

The Status
of Thematic-Conceptual Structure
in Language Processing
and Linguistic Structure

Grant James Evans

PhD
University of Edinburgh
2001



Declaration

This thesis has been composed by me, and the work contained within it is my own.

Grant Evans

Contents

1	Introduction	4
2	Thematic Information in Syntactic Theory	6
2.1	Thematic Roles in Generative Semantics	6
2.1.1	Thematic Roles in Case Grammar	8
2.1.2	The Fall of Generative Semantics	10
2.2	Thematic Roles in Principles and Parameters Theory	12
2.3	Thematic Structure versus Argument Structure	15
2.4	Further Developments of Configurational Notions of Thematic Roles	18
2.4.1	An Extended Theory of Thematic Relations	18
2.4.2	Thematic Roles and Incorporation	21
2.5	Decomposition of Thematic Structure: A Continuum	24
2.5.1	Level 1: Canonical Realization of Semantic Structures	24
2.5.2	Level 2: Linking Rules and Lexical Alternations	27
2.5.3	Level 3: A Simplification of the Syntax/Semantics Interface and the Selection of Arguments?	32
2.5.4	Level 4: Thematic-conceptual Structure	34
2.6	Chapter Summary	43
3	Considerations from Language Acquisition	45
3.1	Theories of Learning	46
3.1.1	Connectionism	46
3.1.2	Constructivism	48
3.2	Functional Categories in Language Acquisition	50
3.2.1	An initial view: a thematic based acquisition device	50
3.2.2	The Opposing View: The continuity of functional categories	53
3.2.3	Middle Ground: Grammars constrained by UG, but with bottom-up thematic information as the input to linking rules	54
3.3	Parameter Setting	58
3.4	The Relationship of Ontogeny and Phylogeny	63
3.5	The Starting Point for Language Acquisition	64
3.5.1	Early Capabilities Reflected in Lexical Acquisition	66
3.5.2	General Principles of Language Acquisition	67

3.6	Thematic-Conceptual Structure and Language Acquisition	69
3.7	Parameter Setting and Grammaticalization	75
3.7.1	The Complexity of the Grammaticalization Data	77
3.7.2	A Link between Metaphor and Thematic Conceptual Structure	79
3.8	Chapter Summary and Predictions for Psycholinguistics	80
4	Thematic Information in Psycholinguistic Theory	82
4.1	What Psycholinguistics is Trying to Explain?	83
4.2	Latent Parallelism in Thematic Roles	85
4.2.1	The Discourse Strand	86
4.2.2	The Thematic Recovery Strand	87
4.2.3	Thematic Influence on Initial Decisions	87
4.3	Thematic Roles as a Filter on Syntax	90
4.4	Thematic Roles Determine Parsing Decisions	93
4.5	Thematic Information in Sentence Processing	95
4.6	Thematic Structure in Discourse	96
4.7	The Time-Course of Sentence Comprehension	98
4.8	Language Processing and Language Acquisition	100
4.9	Conclusions	102
5	Psycholinguistics and Thematic-Conceptual Structure	103
5.1	Attachment Ambiguities with <i>with</i>	103
5.2	Sentence-Level Priming is Not Dependent on Syntactic Structure	108
5.2.1	Thematic Structure in Syntactic Priming	110
5.2.2	Procedures in Production and Comprehension	117
5.3	Future Research on Sentence-Level Priming	121
5.4	Recent Experiments in Thematic-Conceptual Sentence Priming	127
5.4.1	Cross-linguistic experiments	127
5.4.2	Mono-lingual experiments	130
5.5	Theoretical and Representational Aspects Of Sentence Level Priming	131
5.6	Summary	132
6	Conclusions	135
6.1	Thematic-Conceptual Structure and Phylogeny	136
6.2	Grounding Thematic-Conceptual Structure in Evolutionary Terms	137
6.3	A Brief History of Humankind	139
6.4	The Evolutionary Status of Language	143
6.5	Scenarios of Language Evolution	149
6.6	Primate Behaviour and Primate Cognition	156
6.7	Teaching Human Language to Apes	159
6.8	Cultural Learning and the Cultural Evolution of Language	163
6.9	Chapter Summary	167

Abstract

After a discussion of the role of thematic information in each of language acquisition, syntax and language processing, thematic-conceptual structure is presented as a common representation that serves each of these sub-disciplines well. Thematic-conceptual structure is a decompositional representation of thematic-information that is constrained to include only those semantic elements that are grammaticalized in some language in the world.

Thematical-conceptual structure is then shown to be a representation that allows for the explanation of previously intransigent sentence processing patterns, as well as providing a representation that best explains sentence level priming not only in comprehension, but also in production. Some novel experiments are proposed, and the initial results of the experimental programme laid out in this thesis indicate that thematic-conceptual structure is indeed the correct representation to explain the results of the experiments.

Throughout the thesis it is noted that thematic-conceptual structure is influenced in its form both phylogenetically and ontogenetically. This idea is somewhat speculative, but provocative nonetheless. In the conclusion of the thesis, these speculative aspects are discussed in depth.

Acknowledgments

Writing this PhD has certainly been a long haul. I now have the opportunity to thank the people who have kept me going over this time. On an academic level, my supervisors Richard Shillcock and, in the early years, Matt Crocker have always been available to encourage me to believe in my ideas.

Personally, I would like to thank my wife for putting up with all the weekends and evenings I have spent working in the thesis instead of doing the things we could have been doing together.

I would also like to thank all the people in my post-academic life who have encouraged me to get the thesis complete. These include my colleagues at Lidl, Perseco and McDonald's as well as my family and friends.

This thesis is dedicated to my father who died in 1997 and always believed that I would complete the thesis one day.

Chapter 1

Introduction

Thematic roles are widely regarded as being an essential integral part of linguistic theory. However, there seems to be little agreement as to the form the theory of thematic roles itself should take. The part played by thematic roles differs from sub-discipline to sub-discipline.

This thesis takes this situation as its starting point and has as its aim a coherent theory of thematic information which is applicable to a number of sub-disciplines, namely syntactic theory, psycholinguistic theory and language acquisition. I shall also discuss what the evolution of language may tell us about the nature of the theory of thematic roles we should propose in a more speculative discussion that forms part of the conclusion of this thesis.

The thesis covers a large number of diverse areas within linguistics and cognitive science and contains a large amount of background information and discussion that is predominantly discursive in nature.

In Chapters 2, 3 and 4 I present how thematic information has been used to explain syntactic, language acquisition and psycholinguistic phenomena respectively. In these chapters, I will present empirical syntactic and psycholinguistic data that will need to be adequately captured by any theory of thematic information. In the course of these chapters I will explain how various theories of thematic information have attempted to capture segments of this data. I identify the strengths and weaknesses of these attempts.

In the course of these chapters I come to the conclusion that thematic-conceptual structure is a representation that can be used to successfully capture phenomena from all of language acquisition, syntax and language processing. A number of psycholinguistic predictions follow from this discussion and in Chapter 5 these are discussed in more depth.

Thematic-conceptual structure is then shown to be a representation that allows for the explanation of previously intransigent sentence processing patterns, as well as providing a representation that best explains sentence level priming not only in comprehension, but also in production. Some novel experiments are proposed, and the initial results of the experimental programme laid out in this thesis indicate that thematic-conceptual structure is indeed the correct representation to explain the results of the experiments.

Throughout the thesis it is noted that thematic-conceptual structure is influenced in its form both phylogenetically and ontogenetically. This idea is somewhat speculative, but provocative nonetheless. In the conclusion of the thesis, these speculative aspects are dis-

cussed in depth.

Chapter 2

Thematic Information in Syntactic Theory

In this chapter I will investigate the status of thematic information in syntactic theory. There are three main aspects to consider here. First, in the beginnings of research into thematic information in language, Generative Semantics proposed a decompositional analysis of thematic information with its own constituent structure. A number of empirical difficulties led to the downfall of Generative Semantics and with it the decompositional approach.

Second, as Chomskyan syntax became the prevalent paradigm for syntactic discussion, thematic information was reduced, in effect, to theta-roles, which had an important configurational role, but were stripped of most of their semantic origins.

Third, decompositional theories have begun to re-emerge. As decompositional theories had originally been felt to have failed because of the unconstrained nature of their representations, these new decompositional approaches have sought various ways to constrain the components of the representation.

I will discuss each of these aspects in this chapter. I conclude that decompositional theories should be developed further as they avoid the syntactocentrism of theta-roles in PPT and, as chapters 3 and 4 will show, decompositional representations of thematic-conceptual information can be applied successfully to psycholinguistic and language acquisition data, and can also capture best the new psycholinguistic evidence presented in Chapter 5.

2.1 Thematic Roles in Generative Semantics

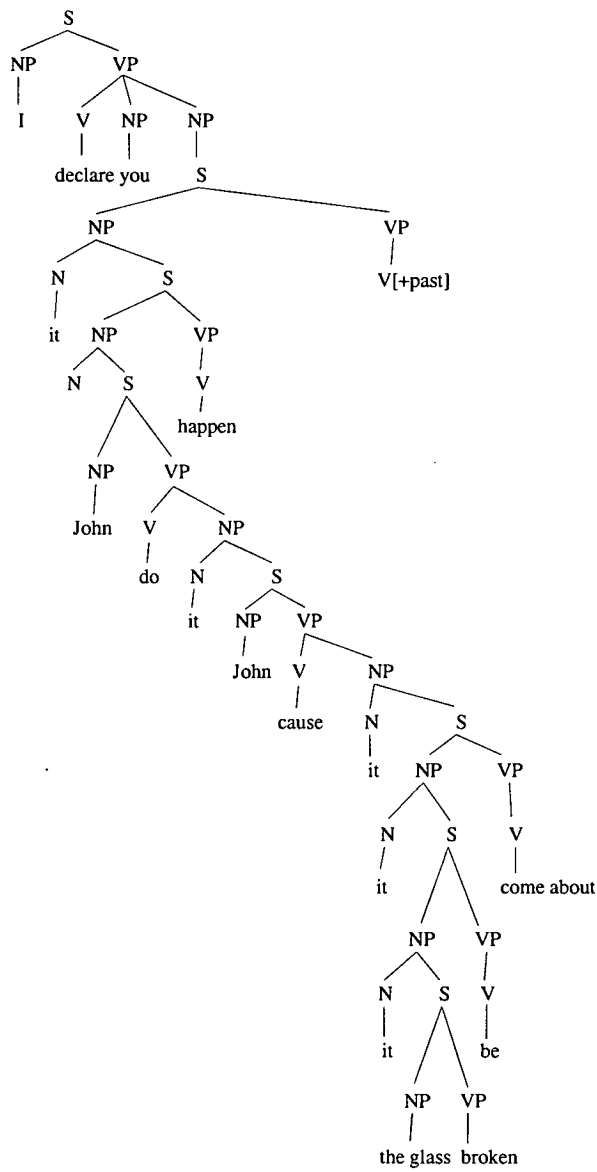
I believe that decompositional approaches to thematic information have been mistrusted and resisted for so long because they are connected in the minds of many researchers with the research on Generative Semantics. Indeed, the concept of thematic roles was born at the time of Generative Semantics and after its demise thematic roles played very little part in linguistic discussion for around a decade.

Central to the theory of Generative Semantics was the Katz-Postal Hypothesis which stated that transformations do not affect meaning. Any difference in meaning had to be related

directly to lexical items and to sentence type markers (Imperative, Interrogative, etc), and merely moving elements around within a sentence could effect no change in meaning.

Over the time Generative Semantics had its influence there was a gradual abstraction of Deep-Structures into semantic terms, leading almost inevitably to the abandonment of Deep-Structure as a theoretical concept.

The relevant abstractions took their form mainly in the lexical decomposition of verbs. It was suggested that the set of linguistic categories could be reduced to the lexical categories of N, V, and A. This seemed attractive as these categories corresponded in a natural way to those of truth-functional logic. However, this abstraction took place within the syntax, resulting in very complex syntactic structures such as the following:



- (2.1) a.
b. John broke the glass

A large number of transformations were necessary both for lexical insertion and conversion to surface structure. The theoretical motivation was originally that projection rules, which

project the deep structure meaning onto the surface structure form, did not supply any additional aspect of meaning, and therefore all meaning had to be contained in Deep-Structure. This proliferation of transformations from Deep-Structure to surface structure is now regarded as the major weakness of Generative Semantics.

Generative Semantics saw the form-content link as direct. That is, the semantic content of the sentence was necessarily linked to a particular surface syntactic form. This conclusion was probably reached due to the concentration on English as the subject of study. In opposition to this view, I will argue in Chapter 5 that the linking between some form of thematically based semantic structures and the surface structure is parameterized for each language and will allow for differing patterns on the surface to represent the same underlying content, or, alternatively, two underlying content relations could be expressible by only one surface syntactic form.

In Chapter 5 in my consideration of some new psycholinguistic evidence it is shown to be arguably the case that decompositional representations of thematic information can be linked in a variety of ways to surface syntactic structure, and that the form-content link is therefore not direct but can vary from language to language.

2.1.1 Thematic Roles in Case Grammar

The theory of Case Grammar (Fillmore (1968)) can be identified as the source of the theory of thematic roles, although Gruber (1965) had already discussed what the possible identity of such thematic roles may be. Fillmore argued that at the deepest syntactic level a sentence consists of a verb and an unordered set of semantically characterizable thematic roles. The motivation for this was that the Aspects model of Chomsky (1965) could not capture categorical and thematic information simultaneously. For example *into the room*, *towards the door*, *on the next day* are all categorically PPs but are thematically LOCATION, DIRECTION and TIME respectively. Fillmore argued that meaning was motivating syntactic structure to some extent.

Behind Fillmore's reasoning was the idea that there is a universal base of syntactic relations (Halliday (1966)) founded on a universal base of covert categories. This idea is implicit in Fillmore's second assumption. This states that covert categories are of central importance. Or in other words, all languages are similar in their expressive capacity; they are just encoded differently.

This assumption is vital and should be borne in mind throughout this thesis as it is also central to my thinking; the conceptual relations encoded in language are language independent.

Fillmore argues that there is a parallel between what he terms *Cases* (thematic roles) and the choice of prepositions for verbs:

It seems to me that the discussion of case could be seen in somewhat better perspective if the assignment of case forms were viewed as exactly analogous to the rules for assigning Prepositions in English (Fillmore 1968: 15)

This is our first example of what is commonly termed *linking*. That is, the conceptual relation to be expressed is associated (or linked) with a language-specific rule that determines the surface expression of the underlying conceptual relationship in that language.

Fillmore distinguishes between *pure* and *labeled case*. *Pure* cases are configurational relations such as Subject and Object and *labeled* cases are cases assigned from prepositions or from the properties of governing verbs. Examples of labeled cases are Objective and Instrumental.

Fillmore argues that Subject is an incoherent notion and that therefore labeled cases should be assumed. It is often proposed that Subject and Object are derived notions, derived either from thematic roles via a thematic hierarchy (Section 2.3), or are determined by sentential position.

Fillmore also argues for a ‘conceptual framework’ interpretation of case systems. Only noun phrases representing the same case can be conjoined. This is also an indication of how grammaticality can be determined by a failure of interpretability.

- (2.2) a. *John and a hammer broke the window. (AGENT + INSTRUMENTAL)
 b. John broke the window.
 c. A hammer broke the window.
 d. John broke the window with a hammer.

Fillmore concentrates on the further investigation of the proposition. The proposition is made up of a verb and one or more case categories, these case categories being what contemporary linguists would term thematic roles.

Insertion of verbs depends on the case frame of the sentence. Thus subcategorization is in terms of thematic classes.

(2.3) *run* can be inserted into [__Agentive]

(2.4) *open* can be inserted into [__Objective + Agentive]

The initial notation means that *run* can be inserted in an environment where there is an Agentive Case role. The second notation states that *open* requires both an Agentive Case role and an Objective Case role to be present for the sentence to be properly formed.

Thus lexical entries will list the environments where the verb can occur.

