, the constraints of possible meaning can override other verbs and turn them from object control into subject control. For example, substituting a passive complement of *beg* in (4.60) means the verb must change from object control to subject control.

- (4.60) a. The boy begged the teacher_i to leave_i the room.
 b. The boy_i begged the teacher to be allowed to leave_i the room.

And in a similar fashion, giving *promise* a passive complement can enforce a switch to object control.

(4.61) ?She promised him to be allowed to leave on time.

The conclusion from all of this is that the schema for the complementizer *to* needs to be further generalized. Instead of taking the nearest argument as subject of the complement

²⁴It should be mentioned that it has been largely overlooked that the construction in (4.59) often sounds distinctly awkward and that a large number of speakers do not accept it at all. A survey of unrepresentative and self-selecting subscribers to the newsgroup linguist-list (9.651, 5 May 1998) found that native speakers ruled against sentences like (4.59) 52 to 34. Many noted that it was better in the negative: *She promised me never to do it again*. It seems that for speakers who reject this construction the syntactic constraint on control is absolute and does not allow itself to be overridden by semantic factors.

proposition, it must take the nearest argument *that can meaningfully be the subject* of that clause. It is interesting to speculate at what point in a sentence like (4.61) the hearer decides that object control is preferable to subject control. Similar effects, termed *coercion* in Pollard & Sag (1994), will be seen in the following section on the passive construction. Together they offer strong support for the view, defended throughout this thesis, that the construction of meaning purely through syntactic operations that cannot access the meaning itself is untenable. Pollard & Sag (1994) admit as much by dividing the verbs into three classes along semantic grounds, in order to predict the control relations. But putting such semantic criteria in the syntax does not seem as natural a solution as abandoning the idea that syntax can be modelled without respect to the meaning being constructed.

4.2.3 Small clauses

In this final section on control relations I shall briefly consider the relationship between the two sentences in (4.62). The first, (4.62)a, is an equi construction of the type we have just examined, where *this dog* fills the null first argument position of *consider*. In the synonymous (4.62)b, however, there is no explicit infinitival copular and we have what is generally referred to as a *small clause* as object.

- (4.62) a. He considers this dog to be exceptional.
 b. He considers this dog exceptional.

I shall take the minimal assumption that both the sentences above are allowed by the single lexical entry already given, and the task once again is to arrive at the same basic interpretation for both sentences.

I shall assume that the propositional argument of *consider* in the small clause sentence (4.62)b is created by a free transition, that is one not associated with a particular word in the string, but rather with a position in the interpretation, as is the rule for adding finite propositional arguments. The rule will apply to fill an empty propositional argument slot with a copular proposition, whose subject is the next most oblique argument of the active constituent, and thus will be formalized as follows.

transition rule	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
reduced predicate		(X, arg:Y)	(N, prop, be, content: , [+Y]) (X, arg:Y, arg:N)

It is clear that this rule must be restricted in its application, for even verbs very similar in meaning do not allow it. Obviously some condition must be placed on the head of (X) in the above schema to rule in *consider* and rule out *know*. If the verbs that allow small

clause complements do not share any semantic characteristics then these must simply be listed in the lexicon.

- (4.63) a. He knows this dog to be exceptional.
 b. *He knows this dog exceptional.

The reduced predicate complement can be the sole external argument if the verb is passivized, as shown in (4.64)a, and the first argument can undergo heavy-movement, as in (4.64)a. Both of these facts support the view that the small clause consists of two separate arguments on the argument list of *consider*, as is the case for *consider X to be Y*.

- (4.64) a. This dog is considered exceptional.
 b. He considers exceptional any dog that can sort bones with its nose.

There are various semantic restrictions on the relation between the two arguments of the small clause, for instance it appears that it must be an equative construction, rather than an existential one.

- (4.65) a. He considers this to be a serious problem.
 b. He considers this a serious problem.
 c. He considers there to be a serious problem.
 d. *He considers there a serious problem.

Considerations of this sort, however, await a comprehensive treatment of copular constructions and are beyond the scope of this thesis.

Here is the derivation of (4.62)b. The reduced predicate transition rule applies at the fifth transition, creating the propositional complement, and the adjective is immediately added as its content, inheriting its subject value.

- (1) prop⁰, STATE², pres, **consider**, exper: 2, \emptyset :3, theme:4, [+2†]
 (2) MALE¹, sg
 (4) prop⁵, STATE, be, content:5, [+3*]
 (5) QUALITY⁵, **exceptional**, theme:3, [+3†]
 (3) def³, proximate, sg, **dog**⁴

(4.62)b: **He considers this dog exceptional.**

Again the coordination data are broadly in support of the analysis. The reduced predicate argument can coordinate with an infinitival clause, as in (4.66)a. Also the two arguments making up a small clause can coordinate with two arguments where the second is an infinitival clause, as in (4.66)b. Both of these facts support the view that *consider* is using one and the same lexical entry in both constructions, and expressing identical subcategorization requirements.

- (4.66) a. He considers this dog exceptional, but still to be in need of further training.
 b. He considers the cat a potential champion, but the dog to be in need of further training.

4.2.3.1 Predicate complements

The reduced predicate transition introduced above to model small clause arguments will also apply to raising verbs to allow them to have predicate complements with identical meanings to infinitival copular ones. This pattern is illustrated in (4.67).

- (4.67) a. He seems restless.
 b. He seems to be restless.

The derivation is similar to that of the small clause given above, with the rule introducing the reduced predicate being applied at transition 3.

- | |
|--|
| <p>(1) prop⁰, STATE², pres, seem, \emptyset:2, theme:3, [+2*]
 (3) prop³, STATE, be, content:4, [+2*]
 (4) QUALITY³, restless, theme:2, [+2†]
 (2) MALE¹, sg</p> |
|--|

(4.67): **He seems restless.**

Again this must be restricted, semantically or lexically, so that it cannot apply to all raising verbs.

- (4.68) a. He happens to be restless.
 b. *He happens restless.

Coordination is possible between a bare predicate complement and an infinitival complement, although it is asymmetric.²⁵ Given the assumptions of this thesis, even coor-

²⁵It is commonly found that infinitival complements are better in the second conjunct than the first.

dination one way is enough to conclude that the subcategorization requirements of *seem* must be the same in both constructions.

- (4.69) a. He seemed restless and to be trying to leave.
 b. ?*He seemed to be trying to leave and strangely restless.

4.3 Passive

The control-shifting constructions considered so far in this chapter have been lexically-dependent in nature, relating to the lexical entries of a small subset of words. The passive construction in English is much wider in its scope and in this section I attempt to account for it in terms of a productive syntactic transition rule.

The general pattern is familiar to all: the surface subject in a passive construction receives the role given to an object by the same verb in an active one. An argument with the role given to an active subject is only optionally present in the passive, in which case it must be marked by the preposition *by*.²⁶

- (4.70) a. The dog bit him.
 b. He was bitten (by the dog).

The transition rule *add-passive-verb* must then perform the transition schematized in the following diagram.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{verb, past-participle}	(X, prop, be, content: , [+Y])	(N, passive, arg1: α , arg2:Y, [+Y†]) (X, prop, be, content:N, [+Y*])

It should be noted that the first argument of the verb is not left unfilled, but rather filled by the placeholder α . The constituent is therefore saturated and it follows that the derivation can end here. The placeholder α should then be read as some unknown entity not stipulated in the interpretation, but upon which semantic constraints will be placed by the verb.

- | |
|---|
| <p>(1) prop⁰, STATE², past, be, content:3, [+2*]
 (3) ACTION³, ppart, bite, passive, agent:α, patient:2, [+2†]
 (2) MALE¹, sg</p> |
|---|

(4.70)b: **He was bitten...**

²⁶Except in the case of finite clauses, see below for discussion.

Considering the action of the optional *by*-phrase, we must take into account the fact that *by* here is semantically vacuous and that the phrase appears in the position of a postmodifier, that is it generally adds information to saturated constituents. This can be captured in the following transition schema.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ by }	(X, passive, larg: α)	(N, by , content:) (X, passive, l-arg: α/N)

The preposition *by* then can in effect reopen the first argument role and its semantic vacuity allows its complement to be interpreted with this role, as is illustrated in the complete derivation of (4.70)b. The interpretation is semantically equivalent to that of the active sentence (4.70)a.

- (1) prop⁰, STATE², past, **be**, content:3, [+2*]
 (3) ACTION³, ppart, **bite**, passive, agent: $\alpha/4$, patient:2, [+2†]
 (2) MALE¹, sg
 (4) **by**⁴, content:5
 (5) def⁵, **dog**⁶, sg

(4.70)b: **He was bitten by the dog.**

A number of consequences follow from this account of the passive construction. Firstly it was stipulated that the surface subject be interpreted as the second argument of the verb, which agrees with the data.

- (4.71) a. The man gave the dog a bone.
 b. The dog was given a bone.
 c. *The bone was given a dog.

Secondly, we have the strange behaviour of the apparent passivization of the subject from finite clausal complements as witnessed by the data in (4.72).

- (4.72) a. He said she was clever.
 b. *She was said was clever.
 c. *He said her to be clever.
 d. She was said to be clever.

The first question this data raises is why it is impossible to form a passive like (4.72)b given the active form (4.72)a. This is not hard to answer, for the add-passive-verb rule given above cannot start a new finite proposition, which is what is needed for (4.72)b to be derived.

The second question is more tricky however, that is how is it possible to have the perfectly fine passive construction in (4.72)d when the corresponding active construction (4.72)c is not available? The explanation seems to be that in the derivation of (4.72)d, when the passive verb *said* is reached, the surface subject *she* is not interpreted as the theme of *say* (which would result in a type-mismatch and the failure of the derivation), but is instead left without a semantic role. At the next transition the complementizer *to* picks up the subject (it being the most accessible constituent) and installs it as the subject of a new complement clause: an entirely standard transition for *to* and one which results in the subcategorization requirements of *say* being filled.

It is debatable whether the general passive rule should be adapted to make the filling of the second argument role by the subject optional rather than obligatory, or whether a separate rule should be posited, slightly modified from the original, to deal with these cases and applying only if the verb in question takes a single clausal object. Either way it goes a long step towards explaining this puzzling data.²⁷

The final piece of data to be accommodated is Visser's generalization, that subject control verbs do not passivize, as seen in (4.75)b.

- (4.75) a. He was persuaded to go to the shops.
b. *He was promised to go to the shops.

If we imagine the course of the derivation of (4.75)b, the passive verb will interpret the subject as its second argument, the beneficiary of the action, and cancel the first argument. The following complementizer *to* will seek to place the most accessible argument, this being *he* as the only one available, as the subject of the new clause. However, this contradicts

²⁷Although no analysis of such examples will be attempted here, the following data are also relevant.

- (4.73) a. She made him (*to) clean his shoes.
b. He was made *(to) clean his shoes.

In van Ek (1966) it is noted that *let* is the only verb that retains its normal bare complement in the passive *he was let go*, but it seems barely productive. In fact *let* appears to have no usable passive form and *allowed* must be used instead.

- (4.74) a. They let him win.
b. ?*He was let win.
c. *He was let to win.
d. He was allowed to win.

This data also raises the question of how the absence of a passive form can be encoded in the lexicon.

the semantics of the construction which disallows the beneficiary from being the one doing the action.

This analysis is backed up by the felicity of (4.76), in which the meaning of the construction does allow the beneficiary to be interpreted as the subject of the complement clause.

(4.76) He was promised to be allowed to go to the shops.

4.3.1 *ing*-passives

A little consideration of the construction in (4.78) reveals that it parallels the standard passive construction above: the first argument of *feed* is left unfilled and the sentence subject, which is not given a role by the raising verbs *need* or *want*, is interpreted as the object of *feed*.²⁸

(4.78) This dog needs/wants feeding.

The lexical entry for *need* and *want* as used in this construction has simply to state that they are raising verbs and can take a propositional complement to allow (4.79).

(4.79) This dog needs to be fed.

An additional transition is needed for the *ing*-verb to allow the passive construction.²⁹

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{verb, ing }	(X, need , arg1:Y, theme:)	(N, passive, ing, arg1: α , arg2:Y) (X, need , arg1:Y, theme:N)

Perusal of the transition introducing a passive *by*-phrase, given above, will show that it may also apply here, which is borne out by (4.81).

(4.81) This door needed painting by a professional.

²⁸It should be noted that this is not a variety of *tough*-movement, similar to the ungradable adjective *worth* (discussed in the next section), which involves use of the store and hence allows unbounded movement. This is shown by the following contrast.

- (4.77) a. *This dog needs telling someone to feed __ .
b. This dog isn't worth going to the trouble of telling anyone to feed __ .

²⁹Note that Scottish English uses here the standard "passive" (ie. past-participle) form with these verbs, which appears more systematic.

(4.80) %This dog needs/wants fed.

Not surprisingly, given that it agrees with the common passive, it is a highly contagious construction, and appears to be quickly acquired by English immigrants to Scotland.

4.3.2 Prepositional passive

Mention should also be made of the prepositional passive construction, also referred to as the pseudopassive. It is illustrated by (4.82), in which the subject of the passive somehow manages to receive an interpretation not in an argument position of the verb, but as an argument of an argument of the verb. It seems unlikely that the account of passive given above will also allow such a construction, but with the addition of a few plausible assumptions it will be seen that it may.

(4.82) This bed has been slept in __.

Examining the course of derivation of (4.82), it is clear that when we arrive at the passive verb *slept* the subject can in no way be interpreted as the object of the intransitive verb *sleep*, as it has none.³⁰ We will assume that it is therefore left momentarily uninterpreted, as in the “said to be clever” construction above, and is then taken at the next transition as the complement of the locative *in*.

The assumption being made is that constituents whose only connection to the interpretation is as surface subject of the verb can be recovered, having failed to be interpreted as an argument by the verb.

As has often been noted, there is also a semantic restriction that the passive subject in these constructions must be affected in some way or else play a central semantic role in the clause.

The claim has often been made that pseudopassives are explained by the verb and preposition forming effectively a compound verb assigning a single thematic role. But this is hard to square with the available evidence, for example that the prepositional phrase can be fronted, as in (4.83).

(4.83) In which bed did she sleep __?

There is no doubt in (4.83) that *which bed* is independently assigned a thematic role by *in*. However, I will not examine here the factors that are involved in successful prepositional passive constructions.

4.4 *Tough-movement*

This construction, whose intricacies have made it a favourite subject of syntactic analyses, involves unbounded movement and may seem out of place in this chapter. However, as its

³⁰The fact that despite this the derivation is possible is good reason in itself for dealing with passive as a productive syntactic rule, rather than linked to information based in the lexicon.

analysis also involves consideration of shifting control relations, it has been necessary to leave it to here. The basic pattern can once again be illustrated by a group of sentences which are syntactically quite different and yet end up having the same meaning.

- (4.84) a. Dogs are easy to understand __ .
 b. To understand dogs is easy.
 c. It is easy to understand dogs.

Sentences (4.84)b and (4.84)c will be generated by the following lexical entry for *easy* together with the previous rule for clausal extraposition (or cataphora) and nothing further need be said about them.³¹

sem: ⟨ STATE, **easy**, [theme: EVENTUALITY] ⟩

It is clear that the construction in (4.84)a, however, will not be allowed by this entry and requires something quite different to anything we have seen so far. It is restricted to a small class of adjectives strongly restricted to the semantic field of “degree of difficulty”.³² As the subject has no direct semantic relation to the adjective we might suppose that the lexical entry allowing this construction should be a raising one with a first null argument as we have seen before.

sem: ⟨ STATE, **easy**, [∅ : NULL, theme: EVENTUALITY] ⟩

Before we employ this transition schema in the derivation of (4.84)a, it may be first helpful to show the derivation of the standard example (4.84)b as this introduces a new transition for the complementizer *to*. This involves starting a constituent of type EVENTUALITY and giving it the arbitrary placeholder α as subject to indicate that the actual subject is left unstated. It is the first transition in the following derivation.

- | |
|--|
| <p>(1) prop⁰, STATE⁴, pres, be, content:4, [+2*]
 (4) QUALITY⁵, easy, theme: 2, [+2†]
 (2) inf, STATE², understand, exper:α, theme:3, [+α†]
 (3) dog³, plur</p> |
|--|

(4.84)b: **To understand dogs is easy.**

³¹Note that the semantic type EVENTUALITY is intended as a supertype which includes STATE, EVENT and ACTION.

³²And also seemingly any noun within that same field.

(4.85) This dog is a cinch/doddle/bugger/bitch to understand.

Turning now to the derivation of (4.84)a, the crucial transition is that where the infinitival complement clause is created. In order for the subject of the tough-adjective to be interpreted indefinitely far away, it is necessary for it to be inherited as the non-subject store of the infinitive as in the following transition.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ to , inf-complementizer }	(X, tough-type, [+Y])	(N, eventuality, inf, [+ α , -Y]) (X, tough-type, arg:N, [+Y*])

The stored constituent $\widehat{\text{dogs}}$ will eventually be interpreted as the object of *understand*, giving a final interpretation which is identical to that of (4.84)b above.

(1) prop ⁰ , STATE ² , pres, be , content:3, [+2*]
(3) QUALITY ³ , easy , \emptyset :2, theme:4, [+2*]
(4) STATE ⁵ , understand , exper: α , theme:2, [+ α †, -2†]
(2) dog ¹ , plur ²

(4.84)a: **Dogs are easy to understand.**

4.4.0.1 Extraction from embedded finite clauses

It has been noted on many occasions that in the *tough*-construction extraction from the subject position of a *finite* embedded clause is disallowed, as is shown by the following contrast.

- (4.86) a. Shearer is difficult to imagine $_$ wanting to move to Hibs.
b. *Shearer is difficult to imagine $_$ wanted to leave.

As far as I am aware, this has never been explained satisfactorily.³³ If the subject of *difficult* really is on the store of the embedded clause then it should be possible to perform a switching transition and install it as the subject as an embedded finite clause.

There is also something not quite kosher about extraction from object position of finite embedded clauses in the *tough*-construction, as (4.87)b testifies.

- (4.87) a. Shearer is difficult to imagine the manager wanting to sell $_$.
b. ??Shearer is difficult to imagine the manager would sell $_$.

³³The GPSG approach to the problem (Gazdar *et al.*, 1985: 150), is to stipulate that the slashed category is -NOM, thereby blocking (4.86)b, but without any insight being gained as to why this should be the case. This approach is borrowed by HPSG (Pollard & Sag, 1994: 168).

However, this judgement does not appear to be generally accepted: both Gazdar *et al.* (1985: 151) and Pollard & Sag (1994: 168) have unstarred sentences parallel to (4.87)b, and Steedman (1996) also assumes that extraction from the object position of finite clauses is unproblematic. However, at least to my mind, the following sentences do appear to show a robust contrast, (4.88)a being *wh*-extraction from the complement clause of the standard non-*tough* construction and (4.88)b extraction from the *tough*-construction.

- (4.88) a. Who is it difficult to imagine the manager wants to sell _ .
 b. ??Who is difficult to imagine the manager wants to sell _ .

Grover (1995), meanwhile, judges that extraction from both subject and object position of a finite clause in these constructions is irredeemably ungrammatical and argues that this is evidence *tough*-movement is not an unbounded dependency construction at all. Grover argues that *tough*-movement should instead be analyzed as essentially a local dependency, such as equi and raising, and that examples of unbounded movement arise from sequences of these local movements.

Although the insight that there is something strange about extraction of the objects of finite clauses does appear to be valid and has long been overlooked, complicating the grammar by introducing new mechanisms for movement should be avoided as far as possible. An approach such as Grover's also does not explain why many people appear to have no particular problem with extraction from object position of a finite clause, and why for those who do find it odd, it is nevertheless much better than extraction from subject position.

One apparent solution to this problem hinges on the slippery and ill-understood nature of syntactic case. The constituents on the store in *tough*-complements are distinguished from any others that appear in store in that they started out as subjects of another clause. Whether they are explicitly marked as nominative, or whether they still are linked to that original subject position, it is not too far-fetched to imagine that either of these will cause problems for the switching transition, which should convert them from being non-subjects into subjects. This might have some relation to the fact that subject extraction is impossible. Similarly, if the constituent is to be inherited into a finite clause, this clause will end up with having two nominative constituents, which some but not all people appear to find problematic. It would be simple to model either of these restrictions in the present model and yet still retain *tough*-movement as essentially an unbounded movement construction.

Evidence for this explanation of the data comes from the otherwise surprising fact that finite clausal subjects of *tough* constructions appear to allow to a much greater degree extraction from subject position, as (4.89)a shows.

- (4.89) a. ?That Jones was an informer was difficult for us to believe _ might really be true.
 b. *His jewels were difficult for us to believe _ might really be false.

This may be connected with the fact that finite clauses are not able to bear case in the same way that nominal constituents are. Clearly further research needs to be done on this construction to establish a consensus on what the data is that has to be accounted for.

4.4.1 Attributive usage

The construction with *easy*-type adjectives in predicate position discussed above has been well-studied, but the same construction with the adjective in prenominal attributive position, (4.90), less so.

- (4.90) He is an easy man to live with _ .

The status of the clause “*to live with*” is not obvious at first glance. That it is not a reduced relative modifying the noun (as in “a wine to enjoy”), but rather the complement of the adjective *easy* is shown by the fact that (4.90) has exactly the same interpretation as (4.91)a, and also by the fact that the sentence is impossible without this clause. This would not be predictable were it an optional modifier of *man*.

- (4.91) a. He is a man who is easy to live with _ .
 b. *He is an easy man.

Also the infinitival relative clause typically carries the additional meaning of *should/ought to*. So (4.92)a and b are practically synonymous, while (4.90) does not mean the same as (4.92)c.

- (4.92) a. He is the man to live with _ .
 b. He is the man who you should live with _ .
 c. He is an easy man who you should live with _ .

The fact that the complement of a prenominal modifier appears separated from it, creating a discontinuous phrase, is quite normal, indeed it is generally impossible to place these before the noun head, as demonstrated in (4.93).

- (4.93) He is the best man at solving puzzles.

There also seems to be no doubt that it is a genuine case of unbounded movement as the extraction site may be indefinitely far off, for although awkward, (4.95) appears to be grammatical.³⁴

(4.95) ?He is the easiest man to imagine Mary wanting her mother to marry __ .

Let us suppose that the transition introducing attributive *easy*-adjectives is exactly the same as that shown above for their use in predicate position. The derivation of (4.90) will then be as follows.³⁵

- (1) prop⁰, STATE², pres, **be**, content:3, [+2*]
 (2) MALE¹, sg
 (3) indef³, sg, ref:2, mod:4, MALE⁵, **man**, [+2†]
 (4) QUALITY⁴, **easy**, \emptyset :3, theme:5, [+3*]
 (5) prop⁶, inf, STATE⁷, **live**, exper: α , mod⁸:6, [+ α †, -3*]
 (6) COMITATIVE⁸, **with**, theme:3, [-3†]

(4.90): **He is an easy man to live with.**

One thing to notice in the analysis of the attributive *easy*-construction is that, in contrast to the predicate construction, the constituent in store is not simultaneously the subject of a finite clause. Therefore these constructions should be a test of the previous conjecture that this conflict may be a factor in the general unacceptability of the extraction of *easy*-subjects from finite clauses. The prediction is that the attributive construction should not show such an effect. Considering the data in (4.96), it certainly seems to be the case that both object extractions (4.96)b and even more noticeably subject extractions (4.96)a (completely impossible with the predicate construction) are much improved.³⁶

- (4.96) a. ?He is the easiest to imagine __ might want to marry Mary.
 b. ?He is the easiest to imagine that Mary might marry __ .

As pointed out by Jespersen (1940: 273-275), the complement clause can contain an initial topicalization and a pied-piped relative pronoun as seen in (4.97)a. At first sight

³⁴And it improves markedly if the noun head is left out of the phrase as in (4.94). Seemingly this removes the interference caused by an initial analysis of the complement clause as a possible reduced relative modifying the noun.

(4.94) He is the easiest to imagine Mary wanting her mother to marry __ .

³⁵I have tentatively assumed that nominal predicate constructions such as this involve a sharing of reference, rather than the assignment of theta roles. This will be examined in more detail in chapter 7. The question of how the discontinuous phrase is handled is treated in chapter 6.

³⁶My intuitions here are not at all sharp, but they are independently confirmed by Calcagno (1998).

such a construction appears to cast into doubt the assumption made above that the final clause here is not an infinitival relative, but rather a normal clausal complement of *easy*. For as (4.97)b shows, such pied-piping in the complement in a predicate construction is impossible.

- (4.97) a. He is an easy man with whom to live __ .
 b. *He is easy with whom to live __ .

However, returning to the treatment of pied-piping in relatives in 3.2.2, the relative pronoun was taken to be a species of resumptive pronoun whose appearance depended solely on the configuration of the extraction site, ie. that the address was at that point uninterpreted. This configuration is precisely paralleled here, so even though the clause is not a relative clause, a relative pronoun is predicted to appear.

Conversely, if relative pied-piping were captured by some feature *rel* introduced solely for relative clauses, it would be difficult to explain how it appears in (4.97)a if we accept the conclusion that it is a complement clause parallel to that in (4.90).

4.4.2 Multiple extractions

In the previous chapter it was seen that the restriction on having multiple fronted constituents, and the difficulties (apparently due to information packaging reasons) of extracting from inside relative clauses and interrogative clauses, generally meant that the store feature seldom held more than one constituent at a time. However, the *tough*-movement construction easily allows sentences with more than one gap to be constructed, as in the following well-used example (4.98).

- (4.98) Which violin_{*i*} is this sonata_{*j*} difficult to play ___{*j*} on ___{*i*} ?

Not only can we deduce the need for the store to hold the address of more than one constituent, but we can also see that it seems to operate as a last-in first-out stack, as the following example with crossing dependencies demonstrates.

- (4.99) *Which sonata_{*i*} is this violin_{*j*} difficult to play ___{*i*} on ___{*j*} ?

Indeed, (4.100) shows that it is possible to construct an example where there are three constituents simultaneously in store³⁷. Though this sentence is undeniably difficult to interpret, it still sounds smoothly grammatical.

³⁷By way of *wh*-movement, *tough*-movement and an infinitival interrogative complement construction.

- (4.100) Which problems_i are these students_j difficult to know what_k to say _k to _j about _i ?

Strangely it seems to have been standardly assumed that such multiple extractions are ungrammatical in English and only possible in Scandinavian languages. Engdahl (1985: 133-137) gives an example of a triply-extracted sentence in Swedish and makes the argument that although difficult to process they are grammatical, making them akin to multiply centre-embedded constructions, and that placing a limit on the number of extractions allowed by the competence grammar would be arbitrary and unrevealing. I shall adopt such a position here.

The derivation of (4.100) involves the standard transitions for *wh*- and *tough*-movement. The analysis of the infinitival interrogative clause is not something to be discussed in this thesis and is quite tentative.

- (1) prop⁰, Q¹, STATE³, pres, **be**, content:4, [+3*, -2*]
 (4) QUALITY⁶, **difficult**, theme:5, [+3*, -2*]
 (5) STATE⁸, **know**, exper:α, theme:6, [+α†, -3*, -2*]
 (6) ACTION¹¹, **say**, agent:α, theme:7, benef:8, mod¹³:9, [+α†, -7⁹†, -3*, -2*]
 (7) NON-HUMAN⁹, ref:-
 (8) **to**¹², content: 3, [-3†]
 (3) def⁴, proximate, plur, **student**⁵
 (9) RESPECT¹³, **about**, theme:2, [-2†]
 (2) ref¹:-, **problem**², plur

- (4.100): Which problems are these students difficult to know what to say to about?

Having argued that the capacity of the store in Swedish and English should be considered equal, however, it turns out that the use made of the store is not. In contrast to the clear unacceptability of the English (4.99), Swedish sentences can be constructed which appear to disobey the last-in first-out stack-like nature of the store. As an example of this, we may take the following Swedish sentence, (4.101), quoted in Engdahl (1985: ex.19c = 118).

- (4.101) Strömning_i er den här kniven_j omöjlig at rensa _i med _j .
 herring is this here knife impossible to clean with
 It is impossible to clean herring with this knife.

The conclusion to be taken from this is that is probably a mistake to view the store as a stack even in English. It is preferable to say that in both languages the store contains a

set of constituents in need of an interpretation and that in English there is a syntactic rule that the most recent one is to be taken out first (the date-stamping of constituents will identify which this is), while in Swedish all are accessible and the decision of which one to use depends on how well it combines with the meaning of the sentence. This changes the difference between the languages from a fundamental structural one to a difference in the formulation of syntactic rules.³⁸

4.5 Summary

In this chapter, I have provided accounts of raising and equi constructions that are based on two innovations in the grammar. Firstly, lexical entries containing arguments that are semantically null³⁹, and secondly an analysis of the infinitival *to* as a complementizer that finds for the new proposition the nearest appropriate subject. I argue that this treatment avoids the proliferation of lexical entries, which is important not only for the parsimony of the grammar, but also necessary for modelling the coordination possibilities.

I have further argued that the preparatory-*it* in extraposition constructions should be viewed not as a genuine “dummy”, but rather as a standard pronoun, standing in a cataphoric relation to the extraposed phrase, thus explaining many of the peculiarities of these constructions.

Following the analysis of Pollard & Sag (1994), I have shown that it is problematic to express explicitly control relations in the syntactic transition rules themselves, as they are open to semantic factors. I argue that such a fact is best handled by a model of grammar, such as the present one, in which there is direct access to the growing semantic interpretation of the sentence, and not in one which syntax acts autonomously.

Finally, I argued that it is possible to analyze the *tough*-movement construction as an unbounded-dependency construction, and that the problematic interaction of such movement within finite clauses may be due to the special nature of the *tough* subject, namely that it is simultaneously the subject of another clause.

³⁸It will also be seen in Chapter 7 that in the analysis of verb-final languages with explicit case-marking the order of constituents on the store loses a great deal of its significance.

³⁹Seen before in the analysis of governed prepositional constructions in 2.5.2.

Chapter 5

Coordination

The discussion in the previous chapters has been sprinkled with preliminary glimpses of the treatment of coordination in the dynamic model. In particular I have been assuming that to begin the second conjunct of a coordination, we must effectively return to a previous state in the derivation and proceed from there, an assumption which has served as an important diagnostic tool. In this chapter I will look in detail at coordinated structures and show how they may be modelled in the dynamic grammar.

The problems that coordination phenomena pose for constituency-based grammars have long been apparent. Dealing with “non-constituent coordination” has been one of the central motivations for the development of varieties of categorial grammar with non-standard constituents.¹ However, recent work by Milward² has argued that the adoption of explicitly dynamic grammars allows simple and general solutions to many of the problems posed. The account of coordination in this chapter relies heavily on the insights of Milward’s work.

5.1 Basic coordination

In models of syntax based on constituency structure, the basic assumption that has been made about coordination is that “like constituents coordinate”. To see why this might be thought a natural assumption, we may consider the coordinated structures in (5.1). If we coordinate the constituents in brackets (in constituency terms, S, VP, NP and N, respectively) to create a conjoined constituent of the same type, then these sentences will be allowed by the same rules that allow simple sentences involving no coordination.

¹For example, Moorgat (1988) and Steedman’s *Combinatory Categorial Grammar* (Steedman, 1996).

²Milward (1990, 1992, 1994a & 1994b)

- (5.1) a. ⟨The dog barked⟩ and ⟨the man wept⟩.
 b. The dog ⟨barked⟩ and ⟨howled⟩.
 c. The man brushed ⟨the dog⟩ and ⟨the cat⟩.
 d. The man brushed the ⟨dog⟩ and ⟨cat⟩.

However, the examples in (5.1) far from exhaust the coordination possibilities. For example, the following coordinations in (5.3) are completely permissible, do not have a noticeable difference in intonation pattern to the above, and yet none of the conjuncts is standardly taken to be a syntactic constituent.³

- (5.3) a. He gave ⟨the dog a bone⟩ and ⟨the cat a biscuit⟩.
 b. He fed ⟨the dog yesterday⟩ and ⟨the cat today⟩

In (5.3)a, the conjuncts are made up the two objects in spite of the fact that they do not form a syntactic (or indeed semantic) constituent, while in (5.3)b they are formed from an object and an unrelated adjunct.

Adopting non-standard constituents as in categorial grammar may be one way to save the constituent coordination principle and still derive some of these examples. I will argue, however, that there is a simpler way to characterize all of the coordinations above, if one has a dynamic grammar in which one can in some sense, which will be made clear in due course, return to a previous state in the derivation of the sentence. It will be seen that in all of the examples in (5.1) and (5.3) if upon completion of the first conjunct we return to where it began, the sentence can continue with the second conjunct.⁴

Not surprisingly, this simple observation about the nature of coordination, as being akin to two branches stemming from a common root, has been made many times before. Goodall (1987) proposes that coordinations be represented by parallel phrase structures sharing part of their structure in a three-dimensional representation. Similarly, the later Prague school of dependency syntax⁵ hold that the representation of coordination in dependency structures must resort to three dimensions. As mentioned in section 2.6.2,

³I will restrict the discussion in this section to patterns of coordination in which the conjuncts have been completed (which for want of a better term I call *basic coordination*). Coordination is also possible when the conjuncts have not been completed, and are then both completed by an additional string to their right. This construction, illustrated in (5.2) and which has come to be termed *right-node raising*, is typically given a distinctive intonational pattern and will be dealt with separately in the next section.

(5.2) ⟨The man brushed⟩, and ⟨the woman combed⟩, the cat.

⁴There are obvious instances where this is not the case, for example, (5.4).

(5.4) *The ⟨man brushed the dog⟩ and ⟨woman combed the cat⟩.

The reasons why such examples will not be allowed by the approach taken here are discussed towards the end of this section.

⁵For example, Sgall *et al.* (1985).

Hudson's dependency-based theory Word Grammar (Hudson, 1990) resorts instead to the exceptional use of constituent structures for the same purpose. In either case, the representation of coordination cannot be achieved without major changes to the basic formalism.

The same insight has also been incorporated into processing models, perhaps the best known of these being the SYSCONJ of Wood's ATN (Woods, 1973). On reaching the second conjunct, the parser can return to a previous parse configuration contained in the parse history and continue to parse the second conjunct from this point. A similar approach is taken from the perspective of a DCG-based grammar in Dahl & McCord (1983).⁶ However such approaches again run the danger of conflating grammar and processing, as Milward points out:

“Although processing accounts can provide reasonable coverage of the coordination data, the exact predictions often require detailed examination of the code. This suggests the need for the more abstract level of description which dynamic grammars provide.” Milward (1994a: 5).

Having motivated the general approach to coordination, I will now consider how the precise formulate of the syntactic rule that will allow it, and the consequences that it will have. The first point to decide is at which point in the derivation the coordinate structure is constructed. It is clear that given an initial proposition, such as *the dog barked* in (5.1)a and b, there is no way we can predict that this is going to be the first part of a coordinated sentence, or indeed which part of it is to be coordinated. We must assume then that this initial proposition is interpreted in precisely the same way as it would be in isolation. This implies that the coordinate structure must be directly introduced by the conjunction itself, which here is *and*.

The active constituent at this point in the derivation of (5.1)a will be a saturated proposition, corresponding to *the dog barked*. The conjunction must start a new proposition, which I will take to be attached to the first constituent by way of a new feature **coord**. This feature takes two values: the logical type of the conjunction (& or **or**) and a list of conjuncts ⟨A,B,C...⟩.⁷

⁶These approaches are reviewed in Milward (1994a).

⁷A previous approach used was to relate the two conjuncts by way of a third coordinate proposition containing their two addresses. Although it appears closer to the logical representation of coordination (ie. A, B and A & B are three separate propositions) it has the disadvantage of creating an ungainly analysis and being hard to square with the assumption of monotonicity. The present representation does allow a simple mapping to such a logical representation as will become evident. It might be objected that the proposition A and the coordinated proposition A & B should not be contained in a single constituent as they can be referred to separately. However, as we saw in the discussion of proposition-modifying relative clauses in 3.2.2.1, separate parts of propositions can be picked out and referred to even when not coordinated.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ conjunction, type }	(X, prop)	(N, prop) (X, prop, coord:{ type , (N)})

The result of applying this rule at the transition for *and* in (5.1)a will be as follows.

(1) prop ⁰ , ACTION ³ , past, bark , agent:2, [+2†], coord ⁴ :{&, (3)}
(2) def ¹ , dog ² , sg
(3) prop ⁴

(5.1)a: S_4 **The dog barked and** || the man wept.

The second proposition can now be added in a perfectly standard way to give the following complete derivation.

(1) prop ⁰ , ACTION ³ , past, bark , agent:2, [+2†], coord ⁴ :{&, (3)}
(2) def ¹ , dog ² , sg
(3) prop ⁴ , ACTION ⁷ , past, weep , agent:4, [+4†]
(4) def ⁵ , MALE ⁶ , sg, man

(5.1)a: **The dog barked and the man wept.**

However, the transition introduced above will only allow full propositions to be coordinated, that is for the derivation effectively to return from the end of the sentence right to the beginning and start again. To capture the intuition that one may return to other points in the derivation, I will introduce the notion of the *reduction* of a constituent. This is defined as the result of removing from the constituent information added within the last n states. Using this we may replace the transition schema given above with the following more general schema.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ conjunction, type }	(X)	(N) (X, coord:{ type , (N)})

where X is saturated, and N is a *reduction* of X.

This will allow the sentential coordination (5.1)a given above, for in this case the active constituent is “reduced” to its original state of (*prop*), that is all information is removed which was added after state 0. Turning now to the example with “verb phrase coordination”, (5.1)b, we require the conjunction to reduce the proposition down to (*prop*, [+2]), that is a proposition with a subject already given, and start the second conjunct with

this as active constituent. As the following derivation shows, not only do the verbs end up sharing the subject, but the initial proposition is interpreted in precisely the same incremental way as in (5.1)a.⁸

- (1) prop⁰, ACTION³, past, **bark**, agent:2, [+2†], coord⁴:{&, ⟨3⟩}
 (2) def¹, **dog**², sg
 (3) prop⁴, ACTION⁵, past, **howl**, agent:2, [+2†]

(5.1)b: **The dog barked and howled.**

Note that the tense does not have to be shared by the two conjuncts, as shown in (5.5), and this follows immediately from the analysis given above. The problem faced by constituent structure approaches of deciding what tense the coordinated VP should bear is avoided here as no such object is constructed.

(5.5) The dog ⟨barked⟩ and ⟨will now howl⟩.

Example (5.1)c exhibiting “noun phrase coordination” will similarly be allowed by the schema, analyzed again as a coordination of propositions. This time the only information removed from the original proposition when it is copied will be the value of its patient role, resulting in the following derivation.

- (1) prop⁰, ACTION³, past, **brush**, agent:2, patient:3, [+2†], coord⁶:{&, ⟨4⟩}
 (2) def¹, MALE², sg, **man**
 (3) def⁴, **dog**⁵, sg
 (4) prop⁶, ACTION, past, **brush**, agent:2, patient:5
 (5) def⁷, **cat**⁸, sg

(5.1)c: **The man brushed the dog and the cat.**

In fact it appears that we must consider example (5.1)c to be ambiguous, for there will also be an interpretation in which the constituents *the dog* and *the cat* coordinate with each other to form a combined entity. The difference between these alternatives shows up in the negative, the conjunction *or* replacing *and* except where it is an instance of the latter direct coordination, as in (5.6)c.

⁸It was argued in 2.6.1, that this was a fundamental problem for categorial grammar approaches to incremental interpretation.

- (5.6) a. He liked ⟨fish⟩ and ⟨chips⟩.
 b. He didn't like ⟨fish⟩ or ⟨chips⟩.
 c. He didn't like ⟨fish⟩ and ⟨chips⟩.

The original coordination schema will also allow nominal coordination for example (5.1)d, copying the **def** feature of the first conjunct into the second.⁹

- (1) prop⁰, ACTION³, past, **brush**, agent:**2**, patient:**3**, [+2†]
 (2) def¹, MALE², sg, **man**
 (3) def⁴, **dog**⁵, sg, coord⁶:{&, ⟨4⟩}
 (4) def⁶, **cat**⁷, sg

(5.1)d: **The man brushed the dog and cat.**

It may be useful here to clarify the approach that we have made so far. It is true to say that we are not returning to a previous state in the parse, in the sense of Wood's SYSCONJ. The syntactic rule allowing coordination is similar to other rules in that it looks at the active constituent on top of the stack and adds a new active constituent to the stack, linked to the previous one. The new active constituent is formed by stripping off various amounts of information from the present one.¹⁰

Let us now consider the sentence we noted previously, (5.4) repeated here, which does not allow coordination, although it agrees with our pre-theoretic notion of “returning to a previous position in the string and continuing”.

- (5.4) *The ⟨man brushed the dog⟩ and ⟨woman combed the cat⟩.

If we consider the options of reducing the active constituent, the initial proposition, when the conjunction is reached, we will see that the only distinct positions that can be achieved are before the first *the* (by removing all the added info), before *brushed* (by removing the info added by the verb) and before the second *the* (by removing the address

⁹I shall not go into nominal coordination, as it appears to involve many problems with reference that are outside the scope of the present model. Later in this section I will argue that this may also be taken to be an example of propositional coordination.

¹⁰Indeed, although for present purposes I will assume that this stripping is based on “recency of information”, it may well be the case that we need to strip in other ways. For example, one possible rule could perhaps be better characterized as “strip everything but the subject”. This would then allow the well-known pattern, exemplified in (5.7).

- (5.7) “Watch out!” said John, and grimaced a little to himself.

Although perhaps marginal in English, such constructions are the norm in verb-second languages such as German. I will such issues to future exploration, however, and assume that information is stripped on the basis of recency.

of the patient). There is no way to get to the position the derivation was in after the initial *the* since there is no information on the proposition added at the second transition, for at this time the subject was the active constituent not the proposition. So it transpires that the formalization we have given is more restrictive than the original pretheoretic insight, and has a better fit with the data.¹¹

The two examples of non-constituent coordination given above in (5.3) will both be derivable with the same schema used for constituent coordination. For example, the derivation of (5.3)b is similar to that of (5.1)b above, but here both the object and the time adjunct are stripped off to create the second proposition.

- (1) prop⁰, ACTION², past, **feed**, agent:2, patient:3, mod⁵:4, [+2†], coord⁶:{&, (5)}
- (2) MALE¹, sg
- (3) def³, **dog**⁴, sg
- (4) TIME⁵, **yesterday**
- (5) prop⁶, ACTION, **feed**, agent:2, patient:6, mod⁹:7
- (6) def⁷, **cat**⁸, sg
- (7) TIME⁹, **today**

(5.3)b: **He fed the dog yesterday and the cat today.**

It is apparent that there is also a general tendency in coordinate structures for the conjuncts to be symmetrical as in the examples considered so far. For example, there is nothing in the grammar itself that will disallow the following examples.

- (5.8) a. He fed ⟨the dog yesterday⟩ and ⟨the cat⟩.
 b. He fed ⟨the dog⟩ and ⟨the cat today⟩

Indeed, (5.3)a is perfectly natural with the interpretation that both animals were fed yesterday. The reading of (5.3)b given here is not so natural, the adjunct being typically taken to refer to both actions, but may be obtained with the appropriate intonation. It will be noted that both of the above sentences will cause further problems for standard “like constituent” approaches.¹²

This analysis will also allow one of the conjuncts to reverse the order of arguments by heavy-element shift, although again it will run up against the symmetry metric.

- (5.9) He placed ⟨a vase on the table⟩ and ⟨on the sideboard a figurine he had inherited from his great aunt⟩.

¹¹ An extension of the grammar which, potentially, will admit such examples is discussed in section 5.1.7.

¹² Though they will not cause problems for a flexible categorial approach such as the CCG of Steedman (1996).

5.1.1 Null conjuncts

In the definition given above it was stated that one could strip off n layers of information to arrive at the *reduction* of a constituent. There was no constraint that this n may not be 0, in which case the constituent and its reduction will be identical. This will then allow the derivation of sentences such as (5.10), where the second conjunct appears to be tagged on to a copy of the first.

(5.10) He does give bones to dogs ⟨ ⟩, but ⟨only after consulting his wife⟩.

Such examples are generally best with the contrastive conjunction *but*. If there were no such contrast involved, information in the second conjunct could be added to the sentence without the need for coordination. However, such sentences can be fine with non-contrastive *and* as long as two independent propositions are being expressed.¹³

(5.11) There *are* trains that go to North Berwick ⟨ ⟩, and ⟨quickly⟩!

Of course, the second conjunct cannot be missing as then the second proposition would be expressing nothing. Therefore, null coordination is inherently asymmetric. Such examples of null coordination fall out naturally from the dynamic analysis of coordination given here, but again present problems to constituency based ones.¹⁴

5.1.2 Coordination of unlike constituents

We have already met in the previous two chapters numerous examples of coordination in which the conjuncts are of distinctly different syntactic (and conceptual) types. A standard example given of this phenomenon is that of predicates coordinating with other predicates of a different category, for example a noun phrase and an adjectival phrase in the well-known example (5.12).

(5.12) He is ⟨a Republican⟩ and ⟨proud of it⟩.

¹³Note the possible reading of (5.11) is synonymous with “There are trains that go to North Berwick and go quickly” rather than “There are trains that go to North Berwick, and there are trains that go to North Berwick quickly.” The latter, impossible, reading is not derivable with the transition rule formulated above.

¹⁴Of course there must also be some *necessity* for coordination, ie. some necessity to express the information in two propositions instead of one. In the case of different arguments of the proposition there is no alternative, but with non-contradictory adjuncts this is not the case, for example *today* and *in hospital* can both follow *he died*, but it is pretty strange, though surely not ungrammatical, to say *he died today and in hospital* when the same content can be expressed perfectly well by *he died today in hospital*.

A common approach¹⁵ is to say that the two conjuncts share some feature PRED, and that they are therefore ‘like’ enough to coordinate. This contrasts with the present approach where the single general coordination schema will predict the coordination of arguments, on condition that the second can continue from the point at which the first was added.

The data in (5.13) appear to support this latter analysis. Given the former underspecification analysis, we must suppose that the adjectival phrase *restless* shares some common type with the infinitival clause in (5.13)a and with a verb phrase headed by an *ing* verb in (5.13)b. If this is so, and the adjectival phrase seems to be of the same conceptual and syntactic type in both instances, then it is difficult to explain why they cannot coordinate with these same constituents in (5.13)c and (5.13)d.

- (5.13) a. He seemed ⟨restless⟩ and ⟨to be trying to leave⟩.
 b. He is ⟨restless⟩ and ⟨trying to leave⟩.
 c. *He seemed ⟨restless⟩ and ⟨trying to leave⟩.
 d. *He is ⟨restless⟩ and ⟨to be trying to leave⟩.

The only way out for this approach would be to suppose that the phrase *restless* is of two different types: a type that can combine with *seem* (shared by infinitives) and a type that can serve as a predicate (shared by *ing*-verb phrases).¹⁶ Instead it is much simpler to suppose that (5.13)c and (5.13)d are ungrammatical precisely because (5.15)a and (5.15)b are ungrammatical.¹⁷ This follows directly from the dynamic view of coordination adopted here.

- (5.15) a. *He seemed trying to leave.
 b. *He is to be trying to leave.

It should be clear that it is not sufficient for a successful coordination that the conjuncts could appear after the initial shared string. The initial string must also be interpreted the same way in both cases. For example, in (5.16) the strings *a cake* and *iron his shirt* can both follow the string *he made her*, but its interpretation would be different in each case.

¹⁵For example, the one assumed in Gazdar *et al.* (1985), and Pollard & Sag (1994).

¹⁶Although it has been claimed that the *ing* verb phrases in (5.13)b are predicates, which allows them to coordinate with other predicates, this seems doubtful. Given standard assumptions about predicates, (5.14)a is a predicate, but (5.14)b is a standard progressive construction.

- (5.14) a. The problem is getting there.
 b. The bus is slowly getting there.

¹⁷Quite how a grammar might rule out (5.15)b is another question, though of no relevance here.

In (5.16)a *make* has a lexical entry requiring a beneficiary and a completed theme, while in (5.16)b it is an object-raising causative.¹⁸

- (5.16) a. He made her a cake.
 b. He made her iron his shirts.
 c. *He made her ⟨a cake⟩ and ⟨iron his shirts⟩.

5.1.3 Negative conjunctions

Mention should also be made of the negative conjunctions *nor* and *neither*, which trigger characteristic subject-auxiliary inversion.¹⁹ These appear to be limited largely to full propositional coordination.

- (5.17) a. She doesn't drink, nor does she dance.
 b. She doesn't drink, neither does she dance.
 c. *?She doesn't drink, nor dance.
 d. *She doesn't drink, neither dance.

The transitions they permit may be characterized in the following transition schema.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ conjunction, neg, & }		(N, prop, neg)
	(X, prop, neg)	(X, prop, neg, coord:{&, ⟨N⟩})

This allows the following derivation of (5.17)a.

<p>(1) prop⁰, pres¹, neg, content:3, [+2*], coord⁴:{&, ⟨4⟩} (3) ACTION³, drink, agent:2, [2†] (2) FEMALE¹, sg (4) prop⁴, neg, pres⁵, content:6, [+5*] (6) ACTION⁷, dance, agent:5, [+5†] (5) FEMALE⁶, sg</p>
--

(5.17)a: **She doesn't drink, nor does she dance.**

¹⁸The reason that in such constructions are potentially humorous is that they force us to revise our initial assumptions about the analysis of the first string before we can understand the second.

¹⁹I will not treat the conjunction *not* which attaches a negative proposition to a positive one. For discussion of this construction and its potential for introducing gapping, see Oerhle (1987).

5.1.4 Correlative constructions

It is not always the case that coordination is not predictable before the second conjunct is reached. The point in the derivation from which the second conjunct is to start may be explicitly marked by one of what I shall term the *pre-conjunctions*: *both*, *either* and *neither*, which together with the later conjunct are referred to as correlatives, and illustrated in (5.18).

- (5.18) a. He wanted both ⟨to succeed⟩ and ⟨to be happy⟩.
 b. Either ⟨he goes⟩ or ⟨I go⟩.
 c. He had neither ⟨seen her⟩, nor ⟨spoken to her⟩.

Unlike *either* (and more marginally *neither*), *both* cannot be used before a proposition.

- (5.19) a. *Both he danced and she sang.
 b. ??Neither could he dance, nor could he sing.

The transition for these pre-conjunctions will have to specify the **type** of the coordination that is to come, that is **&** or **or**, but the address of the first conjunct cannot yet be identified.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ pre-conjunction, type }	(X, prop)	(X, prop, coord:{ type , ⟨ ⟩})

Obviously this will rule out the second conjunct never appearing, as in (5.20), since the proposition will remain unsaturated.

- (5.20) a. *He wanted both to succeed.
 b. *Either he goes.

The transition schema for the conjunctions will then have to be relaxed somewhat to allow the type of the conjunction to be either supplied (as in the previous cases) or merely checked, as at S_4 in the following derivation of (5.18)b.

- | |
|---|
| (1) prop ⁰ , EVENT ³ , pres, go , theme:2, [+2†], coord ¹ :{ or , ⟨3 ⁴ ⟩} |
| (2) MALE ² , sg |
| (3) prop ⁴ , EVENT ⁶ , pres, go , theme:4, [+4†] |
| (4) 1ST-PERSON ⁵ , sg |

(5.18)b: **Either he goes or I go.**

5.1.5 Embedded coordination

In the same way that when we begin a simple sentence we cannot predict that it might be the first part of a coordinate structure, when we start a coordinate structure we do not know if it will turn out itself to be a conjunct, as in (5.21).

(5.21) The dog barked and howled or cowered and whined.

Assuming binary coordination, the sentence has five possible interpretations, which correspond to the following bracketings.²⁰

A(B(CD)) (AB)(CD) ((AB)C)D (A(BC))D & A((BC)D)

The simple coordination schema that we have will allow all of the five interpretations to be derived. As an example, let us consider the derivation of the second of these, as shown in (5.22).

(5.22) The dog ⟨barked and howled⟩ or ⟨cowered and whined⟩.

The first coordination *barked and howled* will be interpreted in all these sentences the same way as above, with a new proposition attaching to the initial one at S_4 . For reading (5.22) the second conjunction *or* must also attach a new proposition to the initial one, *the dog barked*.²¹ The final conjunction *and* will attach a new proposition to the previous one, *the dog cowered*.

The resulting derivation of (5.22) is as follows.

- (1) prop⁰, ACTION³, past, **bark**, agent:2, [+2†], coord⁴:{&, ⟨3⟩}, coord⁶:{or, ⟨4⟩}
 (2) def¹, **dog**², sg
 (3) prop³, ACTION⁵, past, **howl**, agent:2, [+2†]
 (4) prop⁶, ACTION⁷, past, **cower**, agent:2, [+2†], coord⁸:{&, ⟨5⟩}
 (5) prop⁸, ACTION⁹, past, **whine**, agent:2, [+2†]

(5.21): **The dog barked and howled or cowered and whined.**

The notation of allowing successive conjunction features on the same constituent may look non-standard, but it is quite consistent: $A,+B,+C$ is equivalent to $(A+B)+C$ and $A,+(B,+C)$ to $A+(B+C)$.

²⁰Allowing n -ary coordination of like conjunctions, it has another reading: $A(BC)D$. See the next section for a discussion of n -ary coordination.

²¹This will also be the case for the third reading listed above.

The five possible bracketings of $A + B \vee C + D$ given above will then be represented in the following manner.

$A,+(B,\vee(C,+D))$ $A,+B,\vee(C,+D)$ $A,+B,\vee C,+D$ $A,+(B,\vee C),+D$ & $A,+(B,+C,\vee D)$

It should be noted that this analysis of multiple coordination, all of it following from the one transition schema, retains maximally incremental interpretation. For example, the constituent *the dog* is interpreted as the first argument of all of the verbs as soon as they are met. This is important, for such constructions are not hard to process, that is the syntax is clear and any difficulty is with the complexity of the meaning itself. Therefore, they should not involve any need for the revision of existing interpretations.

5.1.6 Non-binary coordination

The multiple coordinations considered in the last section only involved binary coordination. However, in a sentence such as (5.24)a any hierarchical arrangement of the conjuncts does not agree with our intuitions that there is one three-way coordination, rather than any hierarchy.²² In contrast, coordination without a conjunction in front of the last conjunct, as in (5.24)b is at best a very marked construction in English.²³

- (5.24) a. The dog barked, howled and whined.
 b. *?The dog barked, howled.

The second conjunct is not introduced by any lexical item. I will assume that it is initiated on the transition of the first word of the conjunct and added to the interpretation before the transition associated with the word itself. Such a transition will thus be possible at any transition, although in written English it is marked with a comma, and in spoken English a break in the intonation contour. The following transition schema to introduce an unmarked conjunct is identical with the one associated with explicit conjunctions, except for the fact that no conjunction type is specified.²⁴

²²In fact the necessary relaxing of the rule for explicit conjunctions appearing after pre-conjunctions, which allows them to appear where the conjunction has been explicitly marked in advance, will also allow a flat reading for (5.23) as well as the two readings with bracketed binary coordination.

(5.23) The dog barked and howled and whined.

These seems to agree with the intuition that such a sentence can be bracketed, but need not necessarily be.

²³At least in standard styles of written English. In speech and more poetic styles of written English it is perhaps more commonly met with. For present purposes I will assume it to be ungrammatical.

²⁴Indeed, it may be possible to generalize the previous transition rule to include unmarked transitions as a special case.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
	(X)	(N) (X, coord: { , ⟨N⟩ })

where X is saturated, and N a *reduction* of X.

The type of the coordination will typically be filled in when the explicit conjunction is met introducing the final conjunct. It will be remembered that the *coord* feature was set up to hold an indefinite number of conjuncts.

The following derivation will thus be allowed, representing a flat coordination of the three conjuncts.

- (1) prop⁰, ACTION³, past, **bark**, agent:2, [+2†], coord⁴:{&⁵, ⟨3,4⟩}
 (2) def¹, **dog**², sg
 (3) prop⁴, ACTION⁴, past, **howl**, agent:2, [+2†]
 (4) prop⁵, ACTION⁶, past, **whine**, agent:2, [+2†]

(5.24): **The dog barked, howled and whined.**

The final conjunction could alternatively have been the disjunctive *or*, or the contrastive *but*. The analysis seems satisfactory in that there appears to be no prediction made as to what conjunction is to come. This can be seen as from the fact that the list of conjuncts can be indefinitely long without the addition of the final conjunction causing surprise, as evidenced in (5.25).

(5.25) The dog would bark, howl, screech, whine, growl or/and gnash its teeth.

5.1.6.1 Unequal coordination

The two sentences below, (5.26)a and b, are evidence that multiple coordinations are often unequal, in that the second conjunct returns to one place in the derivation, and the final conjunct to another. For example, the second conjunct in (5.26)a strips back to after *has*, while the third strips back to before *has*.