- (2.5) a. The door opened. [__Objective]
 b. John opened the door. [__Objective + Agentive]
 c. The wind opened the door. [__Objective + Instrumental]
 d. John opened the door with a chisel. [__Objective + Instrumental + Agentive]

and the frame for *open* will be:

(2.6) +[__Objective (Instrumental) (Agentive)]

The semantic interpretation of any particular preposition introduces all the information provided by the specific case relationships represented in the preposition, allowing such information to be left out of the semantic representation of the verb.

Fillmore suggests there are a number of ways that the cases may be represented at the surface: selection of overt case forms by suppletion, affixation, adding of P, subjectivization,

objectivization, sequential ordering and nominalization. These are all further examples of *linking rules*.

Fillmore claims semantic and syntactic benefits for his theory. The semantic benefits can be seen in sentences (2.7). In all the sentences *the door* bears the same relation to the verb, but is considered a surface-structure subject in (a) and a surface-structure object in (b) and (d). Also the *door* in (a) and *John* in (b) and (d) and the *wind* in (c) all manifest different semantic relations with respect to the verb but all are treated as surface subjects. Here, then, we have an example of a semantic function being expressible by a number of different syntactic functions.

- (2.7) a. The door opened.
 b. John opened the door.
 c. The wind opened the door.
 d. John opened the door with a chisel.

The syntactic benefits are seen in Fillmore's concept of a thematic hierarchy, which allows for the identity of the subject of a sentence to be predicted on the basis of its deep-structure case. Fillmore claims that if there is more than one NP in a sentence, the highest on the thematic hierarchy will become the subject. And if there are exceptions, these are marked by special morphological elements. In section 2.3 we will see the concept of thematic hierarchies being developed.

The direct result of Fillmore's work has been that most syntacticians agree that syntactic theory must include a characterization of thematic roles. However, the association of thematic roles with Generative Semantics (unfairly, as Fillmore never counted himself as part of the GS movement) led to much of this work not being properly developed for some time after the acrimonious fall of the Generative Semantics theory.

2.1.2 The Fall of Generative Semantics

A number of theoretical developments played a part in the demise of Generative Semantics. First, the lexicalist hypothesis (Chomsky (1970)) stated that there should be no category-changing transformational rules. This bans rules changing, say, a verb into a noun, or a noun into an adjective, and so on. The motivation was that derived nominals exhibited different syntactic behaviour to the verbs they were derived from; for example:

- (2.8) a. John believed that Bill was a fool. →
 John believed Bill to be a fool.
 b. John's belief that Bill was a fool. →
 *John's belief (of) Bill to be a fool.

In addition, it was argued, the rules for deriving nominals were not productive; not all verbs have a derived noun and often when there is a derived nominal, it is idiosyncratic (*do/deed*).

Second, many examples were found by Jackendoff (1972) where Surface Structure did make a difference to meaning. These examples thus refuted the Katz-Postal hypothesis¹.

¹These are also counterexamples to Baker's theory as presented in Section 2.4.2

- (2.9) a. Many arrows did not hit the target.
- b. The target was not hit by many arrows.
- (2.10) a. Only John reads books on politics.
- b. John only reads books on politics.
- c. John reads only books on politics.

However, this remains a problem only if it is assumed that the decomposition of the meaning of these sentences has to be achieved within the syntax and the relationships between the elements have to be characterized by transformations. If, on the other hand, we allow separate levels for the thematic-conceptual relationships and the linking rules that map these onto surface structure, these problems can be avoided as I believe will become clear in this thesis. Thus the thematic-conceptual structure for each of these sentences would map in a one-to-one fashion onto the surface structure via linking rules.

Third, McCawley (1968) and Jackendoff (1972) showed that selectional restrictions had to be based on semantics and not on syntax. So, for example, paraphrases have the same selectional restrictions. And when grammatical and semantic features conflict in a language, selectional restrictions occur based on the latter, not the former. In German, for example, a number of verbs occur only with semantically female subjects, but there are no verbs that take only a grammatically female subject.

Jackendoff also argues that subcategorization cannot be based on semantic restrictions as *eat*, *devour* and *dine* have the same meaning, but are respectively optionally transitive, transitive and intransitive. However, I believe that there are differences between these verbs in semantic terms at some level or another of the representation as I shall be following Clark's (1993) principle of *Contrast*, which states that speakers do not tolerate complete synonyms. (See Section 3.5.2).

Fourth, it was found difficult to conflate all elements into A, N and V. This is because there is always a difference in the behaviour of nouns, verbs, adjectives, adverbs, quantifiers and prepositions in surface structure, regardless of the transformational levels that underlie them.

Fifth, Lexical Decomposition failed as it proved very difficult to decompose every word into a limited number of semantic primitives. As Fodor (1970) argued, if lexical decomposition is a true representation of linguistic facts and *melt* can be decomposed into *cause to melt*, then the following should be paraphrases, which they clearly are not.

- (2.11) a. Floyd caused the glass to melt on Sunday by heating it on Saturday.
- b. *Floyd melted the glass on Sunday by heating it on Saturday.

There have been two main themes in this first section. The first is the birth of thematic roles as a theoretical construct and their definition. The second is the blossoming and subsequent death of the theory of Generative Semantics. The two are linked because, although Fillmore's theory of thematic (case) roles was not really part of the Generative Semantics movement, it came to be associated with it, so that, when Generative Semantics collapsed, Fillmore's theory and observations were largely ignored for over a decade.

However, this is not to say that many of the ideas of Generative Semantics will not be useful to my aims here. This will be the case as long as lexical decomposition and linking rules are assumed to be architecturally distinct from the syntax. Decomposition at the level of conceptual structure will allow us to develop a more flexible internally structured representation of thematic roles. The separation from syntax will allow for the continued observation and explanation of purely syntactic phenomena at the level of syntax.

From Fillmore's work we can take away the early characterization of thematic roles and the idea that grammatical function (Subject, Object) and semantic function (AGENT, THEME etc) must be separated. He proposes a thematic hierarchy to provide the mapping between them; an idea we will see developed in the field, but that I will ultimately discard due to its inflexibility.

Also of central importance to the theory developed in this thesis is the idea that thematic roles are in some way based on a universal set of covert categories. I will argue that thematic roles are shorthand for frequent combinations of covert conceptual relations, and although useful as shorthand, they should not replace this underlying conceptual structure as the level of representation to be used in the explanation of a range of linguistic phenomena in syntax, language acquisition and psycholinguistics.

2.2 Thematic Roles in Principles and Parameters Theory

Jackendoff (1972) defined the following thematic roles in semantic terms: AGENT, GOAL, SOURCE, LOCATION, EXPERIENCER, RECIPIENT, INSTRUMENT, BENEFACTIVE, THEME, PATIENT and PERCEPT.

These are the roles that Jackendoff said could be applied to noun phrases, but, he argued, were not predictable from the structural position of the NP. It may be the case, however, that the converse is true; the structural position of the NP being predicted by the thematic relations to be expressed.

Principles and Parameters Theory has nominally used these roles, but it is essential to note that PPT uses thematic roles in a purely structural way, and the semantic motivation of the roles is largely ignored. Thematic roles within PPT are analogues of argument structure merely expressed in other terms. The main contribution they make is that in a movement chain, there is always one Case position and one theta (thematic)-position. Thus only one theta-role can be assumed to be assigned to each NP, and this is made explicit in the Theta-Criterion.

(2.12) *Theta-Criterion* Chomsky (1981).

- a. Each argument is assigned one and only one theta-role.
- b. Each theta-role is assigned to one and only one argument.

With this criterion in place, the interaction of Case-theory, Move- α and the Projection Principle accounts for the movement properties of language.

(2.13) *Case Filter*

Every overt NP must be assigned abstract case.

(2.14) *Move- α*

Any NP can move anywhere as long as it meets the constraints of PPT.

(2.15) *Projection Principle*

Lexical information is syntactically represented.

A good starting point for examining the movement properties of PPT is the summary given in Haegeman (1994) and reproduced in (Table 2.1)

Table 2.1: Movement Properties of PPT

	NP-trace	wh-trace
Moved category	NP	XP(NP, PP, etc)
Landing site	A-position by substitution	A'-position by substitution or adjunction
	NP-position	[Spec, CP] or adjoined position.
Properties of antecedent		
Case	Yes	No
Chain	A-chain	A'-chain
Properties of trace		
Features	[+Anaphor] [-Pronominal]	[-Anaphor] [-Pronominal]
Binding Theory	A	C
Theta-Role	Yes	Yes
Case	No	Yes (when target = NP)
Governed	Yes	Yes

These properties of movement are predicted and explained by the interactions of the principles above. An NP moves when it is assigned no Case in its original position to a position where it can get Case. This position where it can get Case is an NP position. This NP position is an argument position. Each NP has to have one and only one theta-role, so as the NP has moved from a position where it already has a theta-role, it must move to a NP position where no theta-role is assigned. In sum, the chain of movement includes one Case-position and one theta-position. Such movement is manifested in the passive and raising constructions, for example.

An example of a raising verb is *seem*. This verb is a one-place predicate that takes a clausal complement. We can see it is a one-place predicate because the subject position is assigned no theta-role by *seem* and this position has to be filled with pleonastic *it*.

(2.16) It seems that John dislikes Tony.

The lexical items in (2.17) are the same so we can assume the same thematic relations are in operation.

(2.17) John seems to dislike Tony.

We note that the external argument (*John*) of *dislike* has been moved to the subject position of *seem*. We know that *John* has received its theta-role from *dislike* so it has not moved to receive a theta-role. *John* has moved to a position where it will be assigned no second theta-role, thus fulfilling the theta-criterion. In the subordinate clause, however, there is a non-finite verb and in the main clause there is a finite verb. Non-finite verbs cannot assign Case, whereas finite verbs can do so. According to the Case Filter overt NPs must receive Case, thus *John* has moved from a position where it can receive no Case to a position where it does receive Case.

The D-structure is (2.18) and the S-structure is (2.17). The chain of movement contains one Case position and one Theta-position.

	[_{IP} e	seems	[_{IP} John to dislike Tony]]
(2.18)	+Case		-Case
	-Theta		+Theta

The analysis of the passive follows the same lines, but relies on the assumption that the passive can assign no Case and that it loses one of the theta-roles it had been able to assign as an active verb through a process of argument suppression.

Turning now to *wh*-movement, we find that the XP that has moved has Case in its original position, at least if it is an NP. Thus it must not move to another position that is assigned Case. It must therefore move to a non-argument position. The original position was also a position assigned a theta-role, so the moved element may not be moved to a position where it would be assigned another theta-role. Once more the chain of movement includes one Case position and one Theta-position.

We have seen then that the Projection Principle states that lexical information is projected into the syntax, but when we examine what this actually means in PPT with respect to thematic roles, we see that, in effect, it is only the argument or subcategorization structure of the lexical item that is projected. There is no real element of content to the thematic roles; they may as well be numbered arguments.

It is also the case that certain constructions lead researchers to impute certain properties to the lexical items that take part in these constructions. Thus raising predicates must assign no Case, and the passivization process is taken to absorb the Case marking properties of the verb that is passivized. The fact that passivization is not a process that can be applied to any verb is not explained in any way. A number of decompositional approaches that explain when verbs do and do not alternate are to be found in the course of this chapter.

What we need, then, is a proper conception of the Projection Principle that really addresses the question of what aspects of the lexical-conceptual-structure have what effects on the syntactic structure. By the end of this chapter we will have been given some clues as to the direction our enquiry should take. The following chapter will add to these clues by examining the part thematic roles play in language acquisition and psycholinguistic theory.

In the next section I will further investigate this configurational notion of thematic roles.

2.3 Thematic Structure versus Argument Structure

The first modification of the PPT theory of thematic roles also takes a course where thematic roles are functionally equivalent to argument structure. Behind this is the assumption of Canonical Structure Representation (CSR) Grimshaw (1988). CSR states that each named thematic role projects onto a particular type of constituent in all cases. This is a small step away from merely numbered arguments, but the inflexibility and coarse structure of the roles limits the success of the projection to constituent structure.

So, for example, AGENT will map onto NP and LOCATION will map onto PP. The aim of CSR was to remove the need for both subcategorization and thematic information in the lexical entry. CSR allows for subcategorization facts to be derived from the thematic information in the lexical entry.

Grimshaw (1990) also assumes a direct linking of thematic structure to argument roles, but she also wants to predict the syntactic behaviour of various classes of verbs by basing this linking on a hierarchy of thematic roles in conjunction with a representation of the event structure of the verbs in question. In this respect, her theory is very similar to that of Fillmore's but with the additional feature of an event structure representation.

This approach is termed a *prominence theory of A-structure*. A-structure represents the prominence relationship between arguments. Prominence is determined by thematic properties of the predicate and its aspectual properties. A-structure, she argues, is blind to named theta-roles. The external argument for Grimshaw is the most prominent argument in the A-structure on both the aspectual and thematic hierarchy.

Grimshaw studies the invariant linking of arguments based on the thematic hierarchy of:

(2.19) AGENT<EXPERIENCER<GOAL/SOURCE/LOCATION<THEME

This hierarchy should be seen as descriptive rather than explanatory as the ordering of the hierarchy reflects the syntactic behaviour of arguments. The reason why arguments linked to the levels of the hierarchy should act in a particular way based on that linkage is not explained. I believe that such an explanation can only be found if we look at the conceptual relations for which the thematic roles of the hierarchy are a shorthand version. However, it is still useful to examine how it is that this hierarchy reflects syntactic behaviour.

The arguments are theta-marked in a particular order: (1) theta-marking in the NP precedes theta-marking in the clause; or in other words NP internal arguments are theta-marked before clausal arguments, (2) the argument lowest on the thematic hierarchy is marked first, the next lowest on the hierarchy is then marked and so on. Thus, for *give* which has the following A-structure

(2.20) (x (y (z)))
AGENT GOAL THEME

The THEME must be theta-marked first, followed by the GOAL, predicting the following pattern correctly

- (2.21) a. Gift-giving to children.
 b. *Child-giving of gifts.

As an explanation of syntactic facts this seems to be quite adequate, although how this constraint is to be expressed with respect to language production or comprehension is unclear.

However, *fear* and *frighten* violate the theory as presented up to now, as the EXPERIENCER is realized as subject for one but not for the other. Grimshaw claims that this is due to aspectual differences in the semantics of the verbs. This approach contrasts with the *ad hoc* analyses of thematic reanalysis or stipulated lexical case marking. Thus it is this type of verb that justifies the next step in Grimshaw's argument: the postulation of a separate level of causal structure.

The aspectual difference in question is that *frighten* has a causal meaning not shared with *fear*. So a second *causal* hierarchy is added to the thematic hierarchy. Within the causal hierarchy the element *CAUSE* is always the most prominent:

- (2.22) (Cause(other(...)))

In *frighten* the causal and thematic hierarchies conflict:

- (2.23)
- | | | |
|-----------------|--|-------|
| <i>frighten</i> | (x | (y)) |
| | Exp | Theme |
| | <div style="position: relative; height: 10px;"> </div> Cause | |

Grimshaw goes on to argue that this causal analysis has an aspectual origin. This is interesting since Fillmore (1968) also saw sentences as being made up of a propositional part and an aspectual part. He did not examine the aspectual side of his representation. Grimshaw has done this and found it to be a vital part of her descriptive vocabulary.

The Cause is always related to the first sub-event in the following event structure:

- (2.24)
- | | | |
|--|--|-------|
| | event | |
| | <div style="position: relative; height: 10px;"> </div> | |
| | activity | state |

The Cause will always be more prominent than the entity whose state is changed. The aspectually most prominent argument is selected as subject. For the major verb groups the interaction of thematic and aspectual hierarchies is as follows.