- (5.26) a. He has ⟨eaten all the trifle⟩, ⟨drunk all the wine⟩, and ⟨will now start on the nuts⟩.
 b. The spy was ⟨in his forties⟩, ⟨of average build⟩, and ⟨spoke with a slightly foreign accent⟩. (Quirk *et al.*, 1985: 972)

Although such examples might be considered on logical reflection to be “sloppy English”, they are common in both the spoken and written language and create no effect of ungrammaticality even when one is conscious of their “irregularity”.

It will be noted that there is nothing in the dynamic treatment of non-binary coordination adopted here that forces the remaining conjuncts to start from the same place. The principle of symmetry, however it is to be formalized, will argue against such constructions, but they will nevertheless be allowed by the simple transition schema given above.

The derivation of a simplified version of (5.26)b will be as follows.

- (1) prop⁰ STATE², past, **be**, content:3, [+2*], coord:{&⁵, <4,6>}
 (3) QUALITY³, **old**, theme:2, [+2†]
 (2) MALE¹, sg
 (4) prop⁴, STATE, past, **be**, content:5, [+2*]
 (5) QUALITY⁴, **well-dressed**, theme:2, [+2†]
 (6) prop⁵, ACTION⁶, past, **speak**, agent:2, mod⁷:7, [+2†]
 (7) MANNER⁷, **funny**

(5.26): **He was old, well-dressed and spoke funny.**

Once again, such coordinations would appear to pose a serious challenge to approaches to coordination based on the principle of “like constituents coordinate”.

5.1.7 Multiple copying

I mentioned previously that in examples such as (5.27)a, the coordination of the nominal heads could be analysed by the coordination schema as forming a compound nominal expression. However, it appears that it is also possible to interpret it as a coordinated proposition. This is backed up by (5.27)a, which appears only to have a reading as a coordinated proposition since the two adverbs are clearly modifying separate propositions.

- (5.27) a. He fed the <dog> and <cat>.
 b. He fed the <dog yesterday> and <cat today>.

However, neither of these examples are derivable according to the coordination rule I have given. We cannot return to the state after *the* since at this point in the derivation the proposition was not the active constituent. The only way they may be derived is to assume that instead of copying a single constituent, the conjunction transition copies the present active constituent plus part of one of its dependents, placing it on the active stack. In the case of the above examples, this will mean that the first part of the object constituent *the dog* will be copied.

The derivation that this assumes for (5.27)b is the following.

- (1) prop⁰, ACTION², past, **feed**, agent:2, patient:3, mod⁵:4, [+2†], coord⁶:{&, (5)}
- (2) MALE¹, sg
- (3) def³, **dog**⁴, sg
- (4) TIME⁵, **yesterday**
- (5) prop⁶, ACTION, **feed**, agent:2, patient:6, mod⁸:7
- (6) def⁶, **cat**⁷, sg
- (7) TIME⁸, **today**

(5.3)b: **He fed the dog yesterday and cat today.**

If we introduce a new rule to allow such constructions, then we must be careful not to allow (5.4), repeated here, which we held to be ungrammatical and which is not allowed by the simple coordination schema.

(5.4) *The ⟨man brushed the dog⟩ and ⟨woman combed the cat⟩.

One obvious difference between the grammatical examples, such as (5.3)b, and the ungrammatical examples, such as (5.4), is that in the latter the analysis returns to a position before the verb. It would be possible to include such a prohibition on the rule, although it is not clear why this should be the case.²⁵

Finally, we should note that it can be demonstrated that such multiple copying of constituents will not be limited to just two constituents. If we really wish to derive the following sentence, (5.28), where the time adjuncts are taken to modify *wanted*, then we have to assume that the conjunction copies part of the object constituent, *the...*, the event *see*, as well as the active constituent the propositional state *want*.

(5.28) The boy can't make up his mind about what to see at the zoo... He wanted to see the ⟨tiger today⟩ and ⟨lion yesterday⟩.

The analysis of such examples clearly requires further investigation.

5.2 Right-node raising

One condition placed on the coordination transition schema so far has been that the first conjunct, ie. the active constituent, should be complete before the second conjunct is

²⁵It seems that we are far from any understanding of what is going on here. One has a feeling that there may be some affinity between the constructions here and the gapping construction discussed later in the chapter, but this again is a mysterious construction.

attached to it. However, the sentences in (5.29) show that coordination is possible even if the first conjunct is unfinished, as long as the missing material is added afterwards to simultaneously complete both conjuncts.

- (5.29) a. He ⟨brushed _⟩ and ⟨fed _⟩ the dog.
 b. ⟨He brushed _⟩, and ⟨she fed _⟩, the dog.
 c. He ⟨gave a shilling _⟩ and ⟨lent a pound _⟩ to the old man.

I will assume that the transition schema for the conjunction will be the same as the one we have already, and we drop the condition that the active constituent be complete.

This will allow the following partial derivation of (5.29)b.

- (1) prop⁰, ACTION², past, **brush**, agent:2, patient: , [+2†], coord³: {&, ⟨3⟩}
 (2) MALE¹, sg
 (3) prop³, ACTION⁵, past, **feed**, agent:4, theme: , [+4†]
 (4) FEMALE⁴, sg

(5.29)b: **He brushed, and she fed, ||** the dog.

We then have to enforce the fact that the final constituent, *the dog*, must be interpreted simultaneously in the two incomplete coordinated constituents. To do this it appears necessary to introduce a generalization of our earlier transition rule, **add-argument**, which stipulates that an argument can be added to any number of coordinated constituents on the stack. The rule for two conjuncts, as required in (5.29)b, will be as follows.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
		(N)
add-argument	(X1, arg:) (X, arg: , coord:{, ⟨X1⟩})	(X1, arg:N) (X, arg:N, coord:{, ⟨X1⟩})

With this schema applying at the fifth transition the derivation may be completed in the following fashion.

- (1) prop⁰, ACTION², past, **brush**, agent:2, patient:5, [+2†], coord³: {&, ⟨3⟩}
 (2) MALE¹, sg
 (5) def⁶, **dog**⁷, sg
 (3) prop³, ACTION⁵, past, **feed**, agent:4, theme:5, [+4†]
 (4) FEMALE⁴, sg

(5.29)b: **He brushed, and she fed, the dog.**

It should be pointed out that this is only an exploratory analysis of right-node raising, and there remains a great deal of work to be done to assess if it is a viable approach.

This construction can also be used to investigate the nature of case checking, which was discussed previously in the section on free relatives, 3.2.4.1. The same phenomenon, that in German mismatched cases are allowed as long as they have the right phonological form, also applies to the final constituent in right-node raised constructions.²⁶

(5.30) *Sie ⟨findet _⟩ und ⟨hilft _⟩ Männer/Männern.
 she finds_{acc} and helps_{dat} men_{acc}/men_{dat}

(5.31) Er ⟨findet _⟩ und ⟨hilft _⟩ Frauen.
 he finds_{acc} and helps_{dat} women_{acc/dat}

Any approach which assumed that the verbs could be conjoined before the arguments were added would need some system for representing the conflicting case demands on the missing argument of the conjunction. The derivation is unproblematic given the present approach provided that we allow constituents to have disjunctive case values.

5.3 Gapping

The term *gapping* has been used previously to refer to various non-constituent coordination constructions, but I will adopt the usage of Hudson (1990: 405) that “gapping is possible only if the ‘gap’ contains a verb (or other word that has a subject).” The construction is illustrated by the examples in (5.33), which resemble standard coordinations, except for the fact that a different subject is supplied in the second conjunct.²⁷

- (5.33) a. Peter likes dogs and Mary _ cats.
 b. The linesman was quick to flag and the referee _ to whistle. (Radio broadcast)

The problem posed by the construction is how to map the string into the same meaning as the non-gapped alternatives. One fundamental choice to be made is whether the gap in the second conjunct is predicted from the conjunction, or whether the second conjunct

²⁶The data is taken from Ingria (1989), who in turn cites Pullum & Zwicky (1986).

²⁷I do not consider here examples where the contrasted item is not the subject, but an initial modifier or topic as in (5.32).

- (5.32) a. Yesterday he fed the dogs and today _ the cats.
 b. The cats, he fed today, and the dogs _ yesterday.

Obviously, these should be taken into consideration in a more complete account.

is started as a normal proposition and the missing material reconstructed at the gap site. I shall tentatively assume the former of these approaches.²⁸

To predict the gap at the conjunction, the transition must stipulate that the subject of the first conjunct be temporarily filled by a placeholder in the copy that is made and attached as the second conjunct. The original transition schema for conjunctions is thus modified into the following gapping schema.

transition	lexical entry	top of stack (S_n)	top of stack (S_{n+1})
gapping	{ conjunction, type }	(X, [+Y])	(N, [+α]) (X, [+Y], coord:{ type , ⟨N⟩})

where N is otherwise a *reduction* of X.

So with this transition for the conjunction in (5.33)a, the interpretation at S_4 will be as follows.

- (1) prop⁰, STATE², pres, **like**, exper:2, theme:3, [+2†], coord⁴:{&, 4}
 (2) MALE¹, sg, **Peter**
 (3) **dog**³, plur
 (4) prop⁴, STATE, pres, **like**, exper: , theme: , [+α]

(5.33)a: S_4 **Peter likes dogs and** || Mary cats.

The constituent added after the conjunction substitutes for the placeholder, as seen in the heavy-shift construction, becoming the subject of the new proposition and also the experiencer of *likes*. The derivation then proceeds as normal with the addition of the theme, as shown in the following diagram.

²⁸The latter approach does have the advantage that it decreases the local ambiguity at the transition for the conjunction. However, it would also predict that there should be no worsening in the processing of the construction with increased length of the subject. I am not sure whether (5.34)b confirms this prediction or not.

- (5.34) a. The pretty lady in green likes dogs and the handsome man I met yesterday likes cats.
 b. ??The pretty lady in green likes dogs and the handsome man I met yesterday __ cats.

- (1) prop⁰, STATE³, pres, **like**, exper:2, theme:3, [+2†], coord⁴:{&, 4}
 (2) MALE¹, sg, **Peter**
 (3) **dog**³, plur
 (4) prop⁴, STATE, pres, **like**, exper:5, theme: 6, [+α/5†]
 (5) FEMALE⁵, sg, **Mary**
 (6) **cat**⁶, plur

(5.33)a: **Peter likes dogs and Mary cats.**

The supposition that *Mary* here genuinely fills the subject position of the second conjunct is confirmed by the observation that where this is a pronoun the form used is the nominative, not the accusative, which would be the case were it merely filling an argument position.²⁹

- (5.35) a. He likes dogs and she _ cats.
 b. *He likes dogs and her _ cats.

As formulated above, the gapping transition rule will also allow cases where the second conjunct is a *null reduction* of the first, that is where it is identical except for the removal of its subject value. This will then allow sentences such as those in (5.36).

- (5.36) a. Peter went away and Mary _ too.
 b. Peter will go or Mary _ .

As in standard coordination, the number of constituents copied into the second conjunct may be greater than one. This is illustrated by the examples in (5.39).³⁰

- (5.39) a. Peter wanted to write novels, and Paul _ plays.
 b. Peter wanted to be allowed to write novels, and Paul _ plays.

²⁹However, this could also be taken as evidence for the alternative approach, discussed above, in which the gapping was not predicted from the conjunction.

³⁰The badness of the examples in (5.37) follows directly from the transition rule, for only the subject can be replaced, and therefore that is the only constituent that can appear to the left of the gap.

- (5.37) a. *Peter wanted to write novels, and Paul wanted _ plays.
 b. *The linesman was quick to flag and the referee was _ to whistle.

There is a similar construction, *pseudogapping*, illustrated in (5.38).

- (5.38) Peter liked dogs, and Mary did cats.

Although I offer no analysis of it here, it appears that the pro-verb *did* must somehow resurrect the missing material.

Replacing the subject value with the placeholder when the second conjunct is initiated will replace all instances of it, including in its position as subject of the embedded clause. Subsequently, adding the subject into the second conjunct will replace all of these. The derivation of (5.39)a will thus be as follows.

- (1) prop⁰, STATE², past, **want**, exper:**2**, theme:**3**, [+2†], coord⁶:{&e, <5>}
 (2) MALE¹, sg, **Peter**
 (3) ACTION⁴, **write**, agent:**2**, theme:**4**, [+2†]
 (4) **novel**⁵, plur
 (5) prop⁶, STATE, past, **want**, agent:**7**, theme:**6**, [+α/7†]
 (7) MALE⁷, sg, **Paul**
 (6) ACTION, **write**, agent:**7**, theme:**8**, [+α/7†]
 (8) **play**⁸, plur

(5.39)a: **Peter wanted to write novels and Paul plays.**

Just as we saw with the standard coordination in 5.1.6.1, there seems no way in this approach to gapping to stipulate that multiple conjuncts should be equal, that is that gapped conjuncts should not be coordinated with non-gapped ones. This seems to agree with the data, as shown in (5.40).

(5.40) The man likes dogs, the woman _ cats and the daughter simply adores horses.

The analysis also predicts that it should be possible to combine gapping with right-node raising, and this is indeed the case.

- (5.41) a. ⟨Peter wanted to write _⟩, and ⟨Paul _ to compose _⟩, a major work about dogs.
 b. ⟨Hugford was among the earliest to forge _⟩, and ⟨Patch _ systematically to record _⟩, early Tuscan frescoes. (Matthews, 1981: 202)

A partial derivation of (5.41)a will be as follows.

- (1) prop⁰, STATE², past, **want**, exper:2, theme:3, [+2†], coord⁵:{&, <4>}
- (2) MALE¹, sg, **Peter**, +nom²
- (3) ACTION⁴, **write**, exper:2, theme:7, [+2†]
- (7) indef⁹, sg
- (4) prop⁵, STATE, past, **want**, agent:5, theme:6, [+α/5†]
- (5) MALE⁶, sg, **Paul**
- (6) ACTION⁸, **compose**, agent:5, theme:7

(5.41)a: (*S*₉) **Peter wanted to write, and Paul to compose, a || major work about dogs.**

5.4 Summary

In this chapter I have extended the model to deal with coordination phenomena of various kinds. This has been done in the main part by adding a transition for conjunctions which creates a partial copy of the active constituent and attaches it to this constituent via the relation *coord*. This simple rule handles constituent coordination, as well as many types of non-constituent coordination alike. It also predicts that coordination can occur when the conjuncts are mismatched in terms of syntactic types, or the number of constituents in each one.

Although to a certain extent this rule captures Milward's insight that coordination involves returning to a previous state in the derivation, it is more restricted in that it cannot return to point where the current active constituent was not on top of the active stack. This latter type of construction can, however, be modelled by a more complicated rule, involving the copy of multiple constituents.

One major advantage claimed for the model is that interpretation is maximally incremental. Even in cases of multiply-embedded coordinations, the coordinate structure does not have to be predicted in advance, nor do constituents have to wait for their interpretation until the coordination has been resolved.

The analysis is extended tentatively to give analyses for right-node raising and gapping.

The constructions discussed here are far from being exhaustive and much further work needs to be done. However, I hope that the discussion has been sufficient to show the potential of dynamic grammars of the present type for offering simple and general solutions to constructions that have hitherto been highly problematic for approaches based on syntactic constituent structures.

Chapter 6

Discontinuous constituency

It may seem paradoxical in an approach which does not employ syntactic constituents, and where conceptual constituents have no relation to the order of words, to devote a chapter to constructions exhibiting “discontinuous constituency”. However, as was shown in chapter 2, we can define in the model a derivative concept of “syntactic phrase”, in terms of “the span of words during which a particular constituent is present on the active stack”. In the present model discontinuous phrases will then be those that involve moving inactive constituents back onto the stack, or modifying non-active constituents.

In this chapter, I examine a range of constructions which exhibit this behaviour, or else have been held to be discontinuous because of the problems they pose for constituency approaches.

6.1 Extraposition

6.2 Optional extraposition

In the analysis of relative clauses in chapter 3, the assumption was made that these had to modify the active constituent and thus appeared directly following the constituent they modified. However, if we consider a sentence such as (6.1)b, we can see that this assumption does not necessarily hold, for the relative clause appears to have moved away from the modified constituent. In this section I shall investigate how the model should deal with such cases.

- (6.1) a. A dog which had no nose came in.
 b. A dog came in which had no nose.

Such dislocations are collectively referred to with the term *extraposition*. It has long been established that extraposition is clause-bounded, as shown by the following examples from Postal(1971: 194).

- (6.2) a. That it was disgusting that Harry came late is obvious.
 b. *That it was disgusting is obvious that Harry came late.

The same pattern holds for extraposed relative clauses.

- (6.3) a. That a dog should come in which had no nose was shocking.
 b. *That a dog should come in was shocking which had no nose.

To formulate the transition rule, therefore, that will allow such extrapositions it is clear that it must refer to the current active constituent as other rules do. If it were to be set up to look blindly in the deactivated constituents for a suitable antecedent there would be no way to model this clause-bounded nature. Therefore finding the antecedent must involve referring first to the active constituent and then looking at its dependents. The transition introducing subject relative clauses will thus be modified to give the following schema.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
extraposed relative clause	(X, arg:Y)	(N, prop, [Y]) (Y, mod:N) (X, arg:Y)

Applying this rule in the fifth transition of the following derivation of (6.1)b allows us to obtain a final interpretation which is semantically identical to that of the non-extraposed (6.1)a above.

(1) prop ⁰ , EVENT ³ , past, come , theme:2, goal:3, [+2†]
(2) indef ¹ , sg, dog ² , mod ⁵ :4
(4) prop ⁵ , STATE ⁶ , past, have , benef:2, theme:5, neg ⁷ , [+2†]
(5) indef ⁷ , nose ⁸ , sg
(3) DIRECTION ⁴ , in

(6.1)b: **A dog came in which had no nose.**

This transition will apply equally well where the antecedent is in object position but not active, to allow the extraposed relative clause in (6.4).

(6.4) He saw a dog yesterday which had no nose.

Interestingly the rule given above will not allow an example such as (6.5)a¹ where the extraposed clause appears before the constituent it modifies has been interpreted by the main verb, since it must be in an argument position of the active constituent.

- (6.5) a. *Children reluctantly who came from far away arrived.
 b. Children reluctantly arrived who came from far away.

This agrees with the observation that (6.5)b, where the modified constituent has been interpreted, is significantly better.

It has often been noted that nominal phrases do not have obligatory arguments, but rather that even where semantically selected these complements are non-obligatory, and hence more akin to modifiers from the point of view of syntax. It would be expected therefore that these too will allow extraposition and this is indeed a particularly common pattern.

- (6.6) a. A decision was made to tell the police.
 b. The marriage took place today of Paul and Linda.

It should be noted that the transition rules for extraposed modifiers (actually all modifiers), which will parallel that given for subject relative clauses above, should not compel the extraposed modifier to be placed at the end of the clause. This is a common position for finite clauses, being the “heaviest” constituents, but other extraposed modifiers are often seen non-finally, resulting in an intertwining of dependencies, as in the following sentences.

- (6.7) a. A decision was made to tell the police with unanimous agreement.
 b. The marriage took place of Paul and Linda today.

These present no particular problem for the present approach.

Further support for the analysis of extraposition comes from consideration of Swedish data involving extraction from the extraposed relative clause. As (6.8) shows, extraction from complex noun phrases in Swedish is generally less problematic than in English.²

¹From Milward (1994b: 597).

²Although as shown in chapter 3, extraction from a subject relative clause in English is far from impossible, especially given a light head as in the Swedish example.

- (6.8) Per_i kjenner jeg ikke noen_j som _j liker _i .
 Peter know I not anyone who likes
 As for Peter_i, I don't know anyone who likes him_i.

Noting this, Maling and Zaenen (1982) have nevertheless no explanation for the contrast in grammaticality between the Swedish sentences (6.9)a and b, which show that extraction is only possible from the extraposed relative clause, but not from the one in canonical position.

- (6.9) a. Per_i slipper jeg ikke noen_j inn som _j liker _i .
 Peter let I not anyone in who likes
 As for Peter_i, I'm not going to let in anyone who likes him_i.
 b.???Per_i slipper jeg ikke noen_j som _j liker _i inn.

However this behaviour follows directly from the present treatment of extraposition, coupled with the condition (discussed in 3.1.0.3) that store values cannot be inherited if the active constituent is unsaturated (ie. Kuno's Clause Nonfinal Incomplete Constituent Constraint). The only way to derive (6.9)b is to lower the stored topicalized constituent *Per* into the relative clause even though *slipper* is unsaturated, having its second external argument position unfilled. In (6.9)a, however, the extraposed relative clause can inherit the store value as the constituent is now saturated. Approaches which assume that the extraposed clause is in some sense present in the first argument will not predict this difference.

The following derivation of (6.9)a shows that by the time the store value is inherited by the relative clause, at S_7 , the arguments of *slippe* have already been filled.

(1) prop ⁰ , subj: 3 , ACTION ² , pres, slippe , agent: 3 , theme: 4 , goal: 5 , neg ⁴ , [-2*]
(3) 1ST-PERSON ³ , sg
(4) HUMAN ⁵ , sg, mod ⁷ : 6 , [-2*]
(6) prop ⁷ , STATE ⁸ , pres, like , exper: 4 , theme: 2 , [+4†, -2†]
(2) MALE ¹ , sg, Per
(5) DIRECTION ⁶ , inn

(6.9)a: **Per slipper jeg ikke noen inn som liker.**

6.2.1 Obligatory extraposition

There are also constructions in English in which extraposition is not only possible, but necessary. This is typically the case for pre-nominal modifiers, whose arguments are not

allowed to appear before the head, and so must appear after it, thus making a discontinuous modifying phrase. These argument-taking pre-nominals are typically comparatives, degree-modifiers (*so* etc), and *tough*-adjectives. It is the last of these that I shall consider in this section.

Let us consider the problems to be faced in specifying a derivation of (6.10)a, and ruling out (6.10)b.

- (6.10) a. Butch was a difficult dog to train.
 b. *Butch was a difficult to train dog.

Firstly, I argued above in 4.4.1 that the infinitival phrase following a *tough*-adjective is an argument. Therefore, I will assume that the premodifier transition must obligatorily remove the constituent headed by *difficult* off the active stack, to allow the head to be added at the next word. However, if we allow the infinitival argument to be added to an unsaturated constituent which is *off* the active stack then we cannot stop the transition applying immediately it is taken off, that is before the head noun, which will permit (6.10)b.

It seems the only way around this is to assume that to receive its infinitival complement the unsaturated adjectival constituent must be on the active stack. There then must be a rule for moving this unsaturated constituent back onto the stack, and it cannot apply if the active constituent has not been given its head. The rule can be formalized in the following schema.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
reactivating		(Y, arg:)
unsaturated constituent	(X)	(X)

where X is saturated and Y was deactivated

If we apply this transition before the complementizer *to* is added at transition 6, we will allow the following derivation for (6.10)a.

- (1) prop⁰, STATE², past, **be**, content:3, [+2†]
 (2) DOG¹, sg, **Butch**
 (3) sg³, indef, ref:2, mod⁴:4, **dog**⁵
 (4) QUALITY⁴, **difficult**, \emptyset :3, theme:5, [+3*]
 (5) inf⁶, ACTION⁷, **train**, agent: α , theme:3, [+ α , -3†]

(6.10)a: **Butch was a difficult dog to train.**

The changing states of the active stack during the course of this derivation can still be reconstituted from the diagram, if we assume that subcategorized complements can only be added to active constituents.

This approach is also compatible with the following data showing the positions where the obligatorily-extraposed complement may occur. According to the transition schema given above, the unsaturated constituent *how difficult* may be replaced on the stack only when the active constituent is saturated. This is the case in (6.11)a, where the nominal phrase is saturated, and in (6.11)b, where the proposition is saturated. It is not the case, however, in (6.11)c where the active constituent proposition is unsaturated.

- (6.11) a. How difficult a dog to train was Butch?
 b. How difficult a dog was Butch to train?
 c. *How difficult a dog was to train Butch?

6.3 Discontinuous word-orders

Although, as we have seen, English does provide constructions involving discontinuous phrases, or *non-projectivity*, it is generally taken to be rather a peripheral problem. There are however languages which pose even more fundamental problems to phrase structure approaches to syntactic modelling, and aspects of three of these will be briefly reviewed here.

6.3.1 German

Reape (1994) presents the following example of a German subordinate clause, (6.12), with its typical verb-final word order. The ordering of the preverbal arguments, in what is known as the *Mittelfeld*, makes any kind of immediate constituent analysis exceedingly problematic, as each of the arguments is separated from its dependent-head by other constituents.

- (6.12) ... daß es ihm jemand zu lesen versprochen hat.
 ... that it(acc) him(dat) someone(nom) to read promised has
 ...that someone promised him to read it.

In fact, as Reape shows, the pre-verbal arguments may be arranged in any order.

- (6.13) ...daß es ihm jemand
 es jemand ihm
 ihm es jemand zu lesen versprochen hat.
 ihm jemand es
 jemand es ihm
 jemand ihm es

The problem of accounting for the syntax of examples such as (6.12) is tackled by Reape by assuming that constituent structure trees do not determine the order of the words which are their terminals. Instead he proposes that word order is determined by an additional level of representation termed a *word order domain*, which is composed compositionally from its daughter word order domains.³ Proposals such as Reape's have gained wide acceptance in HPSG circles, for example such a divorce of constituent structure from word order is assumed in Sag (1997).⁴

In the analysis that follows, I will argue that instead of increasing the grammatical machinery available by introducing an additional level of representation, a simpler solution to this problem is available which avoids the use of constituent structure altogether.

The main assumption that we have to make about German subordinate clauses is that all verbs must have their arguments in place when they are reached, that is their arguments appear to their left. I shall also assume that, like English, German has a syntactic relation of subject, +, and unmarked object, -, and in addition the relations dative, ~. I shall assume then that the state after the first three arguments have been added is as shown in the following diagram, with the addresses of all three arguments in store. This would be the state given any ordering of the arguments, modulo the position on the store, which I shall take to be irrelevant for these case-marked constituents.

- | |
|--|
| <p>(1) prop¹, [+4,~3,-2]
 (4) HUMAN⁴, indef, sg
 (3) MALE³, sg
 (2) NON-HUMAN², sg</p> |
|--|

(6.12): ... daß es ihm jemand || zu_lesen versprochen hat.

I will assume that the infinitive *zu lesen* functions as a single word in German, unlike the corresponding English *to read*, and provides an action constituent which is itself placed

³See Reape (1994, 1996) for the details of the analysis.

⁴It should be noted that this goes further than the assumption that linear precedence and immediate dominance relations should be kept separate, a proposal which goes back at least to GPSG, as this was restricted to determining the order *within* constituents. Under Reape's analysis the constituents themselves interleave with one another.

on the store, inheriting at the same time the constituent *es* which fills its *theme* argument role. I will also assume that the infinitive fills its subject position from one of the other available constituents, the most likely one being the matrix subject constituent **4**, which also turns out to be correct choice in this derivation.⁵

- | |
|---|
| <p>(1) prop¹, [-5,+4*,~3,-2*]
 (4) HUMAN⁴, indef, sg
 (5) ACTION⁵, inf, lesen, agent:4, theme:2, [+4†,-2†]
 (2) NON-HUMAN², sg
 (3) MALE³, sg</p> |
|---|

(6.12): ... **daß es ihm jemand zu_lesen** || versprochen hat.

The past-participle *versprochen* is then added to the store, inheriting the two uninterpreted items on it, the dative *ihm*, with which it fills its beneficiary slot, and the infinitival clause which becomes its theme. As this is a subject control verb, the subjects of both clauses will be constrained to be identical (as discussed in 4.2.1).

- | |
|---|
| <p>(1) prop¹, [-6,-5*,+4*,~3*,~2*]
 (4) HUMAN⁴, indef, sg
 (6) ACTION⁶, ppart, versprechen, agent:4, benef:3, theme:5, [+4†,-5†,~3†]
 (3) MALE³, sg
 (5) ACTION⁵, inf, lesen, agent:4, theme:2, [+4†,-2†]
 (2) NON-HUMAN², sg</p> |
|---|

(6.12): ... **daß es ihm jemand zu_lesen versprochen** || hat.

There is now only one uninterpreted item on the store of the main clause, the past-participle clause. At the next transition this can be interpreted as the content of the perfective auxiliary *hat*. Being a raising verb, the subject of the auxiliary is constrained to be identical to the subject of its content clause.

⁵If the infinitive does not guess a subject at this point, the placeholder α may be used until and later replaced by the main clause subject. There will be occasions where another choice has to be made. For example, in (6.14) the subject of the infinitive turns out to be the dative constituent.