- (2.25) a. *Transitive Agentive*
 (Agent (Theme))
 1 2
hit
- b. *Ditransitive*
 (Agent (Goal (Theme)))
 1 x x
give
- c. *Unergative*
 Agent
 1
- d. *Psychological State*
 (Exp (Theme))
 1 2
fear
- e. *Psychological Causative*
 (Exp (Theme))
 2 1
frighten
- f. *Agentive Psychological Causative*
 (Agent (Exp))
 1 2

To account for the linking of the *fear* class, Grimshaw is left with no option but to stipulate that the EXPERIENCER is the most prominent in the aspectual hierarchy.

The external argument is always the last to be theta-marked, therefore there can only be one external argument. In the *frighten* class there is a conflict of prominence across the hierarchies. For this reason, Grimshaw assumes that the *frighten* class verbs have no external argument and therefore cannot take part in the causative/anticausative alternation (*Someone frightened John*/**John frightened*).

This in turn means that a subject is not always an external argument because *frighten* verbs have a d-structure subject but no external argument.

Grimshaw needs to make a further stipulation to explain Unaccusatives which have no external argument but are monadic, to explain why the single argument is not the most prominent in the hierarchies. She proposes two possible explanations: (1) there is a hidden lexical-conceptual-structure CAUSE argument taking up the external slot. (2) the THEME cannot be most prominent as only arguments participating in the first sub-event (the activity) can be most prominent (We will see an alternative approach to Unaccusatives in section 2.5.2).

Grimshaw has a theory that empirically is fairly accurate, but for this accuracy it relies on a number of lexical stipulations. For a theory of this kind to be attractive we have to find

some type of explanation as to why certain lexical items have the properties they do, i.e. are there features they have in common?

Grimshaw identifies sets of verbs that behave in a certain way syntactically. However, identifying the sets of verbs does not constitute an explanation of why they should include the particular verbs they do. Commonalities between the individual verbs in a set should be identified, and this identification will probably be in terms of conceptual structure.

Indeed, when of the elements Grimshaw identifies as being of relevance is *cause*. This is clearly a conceptual notion and is adopted by almost all decompositional approaches.

The inflexibility of the description proves to be a problem for languages with freer word order than English. English is very fixed in using word order as the main method of signaling grammatical relations. Languages with richer Case systems, such as German, are frequently freer with respect to their word order. Thus the subject of a German sentence can be the Patient of the verb. Whenever this is the case, a rigid hierarchy of the type Grimshaw proposes is not easily applicable. We need to find a more malleable way of linking the conceptual relations expressed by a proposition to the surface expressions and structures used by a particular language.

2.4 Further Developments of Configurational Notions of Thematic Roles

In this section I will present two further developments of the configurational view of thematic roles. The first of these comes from Williams (1994) where a reanalysis of the assignment of thematic roles allows for the removal of movement from the analysis of raising constructions. This is provocative as it opens up the possibility that many syntactic processes actually reflect an underlying passing of thematic information. It may be the case that much of the movement aspects of syntactic theory can be reanalyzed in purely thematic(-conceptual) terms.

The second extension of the theory comes from Baker (1988) and takes this development one step further by maintaining that thematic information is maintained during syntactic processes.

2.4.1 An Extended Theory of Thematic Relations

Williams (1994) does not address the problem of thematic-structure/argument-structure linking, but turns his attention to the processes by which thematic roles are assigned. In so doing he is able to remove movement from the analysis of raising constructions. Williams sees raising as an example of *Function Composition*. This is where an element transmits its complement's theta-role to its subject. A functor thus combines with X, without changing the theta-structure of X.

Williams is arguing for a theta-theory without semantic content as merely the number of arguments of a verb are given together with the realization conditions for those arguments². Each lexical item includes the linking relations to surface structure within its lexical entry. This seems to be missing out in two important ways: there may well be generalizations in

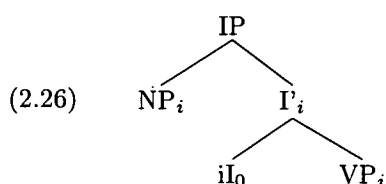
²In effect a CSR for each lexical item.

the realization conditions in semantic terms across lexical items, and the desemanticization of thematic structure ignores the insights of Jackendoff (1972) and Fillmore (1968).

Williams sees the construction of the clause as akin to coordinate structure. Williams gives two requirements that the clause structure must fulfill.

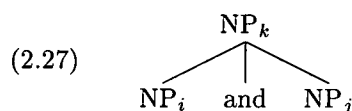
1. That Tense is head of S, as complementizers select for Tense, and only heads are selected for.
2. The subject-predicate relationship must be well formed. That is, the index of the external argument should percolate to VP, and the subject must be sister of VP_{*i*}.

Fulfilling these requirements would result in the structure:

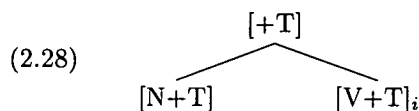


However, this structure violates the idea of external argument because although the index *i* of V is realized external to VP, it is realized *internal* to IP since it is represented on the head of IP. But it should be external to IP. Externality is required because, as has been shown (Williams (1984)) the interpretation of the external argument can only be ascertained once the internal arguments have been discharged and interpreted. Thus, a representational position that allows for this internal/external interpretational difference should be respected.

To solve the problem of maintaining the externality of the subject argument Williams borrows the double-headedness of the nominal conjunct structure

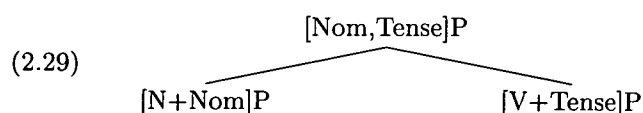


and continues to assume that S is a projection of Tense, but contrary to the usual analysis he assumes that Tense is realized on both subject and predicate.

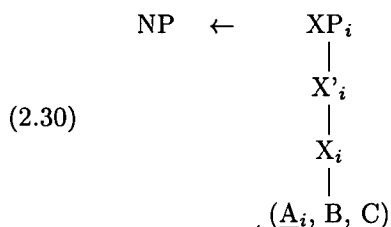


This structure is double-headed, and each half of the structure is maximal. It also allows him to claim that Nominative Case is merely the realization of Tense in a Nominal structure. This, for the first time, gives a theoretical reason for the Nominative case. The Nominative case is associated with the Subject, as it is the Subject that is most closely identified with the Cause side of the event in Grimshaw's terms. We have here a further conceptual link between the surface linguistic expression and the semantic content that is being expressed.

Thus the final structure for the clause Williams assumes is:



The external argument must be the sister of the maximal projection of the verb. This maximal projection is a one-place predicate, and the theta-role it assigns is the external role of the verb that is its head. The binding of the external argument of the verb by the maximal projection is simply the X-bar projection of the index of the external argument.

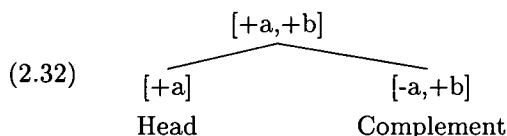


Thus a theta-role is assigned to the external argument by predication, which is defined by the verb and its internal arguments. The external argument of the argument structure is made the head of the argument structure to ensure that it percolates beyond the VP to the external argument NP.

If we assume that there is an absolute head, then *head with respect to F'* is defined as:

- (2.31) X is the head with respect to F of Y if X is marked for a value of F and either X is the absolute head of Y, or the absolute head of Y is not marked for F.

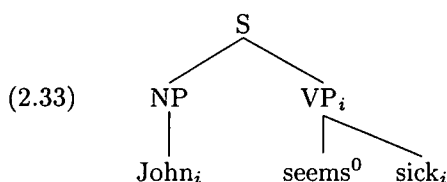
In cases of conflict the absolute head will always win: e.g.



With the notions of clause as coordinate structure and relativized head now explained we can move on to a discussion of the analysis of raising without the need for NP movement.

Williams departs from the standard PPT approach to raising. His approach depends on some assumptions regarding the structure of clauses and the introduction of the concept of *relativized head*. The advantage of the approach is that it removes NP movement from the analysis. NP-movement appears to be the least robust type of movement psycholinguistically (Fodor (1989)). And as we shall see in the Chapter 4, it seems to be the case that interpretation of the construction takes place not at the gap but at the verb. As thematic roles are associated very closely with the verb, any mechanism that allows interpretation at this level and without positing gaps should be preferred to a movement-based argument.

In raising structures, raising predicates assign no external theta-role and thus the complement is free to pass up its external theta-role via the mechanism of X-bar percolation and the notion of relativized head. In the structure below, *seems* is the head of its maximal projection, but it has no external argument associated with it. The notion of relativized head allows the non-head to pass its external argument through the maximal projection of *seem*.



There is a parallel between *seems* and the copula. These verbs seem to be dummy case markers in the same way that *it* is a dummy argument. The difference between them is the aspect of their meaning that describes the attitude of the speaker to the main proposition expressed by the non-case-marking predicate.

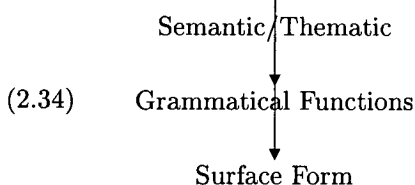
The relativized head mechanism and its concomitant approach to raising allows for the distinction between A'-arguments and A-arguments to be removed from the theory as there is no longer any requirement for A'-arguments to move, as they can now be assigned the theta-role they require via theta-role percolation.

2.4.2 Thematic Roles and Incorporation

I will use the term *incorporation* here for a number of syntactic processes that Baker (1988) has grouped together under that term. These syntactic processes had not been previously captured within one framework.

The syntactic processes in question are Noun Incorporation, Verb Incorporation, Preposition Incorporation and Passive Incorporation. Incorporation theory is predominantly a theory of grammatical function changing. The grammatical functions that are most commonly listed are Subject, Object, Indirect Object, Object of a Preposition and Obliques.

Grammatical functions, on the standard view, are seen as standing between semantic/thematic relationships and the surface form relationships among those phrases. Thus semantic/thematic structure is mapped onto grammatical functions and the grammatical functions are mapped onto the surface form.



With this architecture, languages do not state generalizations of how semantic and thematic relationships correspond to surface form relationships in any direct way. I take issue with this approach as one of the main features of my proposals in this thesis is that it is the direct mapping of semantic/thematic relationships onto surface form that is central to elements of language acquisition and psycholinguistic explanation.

This links very well to Baker's opinion that languages vary in the ways they express semantic relationships in surface form, together with the deeper theme that each language has a coherent system for this representation. So for example the same conceptual relationship is expressed by an SVO structure in English (2.35), by an SOV structure in Japanese (2.36) and by word form and not word order in Basque (2.37).

(2.35) Linda eats cake

Linda ga	okasi o	taberu
(2.36) Linda-NOM	cake-ACC	eat
'Linda eats cake'		

- | | | | | |
|--------|--------------------|------------|-------|-------------|
| | Linda-k | Rover | ikusi | du |
| | Linda-ERG | Rover(ABS) | see | Aux/3sS/3sO |
| | ‘Linda sees rover’ | | | |
| (2.37) | Rover | Linda-k | ikusi | du |
| | ‘Linda sees rover’ | | | |
| | Rover | ikusi | du | Linda-k |
| | ‘Linda sees rover’ | | | |

Baker (1988) regards structures which have previously been assumed to have undergone grammatical function changing in the lexicon as *thematic paraphrases*. This is the central claim of his theory: the same phrases stand in the same thematic/semantic relationships in the two alternative structures, but are merely expressed in very different surface forms.

A simple example of grammatical function changing is the passive:

- (2.38) a. Rover bit Linda.
b. Linda was bitten by Rover.

The grammatical functions that are changed here are the Subject NP becoming an Oblique and the Object NP becoming the Subject. Baker says the interest of these items is that the choice of forms is superfluous given that they have precisely the same meaning. However, I shall argue that these forms are not mere alternatives, but actually are the encodings of differing conceptual relationships.

Often associated with these syntactic changes are morphological changes which can perhaps be seen as signposts of the change having taken place. Baker claims that the ordering of the affixes reflects the ordering of the syntactic changes; he calls this the *Mirror Principle*.

Baker claims that at the heart of all grammatical function changing processes is the movement of a word or a lexical category - i.e. X^0 -movement. Thus any transformational rules given in terms of grammatical functions are regarded as being mere reflexes of X^0 -movement.

To give the notion of thematic paraphrases a more concrete foundation, Baker strengthens the notion of D-structure, so that it is a representation of thematic structure more generally.

(2.39) *The Uniformity of Theta-Assignment Hypothesis (UTAH)*

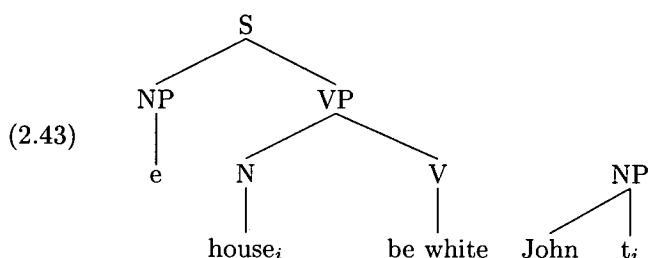
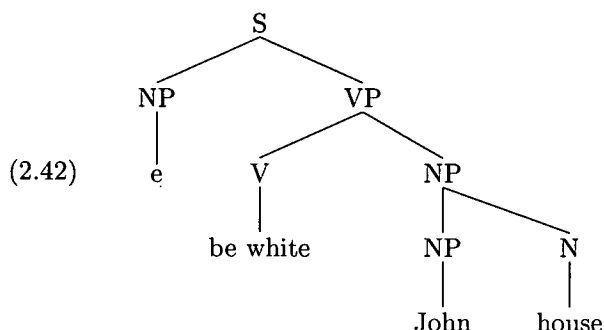
Identical thematic relationships between items are represented by identical structural relationships between these items at the level of D-structure

UTAH gives theoretical motivation for analyses in terms of X^0 -movement. An example of noun incorporation will indicate how the two alternative surface structures are related and that they have the same underlying deep structure.

- | | | | | |
|--------|-----------------------|-----|--------|--------------|
| | Ka-raku | ne | sawati | hrao-nuhs-a? |
| (2.40) | 3N-be white | DET | John | 3M-house-SUF |
| | John's house is white | | | |

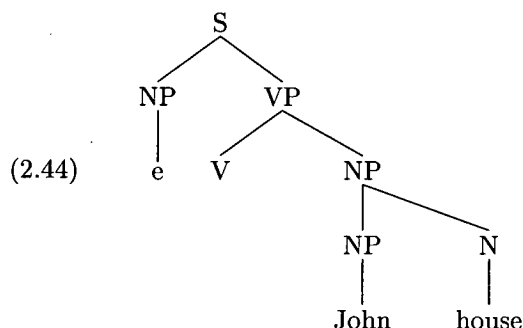
- | | | | |
|--------|-------------------------|-----|---------|
| | Hrao-nuhs-raku | ne | sawatis |
| (2.41) | 3M-house-be.white | DET | John |
| | ‘John's house is white’ | | |

Here (2.40) has an independent verb root *-raku* ‘be white’ and a noun root *-nuhs-* ‘house’ whereas (2.41) combines the two into a larger verb form, sometimes called a complex predicate.



One lexical item is incorporated within another and the morphology indicates this. The object agreement on the verb switches from neuter agreement, matching the thematic argument of the verb, to masculine agreement matching the thematic possessor of that argument. Thus the possessor comes to act like an object of the verb.

UTAH means that the two alternatives above must share the same D-structure thematic relationships:



If Incorporation indeed involves movement, it should have the properties of other movements and Baker presents a complex argument to show that this is the case. However, in my view alternations similar to incorporation in English reflect a difference in meaning in the underlying conceptual structure. I believe that this will be the case for the two alternative structures that Baker proposes for each case of incorporation. The full structure will have a slightly different meaning to the incorporated structure. The underlying conceptual meanings will be linked to their respective surface structures by linking rules.

One of the questions Baker attempts to answer with his theory of incorporation is why are some orders of grammatical function changes grammatical and others ungrammatical in

all languages. Baker formulates his answer in terms of possible and impossible movement operations. Following the other major theme of this thesis, such an explanation should, in my view, be couched in terms of impossibility of interpretation of the underlying thematic-conceptual structure.

In any case, it seems to me that the positing of movement should be kept as infrequent as possible. In the preceding discussion of Williams' work we saw how raising could be reanalyzed in thematic terms without needing a movement operation and this type of analysis should be applied as widely as possible.

2.5 Decomposition of Thematic Structure: A Continuum

We have seen that a decompositional approach to thematic information became unfashionable with the demise of Generative Semantics. Thereafter syntactocentrism was prevalent and thematic information was included, if at all, in the semantically barren notion of theta-roles. As theta-roles were brought into theories of argument selection or linking (Grimshaw) or the assignment of theta-roles was explored more fully (Williams), we see the gradual reintroduction of semantic entities such as *cause* and the discovery that the percolation of thematic information can make syntactic movement processes unnecessary.

Following on from this trend, there has been a redevelopment of theories of thematic information where decomposition plays a role. The amount of decomposition applied is on a continuum. At Level 1, the decomposition is constrained to the lexical categories V, N, A and P and event structure. At Level 2, there is causal structure, predication and such conceptual notions as temporal delimitation, states and directed motion. At Level 3, the linking rules are derived from the verbal entailments associated with the Agent and Patient roles. Level 4, which is the approach I will follow in this thesis, decomposes thematic information into conceptual structures, where the conceptual elements introduced are those and only those that are relevant to surface syntactic expression. I will discuss each of these levels in turn in the remainder of this chapter.

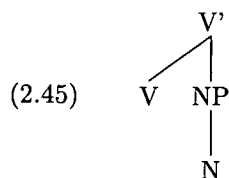
2.5.1 Level 1: Canonical Realization of Semantic Structures

One approach that begins again to regard thematic roles as decomposable in semantic terms in a limited way is found in the work of Hale and Keyser (1992, 1993). They argue that the proper representation of argument structure is itself a syntax, which they term *lexical argument structures* or *lexical relational structures*. It is perhaps not quite accurate to call their approach a thematically based one as they actually state that they are sympathetic to the arguments that thematic structure has no independent status. They see argument structure as linked to surface structure by a process like that of Grimshaw's Canonical Structure Realization. As I argued above, this is in effect a renaming of thematic structure as argument structure. The type of elements that occur in each of these approaches are similar.

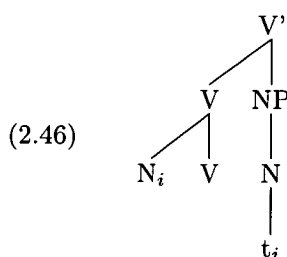
Hale and Keyser were led to their conception of lexical argument structures by considering denominal verbs such as *calve*, *lamb*, *shelve* and *bottle*. It is clear that the derivation is lexical, but this is independent of whether the process is *syntactic*. The syntactic process they

implicate in this derivation is one we have already come across in the discussion of thematic structure: incorporation.

Hale and Keyser contend that verbs like *shelve* and *saddle* are formed by the head-movement component of move- α , or incorporation. Thus for Unergative s, which are the simplest class of denominals *laugh*, *sneeze*, *neigh* the following structure:



is incorporated into



and this incorporation conforms to the Head Movement Constraint, where an X^0 may only move into the Y^0 that properly governs it.

It should be noted that the structures shown above are not D- or S-structures. Rather they are the structures of the lexical items themselves, the so called lexical relational structures. As these structures seem to be constrained by the same type of syntactic rules as syntactic structures, Hale and Keyser contend that lexical relational structures are syntactic in nature.

As discussed above, Hale and Keyser do not feel there is any need for a separate status for thematic structures. One of the reasons they see for this is that the number of thematic roles that have been identified in other frameworks is relatively few. They argue that the relative paucity of thematic roles is due to the fact that all lexical relational structures can be represented using the lexical categories V, N, A and P. This is combined with what they term *unambiguous projection*, where each lexical head determines an unambiguous projection of its category to a phrasal level XP, and an unambiguous arrangement of its arguments as specifier and complementizer.

It should be borne in mind that the attempt to limit Generative Semantics structures to V, N, A and P was what led in part to the elaborate representations of the structures. If the Hale and Keyser paradigm is followed to capture all lexical items, the same problems may occur. All other things being equal, semantic information should be represented in semantic terms and not forced into a syntactic straight-jacket.

What others see as thematic roles are seen by Hale and Keyser as being determined by these four categories and their projections. Thus an AGENT reflects the relational structure of the NP in the upper specifier position. Further, the matrix event of V implicates the subordinate event

(2.47) $e^1 \rightarrow e^2$

This structure represents a causal relation. The external argument of this, the NP in the upper specifier position bears an unambiguous relationship to the V'

$$(2.48) \quad n > (e^1 \rightarrow e^2)$$

Turning to the inner VP of the verb *shelve* we see that the head-complement relations involve V and P, with the latter subordinate to the former. V is a *dynamic event* *e* and P is the *interrelation* *r*.

The dynamic event implicates the interrelation:

$$(2.49) \quad e \rightarrow r$$

and the noun in this relationship is a subject of this change of state interrelationship. This is sometimes called a *theme* or *affected* patient in thematic theories.

$$(2.50) \quad n > (e \rightarrow r)$$

The lexical category A is associated with the notional *state* *s*. An action or a dynamic event implicates a state (again a *theme* in thematic theories)

$$(2.51) \quad n > (e \rightarrow s)$$

The fourth lexical category N is exemplified by the Unergative verbs such as *laugh*, *sneeze* and *sing*. A dynamic event implicates an entity, i.e. the event results in the creation, production or realization of that entity. Hale and Keyser (1992) point out that a mature theory of these elements would not represent parts of speech but would rather represent semantic constructs such as activity, spatio-temporal coincidence and so on, which are canonically realized as V, N, A and P.

$$(2.52) \quad n > (e \rightarrow n)$$

Thus what we see in these representations are the lexical categories V, N, A and P, a representation of events, and such notional concepts as *state*, *entity* and *interrelation*.

Bearing in mind that these structures are not S- or D-structures, but rather a syntactic representation of the underlying lexical relational structures of lexical items, these structures need to be projected into the true syntactic levels. Hale and Keyser argue for Unambiguous Projection and Full Interpretation to achieve this. These elements can be illustrated by considering gaps in the projection of Unergative verbs:

- (2.53) a. *The clown laughed the child.
 b. *The alfafa sneezed the colt.

For this not to be a problem, Unergatives would need not to have a subject in lexical relational structure terms, but this must be more than stipulation. It must follow from some principle. Hale and Keyser propose predication to be such a principle. For the Unergative the complement of the verb is not a predicate and thus no subject is required. If the specifier (internal subject) for the Unergative is excluded by Full Interpretation, then the expressed

subject must be external. This external argument related to the verb by predication appears in the s-syntactic position appropriate for predication. Another problem arises, how does this external argument get the role from the verb? Hale and Keyser argue that the subject is understood as an *agent* in a constructional manner.

There are a number of curious properties of the Hale and Keyser approach that may be its undoing. The fact that they see the LRS representations as canonical realizations of concepts such as activity, states, entities etc, would seem to indicate that there is arguably another level of semantic representation below lexical relational structures. If it can be shown that the realization of this underlying structure is not canonical, then it would make no sense to construct such a representation as lexical relational structure. In our exploration of the continuum of the decomposition of thematic structure we will gain more clues as to whether canonical realization of underlying semantic structures as lexical relational structures is possible.

2.5.2 Level 2: Linking Rules and Lexical Alternations

Regularities in the association of arguments with syntactic expressions are *linking regularities* and the rules that affect these are *linking rules*. Levin and Rappaport (1995) argue that verb meaning is a factor in determining the syntactic structure of sentences. They come to this conclusion in their analysis of two types of intransitive verbs, the *Unaccusatives* and the *Unergatives*.

An Unergative verb takes a D-structure subject and no object, it thus has an external argument, but no internal argument:

$$(2.54) \quad \text{NP}[_{VP} \text{ V}]$$

whereas an Unaccusative takes a D-structure object and no subject and has therefore an internal argument, but no external argument. In D-structure terms they are identical to passives:

$$(2.55) \quad [_{VP} \text{ V NP/CP}]$$

Traditionally there have been two approaches to unaccusativity; the syntactic approach that contends that Unaccusatives are not fully predictable from the semantics of the verb, and the semantic approach that argues that unaccusativity is not encoded syntactically. Levin and Rappaport take the view that Unaccusatives are both semantically determined and syntactically encoded. They assume that each verb has two representations: lexical semantic representation (similar to lexical conceptual structure) and a lexical syntactic representation (argument structure). The positions in argument structure in their terms are not represented by theta-roles, but rather by a purely syntactic representation:

$$(2.56) \quad \text{put } e, x < y, P_{loc} Z >$$

This representation shows an event with an external argument, a direct internal argument and an indirect internal argument with a locative preposition. Levin and Rappaport see

the relationship between argument structure and D-structure as trivial (cf. the discussion of Grimshaw's work above).

As for the lexical semantic representation, Levin and Rappaport propose lexical semantic templates, which are generative entities that are constructed from predicate decompositions and a number of constants representing the roles of the arguments in relation to the verb. Thus the non-causative meaning of *break* as in *The window broke* will be represented as:

(2.57) *break*: [y BECOME BROKEN]

Levin and Rappaport identify a diagnostic for Unaccusative verbs. There is a restriction on all resultative constructions in that the resultative can only be predicated of the direct object of the verb for which the resultative forms the result.. They argue that this restriction can be derived from a linking rule (this rule will be presented in the course of this discussion).

Thus for Unergative verbs where there is no direct object, no resultative phrase should be possible:

(2.58) *Dora shouted hoarse.

This can be rescued by the introduction of a fake reflexive

(2.59) Dora shouted herself hoarse.

The Direct Object Restriction follows from principles of grammar. Levin and Rappaport argue that the lexical representation of verbs in the resultative construction does not differ from the verb in isolation. In Unaccusatives, the post verbal NP in the resultative construction behaves like an argument, the NP in the Unergative version does not behave like an argument of the verb. Thus in Unaccusatives the noun can be externalized by middle formation or adjectival passives, but this is not the case for Unergatives:

- (2.60) a. This table wipes clean easily.
b. *This type of pavement runs thin easily.

To explain the direct object restriction on resultatives, Levin and Rappaport decompose the meaning of the resultative. Resultatives derive accomplishments from activities. Accomplishments in turn consist of an activity and a state. A Resultative denotes a change of state even when the verb in isolation does not denote a change of state:

- (2.61) a. The blacksmith pounded the metal.
b. The blacksmith pounded the metal flat.

As a first approximation, Levin and Rappaport propose the Change of State Linking Rule based on the preceding discussion:

(2.62) *Change of State Linking Rule*

- a) An NP that refers to the entity that undergoes a change of state in the eventuality described in the VP must be governed by the verb heading the VP.
b) An NP that refers to the entity that undergoes a change of state in the eventuality described in the VP must be the direct object of the verb heading the VP.

However, further semantic definition and decomposition needs to be carried out as not all Unaccusatives are compatible with resultative phrases. Specifically, stative verbs and verbs of inherently directed motion (*come, go, arrive*) are not compatible with resultative phrases. The semantic element that defines whether a resultative phrase is compatible or not is temporal delimitation. Resultative phrases serve to act as a temporal delimiter of an eventuality. Verbs of inherently directed motion are already lexically delimited (they describe an achieved change of location) and the addition of a second delimiter specifying a change of state is blocked. All statives are inherently delimited and a second delimiter is again blocked.

Levin and Rappaport then examine the causative alternation to identify further the lexical properties of the Unaccusative verbs with regard to their adicity and their lexical semantic characterization. The causative alternation is said to hold between a transitive and an intransitive version of the same verb where the transitive meaning is understood as ‘cause to V-intransitive’

- (2.63) a. Pat broke the window.
b. The window broke.

The subject of the intransitive verb and the object of the transitive verb play the same semantic role. Levin and Rappaport identify three types of intransitive verbs:

- Unaccusatives where the semantic representation of the causative is relevant. A dyadic verb with a single direct internal argument.
- Unaccusatives not related to the causative construction. A dyadic verb with two internal arguments
- Unergatives which a monadic with a single external argument

Levin and Rappaport assign both variants the same lexical semantic representation. One is NOT derived from the other. The intransitive verbs that alternate are dyadic, those that do not are monadic. The lexical semantic representation assigned is as follows:

- (2.64) [[x DO-SOMETHING] CAUSE [y BECOME BROKEN]]

The representation of both versions includes the predicate CAUSE, and this represents the meaning of such verbs as involving two sub-events. Each of the arguments of the verb are associated with a distinct sub-event, with the passive participant (patient, theme) associated with the central sub-event. The non-alternating intransitive verbs do not include the predicate CAUSE.

However, CAUSE alone does not determine which verbs alternate and which do not. A further semantic differentiation is required. This differentiation is between *internally* and *externally* caused eventualities. For example, verbs such as *break* and *open* are under the control of some external controller. However, verbs like *laugh* and *play* are not like this. They can be controlled only by the person engaging in the activity, control cannot be relinquished. Internal control verbs also have severe restrictions on the nature and identity of their single argument. The relevant representations are

- (2.65) a. Externally controlled: [[x DO-SOMETHING] CAUSE [y BECOME STATE]]
 b. Internally controlled: [x PREDICATE]

Interestingly given the speculative discussion to come in Chapter 6, Levin and Rappaport state that the relationship between the linguistic description of events and the events taking place in the real world is mediated by the human cognitive construal of events and this influences the form of lexical semantic representations. Indeed, the same events could be construed as either external or internal and this is often the source of crosslinguistic variation, according to Levin and Rappaport. This should be borne in mind in the discussion of language learning, parameter setting and grammaticalization to be found in the next chapter.

Externally caused verbs can detransitivize where an externally caused eventuality can come about without the intervention of the agent

- (2.66) a. The waiter cleared the table.
 b. *The table cleared.
- (2.67) a. The wind cleared the sky.
 b. The sky cleared.

Verbs of existence and appearance are all Unaccusative, but do not take part in the causative alternation. This is explained by a further semantic differentiation. These verbs have two arguments, the entity that exists and a location. These are argued to be two internal arguments as the location requirement is required and if it is not present it is understood by default. They are also found in the locative inversion construction and with *there*-insertion.

- (2.68) a. In front of her appeared a fabulous sight.
 b. There appeared a ship on the horizon.

With all these elements in place, Levin and Rappaport go on to define the linking rules. These linking rules are based around the causality of events.

- (2.69) *The Immediate Clause Linking Rule (ICLR)*
 The argument of a verb that denotes the immediate cause of the eventuality described by the verb is its external argument.

This linking rule applies to internally and externally caused verbs and intransitive and transitive verbs. The internally caused verbs are typically Unergatives such as *cough*, *shiver*, *sleep* and *snore*. Verbs of spatial configuration in the maintain position sense also fall under the ICLR. The ICLR also explains why internally caused verbs cannot be causativized. Their single argument is an external argument and a second external argument cannot be assigned.

- (2.70) *The Directed Change Linking Rule (DCLR)*
 The argument of a verb that corresponds to the entity undergoing the directed change described by that verb is its direct internal argument.

This linking rule captures the parallels between the change of state verbs such as *break* and verbs of inherently directed motion such as *fall* and *come*. These should be contrasted with verbs with no inherent direction such as *roll* and *bounce*. The distinctive properties of internal and external causation do not seem relevant to the semantic characterization of verbs of existence. These verbs are Unaccusative as they take the relevant auxiliary in languages such as Italian, they cannot take any object, and they cannot form adjectival past participles.

(2.71) *Existence Linking Rule (ELR)*

The argument of a verb whose existence is asserted or denied is its direct internal argument.