- (6.14) ... daß es ihm jemand zu lesen gegeben hat.
 ... that it(acc) him(dat) someone(nom) to read given has
 ...that someone gave it to him to read.

It should be open to experimental investigation whether the main clause subject is assumed to be the subject of the infinitive in cases such as this.

- (1) prop¹, STATE⁷, pres, perf, content:6, [-6†, -5*, +4*, ~3*, -2*]
 (6) ACTION⁶, ppart, **versprechen**, agent:4, benef:3, theme:5, [+4†, -5†, ~3†]
 (4) HUMAN⁴, indef, sg
 (3) MALE³, sg
 (5) ACTION⁵, inf, **lesen**, agent:4, theme:2, [+4†, -2†]
 (2) NON-HUMAN², sg

(6.12): ... daß es ihm jemand zu lesen versprochen hat.

Reape (1994: 193-194) argues that it is problematic to conflate local “scrambling” movement with unbounded movement, a conflation which is inherent in the above approach as the store is used for both. His evidence is the following contrast.⁶

- (6.15) a. *...daß Bücher er zu lesen nur am Samstag versucht hat.
 that books he to read only on Saturday tried has
 ...that he has only tried to read books on Saturday.
- b. Bücher die er zu lesen nur am Samstag versucht hat.
 books which he to read only on Saturday tried has
 Books that he has only tried on Saturday to read.

However, as Reape himself points out, there are various poorly-understood constraints on the possible order of pre-verbal arguments, relating to focus and information structure, which may independently rule out (6.15)a, but would not necessarily affect (6.15)b. Also Reape notes that some speakers reject unbounded extractions as (6.15)b. This evidence does not appear conclusive enough, therefore, for us to abandon the simplest account, which is to treat local and unbounded movement using the same mechanism, the store.

6.3.1.1 Complex fronting

One of the most discussed areas in syntax over the past decade has been the topicalization of partial verb phrases in German, a phenomenon that is referred to in Uszkoreit (1987) as *complex fronting*.

The sentences in (6.16)⁷ all exhibit the typical constituent order of German main clauses, with an initial topic and the finite verb in second position. Sentence (6.16)b shows that this topic can be an infinitival clause without its arguments, and (6.16)c that

⁶My thanks to Gert Westermann for discussion of these and other examples.

⁷Taken from Uszkoreit (1987).

this topicalized clause may bring with it a number of arguments and adjuncts, leaving the rest in the matrix clause.⁸

- (6.16) a. *Der Brief sollte der Kurier nachher einem Spion zustecken.*
 the letter should the courier later a spy slip
 The courier was later supposed to slip a spy the note.
- b. *Zustecken sollte der Kurier den Brief nachher einem Spion.*
- c. *Nachher einem Spion zustecken sollte der Kurier den Brief.*

The problems that such a construction poses for accounts based on constituent structure are clear. Once again, syntactic constituents seem to have become intertwined with one another, with elements of a single embedded clause being spread on either side of the matrix verb. Hudson (1997) offers a Word Grammar solution to the problem and argues that many of the problems do not arise in a dependency-based approach. I will argue here that a dynamic approach, using the same mechanisms that we have already seen, offers a simple and accurate analysis of the phenomenon.

Let us take as an example the derivation of (6.16)c, which will be seen to follow directly from the approach to subordinate clauses taken above. The first two constituents are placed on the store of the proposition. The infinitive clause is next placed there, inheriting the two constituents already there and interpreting them, ie. giving them the roles of modifier and beneficiary respectively. In this instance the placeholder α must be used to fill the subject position as neither of the topicalized constituents can be interpreted as the subject. At this stage the derivation will be as follows.

- (1) prop⁰, [-4, ~3*, -2*]
 (4) ACTION⁴, inf, **zustecken**, agent: , benef:3, theme: , mod:2, [~3†, -2†]
 (3) indef², sg, masc, **Spion**³
 (2) TIME¹, **nachher**

(6.16)c: (S₄) **Nachher einem Spion zustecken** || sollte der Kurier den Brief.

The infinitival constituent is at this stage the only uninterpreted constituent on the store. The following finite verb *sollte* takes it to fill its content role, and the two subject values are made to be the same, ie. they both are the placeholder α , and the derivation is as follows.

⁸There are various restrictions on which arguments can be included in the topicalized constituents: for example, it is generally impossible for it to contain an a subject if this is an agent. I will leave this matter for further research.

- (1) prop⁰, STATE⁵, past, **sollen**, content:4, [-4†,~3*,-2*]
 (4) ACTION⁴, inf, **zustecken**, agent: , benef:3, theme: , mod:2, [~3†,-2†]
 (3) indef², sg, masc, **Spion**³
 (2) TIME¹, **nachher**

(6.16)c: (S₅) **Nachher einem Spion zustecken sollte** || der Kurier den Brief.

At the next transition the unsaturated infinitival clause must be replaced on the active stack, in a transition identical to the one given above for discontinuous *tough*-adjectival phrases in English. The remaining arguments can then be added to this clause in the normal fashion, to give the following final derivation.

- (1) prop⁰, subj:5, STATE⁵, past, **sollen**, content:4, [+5*,-4†,~3*,-2*]
 (4) ACTION⁴, inf, **zustecken**, agent:5, benef:3, theme:6, mod:2, [+5†,~3†,-2†]
 (5) def⁶, sg, masc, **Kurier**⁷
 (3) indef², sg, masc, **Spion**³
 (6) def⁸, sg, masc, **Brief**⁹
 (2) TIME¹, **nachher**

(6.16)c: **Nachher einem Spion zustecken sollte der Kurier den Brief.**

The essential difference between German and English seems to be then that in English unsaturated constituents can only move off the store if they are adjectival in nature, and not if they are verbal.

- (6.17) a. He saw the vase that had been donated to the museum.
 b. He saw the donated vase.
 c. *He saw the donated vase to the museum.

This results in the ungrammaticality of English sentences with complex fronting.

- (6.18) a. *Given the spy a letter* he certainly had _ .
 b. **Given the spy* he certainly had _ a letter.

The analysis given above also respects the fact that having two constituents in topic position in German is generally impossible, as is shown in (6.19).

- (6.19) ??*Nachher einem Spion* sollte der Kurier den Brief zustecken.

Assuming the present analysis it is natural to see this as a restriction on there only being one uninterpreted constituent in store when the finite verb is added, rather than the store in the matrix clause being in some way limited to only hold item. This is something that we should expect, given that the store in subordinate clauses can take an indefinite number of items.

The topicalized clause may even be embedded inside another, as long as the clauses are non-finite, as shown in (6.20)b.

- (6.20) a. *Peter* wird Maria das Buch geben können.
 Peter will Maria the book give can
 Peter will be able to give Maria the book.
- b. *Geben können* wird Peter Maria das Buch.

Again this is quite predictable given the present approach. The first topicalized constituent is inherited by the second. This in turn becomes the content of the main verb *wird*. The unsaturated constituent *geben* must then be replaced on the stack to have its argument slots filled.

- (1) prop⁰, STATE³, pres, **werden**, content:3, [+4*, -3†, -2*]
 (3) prop², STATE, inf, **können**, content:2, [+4*, -2†]
 (2) prop¹, ACTION, inf, **geben**, agent:4, benef:5, theme:6, [+4†]
 (4) MALE⁴, sg, **Peter**
 (5) FEMALE⁵, sg, **Maria**
 (6) def⁶, sg, neut, **Buch**⁷

(6.20)b: **Geben können** wird Peter Maria das Buch.

6.3.2 Dutch

The word-order in Dutch subordinate clauses has also posed severe problems for phrase-structure based accounts. As shown in (6.21)a,⁹ instead of the nesting structures of German clauses, caused by the constraint that all arguments must appear to the left of the verb, Dutch verbs exhibit the opposite order of verbs, as a result of verbs generally taking nominal arguments to their left and clausal arguments to their right. The resulting *cross-serial dependencies* are especially problematic in that they cannot easily be modelled by context-free grammars.

⁹The data is taken from Steedman (forthcoming: 153).

- (6.21) a. ...omdat ik Jan het lied probeer te leren (*te) zingen.
 ...that I John the song try to teach (*to) sing
 ...because I try to teach John to sing the song.
- b. ...omdat ik probeer Jan het lied te leren (*te) zingen.

The derivation of these sentences differs only minimally from that of the German subordinate clauses in the preceding section. If we take (6.21)a, the first three constituents are added to the proposition, one as subject and the other two in store. The finite verb *probeer* interprets the subject as its first *agent* argument. The following infinitive *te leren* fills the theme role of *probeeren* and inherits the stored constituents, giving *Jan* the role of beneficiary. The final infinitive *zingen* fills the theme role of *leren*, inheriting the final uninterpreted constituent from the store, which it makes its theme.

- (1) REASON¹, **omdat**, theme:2
 (2) prop¹, ACTION⁶, pres, **probeeren**, agent:3, theme:6, [-5*, -4*, +3†]
 (3) 1ST-PERSON², sg
 (6) prop⁷, inf, ACTION⁸, **leren**, agent:3, benef:4, theme:7, [-5*, -4†, +3†]
 (4) MALE³, sg, **Jan**
 (7) prop⁸, inf, ACTION⁹, **zingen**, agent:4, theme:5, [-5†, +4†]
 (5) def⁴, sg, **lied**⁵

(6.21)a: ...omdat ik Jan het lied probeer te leren zingen.

It appears from this analysis that the difference between Dutch and English, is that the former allows much freer use to be made of the store. Not only can more than one item easily be placed in topic position, they can also appear after the subject, and, as shown in (6.21)b, even after the head verb has been added.¹⁰

A further problem to be accounted for is presented in Steedman (forthcoming: 153), and shown in the sentences in (6.22).¹¹ Sentences (6.22)a and (6.22)b show that if the verbal head is separated from its infinitival argument then the infinitival complementizer *te* is both possible and necessary. While if the infinitival argument is adjacent, as shown in (6.21)a and (6.21)b above, then the complementizer is disallowed. As Steedman notes such an effect is “likely to create complications under any theory” and he finds no alternative but to model it without explanation by introducing the appropriately-named *Te Category Brute Force Stipulation*.

¹⁰It might be speculated that such operations are not allowed in English due to the fact that any topicalized elements must be there for some reason, either because they are *wh*-items or they are the topic of the sentence. It appears to be possible to place items on store in Dutch and German without making an explicit statement about their information packaging status.

¹¹Quoted by Steedman, and originally due to Seuren.

- (6.22) a. ...omdat ik probeer Jan te leren het lied te zingen.
 b. ...omdat ik Jan probeer te leren het lied te zingen.

It will be seen that it has been assumed in the derivation of (6.21)a above that the Dutch infinitival complementizer *te* functions in much the same way as the English *to*, initiating a proposition and filling the subject slot with an appropriate constituent. In the derivation of (6.21)a above, I have assumed that there is a transition in Dutch which in a way anticipates the function of the complementizer, installing a proposition in the theme role of *leren* before the verb *zingen* is reached. This is a non-finite version of the “switching” operation which allows subject extractions in English. An explicit complementizer in such sentences would therefore be superfluous as the proposition has already been formed, exactly the same situation as in the *that*-trace effect, discussed in chapter 3.

The impossibility of omitting the complementizer in the sentences in (6.22) follows automatically—if the new propositional argument had been formed at *leren* it would be on top of the active stack and the following constituent *het lied* could not be added to the store of the *leren* constituent. The infinitival argument must therefore be initiated by an explicit infinitival complementizer *te*, as shown in the following derivation.

- (1) REASON¹, **omdat**, theme:2
 (2) prop¹, subj:3, ACTION³, pres, **probeeren**, agent:3, theme:5, [-4*,+3†]
 (3) 1ST-PERSON², sg
 (5) prop⁵, inf, ACTION⁶, **leren**, agent:3, benef:4, theme:7, [-6*,-4†,+3†]
 (4) MALE⁴, sg, **Jan**
 (7) prop⁹, inf, ACTION¹⁰, **zingen**, agent:4, theme:6, [-6†,+4†]
 (6) def⁷, sg, **lied**⁸

(6.21)a: ...omdat ik probeer Jan te leren het lied te zingen.

To summarize, if we assume that Dutch *te* has the same function as an English complementizer initiating a clause and fixing an appropriate subject, and if we assume that the existence of constructions where this complementizer is missing, as in (6.21) above, show that for some verbs an *early switching transition* is available to start this clause, then the pattern of data is predictable, and the explanation has much in common with the *that*-trace effect discussed in section 3.1.3.

6.3.3 Hungarian

In languages with a highly-developed and explicit case system, the order of constituents has typically less of a role in syntax and often expresses discourse functions such as focus.

In this section I will take a brief look at word order in Hungarian, which is a prime example of this type, and explore how a focus-based word order might be approached using the dynamic grammar.

In Kiss (1987), the three word Hungarian sentence “János loves Mária”, plus an obligatory sentence stress, is shown to have 10 meaningful orderings out of the 18 (= 3!.3) logically possible permutations. The sentence structure can be characterized by identifying four distinct positions in the clause: (I) an indefinite number of topics, (II) a single focus position¹² (stressed, but not necessarily filled), (III) the finite verb (stressed if position (II) is empty) and (IV) any other constituents. The permitted sentences are set out in the following table.

I	II	III	IV	meaning
		'szereti	János Máriát	<i>János loves Mária.</i>
		'szereti	Máriát János	<i>János loves Mária.</i>
János		'szereti	Máriát	<i>As for János, he loves Mária.</i>
Máriát		'szereti	János	<i>As for Mária, János loves her.</i>
	'János	szereti	Máriát	<i>It is János who loves Mária.</i>
	'Máriát	szereti	János	<i>It is Mária who János loves.</i>
János	'Máriát	szereti		<i>As for János, it is Mária who he loves.</i>
Máriát	'János	szereti		<i>As for Mária, it is János who loves her.</i>
János Máriát		'szereti		<i>As for János and Mária, he loves her.</i>
Máriát János		'szereti		<i>As for Mária and János, he loves her.</i>

The challenge for phrase structure accounts is not so much to model the data, for this can be done using multiple movement analyses to various topic and focus positions on an extended phrase structure tree. But it is perhaps difficult to motivate the necessity for the complex constituent structures that result. The claim I will make here is that the dynamic model allows us to model all of these possible sentences with the minimum of apparatus.

Let us look at the derivation of the seventh sentence in this list, shown in (6.23), as this includes both a topicalized and a focussed constituent.

(6.23) János 'Máriát szereti.

The topics are placed on the store of the proposition as is the focussed element. Presumably the representation of focus should include the necessary information about the

¹²This is also the position in which incorporated objects and the verbal prefixes appear, but I will not discuss these here.

exclusivity of the focussed element and other elements, but the sake of simplicity I will mark this with a single *focus* feature on the proposition.

It is sufficient to state that there can only be one focussed constituent and that no topics can be added after this position is filled to fully characterize the ten possible sentences given above. The finite verb interprets the constituents in store according to their case marking.

- (1) prop⁰, focus²:**3**, STATE³, pres, **szeret**, exper:**2**, theme:**3**, [-3†, +2†]
 (2) MALE¹, sg, **János**
 (3) FEMALE², sg, **Mária**

(6.23): **János 'Máriát szereti.**

It will be seen that all the permissible sentences will receive the same basic interpretation, modulo the differences in focus and topic.

As in English the topicalized constituent may end up being interpreted in an embedded clause, as seen in (6.24)a. As expected there is no *that*-trace effect when subjects are extracted as Hungarian is pro-drop. Indeed, in general the finite complementizer *hogy* cannot be omitted, as shown in (6.24)b.

- (6.24) a. János hiszem, hogy jön.
 János I-think that come
 János, I think (*that) will come.
 b. *János hiszem, jön.

The derivation of (6.24)a will then be as follows.

- (1) prop⁰, STATE², pres, **hisz**, exper:**3**, theme:**4**, [+3*, +2*]
 (3) 1ST-PERSON¹, sg
 (4) prop³, EVENT⁴, pres, **jön**, theme:**2**, [+2†]
 (2) MALE¹, sg, **János**

(6.24)a: **János 'hiszem, hogy jön**

It is also possible for a constituent in focus position to be interpreted in a lower clause, a construction referred to as *long focusing*. Surprisingly, it is also possible, even preferable, for such a focussed constituent to have accusative case, even when its role in the embedded clause would normally require it to be nominative.

- (6.25) a. 'Jánost hiszem, hogy jön.
 János-acc I-think that come
 It is János that I think (*that) will come.
- b. ?'János hiszem, hogy jön.

This construction appears to support the present analysis which avoids having syntactic case features on the constituents themselves, but rather treats them as syntactic relations.

A final aspect of Hungarian I will consider here is the fact that verbs show both subject and object agreement.¹³ In a similar fashion to the subject raising constructions of English, considered in Chapter 4, Hungarian allows simultaneous subject and object raising as in (6.26).

- (6.26) Foglak látni.
 will-1sg-2sg see-inf
 I will see you.

As there is no requirement for an explicit subject in Hungarian, there is no need for empty argument positions in the lexical entries of raising verbs, and also no need for dummy pronouns. The derivation of (6.26) requires only that *foglak* does not interpret its syntactic arguments (it cannot do so in any case as they are of the wrong type), but instead places them in store. They are then inherited by its infinitival argument, where they receive their interpretation.

- | |
|---|
| <p>(1) prop⁰, STATE¹, pres, fog, content:4, [-3*,+2*]
 (4) EVENT², lát, exper:2, theme: 3, [-3†,+2†]
 (2) 1ST-PERSON¹, sg
 (3) 2ND-PERSON¹, sg</p> |
|---|

(4.1): **Foglak látni.**

A similar approach can be taken to tackle the phenomenon of free word order in Finnish, which is considered by Dowty (1989).¹⁴ The following sentence exhibits a continuous verb phrase *play these-in tennis*.

- (6.27) En minä ole aikonut ruveta pelaamaan näissä tennistä.
 not I have intend start play these-in tennis
 I did not intend to start to play tennis in these.

¹³Apart from the special *I-thee* form considered here, however, object agreement is limited to an indication of the definiteness or indefiniteness of the object.

¹⁴The examples coming originally from Karttunen.

However, the arguments of the embedded clause can with a great deal of freedom be scrambled out of this clause giving rise to sentences such as the following in (6.28), which are identical in meaning to (6.27).

- (6.28) a. En minä **näissä** ole **tennistä** aikonut ruveta **pelaamaan**.
 not I these-in have tennis intend start play
- b. En minä **tennistä** **näissä** ole aikonut ruveta **pelaamaan**.
 not I tennis these-in have intend start play
- c. En minä **tennistä** ole aikonut **näissä** ruveta **pelaamaan**.
 not I tennis have intend these-in start play

If we assume that the scrambled constituents are placed on the store of the active constituent when they occur, and are then inherited down to the embedded proposition in which they are interpreted, then such sentences can be derived in the following manner, which parallels the analysis of the other languages with relatively free word order that we have considered in this chapter.

- | |
|---|
| <p>(1) prop⁰, neg¹, STATE⁴, pres, perf, content:6, [-5*, -3*, +2*]
 (6) STATE⁵, intend, exper:2, theme:7, [-5*, -3*, +2†]
 (2) 1ST-PERSON², sg
 (7) EVENT⁶, start, theme:8, [-5*, -3*]
 (8) ACTION⁷, play, agent:2, theme:5, manner:3, [-5†, -3†]
 (5) sg⁵, tennis
 (3) LOCATION³, loc:4
 (4) plur³, def, proximate</p> |
|---|

(6.28)a: not I these-in have tennis intend start play

6.4 Parenthetical interruptions

Any attempt at a syntactic analysis of natural speech is soon likely to run into the problem that people often seem to interrupt what they are saying to make comments on it. This has been a largely ignored area in syntax, for much of the data we have to thank research made in the early 1970's, especially the wonderfully detailed work of Ross, to whom we also owe the colourful terms with which it is described.

It has been assumed that these common parenthetical constructions pose severe problems for constituent structure approaches. For example, McCawley (1982) argues that sentences such as (6.29) are genuinely discontinuous and should be dealt with in terms of crossing syntax trees.

(6.29) John talked, of course, about politics.

I will attempt to show here that such constructions submit to a neat analysis and simple in the present dynamic framework.

6.4.1 Slifting

The term *slifting*, introduced and described in Ross (1972), refers to a distinctive construction, illustrated by the sentences in (6.30), in which a proposition is interrupted by what seems to be a superordinate proposition, taking the original proposition as a missing argument.

- (6.30) a. John, it seems __ , gives bones to dogs.
 b. John gives, it seems __ , bones to dogs.
 c. John gives bones, it seems __ , to dogs.
 d. John gives bones to dogs, it seems __ .

Ross (1972) uses a thorough array of tests to show convincingly that the apparent gap in the superordinate sentence has the same characteristics as that of other extractions. He also shows convincingly that the superordinate element cannot be treated as some sort of modifier of the original clause and that the underlying interpretation of the sentences must be taken to be the same as that of the non-slifed sentence (6.31).

(6.31) It seems (that) John gives bones to dogs.

The obvious way to model such a construction in the dynamic model is to have a transition introduce a new proposition with the address of the existing active constituent on its store. of the following sort. Here any proposition may be interrupted by placing a new proposition on the active stack with the address of the original proposition on its store, marked as a non-subject.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
slifting	(X, prop)	(N, prop, [-X]) (X, prop)

When the superordinate clause gives an interpretation to the address of the original proposition on its store and becomes saturated, this clause can be removed from the active stack and the derivation of the interrupted original proposition proceeds as normal. It will be seen that the typical dependency relations on the stack are here reversed, with the

superordinate on top of the subordinate, but although unusual this does not give rise to any problems.

Let us take (6.30)c as a concrete example. The first three transitions are as normal. The fourth transition is the slifting transition introducing the superordinate clause. The address of the original proposition is consequently interpreted as the theme of *seems*, the superordinate proposition is removed from the active stack and the derivation of the original proposition continues.

- (4) prop⁴, STATE⁵, pres, **seem**, \emptyset :5, theme:1, [+5†, -1†]
 (5) sg⁴
 (1) prop⁰, ACTION², pres, **give**, agent:2, theme:3, benef:6, [+2†]
 (2) MALE¹, sg, **John**
 (3) bone³, plur
 (6) to⁶, content:7
 (7) dog⁷, plur

(6.30)c: **John gives bones, it seems, to dogs.**

That there is no limit on the depth of extraction is shown by (6.32).

- (6.32) John gives bones, someone is once supposed to have have heard Peter observe __
 , to dogs.

The rule explicitly stipulates that the superordinate proposition must have an object extracted not a subject and this corresponds to the data.¹⁵

- (6.34) a. *John gives bones, __ is true, to dogs.
 b. John gives bones, it's true __ , to dogs.

When there is more than the original proposition on the active stack, “slifting down” to this bottom proposition is much more difficult though perhaps marginally possible, as shown in (6.35). Allowing such sentences would require another constituent to be permitted over the original proposition in the transition schema given above.

¹⁵Subject extraction from an embedded clause in the superordinate proposition should be possible given the above syntactic rule, however. In practise, it is hard to come up with an example that sounds impeccable. The best I can think of is (6.33).

(6.33) ??John gives bones to dogs, he thought __ might be the case.

- (6.35) a. ??He thought John gave, it seems, bones to dogs.
 b. ??John was right underneath, it seems, the table.

It should be clear from the discussion so far that the interruption of the matrix clause is quite different from the interruption caused by inserting a modifier before an argument in a heavy-constituent shift construction. In the heavy-shift case we have assumed that the placeholder α fills the right-shifted argument. This difference is demonstrated by the fact that in the heavy-shifted case, the modifier may include a gap or a pronoun coreferenced with this placeholder and hence with the heavy-shifted constituent as shown in (6.36). For the heavy-shift case both the gap, allowed as an instance of right-node raising as discussed in 5.2, and the coreferenced pronoun are available.

- (6.36) John insulted, by not recognizing $_i$ /him_{*i*}, his uncle_{*i*} from Philadelphia.

The same possibilities are not available in the interrupting clause in the slifting construction, as the examples in (6.37) show. This is accounted for by the fact that the original clause here is genuinely interrupted, and the material to the right of the interruption is not “right-shifted” in any way.

- (6.37) a. *John insulted, it appeared $_$ to $_i$, his uncle_{*i*} from Philadelphia.
 b. ?*John insulted, it appeared $_$ to him_{*i*}, his uncle_{*i*} from Philadelphia.
 c. John insulted, it appeared $_$ to Mary, his uncle from Philadelphia.

The approach also agrees with the following data, which show that the restriction on the clause-boundedness of extraposition movement is not compromised by the slifted expression, (6.38)a, whereas it is in topicalized verb phrase constructions such as (6.38)b.

- (6.38) a. The marriage took place today, it was reported, of Paul and Linda.
 b. *Attend the marriage I certainly did, of Paul and Linda.

The reason is that in the topicalized expression the verb phrase must be taken off the active stack when it is finished in order to allow the rest of the sentence to be derived, and so when the modifying phrase is reached, the constituent *the marriage* is not a dependent of the active constituent *did*. This is not true in the slifting example, (6.38)a, where the original proposition remains on the stack throughout the derivation.

Ross (1972: 136) argues that multiple embedded slifting is possible, as in (6.39).

- (6.39) Frogs have souls, Osbert feels, I realise.

It is not clear whether or not we would really wish the grammar to derive a meaning for this sentence. However, the above approach will allow it to be derived since the first superordinate clause will still be on the stack as the second is reached and could therefore be subject to the same slifting transition. It would not, however, allow embedded superordinates where the two superordinate clauses were not adjacent, as in (6.40).

(6.40) ??Frogs, Osbert feels, have souls, I realise.

I have no clear intuitions about what, if anything, such a sentence might mean, or whether it is any better or worse than the previous one.

Finally, the following example (6.41)a, adapted from Ross (1972), shows that the transition schema given above to introduce slifting is correct in not stipulating that the interrupted clause should necessarily be a matrix clause, for here slifting takes place in the subordinate clause. This predicts that a reading is possible with the same interpretation as (6.41)b.

- (6.41) a. They said that there would, they felt, be enough helmets.
 b. They said that they felt there would be enough helmets.

6.4.2 Parenthetical coordination

The parenthetical coordination construction is akin to slifting, involving a parenthetical interruption commenting on the original clause, though it has important syntactic differences. The data are again taken from (Ross, 1972).¹⁶

- (6.43) a. Slim, and even his mother doesn't know this, is proud of his wrists. (= Ross, 1972: ex. 90a)
 b. Peter, and Jane should know this, being his wife, doesn't like pickled eggs.

The transition responsible is a simple variant of that given for basic coordination in Chapter 5. The condition that the original clause be unsaturated will stop simple propositional coordinations having two identical derivations.

¹⁶It is also mentioned in Postal (1971: 195), where it is referred to as *conjunction insertion*. Postal assumes that the inserted conjunct must begin with a noun-phrase co-referential to the first clause subject. However, (6.43)b would appear to show that is not the case. There does appear to be a constraint that the coordinated clause refers in some way to the clause it interrupts, but this is not constrained syntactically (as slifting is). Although the inserted clause often contains a pronoun coreferential to the matrix clause, this is not necessary as shown by the following example.

- (6.42) I, and I am not loquacious, couldn't prevent myself from speaking. (= Postal, 1971: ex. 18.(15)b).

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ conjunction, type }	(X, prop)	(N, prop) (X, prop, coord: { type , ⟨N⟩})

where X is unsaturated.

It is important to note that the proposition introduced by this rule cannot be any reduction of the original proposition, but must be an absolute reduction of it, that is an empty proposition. This is supported by the following contrast.

- (6.44) a. Peter is handsome and (he) knows it.
 b. *Peter, and knows it, is handsome.
 c. Peter, and he knows it, is handsome.

The derivation of (6.43) is given below. To make it more compact, pronouns coreferenced with other constituents in the derivation are not shown as full constituents.¹⁷

(1) prop ⁰ , coord ² : {&, ⟨3⟩}, STATE ⁹ , pres, be , content:6, [+2*]
(6) QUALITY ¹⁰ , proud , exper:2, theme:7, [+2†]
(2) MALE ¹ , sg, Slim
(7) of ¹¹ , content:8
(8) possess ¹² :⟨ref:2⟩, wrist ¹³ , plur
(3) prop ² , pres ⁶ , neg, content:5, [+4*]
(5) STATE ⁷ , know , exper:4, theme:⟨ref:1⟩, [+4†]
(4) possess ⁴ :⟨ref:2⟩, FEMALE ⁵ , sg, mother

(6.43)a: **Slim, and even his mother doesn't know this, is proud of his wrists.**

There is a clear parallel between this construction and the parenthetical placement of subordinating clauses and other modifiers.