Verbs of appearance could fall under the ELR or the DCLR depending on whether the appearance is as a result of directed change. In this way the building up of criteria as to which linking rule will apply for a verb is similar to the idea of Proto-Roles found in Dowty (1991) (This is discussed further in the next section).

For the other monadic verbs there is a default linking rule

(2.72) *Default Linking Rule (DLR)*

An argument of a verb that does not fall under the scope of any of the other linking rules is its direct internal argument.

This rule has the result that a verb will take an internal argument before taking an external argument. This rule will not apply to manner of motion verbs with agentive animate arguments. These are internally caused and will thus fall under the ICLR. It will, however, apply to non-agentive motion verbs such as *bounce*, *roll* and *spin*. These are externally caused and will fall under the DLR, and are thus Unaccusative.

The linking rules are ordered with the DCLR taking precedence over the ICLR. Verbs such as *blush* and *bloom* are internally caused and as they are Unaccusative the DCLR must take precedence over the ICLR. Additionally, verbs of inherently directed motion are consistently Unaccusative, even though they can be used both agentively and non-agentively. This is predicted if the DCLR takes precedence over the ICLR.

With regard to the ordering of the ELR, if agentivity is irrelevant to the linking of a verb of existence, then the ELR has precedence over the ICLR and verbs of existence are consistently Unaccusative.

Some verbs have multiple meaning that allow them to fall under different linking rules depending on which of the verbs meanings is applicable. So for example *bake* can have a change of state meaning as in *bake the potatoes* and a creation meaning as in *bake a cake*. Levin and Rappaport discuss internally caused verbs and point out that some of them can acquire an additional meaning through what they see as the application of a lexical rule. Thus these verbs will fall under the ICLR on the basic meaning and the DCLR on a directed motion meaning.

So, for example, manner of motion verbs are found in some languages such as English, German and Modern Hebrew and not in others such as the Romance languages. In the non-directed motion use these verbs fall under the ICLR and are unergative, when a directional

phrase is added they fall under the DCLR and are Unaccusative. This is reflected by the auxiliaries these verbs take in their past forms in German.

Verbs of sound emission can also become verbs of directed motion, as in *The elevator wheezed upwards*. Sound emission verbs if internally caused are Unergative. If a directed motion phrase is added these verbs become Unaccusative. The sound must be a necessary concomitant of the motion.

Thus what we see here is that agentive verbs of manner of motion and internally caused verbs of sound emission are subject to a lexical rule that can map them into the class of verbs of directed motion. These verbs of manner of motion and sound emission can assume Unaccusative class if the resultative denotes a change of location and not if it denotes a change of state. The explanation of such mappings would require a full theory of possible and impossible meaning shifts. It is provocative that in a crosslinguistic context the languages that allow verbs of sound emission to be used as verbs of directed motion are the same languages that allow agentive verbs of manner of motion to become verbs of directed motion. It may well be the case that meaning shifts reflect the deeper properties of the language's lexical semantic organization. This will be explored further in the discussion of grammaticalization in the next chapter.

2.5.3 Level 3: A Simplification of the Syntax/Semantics Interface and the Selection of Arguments?

Thematic roles in PPT theory are fundamentally syntactic notions. There has, however, always been an underlying semantic origin for these roles. In all cases, thematic roles are creatures of the syntax-semantics interface. The differences in opinion from one PPT proposal to another, in general terms, concern how the thematic roles are implicated in argument selection.

Dowty (1991) has attempted to simplify the rules governing argument selection by returning to the semantic motivations behind the roles. He proposes the adoption of thematic proto-roles. He defines two proto-roles: Agent and Patient. These are defined in terms of the verbal entailments associated with these roles.

(2.73) Contributing properties for the Agent Proto-Role

- volitional involvement in the event or state
- sentience (and/or perception)
- causing an event or change of state in another participant
- movement (relative to the position of another participant)
- (exists independently of the event named by the verb)

(2.74) Contributing properties for the Patient Proto-Role

- undergoes a change of state
- incremental theme
- causally affected by another participant
- stationary relative to movement of another participant

- (does not exist independently of the event, or not at all)

Here the notion of Incremental Theme needs further explanation. A theme is incremental in that in a sentence such as *John ate the bagel*, one can look at the bagel in the event the sentence describes and see how much of it is eaten. One cannot look at John, however, and conclude anything about the completeness of his eating of the bagel.

The individual parts of the definition of the proto-roles should not be confused with feature decomposition as Dowty does not see the boundaries of these entailments as being entirely clear-cut. With these proto-roles in place, argument selection occurs according to the following principles.

(2.75) ARGUMENT SELECTION PRINCIPLE

In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient properties will be lexicalized as the direct object.

(2.76) COROLLARY 1

If two arguments of a relation have approximately equal numbers of entailed Proto-Agent and Proto-Patient properties, then either or both may be lexicalized as the subject (and similarly for objects).

(2.77) COROLLARY 2

With a three-place predicate, the nonsubject argument having the greater number of entailed Proto-Patient properties will be lexicalized as the direct object and the nonsubject argument having fewer entailed Proto-Patient properties will be lexicalized as an oblique or prepositional object (and if two nonsubject arguments have approximately equal numbers of entailed Proto-patient properties, either or both may be lexicalized as the direct object)

(2.78) NON-DISCRETENESS

Proto-roles, obviously, do not classify arguments exhaustively (some arguments have neither role) or uniquely (some arguments may share the same role) or discretely (some arguments could qualify partially but equally for both proto-roles).

The idea that traditional thematic roles can be split into their defining features is one that I agree with and will follow in this thesis. However, I am not going to say that any particular combination of features, or entailments, forms a Proto-Agent or Proto-Patient or Proto-anything else, for that matter. For me, the decomposition of thematic information is an end in itself; it does not need to define any traditional role or proto-role.

Thus Dowty points out that certain combinations of entailments seem to correspond to traditional roles: AGENT is volition + causation (or just volition, or just causation), EXPERIENCER is sentience without volition or causation, INSTRUMENT is causation and movement without volition or sentience. However, each verb is be associated with a lexical entry that defines its meaning in terms of these semantic decompositions, or entailments. We do not need to say that they correspond to anything. We only need to describe how these

elements of meaning (represented as thematic-conceptual structure, in my terms) correspond to the surface realization of the sentence, and how they are implicated in the processing and acquisition of language.

As Dowty himself explains, he is only looking for constraints on what kind of lexical predicates can exist in a natural language. This is an interesting aim, and one I address from a phylogenetic and cognitive setting in Chapter 6. Thus Dowty can say what the characteristics of an Agent or a Patient are, but he cannot predict the ordering of a sentence, or whether a sentence will occur in the active or the passive form, or which verbs can passivize and which cannot.

2.5.4 Level 4: Thematic-conceptual Structure

In this section I will discuss the most elaborate, yet constrained, decomposition of thematic structure developed thus far. I will term this representation *thematic-conceptual structure*. The basis for these representations can be found in the work of Pinker (1989) and Jackendoff (1990). The motivation for the adoption of decompositional representations of thematic information can be found in three main areas: Evidence that thematic hierarchies do not capture some constructions successfully, the failure of the theta-criterion, and aspects of language acquisition. I will lay out the first two of these motivations here, but will only touch on the third. The third motivation and its representational solution in terms of thematic-conceptual structure will be discussed in more detail in the next chapter, where language acquisition is discussed.

There are problems with the use of hierarchies to explain linguistic facts. So for example the following examples show a number of verbs that do not passivize.

- (2.79) a. *One kilogram is weighed by this melon.
 b. *30p is cost by this pen.
 c. *John is resembled by Bill
 d. *20 is equaled by five times four.

This data was first accounted for by an actional/non-actional or stative/non-stative distinction. However, non-actional, stative and abstract passives do exist.

- (2.80) This shop is owned by the news-agent.

- (2.81) The measures were justified by the situation.

It was thought that a constraint on passive formation based on the hierarchy of thematic roles (2.82) could better capture the data. The constraint was that in passives the surface subject must have a thematic role that is higher than the thematic role of the object of *by* or the implicit argument. So in the non-passivizable sentences above, in each case a theme is compared to what is a metaphorical location (as indicated by the locational prepositions).

- (2.82) theme < source/goal/location < agent

- (2.83) a. melons AT one kilogram
 b. pens AT 30p
 c. Bill's similarity TO John
 d. five times four is equal TO 20.

However, there are passives which are not predicted by the hierarchical constraint (2.84), and passives which do not exist, but are predicted to exist by the hierarchical constraint (2.85).

- (2.84) a. The audience was bored by the movie. (*audience* = Goal)
 b. Russia was invaded by a horde of locusts. (*Russia* = Location)
 c. The bed was covered by a blanket. (*bed* = Location)
 d. John was hit by a car. (*John* = Goal)
 e. The mountain was capped by snow. (*mountain* = Location)
- (2.85) a. *The house is had by John. (*John* = Possessor = Location, *house* = Theme)
 b. *A disk is lacked by this computer. (*computer* = Location)
 c. *Water is dripped by this ceiling. (*ceiling* = Source, *water* = Theme)
 d. *Water was contained by the bottle. (*bottle* = Location, *water* = Theme)

It is also of note that the hierarchy assumed to explain passivization facts is different to the one that Grimshaw assumes to explain nominalization facts. If any form of hierarchical explanation is to be adopted, it will have to be the same hierarchy for each of the facts that are to be explained by that hierarchy. In general, it must be said that hierarchies are by their very nature susceptible to counterexamples.

More seriously for PPT in general, as it puts in doubt the Theta-Criterion itself, there are both cases where a NP has more than one theta-role and cases where multiple NPs share a single theta-role. In (2.86) X has the theta-role of both GOAL and SOURCE.

- (2.86) X buys Y from Z
 Y changes possession from Z to X
 money changes possession from X to Z.

It is important to note that the counter-transfer does count. It is the counter-transfer that distinguishes *buy* from *obtain*. In (2.87) John has two theta-roles in conceptual terms, according to Jackendoff.

- (2.87) John carried some books with him.

Pinker (1989) examines the question of how it is that children learn the rules governing lexical alternations given Baker's Paradox. Baker's Paradox concerns the fact that children can learn which verbs have both alternate forms and which verbs do not without the benefit of negative evidence. Pinker examines four alternations: the dative (5), the causative (6), the locative (7) and the passive (8).

- (2.88) John gave the book to the boy.
John gave the boy the book.
- (2.89) John donated a painting to the museum.
John denoted the museum a painting.
- (2.90) The door opened.
John opened the door.
- (2.91) John loaded hay onto the truck.
John loaded the truck with hay.
- (2.92) John hit the ball.
The ball was hit by John.

He proposes that lexical rules act directly on the semantic representation of a verb, transforming it into a new semantic representation. The resulting semantic structure is then mapped onto surface syntactic structure by linking rules. It is only verbs that are members of certain semantic conflation classes that are eligible to be counted as the input to a lexical rule.

This is the basis of the solution as Pinker sees it. However, the full discussion of this topic will be left to the next chapter. Pinker relies on a decompositional representation of thematic information for his explanation and in the remainder of this section I will present how Jackendoff and Pinker have come to the decompositional elements they propose.

The theory of semantic structures must be constrained so that only the aspects of semantics that are visible to the grammar are allowed to be contained within the representations. The most successful approach to this is to look at crosslinguistic evidence. Items which appear in the agreement, classifier or other closed class systems can safely be represented in the semantic structure. The argument is that if some languages encode some of these semantic features directly in the syntax, it is quite likely that these features are relevant to the grammars of all languages, even when they are not directly represented in the grammar of a particular language.

Most of these aspects of verb meaning can be incorporated into the semantic structure of a lexical entry. There are some differences between Pinker and Jackendoff in how they are encoded and I shall indicate these. These differences are not major and it is unclear whether they make different predictions.

The aspects of word meaning that have been identified as relevant are given below. It is interesting to note that in the lesser decompositional theories, many of the same elements have been proposed as having an influence on the syntactic behaviour of verbs.

1. **The main event:** a state or motion, i.e. a position state or change of position or state predicated of a theme. The main event is the backbone of a verbs meaning.
2. **Paths, direction and location:** A verb can specify a particular path of motion of a theme with respect to an object, the speaker or with respect to a specific kind of object (*the ship berthed*).

3. **Causation:** Verbs can specify whether an event has been caused or just occurs (*kill* vs. *die*), and what kind of causation is involved. This can be analyzed in terms of force dynamics.
4. **Manner:** This refers to how an actor acts or a theme changes or refers to something the actor or theme is doing concurrent with the change.
5. **Properties of the Theme or Actor:** Verbs can specify that the arguments have specific properties (*drink*, *fressen*).
6. **Temporal Distribution:** Verbs can describe point-like events (*hit*), boundariless processes (*run*), iterated events (*beat*), entering a state (*sit down*), or point-like events that terminate an ongoing process (*arrive*).
7. **Purpose:** Verbs can encode activity in pursuit of a goal.
8. **Co-referentiality:** In some languages *to comb* means to comb someone else's hair, in other languages it means to comb one's own hair.
9. **Truth Value:** Verbs can express the assumptions of the speaker or some participant concerning the truth of a proposition.

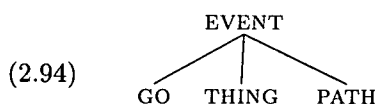
These elements then form the components of representational structures. The first set of elements the theory of semantic structures uses are the basic conceptual and ontological categories: THING, EVENT, STATE, ACTION, PLACE, PATH, PROPERTY, AMOUNT.

Each category is formed from semantic constituents which are specified in formation rules.

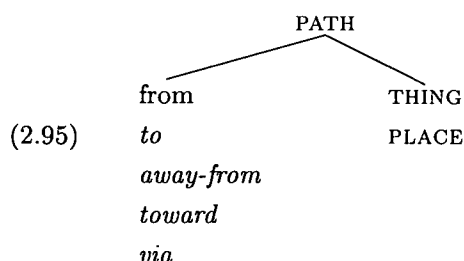
$$(2.93) \quad [\text{PLACE}] \rightarrow [\text{PLACE PLACE-FUNCTION } ([\text{THING}])]$$

Jackendoff uses these flat bracketed structures. Pinker, however, uses tree notations which are much easier to read. For this reason I shall also use the tree notation for most of the structures I discuss below.

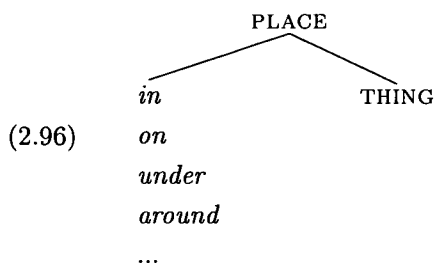
Formation rules as we have seen combine them into more complex concepts



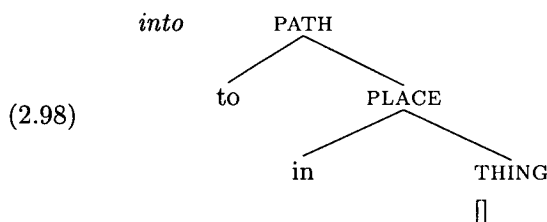
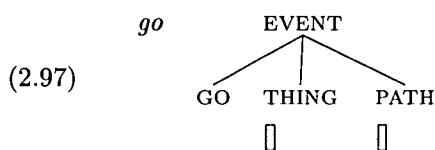
and PATH can be further expanded to



And a PLACE is a region defined with respect to an object, such as its interior, surroundings or surfaces.



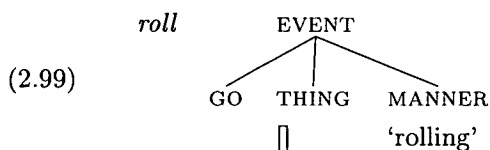
It is necessary to indicate which of these conceptual constituents in a particular semantic structure can serve as an open argument linked to a syntactic role position in the verb's argument structure. These open roles are indicated by open brackets (□)



These open arguments are mapped onto syntactically distinguishable argument types by linking rules. There is a further constraint on the syntactic structure with respect to the semantic structure: Major syntactic phrasal categories must correspond to a complete conceptual category.