- (6.45) a. Peter thinks, although he daren't say it, that Mike is doomed.
 b. Mike thinks, of course, that Peter is too timid.

6.4.2.1 Parenthetical parataxis

There can even be parenthetical interruptions in which the inserted material has no syntactic connection with the original clause, but is a completely independent sentence. Such constructions, illustrated by (6.46), are referred to as *parenthetical parataxis* by Quirk *et al.* (1985: 977).

¹⁷And the word *even* is ignored.

(6.46) I thought Joe (what a great man he is!) might like to read my thesis.

This will be allowed if we introduce the following transition rule, which simply introduces a new proposition with no connection to the first.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
parenthetical parataxis		(N, prop)
	(X, prop)	(X, prop)

Of course, in a practical system one would have to make sure that this transition is not overused, perhaps by making it a transition of very low weight. It may also be identified by the fact that in written language it is typically bracketed off with parentheses or commas, while in speech there is a characteristic break in the intonation introducing and finishing the interruption.

6.5 Summary

In this chapter I have examined a number of phenomena that have been classified as involving discontinuous constituents. The claim is made that the present model is able to avoid many of the problems that these constructions present to constituency-based approaches, because it does not assume syntactic constituents and it does not assume that conceptual constituents must be formed by words which form an unbroken span.

In the first section, I proposed that the non-adjacent modification of constituents (ie. optional extraposition) may be handled by changing the rule for introducing post-modifiers to allow the modification of constituents that were dependents of the active constituent. This is similar to the case of non-adjacent complementation (obligatory extraposition), where the deactivated unsaturated constituent must be returned to the stack before it can receive its arguments.

In the next section, I analysed problems caused by the scrambling of constituents in German and Dutch. I argued that a left-to-right derivation of semantic structures, coupled with liberal use of the store, allows clear and simple characterizations of the constructions that have been most problematic for accounts based on phrase structure, without complicating the model by introducing extra assumptions. I argued that a similar approach to Hungarian could account for word order patterns in this “discourse configurational language” with the greatest economy.

Finally, I turned to some parenthetical constructions, which have received scant attention in recent generative approaches, and proposed simple transition schemas which will introduce them and account for a range of data.

As Pullum (1982b: 209) notes with regard to arguments that languages could be divided into “configurational” and “non-configurational”.

“A proposal that represents languages as being radically differentiated from one another with respect to such a fundamental property as the possession of constituent structure is too pessimistic (i.e. too weak) to be adopted in the absence of compelling justification. A universal linguistic theory should aim to treat superficially different languages in terms of exactly the same theoretical constructs, bringing out underlying similarities in syntactic structure.”

This is surely so. But in this chapter I have argued that another way to bring out the similarities between languages is not to adapt or complicate the concept of syntactic structure, but rather abandon it. What we are left with is that all languages allow the incremental construction of conceptual structures, but that they allow this to be done in different ways.

Chapter 7

Clefts

In this chapter I will consider a family of constructions that serve to rearrange the order of constituents in a sentence in order to bring one of them into focus, the cleft constructions. These pose a number of difficulties for generative approaches and I will explore how they might be tackled in the dynamic grammar.

7.1 *It*-clefts

The first construction to be examined is the *it*-cleft, illustrated by (7.1)a, which has the same propositional content as (7.1)b, but brings the constituent *onions* into focus (conveying at the same time exclusivity).

- (7.1) a. It was onions (that) he liked _ best.
b. He liked onions best.

The finite clause after the focussed constituent contains a gap and looks very much like a relative clause. This observation is supported by the fact that it may also be introduced by *wh*-relativizers as shown in (7.2).

- (7.2) a. It was Peter who he liked _ best.
b. It was Peter whose father got elected.
c. It was the onions which he liked _ best.

If the final clause is indeed a relative clause, as the evidence indicates, the question to be answered is what is its antecedent, that is which constituent is it restricting? It seems that it cannot be the focussed constituent. Firstly because a restrictive relative can follow

constituents, such as proper names that cannot normally be restricted, as in (7.3)a. And secondly because it would then be hard to explain the clear ambiguity of sentences such as (7.3)b, in the non-cleft reading of which the relative clause clearly is restricting the constituent immediately before it.

- (7.3) a. It was Peter he liked best.
 b. *He saw Peter he liked best.
 c. It was the dog he liked best.

The only alternative is that the relative clause is placing a restriction on the initial dummy pronoun *it*. This means that the construction of (7.1)a is essentially the same as that of (7.4), except that in the former there is obligatory extraposition of the relative clause.

- (7.4) The thing that he liked _ best was onions.

Before I attempt to work out the transitions for the derivation of *it*-clefts, it is necessary to consider the interpretation of a non-clefted copular such as (7.5).

- (7.5) John was the winner.

I will assume that the copular does not assign thematic roles, and that the interpretation essentially involves the two arguments coming to share their referential index, giving the following derivation.

- | |
|--|
| <p>(1) prop⁰, STATE², past, be, content:3, [+2*]
 (3) def³, ref:2, winner⁴, sg, [+2†]
 (2) MALE¹, sg, John</p> |
|--|

(7.5): **John was the winner.**

The crucial transition here is the third one. The predicate fills the content slot of the proposition and is coreferenced at the same time with the subject of the proposition.¹ The relevant transition schema may be represented as follows.

¹Again it must be admitted that the notion of reference in the present model is crude and vague, and needs to be investigated further. For present purposes, however, I will assume that such ill-defined notions as "sharing reference" are sufficient to investigate the syntax of the construction.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
add-predicate	(X, be , content: , [+Y])	(N, ref:Y) (X, be , content:N, [+Y†])

Bearing in mind this approach to copular constructions, let us now take a look at the derivation of the *it*-cleft (7.1). The situation after the third transition will be as follows.

- | |
|---|
| <p>(1) prop⁰, STATE², past, be, content:3, [+2*]
 (3) onion³, plur, [+2†]
 (2) sg¹, ref:3</p> |
|---|

(7.1): **It was onions** || that he liked best.

The following relative clause can now restrict the empty constituent in subject position. The fact that this constituent is coreferenced with the predicate *onions* ensures that the final interpretation expresses the same content as does the sentence *He liked onions best*. This coreference will also explain why it is possible to use explicit relativizers, such as *who* and *whose*, which constrain the semantic type of the antecedent.

- | |
|--|
| <p>(1) prop⁰, STATE², past, be, content:3, [+2*]
 (3) onion³, plur, [+2†]
 (2) sg¹, ref:3, restr⁴:4
 (4) prop⁴, STATE⁶, past, like, exper:5, theme:2, mod⁷:6, [+5†,-2†]
 (5) MALE⁵, sg
 (6) MANNER⁷, superlative, good</p> |
|--|

(7.1): **It was onions that he liked best.**

One concern which this analysis might raise is that there is a clash in number agreement between the two coreferential items, since this is generally not possible.²

(7.6) *John was the winners.

One possible solution to this would be to conjecture that the dummy pronoun *it* does not even contain the information that its number is singular, and that the reason it shows agreement with a singular verb is by default, in the same way that a clause or a prepositional phrase does.

²Except in cases where there are singular expressions referring to groups of individuals and so on.

This assumption is also necessary in the derivation of an *it*-cleft with a subject relative clause as in the sentences in (7.7), where agreement appears to be with the focussed constituent *dogs*, rather than with the dummy *it*.

- (7.7) a. It is the dogs that _ bark.
 b. *It is the dogs that _ barks.

The derivation follows that of the object-relative cleft above, but crucially agreement in the relative clause is with a plural subject.

- (1) prop⁰, STATE², pres, **be**, content:3, [+2*]
 (3) def³, **dog**⁴, plur, [+2†]
 (2) ref¹:3, restr⁵:4
 (4) prop⁵, ACTION⁶, pres, **bark**, agent: 2, [+2*]

(7.1): **It is the dogs that bark.**

This derivation will go through upon the assumption that coreferencing allows the dummy element to act as though it had the number, as well as the semantic type, of the coreferential constituent.

A little reflection on the initial two transitions in the derivations of *it*-clefts given above will reveal that they are the same as in an extraposed sentence with a dummy subject as in (7.8).

- (7.8) It was obvious that he was drunk.

This analysis is confirmed by the coordination data, presented in (7.9).

- (7.9) a. It was ⟨obvious that he was drunk⟩, but ⟨only Peter that _ noticed it⟩.
 b. *Obviously he was drunk, but only Peter that _ noticed it.
 c. ?It was ⟨Peter that _ noticed he was drunk⟩, but ⟨obvious to everyone that he had no trousers on⟩.

Once again these constructions will prove problematic for approaches to coordination based on the assumption that “like constituents coordinate”.

The focussed constituent can have the semantic type of a modifier and the gap in the relative clause will therefore, due to the coreferencing, be of that type.

- (7.10) a. It was on the bus that she fed the baby __ .
 b. It's tomorrow he's leaving __ .
 c. It's because of me he left __ .

However these clauses cannot in general be introduced by the appropriate explicit relativizers. This would appear to suggest that the transition initiating relative clauses is distinct from the standard one.

- (7.11) a. ?*It was on the bus where she fed the baby __ .
 b. ?*It's tomorrow when he's leaving __ .
 c. *It's because of me why he left __ .

Further evidence for the distinctness of the clefting relative is the fact that subject relatives in clefted constructions quite often lose their initiating relativizer, which is generally not possible in normal relative clauses.³

- (7.12) %It was Peter __ wanted the coffee, not me.

Interestingly, in inverted constructions this phenomenon appears to become completely acceptable.

- (7.13) a. Do you know who it was __ wanted the coffee?
 b. Who was it __ wanted the coffee?

It might be speculated that in the last case (7.13)b, the subject-gapped relative clause is formed at the word *it*, but that does not explain why (7.13)a is also so good.⁴

The focussed constituent may also be topicalized, as in (7.15).

- (7.15) Susan it was who didn't like dogs.

This is predictable, given that the focussed element is a standard argument, and the derivation is quite straightforward.

³There are of course dialects of English where this does happen on a regular basis, but the same phenomenon in *it*-cleft environments is much more widespread.

⁴To complicate matters, the non-interrogative sentence (7.14), otherwise identical to (7.13)a, sounds as marked with such a contact clause as (7.12) does above.

- (7.14) %Peter it was __ wanted the coffee.

So a simple structural explanation of why the contact clause should be so good in (7.13)a would appear to be difficult to formulate. I leave this question to further research.

- (1) prop⁰, STATE², past, **be**, content:2, [+3*, -2†]
 (2) FEMALE¹, sg, **Susan**, [+3†]
 (3) ref:2, restr:4
 (4) prop⁴, past⁵, neg, content:5, [+3*]
 (5) STATE⁶, **like**, exper:3, theme: 6, [+3†]
 (6) **dog**⁷, plur

(7.15): **Susan it was who didn't like dogs.**

It is predictable given the analysis that it is possible to have *it*-clefts where the focussed item is a *wh*-constituent. This is shown in (7.16) by examples of *wh*-constituents in interrogative, embedded interrogative and free relative constructions respectively.

- (7.16) a. Who was it who was dancing?
 b. He wondered what it might be that she wanted.
 c. He liked whatever it was she liked.

It should be noted however that this is impossible to do with “relative pronouns”. This is support for the conjecture made previously in Chapter 3 that they do not create a separate constituent, but merely initiate a relative clause.

- (7.17) a. *He didn't like the dog_{*i*} which it_{*j*} was _{*i*} that she had fed _{*j*} .
 b. *He didn't trust Mary_{*i*}, who it_{*j*} was _{*i*} that his friend wanted to marry _{*j*} .
 c. *He didn't trust Mary_{*i*}, which it_{*j*} was _{*i*} that _{*j*} shocked his friend.

That this is not due to the constraints of information packaging is shown by (7.18) in which the subject of the relative clause does have its own constituent and where an *it*-cleft is quite possible.

- (7.18) He didn't trust John, whose dog_{*i*} it_{*j*} was _{*i*} that _{*j*} had bit him.

This data, together with the fact that other *wh*-constituents do allow themselves to be *it*-clefted, seems to allow a clear case to be made for treating relative pronouns quite differently from *wh*-pronouns.

The *it*-cleft data often lead one to think that instead of having the dummy constituent, coreferenced as it is with the focus, as the gap in the relative, it would be more straightforward to have the address of the focussed constituent itself as the gap. However, there

is some interesting data from Ghanaian English⁵, which would argue against such an analysis.

In contrast to standard international English, Ghanaian English makes wide use of resumptive pronouns. They standardly fill the gap sites in *wh* and relative clause constructions, as shown in the sentences below. However, in the relative clause of *it*-clefts they are strictly forbidden.

- (7.19) a. The other soldier_{*i*}, that I hadn't seen him_{*i*}/ ? _{*i*} before, came up to us.
 b. Which soldier_{*i*} did you think that he_{*i*} would be there?
 c. ?Which soldier did you think would be there?
 d. It was John_{*i*} that they put the coat on **him_{*i*}/∅_{*i*}.

This difference can be explained as the constituent on store in (7.19)d is not *John* it is the dummy pronoun, which cannot be referred to with the pronoun *him*. It appears that this would not be predictable were the address on store that of the focussed constituent itself.

7.2 Pseudoclefts

The pseudocleft construction is illustrated by (7.20)a. It is a copular construction, where the pre-copular *wh*-phrase appears to be, and I will argue is, a free-relative. Following the usage of Heycock & Kroch (1998), I will refer to the post-copular portion of the sentence as the *counterweight*. The effect of the sentence is to focus the material in the counterweight, but the overall propositional content is equivalent to that of (7.20)c.

- (7.20) a. What he didn't buy is any eggs.
 b. *What he didn't buy is no eggs.
 c. He didn't buy any eggs.

In many respects the counterweight appears to behave exactly as it would do if actually present in the free relative, as evidenced by (7.20)a and b, and the construction is said to show *connectivity effects*.

As argued in Heycock & Kroch (1998), from which most of the revealing data here originates, pseudocleft constructions set serious challenges for any model of syntax.

⁵Quoted in this context by Cotel (1997), and taken originally from Trudgill & Hanna (1982)

“...pseudoclefts provide a real opportunity for rethinking some of our fundamental assumptions about syntactic representations: they show that the representations over which we state syntactic constraints must be much more abstract than we have concluded on the basis of simpler cases.” Heycock & Kroch (1998)

The conclusion is that constraints apply not at constituent structure, but some deeper level of LF. This is consistent with the approach taken here, where there is no surface structure and any constraints must be expressed at the one and only level of conceptual structure.⁶

Pseudoclefts also occur with verbal and adjectival counterweights, and here again connectivity effects are evident as witnessed by the possibility of the reflexives in (7.21).

- (7.21) a. What he did next was shave himself.
b. What he is is proud of himself.

With nominal predicates, they are often ambiguous and very easy to confuse with other constructions. For example, (7.22)a is an interrogative subject clause, a construction that does not exhibit connectivity effects as shown by (7.22)b and which is not synonymous with the non-clefted clause (7.22)c. This is to be expected given the analysis of such constructions in 3.2.4.

- (7.22) a. What he didn't buy is no interest of mine.
b. *What he didn't buy is any interest of mine.
c. He didn't buy no interest of mine.

The sentence (7.23)a does have a free-relative subject, but it is not a pseudocleft as the construction is not an equative one, and is not equivalent in meaning to the non-clefted variant (7.23)b.

- (7.23) a. What he bought is an embarrassment to everyone.
b. He bought an embarrassment to everyone.

Further differences from the free relative construction are that the pseudocleft is restricted to *what* and that no constraints are placed on its semantic type.

- (7.24) a. What he wanted was Peter.
b. ??Who he wanted was Peter.

⁶Although whether a level of conceptual structure deserves to be called “more abstract” than a hypothetical syntactic structure seems to be debatable.

The initial transition for *what* must therefore be different from a standard free relative or we would get pseudoclefts with the other free relative *wh*-items such as *whatever*, which is not the case.

(7.25) *Whatever he didn't buy was any eggs.

The initial transition for a free relative will result in the following first state.

<p>(1) prop⁰, [2] (2) NON-HUMAN, restr¹:3 (3) prop¹, [2]</p>

Whatever...

Let us assume that the first transition in (7.20)a differs from a free relative transition not only in the absence of semantic restrictions, but also in the fact that it places on the store of the relative clause the referent of the item to which it will refer. The resulting first state will then be the following.

<p>(1) prop⁰, [2] (2) restr¹:3, ref:α (3) prop¹, [α]</p>

(7.20)a: (S₁) **What** || he didn't buy is any eggs.

I will argue that all the properties of the pseudocleft can be predicted from the availability of the following transition for the single word *what*, together with the approach to copular constructions as set out in the previous section.

Given this initial transition the derivation of (7.20)a proceeds as normal, with the placeholder α filling the empty theme role of *buy*. When the predicate is added, the equative construction demands that the empty reference of the subject is replaced by the reference of the predicate, resulting in the following state.

<p>(1) prop⁰, STATE², pres, be, content: , [+2*] (2) restr¹:3, ref:α (3) prop¹, past², neg, content:5, [+4*,-α*] (5) ACTION³, past, buy, agent:4, theme: α, [+4†,-α†] (4) MALE², sg</p>
--

(7.20)a: (S₅) **What he didn't buy is** || any eggs.

When the constituent *any...* is added as the content of the copular, its address replaces that of the placeholder. As a consequence the constituent *any...* is simultaneously interpreted as both the content of the copular construction and the theme of *buy*. The latter fact predicts therefore that it must fulfill the requirements imposed by appearing in a negative context.

- | |
|--|
| <p>(1) prop⁰, STATE², pres, be, content:6, [+2*]
 (2) restr¹:3, ref:α/6
 (3) prop¹, past², neg, content:5, [+4*, -α]
 (5) ACTION³, past, buy, agent:4, theme:α/6, [+4†, -α†]
 (4) MALE², sg
 (6) indef⁶, plur, egg⁷, [+2†]</p> |
|--|

(7.20)a: **What he didn't buy is any eggs.**

For (7.20)a, it seems natural to assume that the free relative is in subject position. But, as is well-known, the position of subject and predicate are often subject to reversal with the predicate appearing before the copular and the subject after.⁷ That the same can happen with pseudoclefts is demonstrated by (7.26). Here the agreement is with the counterweight, indicating that it is in subject position.

(7.26) What he didn't buy are any eggs.

Given that reverse copular constructions exist, it should be expected that they are also possible for pseudoclefts, as the identification between the referents of the two arguments is an equative one. The derivation which follows is practically the same as that of (7.20)a above, only with the values of the subject and content roles of the copular reversed.

- | |
|---|
| <p>(1) prop⁰, STATE², pres, be, content:2, [2†], [+2*]
 (6) indef⁶, plur, egg⁷, [+2†]
 (2) restr¹:3, ref:α/6
 (3) prop¹, past², neg, content:5, [+4*, -α*]
 (5) ACTION³, past, buy, agent:4, theme:α/6, [+4†, -α†]
 (4) MALE², sg</p> |
|---|

(7.26)a: **What he didn't buy are any eggs.**

⁷The phenomenon is investigated in Birner (1996).

The supposition that the free relative can appear in subject position is supported by the fact that it can also appear in a raising construction, and that if so then it must have singular agreement.

- (7.27) a. What he forgot to buy seems to have been any eggs.
 b. *What he forgot to buy seem to have been any eggs.

For completeness, we should also note that it is predictable from the above observations that the order of the free relative and the counterweight can be reversed, with again any order of subject and predicate, as demonstrated in (7.28)a. It might be speculated that an *any*-phrase in subject position receives a different interpretation than the one needed as the argument of a negative and therefore (7.28)c is ruled out and (7.28)b sounds better.

- (7.28) a. Eggs is/are what he always forgets to buy.
 b. ?Any eggs is what he always forgets to buy.
 c. *Any eggs are what he always forgets to buy.

I turn now to the constructions where the counterweight is a verbal or adjectival phrase. Strangely the verb phrase can not only be the bare verb phrase we would expect after *do* (*he did laugh*), but also an infinitive (**he did to laugh*).

- (7.29) a. What he did was laugh.
 b. What he did was to laugh.
 c. What he is is lazy.

Application of the raising test shows clearly that of the above examples, only the infinitival clause can appear as the predicate, the other two being necessarily postponed subjects.

- (7.30) a. *What he did seemed to be laugh.
 b. What he did seemed to be to laugh.
 c. *What he is seems to be lazy.

Also the infinitive appears to be blocked from appearing sentence initially, either as subject or predicate.⁸

⁸Consequently it is impossible to tell whether it can appear as a subject at all, since there is no verb agreement data to go on.

- (7.31) a. *To dance is what he does best.
 b. Dance is what he does best.

There are no equivalent *it*-cleft constructions.⁹ This strengthens the argument that there must be something fundamentally different in the way that coreferencing takes place between *it* and pseudo-clefts.

- (7.33) a. *It was (to) feed the dog that he did.
 b. *It was tired of her antics that he was.

Let us look at the derivation of the example with a bare verb phrase counterweight (7.29)a.¹⁰ Initially it will be entirely parallel to that of (7.26)a, with the free relative in store and the counterweight in subject position. The action *laugh* is added simultaneously as the subject of the copular and the content of the semantically empty proposition in the free relative.

- (1) prop⁰, STATE⁵, past, **be**, content:2, [+5†, -2†]
 (5) ACTION⁵, **laugh**, agent:4, [+4†]
 (2) restr:3, ref:α/5
 (3) prop¹, past³, content:α/5, [+4*, -α†]
 (4) MALE², sg

(7.29)a: **What he did was laugh.**

The counterweight plays two simultaneous roles when it is added. I have ordered the constituents above so that its role as the subject of the copular is emphasized rather than that of the content of the free relative clause. This makes clear the cleft nature of the construction, and the way it has been put together. An alternative ordering of the same derivation which emphasises the semantic content would be as follows.

⁹Although Irish English does allow the latter of these, so we do not want to rule them out in any principled way. Also gerunds are possible in *it*-clefts, but they can appear anywhere noun phrases can. It appears that they have no problem functioning without being provided with a syntactic subject.

- (7.32) a. It was dying that he was scared of.
 b. It was Peter dying that had scared him.

¹⁰The analysis of example (7.29)b with the infinitival counterweight presents many unsolved problems and I will leave it to future research.

- (1) prop⁰, STATE⁵, past, **be**, content:2, [+5†, -2†]
 (2) restr¹:3, ref:α/5
 (3) prop¹, past³, content:α/5, [+4*, -α†]
 (5) ACTION⁵, **laugh**, agent: 4, [+4†]
 (4) MALE², sg

(7.21)a: **What he did was laugh.**

Turning now to the derivation of the sentence with an adjectival counterweight, (7.21)c, we must bear in mind again that this is constrained to appear as a postponed subject.¹¹ When the adjectival constituent is added as the subject of the matrix copular, it must simultaneously be interpreted as the content of the free relative copular, this fact ensuring that it inherits the latter's subject. This is shown in the following derivation.

- (1) prop⁰, STATE⁵, pres, **be**, content:2, [+5†, -2†]
 (5) QUALITY⁵, **lazy**, theme:4, [+4†]
 (2) restr¹:3, ref:α/5
 (3) prop¹, pres³, **be**, content:α/5, [+4*, -α†]
 (4) MALE², sg

(7.21)c: **What he is is lazy.**

7.2.1 Gapping and pseudoclefts

It has long been noted in the literature that, unlike other copular constructions, gapping is not allowed with pseudoclefts, and this contrast is shown by the following sentences, where (7.34)b is strikingly ungrammatical.

- (7.34) a. Paul is lazy and Susan _ mean.
 b. *What Paul is is lazy and what Susan is _ mean.

It has already been demonstrated that in this construction the free relative is not the subject, but a preposed predicate. Referring back to the analysis given in Chapter 5, the transition introducing the gapping construction makes explicit reference to replacing the subject of the proposition with a placeholder which has to be filled at the next transition

¹¹It might be speculated that the reason it cannot appear in predicate position is that there would then be confusion about it appearing as the copular predicate in two different clauses simultaneously and having to mark as its theme two different subjects.

by the subject of the second conjunct. These facts on their own are sufficient to predict that the gapping will be impossible.

This predicts that if we reverse the pseudocleft, getting (7.35)a¹², then gapping should be possible. Given that the ungapped sentence is not entirely natural, and gapping is at best a marked construction, the gapped sentence is surprisingly good and certainly much better than (7.34)b above.

- (7.35) a. Lazy is what Paul is.
 b. ??Lazy is what Paul is and mean _ what Susan is.

Heycock & Kroch (1998) introduce a problematic example of gapping, repeated here as (7.36), involving a pseudocleft as the first conjunct and a non-pseudocleft as a gapped second conjunct. As the authors note the problem is that this appears not only to break the restriction on pseudoclefts appearing in gapped sentences, but also to break the constraint that gapping should involve conjuncts with parallel structures.

- (7.36) ?What Bill does is run marathons; (and) his wife _ half-marathons.

As it turns out, this example is predicted to be grammatical by the analyses we have so far. To see this we must consider how it will be derived. The derivation of the first pseudoclefted clause is quite regular. The gapping transition for *and* copies a reduction of the active constituent *run*, with the subject replaced by a placeholder, and places it on the active stack. The remaining two arguments are added as in a normal gapping construction.

- | |
|---|
| (1) prop ⁰ , STATE ⁴ , past, be , content:2, [+5,-2†] |
| (5) ACTION ⁵ , run , agent:4, theme: 6, [+4], coord ⁷ : {&, <7>} |
| (4) MALE ² , sg, Bill |
| (6) marathon ⁶ , plur |
| (7) ACTION ⁷ , run , agent:β/8, theme: 9, [+β/8] |
| (8) possess ⁸ :{ref:4}, FEMALE ⁹ , sg, wife |
| (9) half-marathon ¹⁰ , plur |
| (2) restr ¹ :3, ref:α/5 |
| (3) prop ¹ , past ³ , content:α/5, mod ⁴ :5, [+4,-α†] |

- (7.36): **What Bill does is run marathons, and his wife, half-marathons.**

However, this analysis should also predict that the following example, (7.37), should be better than it is.

¹²This is not so much reversed as in the canonical position with subject first.

(7.37) ??What Paul is is lazy and his wife _ mean.

7.2.2 Propositional arguments

In addition to the problem with the *to*-clause appearing unexpectedly as the complement of *do*, as seen in (7.21)b, there are other mismatches between the counterweight and the slot it fills in the free relative. For example, the following sentences show that the restriction on prepositions taking finite clause complements, shown in (7.38)b, does not seem to hold if the clause is in the counterweight, as in (7.38)a. Indeed, (7.38)c shows that the preposition must be present if the counterweight finite clause is to be interpreted as the complement of *surprised*. In this section I will attempt to explain this behaviour with reference to the assumptions of the dynamic model.

- (7.38) a. What he was surprised at was that she was single.
 b. *He was surprised at that she was single.
 c. *What he was surprised was that she was single.
 d. He was surprised that she was single.

One piece of evidence we should bear in mind is that, as the examples in (7.40) demonstrate, the same phenomenon also occurs with the topicalization of propositions. Therefore we should try to find a common solution.¹³

- (7.40) a. That she was single, he was really surprised at _.
 b. *That she was single, he was really surprised _.

It shows that the transition introducing a proposition can act differently to simply filling an argument role with the address of a proposition.

We established earlier, in chapter 4, that a propositional subject is initiated with a specialized rule, associated solely with the complementizer *that*. I will assume again that propositions are not added as arguments by the standard rule *add-arg*, but have their own individual rule, although in this case it is not limited to the complementizer. As we can see from (7.38)d above this transition can ignore the requirements for a particular governed preposition that are normally placed on the arguments of a verb like *surprise*.¹⁴ The transition schema allowing this transition will therefore be of this kind.

¹³This is not the case however with the infinitival clause arguments discussed above.

- (7.39) a. He said he would meet her and meet her he did _ .
 b. *He said he would meet her and to meet her he did _ .

¹⁴This seems to be an idiosyncrasy of English, for it is not so in the closely-related Scandinavian languages for example.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
add-propositional-argument	(X, arg: [lex: —])	(N, prop) (X, arg:N)

That is, the rule that initiates propositions in argument position can ignore any lexical requirements placed on it by the verb.

Once we have this transition then the behaviour of propositions when the counterweights of pseudoclefts and when topicalized follows automatically.

First consider the topicalization data in (7.40) above. The proposition is created by the rule that creates propositions in sentence-initial position and is directly associated with the complementizer. This constituent will be placed on the store. When it gets to the word *at*, in (7.40), this constituent can interpret it as its argument. There is no restriction on the preposition that it cannot take finite propositions as its argument, the only reason this is normally not possible is that the rule for initiating propositional arguments is so set up.

The stored proposition cannot be interpreted as the argument of *surprised*, as would be required in (7.40)b, since the lexical entry for the verb states that it subcategorizes for a item with lexical value **at**. This requirement can be overridden by the rule *add-propositional-argument*, but not by a constituent being interpreted from store, for which the standard *interpret-store* rule will be used.

The same arguments apply in the case of the corresponding phenomenon in the pseudocleft data in (7.38) above. It would appear that such contrasts must be very difficult for constraint-based approaches to model.

7.3 Existential *there*

This seems an appropriate place to take a look the dummy existential subject *there*, especially as it involves consideration of copular constructions and questions of reference.

It is easily established, by consideration of raising and tag question data, that the dummy-*there* fills the subject position in (7.41)a and (7.41)c. From a functional perspective it moves an indefinite subject from subject position, which is a position more typically filled by a sentence topic, and presents it as new information in the second half of the sentence. As shown by (7.41)c and (7.41)d, it also gets around the fact that the copular is generally a raising verb, a function which it shares with the dummy *it*.

- (7.41) a. There were some dogs in the park.
 b. Some dogs were in the park.
 c. There were some dogs.
 d. *Some dogs were.

The two dummy pronouns are not interchangeable, however, and for want of a better way to distinguish *there*, I will assign its constituent a feature **existential**. I will assume that the *there*-constituent has no inherent number of its own, but that its number can be constrained by the necessary agreement of the verb of which it is subject. We also require a new transition for adding an existential predicate.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
existential predicate	(X, be , content: , [+Y])	(N) (X, be , content:N, [+Y†])

where Y is *existential*, N is nominal, Y and N corefer

The derivation of (7.41) will then be as follows.

- | |
|--|
| (1) prop ⁰ , STATE ² , past, be , content:3, mod ⁵ :4, [+2†] |
| (2) existential ¹ , ref:3, plur ² |
| (3) indef ³ , plur, dog ⁴ |
| (4) LOCATION ⁵ , in , loc:5 |
| (5) def ⁶ , park ⁷ , sg |

(7.41): **There were some dogs in the park.**

The fact that the dummy is coreferent with *some dogs* means that they must agree in number and hence gives the effect of the verb agreeing with the number of the predicate, rather than the subject.

The following example, (7.42), shows subject *there* both appearing in subject-auxiliary inversion and also in position raised out of its logical position.

- (7.42) What problems do there seem to be _ ?

The finite verb *do* is still constrained to agree, via the dummy pronoun, with the topicalized *what problems*.

- | |
|---|
| <p>(1) prop⁰, Q¹, pres³, content:4, [+3*, -2*]
 (4) STATE⁵, seem, theme:5, [+3*, -2*]
 (5) prop⁶, inf, STATE⁷, be, content:2, [+3†, -2†]
 (3) existential⁴, plur, ref:2⁷
 (2) ref¹:-, problem², plur</p> |
|---|

(7.42): **What problems do there seem to be?**

Just like dummy-*it*, dummy-*there* can appear in object position, just as long as the role assigned to it is semantically vacuous as in the following object raising example, (7.43).

(7.43) He expected there to be dogs there.

- | |
|--|
| <p>(1) prop⁰, STATE², past, expect, exper:2, \emptyset:3, theme:4, [+2†]
 (2) MALE¹, sg
 (4) prop⁴, inf, STATE⁵, be, content:5, mod⁷:6, [+3†]
 (3) existential¹, ref:5
 (5) dog⁶, plur
 (6) LOCATION⁷, there</p> |
|--|

(7.43): **He expected there to be dogs there.**

As has often been noted, and frequently complained about, in colloquial use the contracted form *there's* does not constrain the number of the existential phrase, giving rise to the contrast between (7.44)a and (7.44)b.

- (7.44) a. There's two people here to see you.
 b. *There is two people here to see you.

We must suppose therefore that this contracted form is associated with a single transition, and that this transition has only one form which does not constrain the number of the *there*-constituent.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
there's	(X, prop)	(N, existential, ref:) (X, prop, STATE, pres, be , content: , [+N])

This will then allow the subject *there* to agree with either a plural or singular predicate. In (7.44)b the verb *is* constraints its subject to be singular, which leads, via coreference, to the number disagreement.

7.4 Summary

Clefting constructions present some of the most challenging data to for models of syntax. This chapter has done no more than cautiously explore some of the possibilities for analysis available in the dynamic grammar. However, I hope that it has showed that the assumption of a single level of representation also offers new insights into dealing with these challenges.

One thing that has become clear from the treatment of clefts here is that the notion of reference plays a crucial role in the syntax. especially in copular constructions, and further work is needed to place this on more solid ground.

Chapter 8

Further topics

In this chapter I consider aspects of a number of further syntactic constructions, using analyses which build on and develop various aspects introduced in the previous chapters. The constructions here are not chosen to be in any way comprehensive, but hopefully to reveal important insights into the way the dynamic grammar needs to be formulated.

8.1 Some lexically-specific transitions

In the first section I briefly consider a pair of minor constructions, which while not raising any serious difficulties, show the need for tying transitions down to individual lexical items.¹

8.1.1 Subjunctive inversion

This construction is illustrated by (8.1)a. It is clear that the initial inverted clause makes the same contribution to the sentence meaning as the *if*-clause in the synonymous sentence (8.1)b.

- (8.1) a. Had I lied, I would have won.
b. If I had lied, I would have won.

The transition associated with *had* when it initiates the conditional clause must then be the following one.

¹I have already argued the need for this previously in this thesis, for instance for the complementizer *that* and the *wh*-exclamatives.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{had}	(X, prop)	(N, prop, past, perf, content:) (X, prop, condition:N)

With the addition of this transition schema to our stock of syntactic rules, the following derivation of (8.1) will be possible.

(1) prop ⁰ , condition ¹ :2, STATE ⁵ , past, will , content:6, [+5*]
(6) STATE ⁶ , perf, content:7, [+5*]
(7) EVENT ⁷ , ppart, win , theme:5, [+5†]
(5) 1ST-PERSON ⁶ , sg
(2) prop ¹ , past, perf, content:4, [+3*]
(4) ACTION ³ , ppart, lie , exper:3, [+3†]
(3) 1ST-PERSON ² , sg

(8.1): **Had I lied, I would have won.**

This conditional construction is highly restricted in modern English to three word forms: perfective *had*, subjunctive *were* and *should*. It appears that the most economical way to express this restriction is to associate these words directly with transitions of the type given above. There is then no need to specify that the word must bear some feature *inverted* constraining it to appear before its subject, for the transition ensures this and no other derivation of the conditional uses is possible.

Similarly, there is no need to stipulate that the negative forms *hadn't*, *weren't* and *shouldn't* cannot appear in this construction, as is shown in (8.2). Even though they may appear in inverted position, the transition is directly associated with their positive counterparts and no conditional derivation will be possible.

- (8.2) a. Had I not seen it, I wouldn't have believed it.
b. *Hadn't I seen it, I wouldn't have believed it.

8.1.2 *Why go?*

The fact that one can ask questions with *why* followed by a base verb phrase, as in (8.3)a, appears again to be an idiosyncrasy that is best captured by relating the transition responsible directly to the lexical entry for the word in question. It is not possible for other *wh*-modifiers, as (8.3)b and c show.

- (8.3) a. Why stay here?
 b. *How do it?
 c. *When go home?

Any attempt to link this with some semantic property of *why*, for example that *wh*-words normally modify propositions, rather than actions, runs into the problem that the same construction is not possible with expressions which are practically synonymous, as in (8.4).

- (8.4) a. *What stay here for?
 b. *How come stay here?

Another thing to notice about the construction is that *why Verb?* appears to be synonymous with *why should X Verb?*, where *X* is some unspecified subject. This intuition can be captured by allowing *why* to have a transition something along the lines of the following.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ why }	(X, prop)	(N, REASON ¹ , ref:-) (X, prop, mod:N, Q, <i>should</i> , content: , [+α])

This will allow the following derivation for (8.3)a.

(1) prop ⁰ , mod ¹ :2, Q, STATE, <i>should</i> , content:3, [+α*]
(3) STATE ² , stay , theme:α, loc:4, [+α†]
(4) LOCATION ³ , here
(2) REASON ¹ , ref:-

(8.3)a: **Why stay here?**

In this construction the *why* constituent is not placed on the store, but immediately modifies the proposition, which is not usual for a *wh*-constituent. There is evidence for this analysis in the contrast in (8.5). If it is possible to interpret the string (8.5)b, then the only reading is one with modification of the matrix proposition.

- (8.5) a. Why do you imagine he left _ ?
 b. *Why imagine he left _ ?

Finally it should be noted that the construction is restricted to the matrix clause, as shown by (8.6). Whether this has to be stipulated in the transition rule or can be derived from independent considerations is a matter for further research.

(8.6) *I can't imagine why go.

All these properties can be captured if we take the transition to be the basic item of syntax. It is conjectured that the kind of large-scale grammar that needs to be built will have a few transitions that have general application to many lexical items, but many transitions that are lexically linked and whose application will therefore not slow up the processor.

8.2 Reduced relatives

In chapter 3, I restricted the discussion of relative clauses to finite ones. In this section I will consider another common type the *reduced relative* construction, and argue that all examples of this can be characterized by introducing a single new transition into the grammar.

An alternative name for the reduced relative construction is *whiz-deleted* relative, from the observation that they appear to have *who/which is/was* deleted from a standard relative clause. This can be seen by consideration of the following examples.

- (8.7)
- a. The dog *barking* is Fido.
 - b. A guy *liked by everyone* is Charles.
 - c. The scientist *to be given the prize* will be named tomorrow.
 - d. She is the woman *most likely to win*.
 - e. The dog under the table should not be fed.
 - f. This man, *a great friend of mine*, has never played chess.

It is simple to realize this intuition in the dynamic model, and allow all of these sentences with the introduction of a single transition schema which is not linked to a lexical item.²

rule	top of stack (S_n)	top of stack (S_{n+1})
<i>whiz-deleted relative</i>	(X)	(N, prop, STATE, be, content: , [+X]) (X, mod:N)

where X is saturated

²Sag (1997) has the same intuition that all of these constructions should be treated as restricted relatives, however there is no recognition of the clear link they show to finite relatives with *who/which is/was* deleted, and in effect any constituent can have this role. Consequently, the only way that impossible relative constructions such as “*the person stand on my foot...*” can be ruled out is by specifying in the lexicon that they are [MOD none]. The present approach rules them out precisely because **the person who is stand on my foot...* is impossible. The present approach also predicts that subject-gapped infinitival relatives such as (8.7)c, have the same future connotation as *the scientist who is to be given...*, which will not be the case with Sag’s approach.

The derivation of (8.7)a will be as follows, with the above transition applying at the third state to create the relative clause, to which the constituent *barking* is then added.

- (1) prop⁰, STATE⁴, pres, **be**, content:5, [+2†]
 (2) def¹, **dog**², sg, mod³:3
 (3) prop³, STATE, be, content:4, [+2*]
 (4) ACTION³, ing, **bark**, agent:2, [+2†]
 (5) **Fido**⁵, sg, ref:2

(8.7): **The dog barking is Fido.**

One gap in this pattern is that nominal predicates can only be non-restrictive, as in (8.7)f above. This is perhaps to be explained in connection with the reference requirements of predicative nominals, but unless the transition can be constrained in some way the grammar will overgenerate and incorrectly allow (8.8)b, with the same interpretation of (8.8)a.

- (8.8) a. The boy who's the teacher's favourite is leaving.
 b. *The boy the teacher's favourite is leaving.

In addition, non-restrictive clausal apposition will be allowed by the same transition schema. The following example is from Hudson (1998: 87).

- (8.9) Global temperatures are rising: a matter of great concern for us all.

8.3 Left-recursive structures

It is well-known that left-recursive grammar rules, such as $NP \rightarrow NP S[rel]$, constitute a problem for top-down parsers. Steps must be taken to prevent the parser applying this rule repeatedly to its own output resulting in an infinite loop. It has also been pointed out that a similar potential problem exists for incremental grammars of the type proposed here, whereby they have to specify an infinite number of possible transitions, leading to an infinite number of possible next states.³

8.3.1 Clausal subjects

To give a concrete example of where this might be the case, consider the sentences in (8.10). The interpretability of the sentences declines rapidly with the repeated embedding

³This problem is discussed in Milward (1994b).

of the gerundive subject phrase inside itself, indeed it may be argued that even a single such self-embedding, as in (8.10)b cannot be interpreted in a straightforward incremental manner, but only after conscious effort. Nevertheless, it is clear that if the grammar allows (8.10)a, as it must, then the remaining sentences must accord with the rules of the grammar, that is they are grammatical, but practically unusable.

- (8.10) a. Peter throwing up shocked Sue.
 b. Peter throwing up shocking Sue surprised Jim.
 c. Peter throwing up shocking Sue surprising Jim seemed strange to Jock.

I do not offer an analysis of gerundive clauses in this thesis, but the central question at issue here is whether or not the subject of the gerundive clause, *Peter* in (8.10)a, has to be immediately placed as subject of the subject gerundive, or whether it can be initially placed in the store of the matrix clause and reanalyzed. I will assume that to fit in with the rest of the grammar, in particular the coordination facts, the former of these choices must be the case, and that the subject clause in (8.10)a must start at the first word *Peter*.

That is a transition of the following type must apply at *Peter*.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
start gerundive subject	(X, prop)	(N1) (N, prop, gerund, [+N1]) (X, prop, [N])

If this is the case it then follows the only way to derive (8.10)b is if at the transition for the first word the prediction is also made the subject clause now beginning will itself be the subject of a subject clause.

This would be the result of the rule⁴ applying again to form a combined initial transition as follows.

rule	top of stack (S_n)	top of stack (S_{n+1})
start gerundive subject (x2)	(X, prop)	(N2) (N1, prop, gerund, [+N2]) (N, prop, gerund, [+N1]) (X, prop, [N])

And following the initial supposition through, we reach the conclusion that the initial constituent could be the subject of an indefinitely embedded subject gerund, and that the total number of following states is infinite.

⁴Or perhaps combination of rules.

Milward (1994b) concludes from such considerations that in a dynamic grammar there is a necessity for the underrepresentation of the structures produced, analogous to parse packing. However, I will argue that even if the grammar does allow an infinite number of transitions from one state this need not necessarily be problematic, either in terms of its practical use in automatic language processing, or in terms of modelling the human language processor. If we assume that we need a weighted grammar for practical language analysis, then given natural assumptions about the low productivity of the rule “make this constituent the subject of a subject gerundive clause”, the multiply-embedded first states would be so unlikely that they would be beneath consideration. Depending on the processing strategy assumed, they could perhaps be backtracked to when the more likely transitions led to a breakdown, but only with increasing difficulty as the point of breakdown moves further and further away from the initial word as the depth of embedding increases.

Indeed, the situation will only be problematic for a very simple-minded processor which looks at every conceivable next state allowed by the grammar, no matter how unlikely. If we are thinking in terms of modelling the brain, it will be the case that it will quickly run up against limits as to what depth of embedding it can compute without exhausting its resources. We are familiar with the concept of a grammatical sentence that is impossible to process due to the cumulative effect of repeated centre-embeddings, but it appears that this may apply just as well to a single transition, and that the rules for constructing transitions allow some (indeed infinitely many) that are beyond the means of the processor to construct.

It should be pointed out that the difficulty of sentences (8.10)b and (8.10)c does not lie in the complexity of the meaning expressed, for read with the correct intonation, the meanings of the informationally-equivalent sentences in (8.11) are relatively transparent.

- (8.11) a. Jim was surprised that Sue had been shocked at Peter throwing up.
 b. It seemed strange to Jock that Jim was surprised that Sue had been shocked at Peter throwing up.

Nor does it seem that the multiply-embedded sentences in (8.10) place undue load on the memory limitations of the processor, something that will be discussed in the next chapter. It appears to be the case that the reason that they are difficult is because the initial transition needed to allow their derivation is so complex, and of such low probability.

8.3.2 Possessives

The Germanic possessive construction is another which appears to allow the construction of structures exhibiting an indefinite amount of left-recursion, as illustrated by (8.12).

(8.12) Who feeds John's mother's friend's sister's dog?

However, in contrast to the gerundive subjects analyzed above, this does not appear to cause any great difficulty in interpretation. It is true that the intended message is quite complex, but it becomes no easier to process, indeed arguably harder, if we employ a non-left-recursive structure as in (8.13).

(8.13) Who feeds the dog of the sister of a friend of the mother of John?

The conclusion then must be that although the resulting structure of (8.12) is left-recursive, there is no necessity to predict the depth of recursion from the first word, and that instead there is an ongoing reanalysis of the construction.

I shall assume that the combination *word's* occupies a single transition, but that this transition contains two subparts: firstly the normal possible transitions for the *word*, and immediately after this the possessive creates a new superordinate constituent to which the active constituent created by *word* is linked by the relation *poss* for "possessor".⁵ The resulting transition when the new constituent is added as the argument of *feed* in (8.12) is shown here.⁶

(1) prop ⁰ , Q ¹ , ACTION ² , pres, feed , agent:2, theme:3/4, [+2†]
(2) HUMAN ¹ , ref:–
(4) poss ^{3b} :4
(3) MALE ^{3a} , sg, John

(8.12): **Who feeds John's** || mother's friend's sister's dog?

It will be seen that the first part of the transition *John* places a constituent in argument position, while the following possessive immediately replaces this and lowers *John* into a position as its own dependent. If we consider this latter transition in terms of operations on the active stack, it may be represented as follows.

rule	top of stack (S_n)	top of stack (S_{n+1})
{ possessive 's }	(X1)	(N, poss: X1)
	(X, prop, relation: X1)	(X, prop, relation: X1/N)

In this transition the original constituent (X1) (that is the *possessor*) is necessarily removed from the active stack. A transition of this sort, where the address pointed to by

⁵Other functions of the "genitive 's" will be ignored here.

⁶For clarity of exposition, I have added to this diagram here the subscripts *a* and *b* to the transition number to show the sequence of creation.

a relation is changed, may be termed *grammaticalized reanalysis*, that is reanalysis carried out by the transition rules themselves and not requiring backtracking of the processor.

In the derivation of (8.12), the next transition for the word *mother's* is exactly the same. The transition for *mother* adds the head information to the unsaturated active constituent, while the possessive clitic creates a new constituent as before, again updating the address of the theme of *feed*.

<p>(1) prop⁰, Q¹, ACTION², pres, feed, agent:2, theme:3/4/5, [+2†] (2) HUMAN¹, ref:- (5) poss⁴:4 (4) poss³:3, FEMALE⁴, sg, mother (3) MALE³, sg, John</p>
--

(8.12): **Who feeds John's mother's** || friend's sister's dog?

And the entire derivation will proceed in a similar manner as follows.

<p>(1) prop⁰, Q¹, ACTION², pres, feed, agent:2, theme:3/4/5/6/7, [+2†] (2) HUMAN¹, ref:- (7) poss⁶:6, dog⁷, sg (6) poss⁵:5, FEMALE⁶, sg, sister (5) poss⁴:4, friend⁵, sg (4) poss³:3, FEMALE⁴, sg, mother (3) MALE³, sg, John</p>
--

(8.12): **Who feeds John's mother's friend's sister's** dog?

The overwriting of the address of the lowered possessive constituent preserves monotonicity in the strict sense that information is always added. However, it might be thought that this repeated changing of the assumption of which constituent was to fill the argument slot would necessarily result in having to undo decisions, for example concerning how suitable the constituent was as theme of *feed*. It should be noted, however, that since the possessive is a pre-head modifier the address given to a relation is changed before the head and other distinguishing features are added to the constituent. Decisions about whether the added constituent is conceptually suitable for the role can only be made when something of substance is added, which in (8.12) will be when the meaning of the *dog* is added to constituent (7). Before that time, each of the successive constituents in the theme of *feed* will be empty apart from the relation *poss*. Therefore, no decisions as to the suitability of a constituent have to be made and then reassessed.

The only case where this will not be true is where the possessive clitic attaches to a postmodifier, as in (8.14).

- (8.14) I would be grateful if you could bring to staff in your department's attention that work on the new Chemistry building has just commenced. [University of Edinburgh memorandum, 1998]

In this case, *the staff* will first be attached as the beneficiary of *bring*, and it will be another three transitions before this assumption must be unmade. Note that this is not the result of a failed processing decision, for this is the only way the grammar can derive this sentence. Luckily, such constructions do appear to show marked garden-path effects in processing.

8.3.3 Turkish nominalized clauses

In a proposal for an incremental parsing algorithm based on HPSG, Güngördü (1997: 136) presents the following data from Turkish, arguing that the construction presents particular problems for incremental parsers. It will be seen that the nominalized clause which is the theme of the final matrix verb *tell* may have embedded within it other nominalized clauses, and since these clauses appear before the verbs that subcategorize for them, they can result in left-recursive structures. The result is that the initial clause in the sequence may be indefinitely embedded and it is impossible to predict how far this will be. In (8.15) the first clause is doubly embedded, and in (8.16) triply embedded.

- (8.15) Berfu'-ya Güneş'-in uyu-duğ-u-nu ben-im gör-düğ-üm-ü Yasemin söyle-di.
 Berfu_{dat} Güneş_{gen} sleep_{fact,3sPoss,acc} I_{gen} see_{fact,1sPoss,acc} Yasemin tell_{past}
 YASEMIN has told Berfu that I have seen that Güneş' was asleep.

- (8.16) Berfu'-ya Güneş'-in uyu-duğ-u-nu ben-im gör-düğ-üm-ü Mehmet'-in bil-diğ-i-ni
 Berfu_{dat} Güneş_{gen} sleep_{fact,3sPoss,acc} I_{gen} see_{fact,1sPoss,acc} Mehmet_{gen} know_{fact,3sPoss,acc}
 Yasemin söyle-di.
 Yasemin tell_{past}
 YASEMIN has told Berfu that Mehmet knows that I have seen that Güneş was asleep.

It is apparent that the first nominalized clause *Güneş sleeping* could be embedded indefinitely deeply in the final interpretation. Therefore if the parser has to predict at this stage how deeply the clause will end up being embedded it will have an infinite number of choices to choose from. This is the the same decision faced in the analysis of subject gerund clauses in English and which I conjectured was the reason that they were practically unusable when embedded. The Turkish construction, however, presents no problems to

users, beyond that of the complexity of the message itself, so the question is how can the sentence be analyzed without having to predict the depth of embedding.

Turkish is in general verb final and so I will suppose that it makes liberal use of the store in much the same way as was seen for German in the previous chapter. We also need to know that the nominalized factive verbs (marked *fact*) require a subject in the genitive, with which they agree. I will assume that genitive nominal phrases do not “predict” that they are to be subjects of a nominalized clause (by creating the clause), but rather wait on the store of the main clause and are then inherited by the nominalized verb.⁷

The derivation of (8.15) will then be as follows. I assume that all constituents are first placed on the store of the matrix proposition. The nominalized verb *sleep* inherits its genitive subject as it is placed on store. The nominalized verb *see* inherits its genitive subject plus its theme, the nominalized clause *sleep*. Finally the finite verb *tell* interprets the three remaining uninterpreted constituents.⁸

- (1) prop⁰, ACTION⁷, past, **tell**, agent:7, benef:2, theme:6, [+7†, -6†, >5*, -4*, >3*, ~2†]
 (7) **Yasemin**⁶, sg
 (2) **Berfu**¹, sg
 (6) EVENT⁵, fact, **see**, exper:5, theme: 4, [>5†, -4†]
 (5) 1-PERSON⁴, sg
 (4) STATE³, fact, **sleep**, theme:3, [>3†]
 (3) **Güneş**², sg

(8.15): **Berfu**_{dat} **Güneş**_{gen} **sleep**_{fact,3sPoss,acc} **I**_{gen} **see**_{fact,1sPoss,acc} **Yasemin** **tell**_{past}.

It will be seen that no decision has to be made when a clause is reached as to how deeply it is to be embedded. Indeed, it is just left on the stack until it is interpreted by a following constituent. In this way more and more deeply embedded interpretation is built up, but the syntactic operations remain simple and uniform.

There is thus no problem in extending it to another level of embedding, as witnessed by the derivation of (8.16).

⁷This is not an essential assumption for the analysis to succeed, the advantage being that this approach will also the “long-distance topicalization” construction discussed below.

⁸I employ here the previous abbreviations for case relations, adding dative, ~, and genitive, >.

- (1) prop⁰, ACTION⁹, past, **tell**, agent:9, benef:2, theme:8, [+9†, -8†, >7*, -6*, >5*, -4*, >3*, ~2†]
- (9) **Yasemin**⁸, sg
- (2) **Berfu**¹, sg
- (8) STATE⁷, fact, acc, **know**, exper:7, theme:6, [>7†, -6†]
- (7) **Mehmet**⁶, sg
- (6) EVENT⁵, fact, acc, **see**⁵, exper:5, theme:4, [>5†, -4†]
- (5) 1-PERSON⁴, sg
- (4) STATE³, fact, acc, **sleep**, theme:3, [>3†]
- (3) **Güneş**², sg

(8.16): **Berfu**_{dat} **Güneş**_{gen} **sleep**_{fact,3sPoss,acc} **I**_{gen} **see**_{fact,1sPoss,acc} **Mehmet**_{gen}
know_{fact,3sPoss,acc} **Yasemin** **tell**_{past}.

The conclusion is that the derivation of these left-recursive structures in a head-final language such as Turkish is not only possible, but does not even require the kind of syntactic reanalysis, that is a relabelling of constituents, that we saw for the possessive construction in English.

This analysis also allows examples of “long-distance topicalization” out of a nominalized clause. The following example, from Güngördü (1997: 160), shows that the genitive constituent may simply wait on the store until there is an appropriate nominalized verb to interpret it as subject.

- (8.17) Adam-m ben san-a kadın-ı gör-düğ-ü-nü söyle-di-m.
man_{gen} I you_{dat} woman_{acc} see_{fact,3sPoss,acc} tell_{past,1sg}
As for the man, I have told you that he has seen the woman.

The derivation of this sentence is left as an exercise for the reader.

8.4 Reduplication in Chinese

Syntactic constructions involving reduplication of lexical items are far from rare in the world’s languages, but they present particular problems for analyses based on syntactic constituency structure. In this section I briefly examine one such construction in Chinese, and argue that the difficulties of such constructions are removed when we have a syntax which is not autonomous, that is one that can refer to the semantic interpretation that has been constructed so far

The modification of intransitive verbs in Mandarin Chinese is illustrated in (8.18)a.

The particle *de* must necessarily be attached to the intransitive verb *pǎo* in order for the verb to be then modified by an adjectival phrase.

- (8.18) a. *tā pǎo-de hěn kuài.*
 she run-DE very quick
 She runs very quickly.
- b. **tā pǎo hěn kuài.*

I shall assume in the analysis of (8.18)a that the verb plus adverbial particle forms a single unit, and that this introduces a modifier relation, as shown in the following derivation.⁹

- (1) prop⁰, EVENT², **pǎo** [*run*], theme:2, mod:3, [2†]
 (2) THING¹
 (3) QUALITY³, emphatic, **kuài**⁴ [*quick*], theme:1

(8.18): *tā pǎo-de hěn kuài.*

Turning now to transitive verbs, the pattern is complicated somewhat by the fact that verbs with the adverbial particle cannot take objects, as shown in (8.19)b, while as we have seen already modifiers cannot attach to verbs that do not have this particle, so (8.19)c is also out. The construction one has to use to modify a transitive verb, and one that every student of Chinese has to learn, is illustrated in (8.19)a: the transitive verb is repeated after the object, this time with the adverbial particle, and the modifier can then be added.

- (8.19) a. *tā xúe yīngwén xúe-de hěn nǜlì.*
 she study English study-DE very diligent
 She studies English very diligently.
- b. **tā xúe-de yīngwén hěn nǜlì.*
- c. **tā xúe yīngwén hěn nǜlì.*

This is a very simple and general pattern, but attempts to characterize it in terms of phrase structure rules run into difficulties. For example, Huang (1983) proposes the following rule for the modified transitive construction:

$$VP \rightarrow VP [+VN] \quad VP [+COMPL]$$

⁹The question of whether this assumption is tenable or not requires further investigation. I also assume that Chinese has no distinguished subject position and that the subject *tā* is placed in store unmarked, as a sentence topic.