Sentence structures are only well formed if they contain phrases corresponding to the conceptual categories selected by the verb. For example, *put* selects for a preposition which incorporates the direction *to* (*on*, *in*, *by*....) and a PLACE. This has the advantage that the individual prepositions which can occur with a verb do not have to be listed individually if they have something in common which can be represented with the inventory of semantic constituents.

But not everything in the semantics of the verb is visible to the grammar. For example in manner of motion verbs, the manner of motion is listed as another daughter node of the EVENT. The manner of motion has no effect on the syntactic structure of the clause which contains it. We say that the manner of motion is *opaque* to the linking rules. This opacity is notated with quotation marks.



So much for motion, a canonical example of a STATE is a THING being situated at a PLACE.



The machinery for spatial events can be extended to non-locational fields with the proviso that the name of the extension is added to the maximal conceptual constituent. Thus for the possessional field:

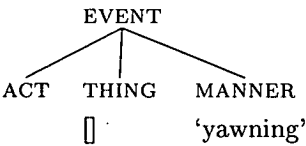


(2.102) *go*: (e.g. *went from sick to well*)



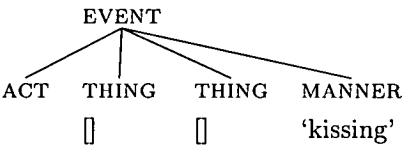
Actional events involve the function ACT, which takes either a single act argument, or two arguments, one being the agent, the other the patient. The simplest actional event has a single argument and can be found in Unergative transitive verbs.

(2.104) *yawn*:



An example of a dyadic ACT event is *kiss*

(2.105) *kiss*:



Jackendoff deals with Actional Events in a more complex way. He uses two tiers of representation to encode the totality of conceptual roles associated with a predicate: a Thematic Tier to deal with motion and location and an Action Tier to deal with Actor- Patient relations. He provides us with tests for Actors and Patients.

A Patient will fit into the frame

(2.106) What happened to NP was....
 What Y did to NP was...

and an Actor into

(2.107) What NP did was.....

So on an intuitive level the contents of the thematic tiers for the following sentences are:

(2.108)	<i>Sue</i>	<i>hit</i>	<i>Fred</i>	
	Theme		Goal	(thematic tier)
	Actor		Patient	(action tier)
(2.109)	<i>Pete</i>	<i>threw</i>	<i>the ball</i>	
	Source		Theme	(thematic tier)
	Actor		Patient	(action tier)
(2.110)	<i>Bill</i>	<i>entered</i>	<i>the room</i>	
	Theme		Goal	(thematic tier)
	Actor			(action tier)

To encode these properly we need a *functional* representation that has Actor and Patient as its argument positions. Thus events are formally elaborated in the Action tier.

- (2.111) EVENT $\rightarrow$...
 AFF(<[THING]>, <[THING]>)
 (Angled brackets encode optional arguments)

AFF is an additional mainstream function. The first argument is the Actor and the second the patient. E.g. for *The car hit the tree*.

- (2.112) INCH [BE ([CAR], [AT [TREE]])]
 Event AFF ([CAR],[TREE])

Jackendoff goes on to use the action tier in his analysis of causation which uses the Talmy (1985) theory of force-dynamics. Force dynamics involves the interaction of two characters: the agonist and the antagonist. The agonist has the tendency toward performing or not performing an action. This tendency is opposed by the antagonist. So in *Harry forced Sam to go away*:

- (2.113) CAUSE([HARRY] [GO([SAM, [AWAY]])]
 AFF(SAM,]
 AFF([HARRY],[SAM])

The antagonist here is the actor. The EFFECT is what he aims to bring about. The inference rule for CAUSE is that Harry was successful in his efforts. However, there are examples where it is not clear whether the act of Causation was successful or not.

- (2.114) a. Harry pressured/urged/goaded Sam to go away (and he did/but he didn't).
 b. Harry impeded Sam's going away (and in the end he didn't/but he did).

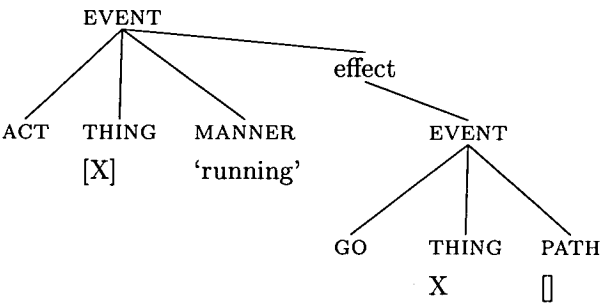
To encode the degree of success of an act of causation the CAUSE function is modified to the function CS with various superscripts:

- CS+ -successful outcome
- CSu -undecided outcome
- CS- -unsuccessful outcome

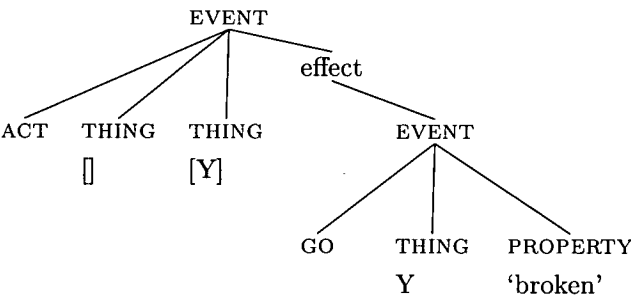
Jackendoff captures inter-event relations by encoding them in different tiers within the representation. On the other hand Pinker uses a small set of subordinating relations that allow one event to be embedded in another.

One subordinating relationship is that of cause and effect, the link between the two is annotated with 'effect'. The co-reference between entities is once again notated by indices. Only one of these positions is the trigger for the relevant linking rule. These relevant position are those in square brackets.

(2.115) *run*



(2.116) *break*:



Given this very different representation of the inter-event relations, Pinker has to propose a different way of encoding the degree of success of the ACT (Jackendoff did this using CS+ /u/-). Pinker claims that two things have to be represented: (i) Whether the focus is on the result of the Actors action or on the action itself and (ii) Whether the Actor is stronger than the Patient. These differences are encoded by introducing a set of subordinating relations. These subordinating relations and the relevant focus and potency are:

Relation	Focus	Potency	Cause	Effect
'effect'	cause	success	yes	yes
'cause'	effect	success	yes	yes
'despite'	effect	failure	yes	no
'but'	cause	failure	yes	no
'let'	cause	success	no	yes
'prevent'	cause	success	yes	no

Instrumentals will play an important role in the psycholinguistic discussion of Chapter 5. Jackendoff proposes the following rule for instrumental adjuncts.

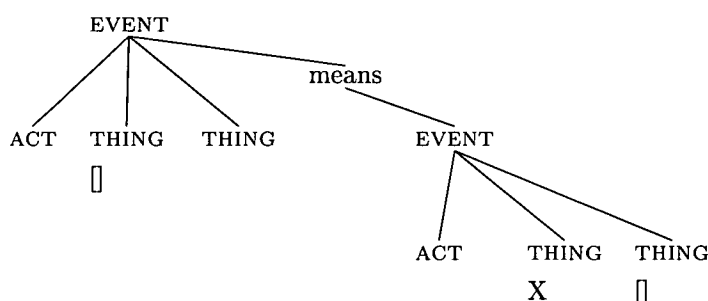
(2.117) *Instrumental Adjunct Rule*

$$\begin{array}{l}
 \text{If } V \text{ corresponds to } \left[\begin{array}{c} \dots \\ \text{AFF}([X],[Y]) \end{array} \right] \\
 \text{and NP corresponds to X, then} \\
 [S..[VP..V[PP \text{ with NP}...]...] \text{ may correspond to:} \\
 \left[\begin{array}{c} \text{AFF}([X]^\alpha,[Y]^\beta) \\ \text{BY} \end{array} \left[\begin{array}{c} \text{CS+}([\alpha],[\text{AFF}([\gamma],[\beta])]) \\ \text{AFF}([\alpha],[Z]^\gamma) \end{array} \right] \right]
 \end{array}$$

This rule would have fairly unconstrained applicability any sentence with a verb with the semantics of one entity acting on another occurring together with a prepositional phrase containing with would be an appropriate entity.

The framework for the instrumental according to Pinker is encoded within the lexical entry for the preposition.

(2.118) *with*:



It is possibly the case that the types of verbs that *with*-instrumentals can be attached to are semantically constrained in the same way that lexical alternations are constrained by the semantics.

The nature of linking rules is not yet clear and the theory has not been worked out in depth. I will have more to say on this in Chapter 5. Jackendoff believes that the linking rules should operate on the Action Tier to derive syntactic structure. Pinker's linking rules operate on the arguments enclosed in square brackets in his representations. This difference may not be important, but if parsing decisions are based in any way on semantic representations, Pinker's method would involve one less step, with the linking rules operating on arguments which are embedded directly within the semantic structure, instead of in a separate tier which is only co-indexed with arguments in the semantic structure.

2.6 Chapter Summary

There are a number of questions that arise from the discussion in this chapter.

First, What is the true nature of the mapping of the semantic aspects of a lexical representation onto the syntactic representation? This aspect will be developed throughout this thesis.

Second, What is the true nature of the lexical-conceptual-semantic representation we need to assume? I argue that the fully compositional structure proposed by Pinker and Jackendoff is the correct route to take. I will back this up in the discussion of language acquisition and psycholinguistics in Chapters 3 to 5. The compositional approach to thematic-conceptual structures makes the correct psycholinguistic predictions for a number of new and several previously intransigent older psycholinguistic findings.

Third, at what level are the typological differences between languages to be explained; the lexical-conceptual-semantic level, the syntactic level or at the level of the linking between the levels? The answer to this question depends largely on the answer to the previous question. In the development of the explanation of the psycholinguistic facts discussed in Chapter 5,

I will propose that thematic-conceptual structures are linked by reversible linking rules to surface syntactic expression and the operation of these rules can be tapped by experiments on sentence-level priming.

Chapter 3

Considerations from Language Acquisition

In the preceding chapter, the role of thematic-conceptual structure in syntax was discussed. I noted that much of the argumentation behind one of the more elaborated theories of thematic-conceptual structure had its basis in considerations of language acquisition. In this chapter I will discuss language acquisition in more depth with the aim of establishing at what level and to what extent thematic-conceptual structure is important to language acquisition.

I begin the chapter by comparing and contrasting the Classical paradigm of learning to the currently more fashionable connectionist view of learning. I will argue that the Classical framework is to be preferred.

I then go on to discuss the differing opinions as to whether functional categories are available to the child acquiring language from the beginning of acquisition (The Continuity Hypothesis) or whether they are acquired by maturation, being preceded by a more thematically based acquisition of language.

I will argue that a middle ground between the no-functional categories approach and the all-functional categories approach can be established. I then go on to discuss the theories of parameter setting in general terms and conclude that they suffer from a number of problems that makes their successful application unlikely.

There then follows an initial discussion of the relationship between phylogeny and ontogeny which includes also an examination of the early cognitive capacities of children in non-linguistic fields and a discussion of the impact these early capacities may have on language acquisition.

I return to the decompositional approach to thematic-conceptual structure that indeed reflects many of these early representations and conceptualizations and very successfully captures and explains the acquisition of a number of lexical alternations.

Rounding off the chapter, I present investigations of grammaticalization and note that much of what diachronically is grammaticalized in syntax reflects conceptual content. It is my view that the acquisition of such grammaticalized conceptual content will also need to be based on a decompositional representation of thematic-conceptual structure.

After discussing the role of thematic-conceptual structure in psycholinguistics in general

in the next chapter, I will go on to test the predictions of this chapter in chapter 5.

3.1 Theories of Learning

3.1.1 Connectionism

The representational type of the thematic-conceptual structure I follow in this thesis can be described as a Classical theory of representations and rules. I have chosen the Classical paradigm as opposed to the more fashionable connectionist paradigm as I am still convinced that much of the criticism of Fodor and Pylyshyn (1988) still holds, in spite of attempts by connectionists to discount them. As the Classical/connectionist paradigms clash most when it comes to learning theory, I shall present the criticisms of Fodor and Pylyshyn in this section.

The issue is not the explicitness of rules - Classical architectures are not committed to the idea that explicit rules mediate behaviour. Equally, the issue is not about the reality of representational states; both connectionists and Classicists are representational realists. Finally, the issue is not the nonrepresentational architecture; a connectionist network may well be able to implement a Classical architecture at the cognitive level.

The differences between the two paradigms occur at a number of levels. There is a difference as to where semantic content is assigned. In connectionism it is assigned to nodes, whereas in Classicist accounts semantic content is assigned to expressions. Causal connectedness is modeled in connectionist systems only as a primitive relation between these nodes, whereas Classicists model not only causal relations but also connections based on *structure*. If the causal relations that need to be explained can only be explained on a number of complexly related levels, connectionist systems with the flatter architecture of nodes may be stretched.

In this same vein, Classicist versions of mental representation have a combinatorial syntax and semantics which allows for the structure sensitivity of processes to be modeled, processes can be modeled according to their *form*. Thematic-conceptual structure should be based on differences, not only in content, but also in form for its operation. Only in this way can the psycholinguistic predictions of Chapter 5 be realized.

Constituent structure can only be said to exist in a system when the individual parts of a semantically evaluable system are themselves evaluable. But in a connectionist neurally distributed representation the relationship between nodes and their constituent units is a between-level relation. They are thus not expressed in the same terms. And micro-features are not the same as constituency in any case. There is no way that the constituency relation of *John* in the sentence *John loves Mary* is of the same type of part-whole relationship or feature-system as, say, *has a handle* could be said to be a feature of *cup*.

However, Chalmers (1994) claims he can refute this type of claim in one sentence. He writes:

If Fodor and Pylyshyn's argument is correct as it is presented then it implies that no connectionist network can support a compositional semantics; not even a connectionist implementation of a Turing Machine or of a Language of Thought. But this is a problem for Fodor and Pylyshyn, as it is well-known that connectionist networks can be used to implement Turing Machines (or at least Turing Machines

with arbitrarily large but finite tape), and it is well-known that Turing Machines can be used to support a compositional semantics.

But this very sort of argument was attacked in the original Fodor and Pylyshyn paper:

we are not claiming that you can't reconcile a Connectionist architecture with a combinatorial syntax and semantics for mental representations. On the contrary, of course you can: All that's required is that you use your network to implement a Turing machine, and specify a combinatorial structure for its computational language. What it appears that you can't do, however, is have both a combinatorial representational system and a Connectionist architecture at the cognitive level.

Connectionism seems in many senses to be associationism that has been improved by adding hidden units that detect the pattern in the visible units and can then provide a more global statistics. But the fact that the statistics are embedded to some extent does not hide the fact that the representation only reflects the statistical properties of the input and does not *explain* the statistical properties in any way.

As the connectionist system does not allow for semantic and syntactic constituency, the semantic/syntactic structure of the representation of the input does not determine the selection of the output as syntactic/semantic structure is not relevant to the representations of connectionist models. Productivity must be based on expressions belonging to a generating set and must be based on a symbol system, Connectionist models are not symbol systems and thus cannot model productivity. For this reason connectionists reject the very idea that productivity exists and for this very reason linguists, who have seen so many examples of productivity, often reject connectionism.

Language is systematic, if you can understand or produce one sentence, you can understand and produce others. Therefore you can learn part of a language without learning the rest, whereas the sentences in a connectionist system are generally atomic. Systematicity, in turn, relies on compositionality. The parts of an expression have the same contribution in different contexts. A bridge to psycholinguistics can be seen here. Language expresses thoughts. The compositionality of language reflects the compositionality of the thoughts. The compositionality of thoughts is represented in this thesis by thematic-conceptual structure and this is joined to the surface expression of language by linking rules. As we have seen above, compositionality is not a feature of connectionist networks, and the meaning of a word in a connectionist system of nodes is never truly compositional as it is always reliant on its co-occurrence with other words.