Here the feature [+VN] is used to force the first VP to contain an object, and the [+compl] to force the second VP to be marked by *de*. As Calder (1986) notes, these innovations are “hard to motivate with respect to the rest of the grammar”.

Furthermore, this rule places no constraint on the construction that the verbs be one and the same. Indeed the only way to do this in a grammar based on syntactic structures would be to introduce into the categories a feature capable of bearing the lexical value of every verb in the language. The consequence is that sentences such as (8.20) are permitted.

(8.20) tā xúe yīngwén pǎo-de hěn kuài.
she study English run-DE very quick

The only thing this string could be interpreted as would be as a coordination of two separate propositions. This is not the interpretation we get for the sentence with reduplication (8.19). Indeed on reflection this might indeed be ambiguous between the usual modification construction and one in which there are two separate coordinated propositions, which happen to have identical verbs, and the intonation patterns for these two interpretations would be distinct. The conclusion is that the modification of transitive construction does appear to be a distinct syntactic construction, which demands that the verbs in both clauses are identical.

In analysing this construction I shall assume that the reduplicated verb + *de* does not start a new constituent at all, but merely adds the modifier relation to the verbal constituent that is already there. Such a transition may be represented as follows.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ verb, head , DE }	(X, head , arg:Y)	(X, head , arg:Y, mod:)

This transition entails that the lexical head of the *verb* + *de* must match that of the active constituent for the modifier role to be added.

The derivation that this transition will allow for (8.19)a will be the following.

- (1) prop⁰, ACTION², **xúe** [*study*], agent:2, theme:3, mod⁴:4, [2†]
 (2) THING¹
 (3) **yīngwén**³ [*English*]
 (4) QUALITY⁵, emphatic, **nǚlì**⁶ [*diligent*], theme:1

(8.19)a: tā xúe yīngwén xúe-de hěn nǚlì.

There is only one proposition produced in this derivation, which agrees with the intuition that the speaker of (8.19)a is making one statement and not two.

This is far from being the only construction in Chinese in which the syntax requires such reduplication of material. According to Chao (1968), the resultative construction, illustrated by (8.21)a is distinct (at least in historical origin) from the manner modification discussed above, although it is very similar in form. Similar too is the durative construction illustrated by (8.21)b. In both these cases, the same sort of transition as the copying rule introduced above will be applicable.

- (8.21) a. tā kàn shū kàn-de hěn lèi.
 she read book read-DE very tired
 She tired herself reading books.
- b. tā chī píngguo chī-le yīge xiǎoshí.
 she eat apple eat-LE one hour
 She ate apples for one hour.

Finally, I turn to a reduplication construction with even wider scope, the *A-not-A* question. This is a type of reduplicative yes-no question, and is illustrated by the examples in (8.22), taken from Calder (1986). They are synonymous questions, each with a different amount of lexical material repeated. This ranges from a single syllable-morpheme in (8.22)a, the whole of the first verb in (8.22)b, and the whole of the predicate in (8.22)c.¹⁰

- (8.22) a. tā xǐ bu xǐhuān hé píjiǔ?
 she li- not like drink beer
 Does she like to drink beer?
- b. tā xǐhuān bu xǐhuān hé píjiǔ?
 she like not like drink beer
- c. tā xǐhuān hé píjiǔ bu xǐhuān hé píjiǔ?
 she like drink beer not like drink beer

Although I will not present an analysis of this construction here, it is tempting to think that this sort of variable length, sub-lexical to phrasal, reduplication can be modelled by the similar rules, which can make reference to the interpretation of the first conjunct.

Radzinski (1990) claims that such yes-no questions can be indefinitely long, which appears to put the construction beyond the range of context-free grammar. Calder (1986), however, came to the opposite conclusion on the grounds that the status of the constructions that could be indefinitely extended is doubtful. Questions of formal power rest as always on whether certain strings are taken to be ungrammatical or grammatical, but unusable. As I argued in chapter 1 this is not a question that can be resolved empirically.

¹⁰This ignores the possibilities for unmatched copying with ellipsis in the second conjunct.

Leaving aside the arguments about formal power, the essential question for a conceptualist approach to grammar is to discover the operations that people are performing in their heads when they use language. The evidence from Chinese is that reduplication constructions do seem to enforce the identity of disconnected words. Such constructions have proved extremely difficult for approaches assuming that syntax can operate without reference to the interpretations being built. I argued in this section that such constructions can be modelled simply and elegantly in the dynamic grammar, where such partial interpretations are available, and are indeed the basis on which the syntactic rules are defined.

8.5 Japanese and other causatives

This brief section is based largely on Manning *et al.* (1998), which with great lucidity and thoroughness examines the status of Japanese causative verbs and proposes an HPSG analysis of them. As noted in the paper, the behaviour of these constructions is problematic for constituent structure accounts since a causative verb, such as *hasir-ase-ta* in (8.23)¹¹, shows all the signs of being a single word, and yet it appears to bestraddle two clauses. The sentence has the same ambiguity as the English translation with the locative phrase modifying either the causative action or the caused event.

- (8.23) Noriko ga Masaru ni gakkoo de hasir-ase-ta.
 Noriko NOM Masaru DAT school LOC run-CAUS-PAST
 Noriko made Masaru run at school.

Manning *et al.* argue against the position that such behaviour entails that causatives should not be viewed as single words, but rather as complex syntactic structures. Instead they provide a lexical solution for this dilemma, arguing that there is a single clause in such constructions, but that the lexical entry for the causative contains embedded argument structure, thus allowing the two readings of (8.23) and other multi-clausal behaviour.

From the perspective of the dynamic grammar presented here, the problem of words somehow belonging to two separate syntactic phrases does not arise. Words here do not occur in constituents at all, they are merely arranged one after the other and cause some meaning to be added to the interpretation. It appears that causatives can add, in effect, two separate, though linked, propositions rather than a single one. This is entirely in keeping with the rest of the grammar, and does not require any change in the architecture to accommodate it.

¹¹Manning *et al.* (1998: ex. 21.)

I will suppose that the transition schema for a causative verb is as follows. It will interpret the nominative constituent as agent of the causative and interpret a dative constituent as first argument of the caused action.

lexical entry	top of stack (S_n)	top of stack (S_{n+1})
{ causative, tense, head , Args: }	(X, prop)	(N, head , Args:) (X, prop, causative, tense, agent:)

The distribution and interpretation of the other elements in store will be carried out by separate principles, as in English.

The derivation of (8.23) in which the locative modifies the event of running is as follows. The locative constituent is inherited by the dependent “caused” event constituent in the normal way, where it is interpreted as a modifier.

- | |
|---|
| <p>(1) prop⁰, ACTION⁷, cause, past, agent:2, theme: 5, [4*, 3*, 2†]
 (2) Noriko¹, nom²
 (5) EVENT⁷, hasiru, theme:3, mod:4, [4†, 3†]
 (3) Masaru³, dat⁴
 (4) gakkoo⁵, loc⁶</p> |
|---|

(8.23): **Noriko ga Masaru ni gakkoo de hasir-ase-ta.**

However the locative constituent could just as well modify the matrix causative proposition, in which case we will have the following derivation.

- | |
|--|
| <p>(1) prop⁰, ACTION⁷, cause, past, agent:2, theme:5, mod:4, [4†, 3*, 2†]
 (2) Noriko¹, nom²
 (5) EVENT⁷, hasiru, theme:3, [3†]
 (3) Masaru³, dat⁴
 (4) gakkoo⁵, loc⁶</p> |
|--|

(8.23): **Noriko ga Masaru ni gakkoo de hasir-ase-ta.**

Of course, there are a great number of other issues that should be discussed with respect to constructions of this type, but this will have to await further research. The main contention being made here is simply that the difficulties with such constructions arise on account of the constituent structure mixing aspects of the two fundamental realities of language, the meaning and the sound.

Such causative constructions are found in many other languages and the same type of analysis will be presumed. In Hungarian, for example, the causative suffix *(t)at/(t)et* can be combined with the potential suffix *hat/het* as shown in (8.24).

- (8.24) Párizsban csinál-tat-hat-om.
 Paris-IN make-CAUS-POT-pres,1sg,def
 I can get it made in Paris.

It is possible to get at least two separate readings for this sentence, with the locative modifying either the action of making, or the potential. It appears to be difficult to find a reading which will split the causative and potential meanings as they tend to be close together in space and time. I will assume, however, that the single word *csináltathatom* provides information for three separate clauses: the potential, the causative and the action itself.

This would allow the following derivation of (8.24), for the reading where the action of making is to take place in Paris.

- (1) prop⁰, STATE², potential, pres, content:3, [2]
 (3) ACTION², causative, agent: ⟨1-sg⟩, theme:4, [2]
 (4) ACTION², **csinál**, agent: α , theme: ⟨def⟩, mod:2, [2+]
 (2) PLACE¹, **Parizs**, loc

(8.24): **Párizsban csináltathatom.**

We may wonder if the same analysis should be followed for English lexical causatives such as *raced* in (8.25), ie. should it be analysed as a single clause or two?

- (8.25) He raced the horses.

With some effort it does seem possible to find an example, (8.25), where there are two readings for the path modifier *from*. It is not clear, however, if these might not be two different ways of modifying the same clause.

- (8.26) He races horses from his headquarters in Suffolk.

Certainly, the lexical causative (8.27)a does not have the ambiguity of modifier attachment which is present in the explicit multiclausal causative construction (8.27)b.

- (8.27) a. He raced the horses for fun.
 b. He made the horses race for fun.

This is not the case for the morphological causative in Hungarian, where the equivalent sentence, (8.28) is genuinely ambiguous.

- (8.28) Szórakozás-ként fut-at-ta a lovak-at.
 fun-AS run-CAUS-past,3sg,def the horses-ACC
 He made the horses race for fun.

So it appears that, unless there is convincing evidence to the contrary, it would be problematic to analyse English causatives as creating two clauses.

8.6 Parasitic gaps

The term *parasitic gap* refers to a construction in which a single filler constituent is seemingly linked to more than one interpretation site in the same sentence. This is illustrated by the two examples in (8.29). It is a rather marginal construction in English, both in terms of its infrequency in usage and its marked nature.¹² As with the *that*-trace effect, the marginal nature of the phenomena makes it all the more revealing for models of syntax, as their analysis should fall out of existing rules, rather than having to be stipulated separately.

- (8.29) a. Which rebel did rivals of $_p$ assassinate $_$?
 b. Which book did he review $_$ without reading $_p$?

It can be established on a number of criteria that these different extraction sites are not of equal status: one of the sites corresponds to the position of a standard extraction and the other is in some way “parasitic” on this one (here marked $_p$). For example, the parasitic gaps may be filled with other constituents, as the following examples show.

- (8.30) a. Which rebel did rivals of Trotsky assassinate $_$?
 b. Which book did he review $_$ without reading the abstract?

Standard gaps however may not be removed, for example (8.31)a breaks the constraint on extraction from subjects.¹³ Extraction from modifiers placed after arguments does occur, although it is generally awkward, as in (8.31)b.

- (8.31) a. *Which rebel did rivals of $_$ assassinate Trotsky.
 b. ?Which book did he review the paper without reading $_$?

¹²It has often been observed that they are an acquired taste, in that their markedness quickly decreases with familiarity.

¹³Which in chapter 3 was argued not to be a separate constraint, but rather due to the constraint on inheriting store values from unsaturated constituents.

As noted by Engdahl (1983), the examples given above fall into two classes, depending on whether it is possible or not to substitute a coreferenced pronoun in place of the gap. As demonstrated in the following examples, (8.29)a above has an *obligatory* gap, in that it cannot be replaced by a pronoun, while (8.29)b has an *optional* gap.

- (8.32) a. *Which rebel_i did rivals of him_i/his_i assassinate _i .
 b. Which book_i did he review _i without reading it_i?

In an attempt to solve a number of problems with the HPSG approach to the phenomenon, Grover (1995) argues against treating parasitic gaps as a unified phenomenon. Instead, she divides them into two separate classes, a-type and c-type, the terms deriving from their being closely related to *anaphora* and *coordination* respectively. The a-type includes all of the obligatory gaps as defined in (1983), plus some optional gaps (Grover, 1995: 159).

This observation translates naturally into the dynamic approach: the two classes being reclassified as *pre-parasitic* or *post-parasitic* gaps, depending on whether the parasitic gaps occur before the standard gap or after it. These will now be considered separately.

8.6.1 Pre-parasitics

Looking at the above example of a pre-parasitic gap, (8.29)a, the question to be solved is that of how the subject can contain a gap, given that the non-clause final condition, discussed in chapter 3, forbids store values to be lowered into subjects and other non-final constituents. One solution would be to assume that store values can be lowered into non-final constituents as long as they are also lowered into the final constituent as well. This essentially parallels the standard GPSG and HPSG approaches to parasitic gaps, but as we shall see gives rise to many problems.¹⁴

However, the dynamic model allows us a second way to characterize pre-parasitic gaps. I shall assume instead that a copy of the stored constituent is not inherited by the non-final constituent, which in the examples discussed above is the subject. If this is the case then when we get to the site of the parasitic gap in the subject, the stored constituent which is to be interpreted here is not on the store of the active constituent, but is instead still on the store of the matrix clause which will be below it on the active stack. We will assume that there is a store-fixing transition that allows the interpretation of this constituent on a remote store, but does not cancel it from the store as would be the case in the normal

¹⁴Including the processing problem that one must then predict whether the non-final argument was going to contain a gap or not, which was used to motivate the restriction against gaps in non-final constituents in the first place.

interpretation of a stored constituent. As we know from crossover effects in binding, it is impossible to use a pronoun to refer a constituent in this position, which makes the parasitic gap strategy the only possibility. The transition schema to capture this will be the following.

transition rule	top of stack (S_n)	top of stack (S_{n+1})
pre-parasitic gap interpretation	(X1, arg:)	(X1, arg:Y)

	(X, [-Y])	(X, [-Y])

To give a concrete illustration of this, we may consider the state of the active stack during the derivation of the pre-parasitic gap sentence, (8.29)a, repeated here.

(8.29a) Which rebel did rivals of $_p$ assassinate $_$?

At the fifth word, the preposition *of*, the state of the active stack will be as follows.

- (4) *of*, content:
- (3) *rival*, theme:4
- (1) prop, past, content: , [+3,-2]

The *wh*-constituent (2), *which rebel*, is on the store of the proposition, but cannot be inherited into the subject as discussed above. The parasitic gap transition schema may apply in this position, filling in address of the *wh*-constituent as the content of the constituent headed by *of*, but without cancelling this constituent on the store of the proposition (1).¹⁵

The derivation of (8.29)a will then be as follows. The *wh*-constituent is interpreted twice, first pre-parasitically at *of* when it cannot be removed from the store, and finally in the standard way by *assassinate*.

- | |
|--|
| <ul style="list-style-type: none"> (1) prop⁰, Q¹, past³, content:5, [+3*, -2*] <li style="padding-left: 2em;">(5) ACTION⁶, <i>assassinate</i>, agent:3, patient:2, [+3†, -2†] <li style="padding-left: 4em;">(3) <i>rival</i>⁴, plur, theme⁴:4 <li style="padding-left: 6em;">(4) <i>of</i>⁵, content:2 <li style="padding-left: 4em;">(2) ref¹:-, <i>rebel</i>², sg |
|--|

(8.29)a: Which rebel did rivals of assassinate?

¹⁵Clearly, if the constituent were to be cancelled from the store of the proposition as it was interpreted here, then it would be impossible to prevent extraction solely from within a subject, or other non-final constituent.

8.6.1.1 Evidence against pre-parasitics being normal gaps

Looking at the evidence that the stored constituent is not simultaneously lowered into the subject, we may start with the fact that they are possible in positions where normal extraction would be problematic, for example from inside a relative clause, as in (8.33)a. This is predictable from the non-lowering account as the store does not have to be inherited down into the relative clause.

- (8.33) a. ?Which woman do men who meet $_p$ usually ask out $_?$
 b. ?*Which woman do you know men who meet $_?$

This is not restricted to subject-gapped relative clauses, which are often quite transparent to extraction in any case, but also applies to object-gapped relatives, resulting in the following pattern.

- (8.34) a. ?Which articles $_i$ had the students $_j$ you talked to $_j$ about $_{ip}$ failed to understand $_i?$
 b. *Which articles $_i$ did I meet the students $_j$ you had talked to $_j$ about $_i?$

Secondly, the pre-parasitic gap seems to be completely impossible to realize as a subject-gap, as witnessed by the contrast between the sentences in (8.35). This again would not be predicted if the pre-parasitic were a genuine gap, but follows from the fact that the subject switching rule introduced previously will not apply here, and there would have to be a special parasitic version of it to get the required interpretation.

- (8.35) a. ?Which woman did his thinking Paul was in love with $_p$ deeply embarrass $_?$
 b. *Which woman did his thinking $_p$ was in love with Paul deeply embarrass $_?$

Finally, there is the observation that the pre-parasitic gap must be nominal, and furthermore must create a discourse referent.¹⁶ Again this is unexpected given the parallel lowering hypothesis, but could be incorporated as a condition of the distant gap hypothesis.

- (8.36) a. *Where do the people who live $_p$ want to stay $_?$
 b. *How do those who built the boat $_p$ want everyone else to build it $_?$

¹⁶An observation attributed by Grover to Cinque (1990).

8.6.1.2 Pre-parasitics in non-subjects

The characterization of pre-parasitic gaps given above does not restrict them to appearing inside subjects, they may also appear inside other non-final arguments as in (8.37). Here the position at *of* entirely parallels that of the pre-parasitic gap positions in the previous examples.

(8.37) Which girl_{*i*} did you send a picture of _{*ip*}/*her_{*i*} to ?

The above example is taken from Engdahl (1983), where it is assumed that the second gap is the parasitic one. If this were the case then a postparasitic gap should also be able to occur in the following sentence, (8.38), contrary to fact.

(8.38) *Which slave did you sell to _{*p*}?

This sentence, (8.38), will not be allowed by the present grammar: the first gap cannot be parasitic, as the constituent in store is directly accessible and thus will be cancelled. If the stored constituent is interpreted here as the beneficiary, then there cannot be a final parasitic gap as this is a position where a reflexive may occur, as shown in (8.39). As we shall see reflexives are in complementary distribution with both pronouns and parasitic gaps.

(8.39) Which slave_{*i*} did you sell _{*i*} to himself_{*i*}/*him_{*i*}?

The approach also predicts that both the subject and a non-final argument should be able to contain pre-parasitic gaps in the same sentence, and although not the most stylish sentence ever written, (8.40) certainly appears to be grammatical.

(8.40) Which rebel leader did rivals of _{*p*} give pictures of _{*p*} to ?

8.6.2 Post-parasitics

Finally, we return to the question of parasitic gaps appearing in positions after the interpretation site of the stored constituent. The earlier example (8.29)b is repeated here.

(8.29b) Which book did he review without reading _{*p*} ?

The state of the active stack after the word *reading* is added will be as follows.

- (6) **read**, agent:3, theme:
 (5) MANNER, **without**, neg, content:6
 (4), **review**, agent:3, theme:2, mod:5, [2†]

If the parasitic gap rule is not sensitive to whether the address on the store of a lower constituent has been interpreted elsewhere then the rule will apply again in this instance. It will thus allow the stored, but now interpreted constituent, to be taken as the theme of *read*. A normal pronoun will also be possible in this position since the constituent *which book* has already been interpreted and can therefore stand as a possible antecedent.

The complete derivation of the sentence will be as follows.

- (1) prop⁰, Q¹, past³, content:4, [+3*, -2*]
 (4) ACTION⁵, **review**, agent:3, theme:2, mod⁶:5, [+3†, -2†]
 (3) MALE⁴, sg
 (2) ref¹:-, **book**², sg
 (5) MANNER⁶, **without**, neg, content:6
 (6) ACTION⁷, ing, **read**, agent:3, theme:2

(8.29)a: **Which book did he review without reading?**

Again this predicts that the parasitic gap may be indefinitely far from the original gap, but the derivation of both versions of (8.41) is identical until the parasitic gap or the pronoun is reached. Therefore there is predicted to be no garden-path effect when the interpretation point is reached, whether the pronoun or the gap is chosen, which is indeed the case.

(8.41) Which book_{*i*} did he agree to review _{*i*} without admitting to anyone that he never actually intended to read _{*i*}p/it_{*i*} ?

As I noted previously a further constraint is that the gap cannot appear in a position where a reflexive pronoun is possible. This not only rules out examples of post-parasitic gaps in argument position as (8.39) above, but also examples where the gap is embedded inside a genuine argument, as in (8.42).

(8.42) Which box_{*i*} did you shove _{*i*} inside itself_{*i*}/* _{*i*} ?

Again the post-parasitic gap appears only to be possible with referential nominal antecedents, which would be unexpected if it was either a genuine gap or else a coordinate structure.

- (8.43) a. *?Where_i do you think John moved to _i without really wanting to live _i ?
 b. Which town_i do you think John moved to _i without really wanting to live in _i ?

Interestingly, the parasitic gap can pick this definite referent out of a constituent and ignore its syntactic form, resulting in the following kind of mismatch noted by Tait (1988).

- (8.44) {To whom_j}_i did Mortimer faithfully continue to write _i after seeing _j only once?

8.6.3 Summary

The central assumption in this analysis of the parasitic gap construction is that the parasitic gap is much the same thing as a standard gap, being a method of referring to a constituent in the discourse which is presently in store. The difference between the two is that in the case of a parasitic gap this is not the store of the present active constituent, but rather one below it on the active stack.

In the case of parasitic gaps which occur before the real gap, which I have referred to as *pre-parasitic* gaps, this will be the only way that reference can be made to the constituent: a pronoun cannot be used since the constituent has not yet received an interpretation. Parasitic gaps which occur after the interpretation of the constituent in store, *post-parasitic* gaps, are held to arise from the same rule as pre-parasitics, but in this position they will be optional, as a pronoun will also be possible. Therefore, in contrast to Grover (1996), I take parasitic gaps to be essentially a single phenomenon.

8.7 Accounting for processing limitations

As mentioned in Chapter 1, there are grammatical constructions that defeat the human processor, no matter which control strategy is employed. In this section I will examine the most notorious such construction, repeated centre-embedding, and attempt to determine if the grammar can be used to pinpoint more precisely why this should be so problematic.

It is clear that the transition rules for relative clauses introduced in chapter 2 will allow the derivation of sentences of the following type, exhibiting an indefinite degree of centre-embedding, or to be more precise self-embedded object-gapped relative clauses. In (8.45) there are two relative clauses and the sentence is already impossible to use for communication.¹⁷

¹⁷It is well-known that other factors, such as increasing the length of the phrases, may ameliorate the

(8.45) The man_i the cat_j the dog chased _j bit _i died.

Let us consider the sequence of states in the derivation of (8.45). After the sixth word *dog* the interpretation state will be as follows.

<p>(1) prop⁰, [2] (2) def¹, MALE², sg, man, mod³:3 (3) prop³, [+4,-2] (4) def³, cat⁴, sg, mod⁵:5 (5) prop⁵, [+6,-4] (6) def⁵, dog⁶, sg</p>

(8.45): S₆ **The man the cat the dog** || chased bit died.

A snapshot of the active stack at this point shows that three nominal constituents have been established, but none of them has been given a semantic interpretation. There are also three propositions that have been initiated, none of which is yet to receive its head finite verb.

- (6) def, **dog**, sg
- (5) prop, [+6,-4]
- (4) def, **cat**, sg, mod:5
- (3) prop, [+4,-2]
- (2) def, **man**, sg, mod:3
- (1) prop, [2]

The remainder of the derivation is as follows. The finite verbs simply fill in the empty propositions, in so doing also completing the nominal constituents which come off the stack.

<p>(1) prop⁰, EVENT⁹, past, die, theme:2, [+2*] (2) def¹, MALE², sg, man, mod³:3 (3) prop³, ACTION⁸, past, bite, agent:4, patient:2, [+4†,-2†] (4) def³, cat⁴, sg, mod⁵:5 (5) prop⁵, ACTION⁷, past, chase, agent:6, patient:4, [+6†,-4†] (6) def⁵, dog⁷, sg</p>
--

(8.45): **The man the cat the dog chased bit died.**

badness of embedded relatives such as (8.45), but they are not directly relevant here as I shall try to compare like with like.

It might seem reasonable to conjecture from this that an increase in the number of semantically unattached constituents is directly related to an increase in the processing difficulty of the sentence. It would seem, perhaps, that three unintegrated constituents is approaching the limit of human tolerance and four quite beyond it. Konieczny (1995: 228) comes to a similar conclusion and enshrines it in the Immediate Semantic Integration Principle, which is summed up as “unintegrated referential objects (refos) induce memory load”.

However, let us consider the derivation of sentence (8.46), a sentence with a *wh*-topic and a subject modified by a single subject relative clause.

(8.46) Which man_i did the cat_j the dog chased $_j$ bite $_i$?

When the word *dog* is reached this sentence also has three uninterpreted referential constituents, one in store and two in subject position.

<p>(1) prop^0, Q^1, past^3, content: , [+3,-2] (3) def^4, cat⁵, sg, mod^6:4 (4) prop^6, [+5,-3] (5) def^6, dog⁷, sg (2) def^1, ref:-, MALE^2, sg, man</p>
--

(8.46): S_6 **Which man did the cat the dog** || chased bite?

The state of the active stack at this point is shown here.

- (5) def , **dog**, sg
- (4) prop , [+5,-3]
- (3) def , **cat**, sg , mod :4
- (1) prop , past , content : , [+3,-2]

If we were to assume that it was the number of uninterpreted referential constituents that led to processing overload, then we would be unable to explain why (8.46) is so much easier to understand than (8.45) above. Indeed, intuitively, it seems little or no worse than a sentence with a single subject relative such as (8.47).¹⁸

(8.47) The cat_i the dog chased $_i$ bit that man.

¹⁸Another reason for doubting this theory is that it does not seem to accord with the facts of languages where three or more such objects routinely appear before the verbs that interpret them, that is in head-final languages such as Japanese or German.

After detailed discussion of the problem, Hudson (1996) concludes that the processing difficulty is not related so much to the syntactic structures involved as to the semantic structures being constructed. In particular, he conjectures that one of the essential problems with sentences such as our first example (8.45) is that the three nominal phrases are unfinished until the verbs are finally added, that is there senses are still being modified (or in the case of the final one have the potential to be modified) by relative clauses. If this is the source of the difficulty it would explain why (8.46) is no more problematic than (8.47), for in neither sentence are more than two referential phrases simultaneously incomplete.

We can test this theory by looking at the derivation of a sentence similar to our original problematic sentence, (8.45), but with the second relative clause being a subject relative rather than an object relative. Hudson (1996) shows from various sources, including his own experiments, that this tends to make the sentence slightly easier to process, though the difference is only marginal. Intuitively it certainly sounds more difficult than the examples with single relative modification, (8.46) and (8.47) above.¹⁹

(8.48) The man_i the cat that chased the dog bit _i died.

- (1) prop⁰, EVENT¹⁰, past, **die**, theme:2, [+2†]
 (2) def¹, MALE², sg, **man**, mod³:3
 (3) prop³, ACTION⁹, past, **bite**, agent:4, patient:2, [+4†, -2†]
 (4) def³, **cat**⁴, sg, mod⁵:5
 (5) prop⁵, ACTION⁶, past, **chase**, agent:4, patient:6, [+4†]
 (6) def⁷, **dog**⁸, sg

(8.48): **The man the cat that chased the dog bit died.**

Below on the left, I show the state of the active stack when the word *dog* is added on this occasion. Again there are three referential constituents that are potentially incomplete, which would certainly be predicted to cause difficulty. This time, however, two of these constituents have received an interpretation via the verb *chased*, which may be a factor that contributes to it being better than (8.45). This also means that three constituents can immediately be removed from the stack at that transition, to leave it in the state shown on the right. In the corresponding place in (8.45), only the top constituent, *dog*, could be removed from the stack, which again may be a factor in the sentence being more acceptable.

¹⁹I find myself that it is understandable on the second reading, that is that I tend to “garden-path” on the first. It would be interesting to find out if this is a common reaction to it.