Counter-arguments against these views of connectionism tend to concentrate on the computational properties of the competing systems and use as their evidence small networks with a very limited range of content relations, whereas the impetus of the arguments given in Fodor and Pylyshyn is the complexity of the explanations that the contentful linguistic data requires.

I accept the argument from Chater and Oaksford (1990) that compositional and logical representation is not everything and that there are indeed aspects of cognition, such as vision, that may well not require a compositional semantics, but on the other hand language is a compositional system *par excellence* and I see no way that a connectionist system can model the complexity of the data.

In addition to the problems discussed above, there are other problems with statistical correlations in learning theories: They must be constrained so that it does not occur that every element is related to every other element in some statistical way. Rules are often subtle and subject to exceptions and how did the process begin in the first place - i.e. how did the statistical properties of the data arise initially. I do not believe that connectionism has, as yet, provided answers to these questions.

3.1.2 Constructivism

Constructivism (Quartz and Sejnowski (1997)) is an attempt to join cognitive development to neural development. Neural development in this view occurs, in part, according to the input the brain receives. The brain structures that this input creates shape the hypothesis space for further input and further neural development.

Constructivism assumes that all that is innate is the constructivist learning mechanism itself. That is, all that evolution has provided us with is a mechanism of abstracting from the environment the information needed to construct all of the mental representations and mechanisms we need to function as humans. There does not seem to be any way to constrain what is picked up from the environment in terms of content, except perhaps the limitations imposed by the structure and function of our senses. There is also no true explanation of the order of acquisition, it has to be assumed that what we can cognize first from the environment is in some way simpler than what we can cognize later.

Another issue in the learning of language and cognitive abilities in general, is the time course of learning with respect to the formation of the brain. The theory describes the learning architecture as nonstationary as it is assumed that the learning properties of the system are time dependent. As input determines structures it is hypothesized that factors such as the order of presentation will affect the order of acquisition.

The nonstationary approach contrasts with theories that assume the Continuity Hypothesis (Pinker (1984)) which argues that adults and children should be seen as qualitatively identical unless forced to abandon this principle by some explanatory failure.

I propose a third view which is analogous to the constructivist view in that it allows for time-dependency in learning. However, in contrast to constructivism I allow for genetic maturational programming to play a role. The development of the brain is genetically determined and the various stages of development will each allow for the emergence of certain cognitive abilities mediated by input. This position seems more realistic than the extreme constructivist position as the constructivists have to assume that children will receive input in fundamentally the same order (or that the order is not critical for some aspects of learning), whereas the maturational account stresses the cognitive abilities of the brain in the range of what proportion of the input it can process at each maturational stage.

The maturational/constructivist accounts are neutral with respect to what forms a large part of the Quartz and Sejnowski argument: the growth of structural complexity in the brain. The constructivist approach predicts that the input determines the neural structure, whereas a maturational account emphasizes that the stages of neural development predict the nature and range of the cognitive capacities that a child will show at a particular maturational level.

Thus the constructivists claim to be following a Piagetian framework, but seem instead to be putting the whole weight of their argument on the nature of the input.

So for example, Quartz and Sejnowski argue that the cortex is largely equipotential and free of domain-specific structure - but this does not explain that normal development results in localization. The argument is that input comes into the brain at certain points and seduces the nearby cortex into forming the appropriate representations.

Changes to the neural substrate in developmental terms are relevant to both the constructivist and the maturational account. The differences are at the level of the processes that regulate these changes in the neural substrate and their correspondence to changes in the form of the representation that is then available. In the constructivist account all changes to the neural substrate are determined solely by the input, whereas in the maturational account there are genetic determinants of neural substrate changes that interact with the input. In constructivism the changes in neural substrate drive the type of representational change, whereas in a maturational account changes in the representation drive the change of neural substrate.

Quartz and Sejnowski have the critical target of neural selectionism (Edelman (1987)) and they make out that neural selectionism is the standard theory in contemporary learning theory. This is not the case, and with respect to language learning, neural selectionism is a minority view. Universal Grammar does not have to be inextricably tied to formal learning theory and its nonstationary character. The seeking of a phylogenetic source of language is exactly what is required and, as I will argue in Chapter 6, predicts the order of development. There is a fundamental difference between selection in the species and neural selectionism, but Quartz *et al* seem intent on conflating these. It is also not the case that a nativist account necessarily implies selectionism in learning.

So Quartz and Sejnowski claim that behind language (or any complex cognitive ability) is no phylogenetic source, but rather increased representational flexibility. However, it does not seem that the brain is as flexible as they would predict given the infinite number of statistical properties of the input that a truly flexible brain would be able to pick up.

They claim richness of structure and not poverty of stimulus is the correct position and use neural networks as the basis of this argument. They do concede that the links that these nets can make have to be constrained by intrinsic properties. This sounds nativist to some extent, but they do not elaborate this point.

To attempt to include a mechanism of localization and a constraint on learning, Quartz and Sejnowski make the proviso that input-determined change must be *local* in nature and must not *undo* previous learning. This still does not explain the order of acquisition and seems to contradict language learning data which often shows children reverting from correct structures to incorrect structures during the process of rule learning.

To allow for critical periods, Quartz and Sejnowski propose limited temporal windows in which activity dependent processes can affect the neural structure. If these limited periods are not genetically determined I can see no way that they could otherwise be instantiated. The fact that representational change must be local seems to allow for localization to some extent particularly with respect to the input devices (eyes, ears) but does not seem to so successfully explain the localization of the speech centres in the brain.

They then turn to the learning properties of the constructivist paradigm. Their starting point is formal learning theory as applied to language acquisition. Gold (1967) has proposed that language learning involves a search through a hypothesis space to a target concept and to make this feasible given time and input constraints an inductive bias is required: some form of Universal Grammar. On the other hand I would argue that what can be described as Universal Grammar has actually been determined in computational terms by the earlier cognitive abilities that language has developed from phylogenetically.

The impasse of bias and variance leads to nativist arguments: high variance can only be overcome by large training sets; and high bias needs to be carefully chosen for the task in hand. The second of these is what I would argue phylogenetic selectionism achieves over evolutionary time.

Although the majority of the discussion of the link between language development and language evolution will be left to Chapter 6, due to its more speculative nature, it is important to stress now that the weight of the evidence points towards a classical theory based on maturation and not towards either a constructivist nor a connectionist account.

3.2 Functional Categories in Language Acquisition

There are currently two opposing camps in the discussion of the role of functional categories in language acquisition. The first camp (Poeppel and Wexler (1993), Deprez and Pierce (1993), Clahsen *et al.* (1993), Wexler (1994), Rizzi (1994), Deprez and Pierce (1994), Harris and Wexler (1996), Hoekstra and Hyams (1998)) see child language as showing ‘continuity’ with the adult grammatical system. In other words, children have the full range of functional categories also shown by adults. The second camp argue that children do not have the full set of functional categories (if any at all) and rely rather on a thematically based system (Radford (1990), Lebeaux (1988), Guilfoyle and Noonan (1992)). For the aims of this thesis the second camp’s arguments are the best fit. However, I see merits in both arguments and would argue rather for a middle way, where the thematic basis of children’s language provides bottom up data that is built into linking rules (setting of parameters) that are constrained by the top-down principles of Universal Grammar.

The nature of the starting point for parameter setting is important for another reason. Parameter setting has to work with something and the nature of the something that forms the starting point will influence and constrain the parameters that can be set and will form the basis of the form of the parameterized system. If such starting points can be established, we might expect that when the parameterized system breaks down, in agrammatism, for instance, the speaker will revert to the forms of the unparameterized system.

3.2.1 An initial view: a thematic based acquisition device

An important work on the starting point for parameterization is found in Radford (1990). Radford argues that early child English lacks what have come to be known as functional categories. These functional categories are the Determiner system, the C system, the I system, and the Case system. He argues that what is left is basically a thematically based system of

language. This suggestion is particularly interesting for the aims of this thesis as it predicts that parameterization will consist of adding other expressive linguistic devices to the thematic system in a language-specific way. This prediction fits in very well with the conceptual-thematic structure and linking hypothesis developed throughout this thesis.

If the determiner system is indeed absent from child speech we would expect that where adults use determiner phrases children will use simple noun phrases. This expectation would appear to turn out to be the case as children use indeterminate nominals in spontaneous speech, in answer to *what*-questions and in imitative speech. It is often argued (e.g. Fukui (1986)) that genitive 's is a determiner and thus in Radford's theory we would expect it to be missing from children's speech in any productive way and Radford gives some empirical data that this is indeed the case.

The determiner system also encodes semantic relations such as the referential properties of nominals and pronominals. If children lack determiners they should also have trouble with the referential properties normally encoded by them. Young children seem not to realize that nominals in adult speech are used only in third person contexts, and will use them to refer to the speaker or the addressee.

Radford notes the similarity in form of the utterances of young children and Small Clauses uttered by adults. Adult small clauses have the schematic form [DP, X']. Given that children seem to lack a determiner system we would expect children to utter clauses of the form [NP, X']. If it is indeed the case that young children are uttering what in effect are small clauses, it suggests that they lack a Complementizer system, or C-system.

Adult small clauses also lack a C-system. They can never be introduced by overt complementizers and do not allow auxiliaries to be preposed to a putative C-position. They also lack a position where *wh*-phrases can be preposed to; thus they lack C-spec. It is also the case that early child complement clauses show the same [NP, X'] patterns they produce in their independent clauses, thus as a complement of *want* small clauses of the form [NP, P'] are common (*want* [hat on]). Also common is the structure [NP, V'] (*want* [Teddy drink]). Children never introduce these complements with overt complementizers.

In full adult speech, the C-constituent provides the landing site for Auxiliary inversion. Given that the C-constituent does not exist in Early Child English, we would expect children not to show any examples of subject-auxiliary inversion in direct questions, or any evidence of *wh*-movement to the equally absent C-spec position. This expectation is confirmed by the evidence. The few examples of *wh*-questions that young children do produce seem to be formulaic and are not productive.

Radford argues that the empirical language acquisition data for English shows no evidence of the acquisition of the following aspects. Given that these aspects appear to be missing Radford concludes that Early Child English lacks an I-system.

- infinitival *to*
- modals
- finite verb inflections (+s/+d)
- *do*-support

- copula *be*
- progressive *be*
- perfective *have*

It is generally assumed that Case is an inherent property of the D-system, as it can clearly be argued to be, as it is in the D-system that case is marked in most languages that are rich in Case. Thus if Radford is correct that children lack a D-system, we would expect them in turn to lack a Case system. If children were to lack a Case-system we would expect children to use caseless NPs in situations where adults use case-marked DPs.

Pronominals are arguably determiners and also encode case directly. It is thus predicted that if young children lack determiners and a case-system, pronominals will likewise be absent from their production data.

Radford argues that these early lexical structures produced by young children might be said to be *thematic* structures in the sense that they directly encode the thematic relations between a head lexical category and its arguments, in so far as all theta-marked arguments of a given head noun are superficially contained within its maximal projection and, I argue, they are later linked to other expressive linguistic devices.

This thematically based structure is certainly promising for the approach taken in this thesis. However, the nature of the thematic stage of child language acquisition has not been characterized in any detailed way by Radford.

Another paper at this end of the spectrum comes from Guilfoyle and Noonan (1992). They suggest a *Structure Building Hypothesis* which makes reference to a distinction between lexical and functional categories. Two grammars are postulated; a lexical grammar (LG) and a functional grammar (FG). Guilfoyle and Noonan suggest that only LG is present at the onset of language acquisition and that FG emerges according to a maturational schedule. This avoids the trigger problem of the Continuity Hypothesis where it is difficult to see how the child can easily identify the criteria that are crucial in order to reset a parameter from the unmarked form, and the question as to why a child suddenly pays attention to a trigger, when the trigger was available all along. The structure building hypothesis is also more constrained than a purely maturational approach, where the stages in language acquisition are related purely to physical maturation. This allows no input for linguistic data and every change can be explained *post-hoc* as maturation, resulting in a loss of explanatory power.

LG is regarded as having a thematic basis and move- α at this stage would always have to be via adjunction as no landing sites would be available. The emergence of functional categories could perhaps be as a result of non-linguistic changes in the child's perception. The grammar will not violate UG at any stage but rules relying on functional categories will hold vacuously.

Thus three word speech is characteristically telegraphic, with the first elements being verbs and nouns. Guilfoyle and Noonan interpret child language acquisition in the following ways: there are few determiners in child speech and it unclear that those that do occur can be analyzed as being in D'. The case filter is not in force in child grammars and this is relevant in the acquisition of IP. They argue that INFL is not present and hence there is no agreement or tense on the verb, and modals and auxiliaries are absent. With no IP, there is no Spec of IP and thus no landing site for movement and the subject remains in the base-generated position

internal to the VP. But as no INFL or case is present the NP can appear as a SPEC of VP without violating UG.

Guilfoyle and Noonan believe that the functional categories appear at different times, with INFL before COMP and DP before IP. However, they also contend that there is crosslinguistic variation in the order of emergence of functional categories depending on the saliency of particular functional categories in the input. This would seem to cast doubt on a purely maturational account.

3.2.2 The Opposing View: The continuity of functional categories

Deprez and Pierce (1993, 1994) note that although there are systematic errors in the placement of the subject with respect to negation (expected if subjects are generated VP-internally), errors with respect to the placement of negation relative to the verb are systematically absent. Deprez and Pierce argue that the early acquisition of verb movement together with the early placement of subjects and negation strongly imply that functional projections are represented in early child grammars.

They make the following predictions:

1. If children fail to raise the subject consistently and the subject is VP-internal, then the order NEG-SUBJ-V is expected.
2. Given main verb raising in French, we expect to find Vfin-NEG-SUBJ orders.
3. The absence of verb raising in English means that it is not expected that we will find postverbal subjects and negatives.
4. If verb movement to COMP is early then we expect to find postverbal subjects and negatives in 2yr-old German.

Deprez and Pierce provide crosslinguistic acquisitional evidence to support their position. There is evidence from English for the early use of NEG-SUBJ-V orders. They argue that the NEG cannot be in COMP as at later stages NOT appears in non-initial position, and there are no utterances where NEG appears to the left of an overt auxiliary. Thus, they argue, the word order pattern reflects the option to leave the subject VP-internal, implying a 'continuity' of structure between adult and child grammars.

For French, in early child language, the NEG is positioned to the right of the finite verb, and the position of *pas* depends on whether or not the verb is finite. Thus, the contrast in word order between French and English is a result of verb movement and not as a result of a difference in D-structure.

For German, it could be the case that V-to-C movement develops simultaneously to V-to-I movement. However, the data suggests that this is a later development. If the child raises verbs to COMP, then the negative element should be postverbal. However, in early child German, preverbal negatives occur, meaning that raising to COMP is not obligatory in early child German. Deprez and Pierce also give evidence from Swedish, French subject-clitic inversion, and embedded questions in English to show that movement to COMP is delayed in these languages as well.

Some have argued that this delay in movement to COMP suggests that the functional category COMP does not exist for early child grammar. Deprez and Pierce disagree with this view as they argue COMP has to exist for the construction to be optional. They suggest the following hypothesis for the delay in using COMP in the adult manner: Parameters involving a unique way of satisfying some UG principle appear to be set at the onset of grammatical development. Parameters involving distinct ways of setting UG principles or language-specific constraints are set after a delay; up to the point at which they are set, both options are manifested.

Poeppel and Wexler (1993) also contend that early child German shows evidence of the major functional categories (including I and C). This evidence is based on a corpus analysis of a 25-month-old child¹. From this acquisition data Poeppel and Wexler hoped to establish and finite verbs occurred in V2 and non-finite verbs occurred in final position. If *-en* occurred in a verb stem it was counted as a non-finite verb. 231 verbs were finite, of these 216 were in verb second position and only 15 were in verb final position. Of the 51 non-finite verbs, 44 were in verb final position, and only 7 were in V2 position. German is a verb final language and head movement from V-to-I-to-C to give V2 in matrix clauses relies on functional projections. Thus as V2 with finite verbs is so strongly established at this stage, Poeppel and Wexler argue that functional categories must be present. To further back this up, they point out that 50% of the sentences had non-subjects in initial position, implying that the V2 phenomenon is indeed in place.