- | | |
|---|---------------------------------|
| (6) def, dog , sg | |
| (5) prop, chase , agent:4, patient:6 | |
| (4) def, cat , sg, mod:5 | |
| (3) prop, [+4,-2] | (3) prop, [+4,-2] |
| (2) def, man , sg, mod:3 | (2) def, man , sg, mod:3 |
| (1) prop, [2] | (1) prop, [2] |

As Hudson (1996) argues, it may also be inferred from the approach to the problem considered here that the situation should be improved if the relative clauses involved are non-restrictive rather than restrictive, for then the sense of the referential object is already complete and the non-restrictive relative is merely adding a separate proposition. This is borne out by sentence (8.49), where instead of self-embedding restrictive relatives, we have a restrictive inside a non-restrictive one. The result, if read with the correct intonation, seems to be hardly worse than sentences with a single relative clause such as (8.46) and (8.47). This is to be expected if the crux of the problem lies in the difficulty of having multiple referential expressions which are semantically incomplete.

(8.49) Peter_i, who the cat_j the dog chased _j bit _i, died.

Another piece of evidence used by Hudson (1996) to support this explanation is the fact that if the third nominal expression introduced in sentences such as (8.45) is a pronoun or other constituent with an obvious referent, then the sentence becomes easier to interpret, as seems to be the case in (8.49) below. This will be because there are in this case only ever two semantically incomplete referential expressions.²⁰

(8.51) The man_i the cat_j I like _j bit _i died.

It is clear, however, that this cannot be the whole story. If it were then we would also expect right-recursive structures such as the succession of embedded subject relative clauses to lead to processing breakdown, which they clearly do not.²¹

(8.52) The man that saw the cat that bit the dog died.

²⁰Including this real-life example, heard by Geoffrey Sampson in a discussion of the difficulty of relative clauses, and quoted by Hudson (1996: 296).

(8.50) ...but don't you find that sentences_i that people_j you know _j produce _i are easier to understand?

Clearly, there are a number of factors in the constructed example (8.49) that make it worse than this real one, but it is still clearly better than the corresponding (8.45).

²¹Indeed, as lovers of insect-ingesting fairy tales will know, the sequence can continue indefinitely without leading to processing difficulties.

Let us look at the derivation of (8.52) and the state of the active stack at *dog* that can be inferred from it.

- (1) prop⁰, EVENT¹¹, past, **die**, theme:2, [+2†]
 (2) def¹, MALE², sg, **man**, mod³:3
 (3) prop³, EVENT⁴, past, **see**, exper:2, theme:4, [+2†]
 (4) def⁵, **cat**⁶, sg, mod⁷:5
 (5) prop⁷, ACTION⁸, past, **bite**, agent:4, patient:6, [+4†]
 (6) def⁹, **dog**¹⁰, sg

(8.52): **The man that saw the cat that bit the dog died.**

- (6) def, **dog**, sg
 (5) prop, **bite**, agent:4, patient:6
 (4) def, **cat**, sg, mod:5
 (3) prop, **see**, exper:2, theme:4
 (2) def, **man**, sg, mod:3
 (1) prop, [2]

The essential difference between the centre-recursive (8.45) and the right-recursive (8.52) is that in the latter the referential items may not be totally complete, in that the restrictive relative clauses are not finished, but at least these referential items have all received semantic roles within these relative clauses. It appears that this counts as being *semantically interpreted* for the purposes of calculating processing difficulty.

8.7.1 Summary

A number of initial conjectures spring to mind when faced with sentences such as (8.45) as to why they are so difficult to interpret. The commonly-given answer is that they exceed memory limitations in some way, but precisely which aspect of them causes the trouble has been a matter of dispute. We have seen from considering the derivation of these sentences with the dynamic grammar that there is no link to the number of constituents on the active stack, nor even to the number of constituents that are not interpreted. Following Hudson (1996), it seems that a good case can be made for processing difficulty being related to the number of referential constituents that have an incomplete sense, in that they are modified, or potentially modified, by relative clauses in which they have received no interpretation.

Using the present dynamic model of syntax, it would be easy to keep track of how many referential objects were incomplete in the above sense. Whether it would be of benefit to a processor to include such constraints is a subject for future experiment, but

it would certainly be interesting to attempt to model the psycholinguistic data and make verifiable predictions about human behaviour.

Chapter 9

Conclusion

9.1 Achievements

In this thesis I have proposed that we “recut the cake” of language interpretation, while respecting the Chomskyan arguments for the modularity of the process and the need to distinguish a syntactic competence from its use in processing. In particular, I have argued that if our model of competence, the system of syntactic rules, takes into account the “left-to-right” nature of language then it leads to simple and general analyses of many seemingly complex syntactic phenomena, as well as simplifying the relation between the grammar and the processor that uses it. To this end, I have outlined a dynamic grammar, where syntactic rules specify how information associated with each word is added to the growing interpretation of a sentence.

One recurrent theme of this thesis has been that the adoption of an explicitly left-to-right model frees the syntax from having to circumvent many problems which arise from the use of constituent structures.

“...the basic defect of a phrase structure grammar is that it confuses information about constructions ... with information about the realization of constructions... In that sense it confuses what is ‘deep’ (the construction) with what is ‘surface’ (the realisation), to the inevitable detriment of the former. But it is not clear that the remedy lies in establishing two levels of phrase structure instead of one.... If we multiply levels of phrase structure we multiply confusions between what is really ‘deep’ and what is really ‘surface’, instead of eliminating them.

No adequate alternative has been developed.” (Matthews 1981, 92-93).

The dynamic model does not confuse the distinction between hierarchical non-linear meaning and linear sound by attempting to link them with a third level which contains aspects of both. Even if constituency is used only to specify the way that meanings can

be constructed and is not held to form a level of representation, it is still unnecessary if one insists, by adopting an appropriate formalism, that meaning can only be constructed word-by-word.

I have thus argued that it has advantages in terms of simplicity and accuracy in many areas where there is a clash between word syntactic constituency. Among these may be mentioned non-constituent coordination, discontinuous constituency, scrambling, and lexical items with multiclausal meaning (eg. causatives). I have also shown that the left-to-right grammar offers a natural interpretation of *wh*-movement phenomena in English, and offers new ways of looking at pied-piping and parasitic gaps. This approach also links in with the work done on crossover effects in a dynamic framework by Kempson and her colleagues.¹

There is no escaping the fact that building models of human syntax is essentially a speculative enterprise, and that the best strategy to follow is to minimize the unobservables of grammar, at the same time keeping it compatible with the observable evidence. As well as being parsimonious in its excising of syntactic constituents, it can also be argued that the model is plausibly “low-tech”, or in Culicover’s terms “concrete minimalist”.

“I sketch out an alternative construal of syntactic theory that is responsive to the dynamical issues. I will suggest that such a theory is “concrete minimalist”, in the sense that it eschews formal devices and representations that do not correspond to perceptual and cognitive distinctions available to a plausible learner.

...the syntactic analysis of an expression in concrete minimalism simply incorporates the smallest amount of structure that a learner needs to assume in order to account for the sound-meaning correspondence...” (Culicover, 1998: 49-50)

The only assumptions made here are a network of semantically-related conceptual constituents, however this be represented, a handful of syntactic features on and syntactic relations between these constituents, a stack to hold constituents at the centre of current attention, and transition rules that may be thought of as routinized operations, sometimes highly lexically-specific, for adding new information to the network.

9.2 Prospects

As I have made clear throughout this thesis, the present dynamic grammar is no more than an exploration of some of the possibilities that such a new perspective on the syntactic process gives. In particular, I have assumed a minimally-expressive representation of

¹For example, Kempson & Gabbay (1998).

“meaning states” in order to focus on the syntax, that is the process by which these states are constructed. As this process of exploration has gone on, it has become increasingly clear that more sophisticated representations are needed in order to be able to specify this syntactic process accurately.

A prime case of such a shortcoming is the need for the representation of reference in the conceptual constituents. This would seem to be crucially important in working out the syntax of copular constructions, including clefts, and will probably also be of great benefit in areas such as relative clauses and other types of modification. A related problem, and one not touched on in this thesis, is the representation of quantification in the conceptual structures. This is a question that has to be addressed by all approaches assuming such a level, and it appears that the representations needed, and indeed the representations possible, will be quite different from those assumed in truth-theoretical approaches.

Another area where much work needs to be done is in embedding the model of syntactic competence presented here in an overall model of language processing. This involves specifying a *control structure*, which determines which path (or paths) to take at any point in the parse and when these paths are to be abandoned. It should be noted that not only does the grammar give the possible moves that can be made, it also specifies the possible “parse states” themselves for they are the basic objects of the grammar: the interpretations of the previous string of words.

Such a paring away of the extra assumptions needed to turn the grammar into a processor is clearly desirable, as Steedman argues in the following passage.

“Competence grammar and performance mechanism are in the end a package deal. Any claim about competence grammar is also a claim about the entire package. The theory of evolution is not mocked, and one day, a reckoning will be demanded. On that day, the linguistic theories that have survived the ordeal of descriptive adequacy will be judged not merely on their purity and parsimony as theories of competence, but on their explanatory value as part of the package. We have already noted that all theories will require *something* more, in the form of a language-independent mechanism for resolving local ambiguity, or grammatical non-determinism, together with a language-independent algorithm and automaton. But if a theory of competence requires much more than that... then these notions will weigh against it. If there is another theory that requires fewer such assumptions, perhaps even *no* further assumptions... then the scales may tilt in its favour.

That day is not yet upon us, but it is good on occasion to meditate upon one’s latter end.” Steedman (1992: 54-55)

Another central objective must be to show that such a dynamic grammar can be used to form a probabilistic language model trained from data. The gauntlet laid down by Brown *et al* (1992), challenging linguistically sophisticated models to make better predictions than

purely statistical ones, must be taken up if language processing and theoretical linguistics are to stay in contact, which I would argue would be a healthy position for both. I have also argued that this is the only way that linguistic models can be assessed in an objective way, and so it remains even more imperative that this step be taken.

Abandoning assumptions is hard, but I hope that this thesis has at least demonstrated that generative grammar, with real formalized models, *is* possible without assuming syntactic structure. Whether any of the particular speculations made in this thesis are heading along the the right lines only time, and experiment, will tell.

References

- Ades, Anthony E. & Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4, 517-558.
- Apresjan, Ju. D. 1973. *Principles and methods of contemporary structural linguistics*. The Hague: Mouton.
- Birner, Betty J. 1996. *The discourse function of inversion in English*. New York: Garland.
- Bod, Rens & Remko Scha. 1996. Data-oriented language processing. An overview. ILLC Technical Report LP-96-13, University of Amsterdam.
- Borsley, Robert D. 1991. *Syntactic theory: a unified approach*. Edward Arnold: London.
- Borsley, Robert D. 1996. *Modern phrase structure grammar*. Oxford: Blackwell.
- Branigan, Holly Parima. 1995. Language processing and the mental representation of syntactic structure. PhD thesis. University of Edinburgh.
- Bresnan, Joan. 1994. Linear order vs. syntactic rank: evidence from weak crossover. In *Papers from the 30th Regional Meeting of the Chicago Linguistic Society, vol. 1*, 57-89.
- Brown, Peter F., Stephen L. Della Pietra, Vincent J. Della Pietra, Jennifer C. Lai, and Robert L. Mercer. 1992. An estimate of an upper bound for the entropy of English. *Computational Linguistics* 18, 31-40.
- Calder, Jonathan. 1986. Reduplicated constructions in Chinese and questions of generative power. *York Papers in Linguistics* 12, 43-66.
- Calcagno, Mike. 1998. Some thoughts on tough movement. Ms. University of Tübingen.
- Cann, Ronnie & Mary Tait. 1989. Free relatives revisited. Ms. University of Edinburgh.
- Chafe, Wallace L. 1970. *Meaning and the structure of language*. University of Chicago Press.
- Chao, Yuan Ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. MIT Press.

- Chomsky, Noam. 1969. Some empirical assumptions in modern philosophy of language. In Sidney Morgenbesser, Patrick Suppes & Morton White (eds.) *Philosophy, science, and method: essays in honor of Ernest Nagel*, 260-285. New York: Martin's Press.
- Chomsky, Noam. 1975. *Reflections on language*. London: Fontana.
- Chomsky, Noam. 1986. *Knowledge of language: its nature, origin and use*. New York: Praeger.
- Chomsky, Noam. 1987. *Language and problems of knowledge: the Managua lectures*. MIT Press.
- Chomsky, Noam. 1991. Linguistics and cognitive science: problems and mysteries. In Asa Kasher (ed.) *The Chomskyan turn*, 26-53. Oxford: Blackwell.
- Chomsky, Noam. 1995. *The minimalist program*. MIT Press.
- Church, Kenneth Ward. 1980. *On memory limitations in natural language processing*. Doctoral dissertation, MIT.
- Cinque, Guglielmo. 1990. *Types of A-dependencies*. MIT Press.
- Cottell, M. Siobhán. 1997. A raising analysis of cleft constructions. Paper given at LAGB Autumn Meeting, University of Hertfordshire.
- Crain, S. & Mark Steedman. 1985. On not being led up the garden path. In David Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.) *Natural language parsing: psycholinguistic, computational and theoretical perspectives*. Cambridge University Press.
- Culicover, Peter W. 1993. Evidence against ECP accounts of the *that-t* effect. *Linguistic Inquiry* 24, 557-561.
- Culicover, Peter W. 1998. The minimalist impulse. In Peter W. Culicover and Louise McNally (eds.) *The Limits of syntax*, 47-78. New York: Academic Press.
- Dahl, V. & M.C. McCord. 1983. Treating coordination in logic grammars. *Computational Linguistics* 9, 69-91.
- Damerau, Frederick J. 1971. *Markov models and linguistic theory*. The Hague: Mouton.
- Davidson, Andrew. 1997. Psychological reality revisited. *SOAS Working Papers in Linguistics* 7, 35-63.
- Dowty, David R. 1996. Toward a minimalist theory of syntactic structure. In Harry Bunt & Arthur van Horck (eds.) *Discontinuous constituency*. Berlin: de Gruyter. [Cataloguing]
- Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6, 5-34.
- Engdahl, Elisabet. 1985. *Constituent questions*. Dordrecht: Reidel.
- Fauconnier, Gilles. 1994. *Mental spaces: aspects of meaning construction in natural language*. Cambridge University Press.

- Fowler, H.W. 1926. *A dictionary of modern English usage*. Oxford University Press.
- Fraser, Norman M. 1989. Parsing and Dependency Grammar. *UCL Working Papers in Linguistics, Vol I*, 296-320.
- Freedman, David. 1972. *Markov chains*. San Francisco: Holden-Day.
- Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* **12**, 155-184.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy. 1989. *Mind, code and context: essays in pragmatics*. Hillsdale, NJ: Lawrence Erlbaum.
- Givón, Talmy. 1990. *Syntax: A functional-typological introduction*, Vol II. Amsterdam: Benjamins.
- Givón, Talmy. 1995. *Functionalism and grammar*. Amsterdam: Benjamins.
- Goodall, Grant. 1987. *Parallel structures in syntax: coordination, causatives, and restructuring*. Cambridge University Press.
- Grimshaw, Jane B. 1977. *English wh-constructions and the theory of grammar*. Doctoral dissertation, University of Massachusetts.
- Grover, Claire. 1995. *Rethinking some empty categories: missing objects and parasitic gaps in HPSG*. PhD thesis, University of Essex.
- Haegemann, Liliane. 1994. *Introduction to Government and Binding theory*. 2nd edition. Oxford: Blackwell.
- Harris, Zellig. 1991. *A theory of language and information: a mathematical approach*. Oxford: Clarendon Press.
- Hausser, Roland. 1989. *Computation of language: an essay on syntax, semantics and pragmatics in natural man-machine communication*. Berlin: Springer-Verlag.
- Hausser, Roland. 1992. Complexity in Left-Associative Grammar. *Theoretical Computer Science* **106**, 283-308.
- Heaton, J. B. 1976. *Writing English language tests*. London: Longman.
- Heycock, Caroline. 1995. *Layers of predication: the non-lexical syntax of clauses*. New York: Garland.
- Heycock, Caroline & Anthony Kroch. 1998. Pseudocleft connectedness: implications for the LF interface level. To appear in *Linguistic Inquiry*.
- Huang, Chu-Ren. 1983. Reduplication reanalysed: a fragment of Chinese PS rules. Cor-

- nell University, ms.
- Hudson, Richard. 1984. *Word grammar*. Oxford: Blackwell.
- Hudson, Richard. 1990. *English word grammar*. Oxford: Blackwell.
- Hudson, Richard. 1996. The difficulty of (so-called) self-embedded structures. *UCL Working Papers in Linguistics* 8, 283-314.
- Hudson, Richard. 1997. German partial VP fronting. Unpublished manuscript. University College, London.
- Hudson, Richard. 1998. *English grammar*. Routledge.
- Ingria, Robert J. P. 1989. The limits of unification. In *Proceedings of the 28th Annual Meeting of the ACL*, 194-204.
- Iwakura, Kunihiro. 1976. Another constraint on sentential subjects. *Linguistic Inquiry* 7, 646-52.
- Jackendoff, Ray. 1990. *Semantic structures*. MIT Press.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. MIT Press.
- Jespersen, Otto. 1927. *A modern English grammar on historical principles*, Vol III. London: George Allen & Unwin.
- Jespersen, Otto. 1940. *A modern English grammar on historical principles*, Vol V. Copenhagen: Munksgaard.
- Kaplan, Ronald M. 1972. Augmented transition networks as psychological models of sentence comprehension. *Artificial Intelligence* 3, 77-100.
- Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In P. Whitelock, M.M. Wood, H.L. Somers, R. Johnson & P. Bennet (eds.) *Linguistic theory and computer applications*, 149-181. London: Academic Press.
- Karttunen, Lauri. 1989. Radical lexicalism. In Mark Baltin & Anthony Kroch (eds.) *Alternative conceptions of phrase structure*. University of Chicago Press, 43-65.
- Kempson, Ruth. 1995. Crossover phenomena: a dynamic perspective. *SOAS Working Papers in Linguistics* 5, 135-186.
- Kempson, Ruth, Wilfred Meyer Viol & Dov Gabbay. 1996. Syntactic computation as labelled deduction: WH a case study. *SOAS Working Papers in Linguistics* 6, 179-218. Submitted to R. Borsley, I. Roberts (eds.) *Syntactic categories*. Academic Press.
- Kempson, Ruth & Dov Gabbay. 1998. Crossover: a unified view. *Journal of Linguistics* 34, 73-124.
- Kiss, Katalin É. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.
- Konieczny, Lars. 1996. Human sentence processing: a semantics-oriented approach. Doc-

toral dissertation. Freiburg.

Kuno, Susumu. 1973. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4, 363-385.

Kuno, Susumu & Takami, Ken-ichi. 1993. *Grammar and discourse principles: functional syntax and GB theory*. Chicago University Press.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar* vol. I. Stanford University Press.

Langacker, Ronald W. 1988. An overview of cognitive grammar. In Brigida Rudzka-Ostyn (ed.) *Topics in cognitive linguistics*, 3-48, Amsterdam: Benjamins.

Langacker, Ronald W. 1991. *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Lencho, Mark. 1992. Evidence that "to" is a complementizer. Handout of talk at the 8th Workshop on Comparative Germanic Syntax, University of Tromsø.

Lyons, John & Roger W. Wales (eds.). 1966. *Psycholinguistics papers: the proceedings of the 1966 Edinburgh Conference*. Edinburgh University Press.

Maling, Joan & Annie Zaenen. 1982. Scandinavian extraction phenomena. In Pauline Jacobson & Geoffrey K. Pullum (eds.) *The nature of syntactic representation*. Dordrecht: Reidel.

Manning, Christopher D., Ivan A. Sag & Masayo Iida. 1998. The lexical integrity of Japanese causatives. To appear in Robert Levine and Georiga Green (eds.) *Readings in modern phrase structure grammar*. Cambridge University Press.

Marcus, Mitchell P. 1980. *A theory of syntactic recognition for natural language*. MIT Press.

Marslen-Wilson, William. 1973. Linguistic structure and speech shadowing at very short latencies. *Nature* 244, 522-523.

Marslen-Wilson, William & Lorraine Komisarjevsky Tyler. 1987. Against modularity. In Jay L. Garfield (ed.) *Modularity in knowledge representation and natural-language understanding*, 37-62. MIT Press.

Martinet, André. 1996. Functional grammar: Martinet's model. Entry in Keith Brown & Jim Miller (eds.) *Concise Encyclopedia of syntactic Theories*, 92-97. Oxford: Elsevier.

Matthews, Peter H. 1981. *Syntax*. Cambridge University Press.

Matthews, Peter H. 1993. *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge University Press.

McCawley, James D. 1971. Tense and time reference in English. in C. J. Fillmore and D. T. Langendoen, eds., *Studies in Linguistic Semantics*, Holt, Rinehart & Winston: New York. pp. 96-113.

- McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* **13**, 91-106.
- McCawley, James D. 1987. Some additional evidence for discontinuity. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.) *Discontinuous constituency*, 185-200. Orlando: Academic Press.
- McCawley, James D. 1988. *The syntactic phenomena of English*. Vols I & II. Chicago University Press.
- Milward, David. 1990. Coordination in an axiomatic grammar. In *COLING-90*, 207-212.
- Milward, David. 1992. Dynamics, dependency grammar and incremental interpretation. In *COLING-92*, 1095-1099.
- Milward, David. 1994a. Non-constituent coordination: theory and practice. In *COLING-94*.
- Milward, David. 1994b. Dynamic dependency grammar. *Linguistics and Philosophy* **17**, 561-605.
- Nanni, D. 1978. Three remarks on pied piping. *Linguistic Inquiry* **9**, 310-318.
- Netter, Klaus. 1991. Clause union phenomena and complex predicates in German. In Klaus Netter & Mike Reape (eds.) *Clause structure and word order variation in Germanic* 1-33. DYANA report, R1.1.B. University of Edinburgh.
- Nuyts, Jan. 1992. *Aspects of a cognitive-pragmatic theory of language: on cognition, functionalism and grammar*. Amsterdam: Benjamins.
- Oerhle, Richard T. 1987. Boolean properties in the analysis of gapping. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.) *Discontinuous constituency*, 201-240. Orlando: Academic Press.
- Pesetsky, David. 1996. Optimality theory and syntax: movement and pronunciation. In Diana Archangeli and D. Terence Langendoen (eds.) *Optimality theory: an overview*, 134-170. Oxford: Blackwell.
- Phillips, Colin. 1996. *Order and structure*. PhD thesis, MIT.
- Pickering, Martin & Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* **6**, 229-259.
- Pinker, Steven. 1994. *The language instinct*. New York: William Morrow
- Pollard, Carl & Sag, Ivan A. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press and CSLI Publications.
- Postal, Paul M. 1971. *Cross-over phenomena*. New York: Holt, Rinehart and Winston.
- Postal, Paul M. 1972. The best theory. In Stanley Peters (ed.) *Goals of linguistic theory*, 131-170. Eaglewood Cliffs, NJ: Prentice-Hall.

- Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* **25**, 63-117.
- Postal, Paul & Geoffrey K. Pullum. 1988. Expletive noun phrases in subcategorized positions. *Linguistic Inquiry* **19**, 635-670.
- Pullum, Geoffrey K. 1982a. Syncategorematicity and English infinitival *to*. *Glossa* **16**, 181-215.
- Pullum, Geoffrey K. 1982b. Free word order and phrase structure rules. *NELS* **12**, 209-220.
- Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* **62**, 751-773.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Radford, Andrew. 1988. *Transformational grammar*. Cambridge University Press.
- Radzinski, Daniel. 1990. Unbounded syntactic copying in Mandarin Chinese. *Linguistics and Philosophy* **13**, 113-127.
- Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.) *German in Head-Driven Phrase Structure Grammar*, 151-198. Stanford: CSLI.
- Reape, Mike. 1996. Getting things in order. In Harry Bunt & Arthur van Horck (eds.) *Discontinuous Constituency*. Berlin: Mouton de Gruyter.
- Ringen, Jon D. 1980. Linguistic facts: a study of the empirical scientific status of transformational generative grammars. In Thomas A. Perry (ed.) *Evidence and argumentation in linguistics*, 97-132. Berlin: de Gruyter.
- Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. MIT Press.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. PhD thesis, MIT. [published as *Infinite syntax!*, Norton, NJ: Ablex (1986)]
- Ross, John Robert. 1969. Auxiliaries as main verbs. *Journal of Philosophical Linguistics*, 1.77-102.
- Ross, John Robert. 1972. Slifting. In Maurice Gross, Morris Halle & Marcel-Paul Schützenburger (eds.) *The formal analysis of natural languages*, 133-169. The Hague: Mouton.
- Sag, Ivan. 1991. Linguistic theory and natural language processing. In Ewan Klein & Frank Veltman (eds.) *Natural language and speech: symposium proceedings, Brussels, November 26/27, 1991*, 69-83.
- Sag, Ivan. 1997. English relative clause constructions. *Journal of Linguistics* **33**, 431-483.

- Schank, Roger C. 1975. *Conceptual information processing*. Amsterdam: Elsevier.
- Schütze, Carson T. 1996. *The empirical base of linguistics : grammaticality judgments and linguistic methodology*. University of Chicago Press.
- Sells, Peter, Stuart M. Shieber & Thomas Wasow (eds.). 1991. *Foundational issues in natural language processing*. MIT Press.
- Sgall, Petr, Eva Hajicova, and Jarmila Panevova. 1985. *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel.
- Slobin, Dan I. 1971. *Psycholinguistics*. London: Scott, Foresman & Co.
- Steedman, Mark. 1987. Combinatory grammars and parasitic gaps. *Natural Language & Linguistic Theory* 5, 403-439.
- Steedman, Mark. 1991. Structure and intonation. *Language* 68, 260-296.
- Steedman, Mark. 1992. Grammars and processors. Report No. MS-CIS-92-53, University of Pennsylvania.
- Steedman, Mark. 1996. *Surface structure and interpretation*. MIT Press.
- Steedman, Mark. Forthcoming. *The syntactic process*.
- Stuurman, Fritz. 1990. *Two grammatical models of modern English*. London: Routledge.
- Suppes, Patrick. 1973. Semantics of context-free fragments of natural languages. In K.J.J. Hintikka, J.M.E. Moravcsik and Patrick Suppes (eds.) *Approaches to natural language*, 370-394. Dordrecht: Reidel.
- Sweet, Henry. 1900. *The history of language*. London: Routledge.
- Tait, Mary Esther. 1988. Multiple PP gaps. Research report EUCCS/RP-41. University of Edinburgh: Centre for Cognitive Science.
- Talmy, Leonard. 1988. The relation of grammar to cognition. In Brigida Rudzka-Ostyn (ed.) *Topics in Cognitive Linguistics*, 165-205. Amsterdam: Benjamins.
- Tesnière, Lucien. 1959. *Elements de syntaxe structurale*. Paris: Klincksiek.
- Trudgill, Peter & Jean Hannah. 1982. *International English: a guide to varieties of Standard English*. London: Arnold.
- Tugwell, David. 1994. Back to basics: *likely to, likely that*, extraposition & analytic syntax. Ms., University of Edinburgh.
- Uszkoreit, Hans. 1987. Linear precedence in discontinuous constituents: complex fronting in German. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.) *Discontinuous constituency*, 405-425. Orlando: Academic Press.
- van Ek, Jan Ate. 1966. *Four complementary structures of predication in contemporary British English*. Groningen: Wolters.

- Wilks, Yorick. 1983. Does anyone still believe this kind of thing? In Karen Sparck Jones and Yorick Wilks (eds.) *Automatic natural language parsing*, 182-189. Chichester: Ellis Horwood.
- Wilks, Yorick. 1987. Bad metaphors: Chomsky and Artificial Intelligence. In Sohan Modgil and Celia Modgil (eds.) *Noam Chomsky: consensus and controversy*, 197-206. New York: Falmer.
- Winograd, Terry. 1977. On some contested suppositions of generative linguistics about the scientific study of language. *Cognition* 5, 151-179.
- Wood, Mary McGee. 1993. *Categorial grammars*. London: Routledge.
- Woods, W. 1973. An experimental parsing system for transition network grammars. In Randall Rustin (ed.) *Natural language processing*. New York: Algorithmics Press.
- Xu, Liejiong & D. Terence Langendoen. 1985. Topic Structures in Chinese. *Language* 61, 1-27.
- Yngve, Victor H. 1996. *From grammar to science: new foundations for general linguistics*. Amsterdam: Benjamins.
- Zaenen, Annie. 1979. Infinitival complements in Dutch. In *Proceedings of the 15th annual meeting of the Chicago Linguistic Society*, 378-389.
- Zwicky, Arnold M. 1978. Arguing for constituents. In *Proceedings of the 14th Regional Meeting of the Chicago Linguistic Society*, 503-512.