Perhaps one of the strongest positions in favour of the Continuity Hypothesis can be seen in Hoekstra and Hyams (1998). Their paper contends not only that children have full command of the finite/non-finite distinction, but that UG constrains the types of DPs that can occur with either of these verb types. Their original claim was that bare nominals are non-finite DPs as they are underspecified with regard to number. Distributionally a finite verb will take a finite DP subject and a non-finite verb will take a non-finite DP subject. They provide evidence from German and English to support this claim.

However, evidence from early child Dutch has caused them to re-evaluate their claim as Det-less DPs were found not only with non-finite verbs as expected, but also with finite verbs, contrary to the hypothesis. They capture this by arguing that the marked form of the verb must be licensed by its specifier and the unmarked form is licit autonomously. As the infinitive is an unmarked form in English, the asymmetry in the distribution of Dutch and English pronouns can also be captured; in Dutch pronouns are almost always found with finite verbs, and in English with both finite and non-finite verbs (licensed autonomously).

3.2.3 Middle Ground: Grammars constrained by UG, but with bottom-up thematic information as the input to linking rules

Rizzi (1994) takes the middle ground. He argues that there seems to be a parsimony, in the functional categories available to children at some stages of their grammatical development. He bases this idea on his findings that early null subjects (see section on Hyams (1986) below) do not have the same properties of null subjects in adult grammars. Thus early null subjects

¹Wexler (1994) seems to tone down this claim to some degree, to a claim that UG is respected, but that functional categories can develop maturationally

occur extremely rarely after a preposed *wh*-element, although this is a valid option for the adult grammar (in Italian, for example). The early null subject is also limited to main clauses. This is not the case for adult grammars where null subjects are allowed. Rizzi argues that early null subjects (and null subjects in diary accounts) are null constants that need to be bound by a null operator. This is not allowed in adult English, as Rizzi argues that it is only allowed in languages that allow *ROOT=CP*, which English parametrically does not. Rizzi argues that it is allowed in early child English as the functional categories are not fully developed at this stage and the *ROOT=CP* restriction cannot hold. Thus the child grammar is consistent with UG, but the full set of functional categories have not yet matured.

Part of Wexler (1994) can also be understood in the same vein. Wexler examines the phenomenon of optional inflections in early child language. Thus, there is a stage in child language where inflections are either omitted or incorrect. Wexler identifies two possibilities for how this can be explained. There is the *Growing Strength (GS)* view where the child has to learn to add inflections, and there is the *Strong UG (SUG)* view where the child is assumed to know that inflection exists, but is unsure of the forms. The problem with the GS view is that it is unclear why it takes the child so long to learn the forms. The problem with the SUG view is that if the child knows the rules of inflection, why is the inflection sometimes optional for the child? Wexler also assumes UG-constrained maturation. That is, although the capacities might mature, the representations are always consistent with UG.

In this regard, Wexler points out that to ask whether a child knows inflection is tantamount to asking whether they know about head movement (or more specifically V-to-V movement). Wexler gives evidence (much of which has been discussed above) from French, German, Dutch, Norwegian, Swedish and Danish to show that an optional infinitive stage exists. In this optional infinitive stage according to Wexler the child:

- knows about the possibility of head movement
- knows that head movement is forced in the finite case
- knows the Principle of Economy (infinitives do not move)
- the existence of Neg outside VP and above I
- does not know that non-finite verbs cannot appear as main verbs

More controversially given the discussion above he also argues that the child:

- knows the existence of I and C

However, the inflection is still optional and given Wexler's position, this needs to be licensed by UG. Normally, the Economy of Derivation would rule out optionality. Thus for optionality to be present, both derivations must be equally complex. Wexler thus sees two equally complex derivations; raising of V-to-I, resulting in the finite form, and the lowering of I-to-V resulting in the non-finite form.

Clahsen *et al.* (1993) cover similar ground to Poeppel and Wexler (1993) (discussed above), but come to slightly different conclusions with regard to the Continuity Hypothesis. They see the acquisition of COMP in a maturational way. In other terms, the functional categories are

posited by the child on the basis of input data. They are not existent from the beginning of grammatical development, but rather mature based on the linguistic input the child receives. Clahsen *et al.* (1993) argue that the child at the first stage has IP, but no CP. V2 requires two functional categories, there are three possibilities as to how child grammar reflects this:

- A full adult tree with two functional categories
- A tree with only one functional category
- A PS tree with only lexical projections

Clahsen *et al.* (1993) ordered their data by MLU. At the two word stage this should not exceed 1.75. Thus they identified two stages. Stage 1, where MLU was less than or equal to 1.75, and Stage 2, where MLU was less than or equal to 2.75.

At stage 1 there was evidence of IP, but no evidence of CP. Verbs occurred before subjects which would not be expected if only PS rules were available to the child. To ascertain CP they searched for lexical complementizers, wh-pronouns and postverbal negation patterns with an overt lexical subject. At Stage 1 these were not evident, but were present at Stage 2.

Atkinson (1996) also takes this middle-ground approach. He complains of the orthodoxy that is being established in which it is proposed that a child embarks on grammatical development with a complete syntactic system in place. This is known as the Full Competence Hypothesis. Atkinson believes that this may be wrong. He proposes rather to follow the Structure building approach of Radford (1990), but this only as far as stating that there is a stage at which children have no functional projections. This therefore means that a different approach to language acquisition is necessary. Maturation seems to be the most likely mechanism.

Methodologically some claim that embracing maturation should be a last resort (Crain (1991), Hyams (1994)). But this approach is qualified as being with all other things being equal. Given that a maturational account is to be followed, it needs to be explained by what mechanisms functional categories can be acquired. For learning one would have to maintain that the input data provides an adequate inductive basis for the learning of a functional category. This could be via hypothesis making and testing. However, from this the triggering problem Hyams (1994) arises; if the child can represent C, why is its consolidation in the grammar protracted.

Atkinson sees this initially as not the concern of psychologists or linguists, but rather that explanations of this process may be sought in biology. However, if the development is maturational and if X occurs before Y (representational capacities), then we would predict this occurs cross-linguistically. Atkinson points out that most biologists would probably see maturation as the norm in the development of complex biological systems.

Lebeaux (1988) sees the acquisition of language as being made up of stages, and at the first stage, the system as he sees it is based purely in theta-assignment. The later stages are geological in their acquisition and the various geological levels correspond to the levels of representation in the synchronic analysis of adult grammar. The levels also correspond in a one-to-one manner to the stages of acquisition.

Lebeaux sees linking rules as being maximally general in associating thematic roles with grammatical functions or abstract case. This is similar to Pinker's idea of semantic bootstrapping. Lebeaux says this early reliance on thematic information is a case of analytic priority.

(3.1) *Analytic Priority*

A set of primitives $a_1, a_2 \dots a_n$ is *analytically dependent* on another set $b_1, b_2 \dots b_n$ iff b_i must be applied to the input in order for a_i to apply.

Lebeaux thus sees the set of theta-theoretic primitives as being analytically prior to the set of case-theoretic primitives. The question then is how the child moves from the one-word thematic looking grammar to full phrasal syntax. Lebeaux argues for a continual application of Project- α to achieve this. Looking at the representation from the head outwards, Lebeaux assumes that the phrase structure enveloping it is projected from lexical argument structure, relying on the theta-subtree. The lexical argument structure proposed is similar in nature to those proposed by Hale and Keyser or Levin and Rappaport in the discussion in Chapter 2. This leads to four questions for acquisitional accounts.

- As the phrase structure is already articulated, does not adding a lexical representation introduce redundancy?
- Assuming Project- α occurs, is all the information projected or just part of the information?
- Is there evidence for pre-Project- α representations in acquisition or elsewhere?
- How does a difference in language typology get represented in Project- α ?

As mentioned above, Lebeaux sees the pre-Project- α representations in acquisition as fully syntactic. This is true at the one-word stage, but becomes intuitively more plausible at the 2-3 word stage. Thus, such two word combinations as *want baby*, *want car*, *want high*, *boot off*, *that Dennis*, *there ball* and *there dog* are seen by Lebeaux as small clauses missing the copula. At this telegraphic stage of acquisition, Lebeaux sees the child as talking in thematic representations. It then falls to the child to insert closed class elements into these representations.

Interestingly, for the discussion of psycholinguistics in the next two chapters, Lebeaux proposes an interesting relationship between the grammar and the parser. When the child's parser fails, it will fall back into another permissible grammar. This lower grammar will also be a lower grammar in the geological acquisition of language. In other words, the parser will fall back to a grammar from an earlier stage of acquisition. Lebeaux terms this the *Property of Smooth Degradation*. Lebeaux states as evidence for this the fact that a telegraphic grammar can also be used by adults in diary reports, for example.

Thus the middle way between the Full Competence Hypothesis and the No-competence hypothesis sees a thematic basis for language acquisition with functional categories developing maturationally.

3.3 Parameter Setting

One of the fundamental tenets of contemporary PPT is that UG can be configured in different ways by different languages. The aim is to find the smallest possible number of parameters that explain the variation in the world's languages. That is, each parameter that is set is assumed to explain more than one variation in structure.

A less developed part of the theory of linguistic parameters is the learning of these parameters. Given the poverty of the stimulus, what can the child use to set, or *trigger* each parameter.

I shall describe the sort of parameters that have been proposed in the literature and also theories of the acquisition of parameters. I argue that syntactic hegemony has once more oversimplified the true state of affairs and has resulted in unrealistic assumptions as to the explanatory potential of syntactic parameters.

The basis of this failure is the belief that syntax is in some way optimally structured and that parameter setting is a purely syntactic phenomenon. I agree that there are undoubtedly syntactic parameters but they are only a small part of linguistic variation. Rather, each language is merely a design solution (and not necessarily an optimal one) to the problem of encoding thematic-conceptual relations into speech. Thus it is not syntax that is parameterized, but rather the linking rules between thematic-conceptual structure and the full range of expressive linguistic devices, including syntax.

Thus parameter setting is in many ways analogous to grammaticalization and has its origins there. As languages will not be optimally designed there is always the possibility of reanalysis and change, and language acquisition will be centrally implicated in this change. Indeed, it often seems to be the case that with respect to syntactic parameters, children are generating parameter values based on often inconsistent triggers rather than picking up parameters from the input (Lightfoot (1991)).

It has also to be explained why parameters are learned in any particular order. It could be the case that the order of acquisition is determined by complexity, as constructivists would presumably argue. On the other hand, it may be the case that the order of acquisition is maturationally determined.

Suggestions for the parameters governing word-order (Travis (1984, 1989); Koopman (1984)) are head-ordering, directionality of case assignment and the directionality of theta-role assignment. Hawkins (1995) claims that this is purely descriptive and gives no account as to why the phenomena have the particular properties they do. This seems to be a largely valid conclusion.

One example set of parameters would be:

	Head	Case-Directionality	Theta-Directionality	Frequency
Japanese	final	Left	Left	Frequent
English	initial	Right	Right	Frequent
Chinese	final	Left	Right	Rare
Kpelle	initial	Right	Left	Rare

And in an even more complex way German changes theta-directionality and case-directionality according to linguistic category.

The explanation does seem to be purely descriptive as it fails to predicts much of the variation shown in Jackendoff (1977) on modifier orders. The relative frequency of the patterns in languages also has to be explained.

One of the first attempts to apply the concepts of PPT to language acquisition was Hyams (1986). The parameter in question has come to be known as the null-subject parameter. This parameter has been posited as languages appear to group into two camps with the following contrastive properties:

	Null Subject	Non-Null Subject
Overt Subject Expressed	No	Yes
Subject Can be Inverted	Yes	No
Subject Can Be Extracted from tensed Subordinate Clause	Yes	No

The extended projection principle states that all clauses must contain a subject. Thus even in null-subject languages, such as Italian, Spanish and Hebrew, a subject must be present in some form or other. As the subject is not overt, it must be an empty category. It cannot be PRO as PRO must appear in an ungoverned position, and the null-subject position is governed by finite I.

Hyams adopts a proposal made by Rizzi (1982) that the empty category is pro. The central idea is that I can be specified as [+pronominal]. When this is the case, this feature absorbs the nominative case normally associated with the subject position. It will also qualify as a proper governor for the Empty Category Principle as it has lexical properties.

The specific parameter Hyams first proposes is:

(3.2) *The Null Subject Parameter*

AG=PRO

AG≠PRO

This requires some explanation. Here AG is analogous to Rizzi's [+pronominal] I. The precise rule Hyams uses is

(3.3) $I \rightarrow (AG), AUX$

Hyams stipulates that if AUX contains lexical material it is the head of I, whereas if AUX is empty, AG heads I. AG encodes person, number and gender features.

The first of the values of the parameter states that languages with this value allow AG to be identified with PRO, licensing pro. This has the result that modal verbs cannot appear in AUX as long as AG=PRO as this would result in PRO being governed.

With regard to acquisition, Hyams takes the continuous approach (Pinker 1984). That is, a child may have an incorrectly set parameter for the language she is learning, but the incorrect language will be compatible with UG. She claims that English children start off with the parameter set to the null subject value.

The question then remains as to how this incorrectly set parameter is then reset to the correct value in English. Italian seems to operate under the Avoid Pronoun Principle (Chomsky (1981)):

(3.4) *Avoid Pronoun Principle*

Avoid lexical pronominal if a null pronominal is possible.

Children faced with English input will be confronted with expletive subjects and unstressed subject pronouns, which must be grammatically motivated. This provides positive evidence to the child that the parameter must be reset.

The problem of giving both sufficient licensing conditions and identification conditions has led Hyams (1987) to rethink her initial proposal. The licensing conditions have been fulfilled in both early English and Italian by the AG=PRO parameter. But whereas the Italian pro can be identified by the system of verbal agreement in that language, children at the proposed AG=PRO stage of English mostly have not developed verbal inflections, meaning that they cannot identify the interpretation of pro.

This is part of a more general problem of associating null-subjecthood with richness of inflection. The starkest evidence for the problematic nature of this correlation is Chinese which is a null-subject language with no agreement features. Conversely, German has a rich inflectional system but does not allow null subjects.

Hyams proposes a new parameter value based on the fact that Chinese is inflected in none of the positions in its verbal paradigm, and Italian is inflected in all positions. Thus inflection in these languages is *uniform*. In contrast English is inflected in only one position.

(3.5) *Uniformity Parameter*

Null subjects are permitted in all and only languages with morphologically uniform inflectional paradigms.

Identification is satisfied in Italian by sentence-oriented agreement features and in Chinese by discourse-oriented topic-identification. German is still a problem; licensing is fulfilled so it must be identification that forms the stumbling block preventing null-subjects being possible. It may be that Tense is in C and agreement in I' following the movement that results in SVO order in main clauses in German, and this separation would ensure non-identification. However, in subordinate clauses this movement has not taken place and null-subjects should still be possible, which they are not.

A further problem is that without agreement features to aid identification, children learning English at a null-subject stage must be using topic-identification. But Chinese allows null-objects and we would thus expect this to be the case for English, which Hyams claims it is not.

Further problems can be found in the fact that neither of these proposed explanations captures the third correlate of null-subjecthood: the possibility of inverted subjects. The choice of AUX or AG as head also seems to be stipulative and *ad hoc*. It also needs to be explained why uniformity/non-uniformity should have the effects it does.

In addition, it seems strange that English children do not reset the parameter earlier, given that they are continuously confronted with expletive subjects and unstressed subject pronouns. Thus we would seem to require an explanation based on the representation that the child has internalized, and that such an early acquired representation predicts and constrains later representations. Or alternatively a maturational account may be called for.

Crucially, perhaps, *pro* is in a functional position (DetP) and its licensing relies on a further functional position I. Above we have seen that Radford (1990) claims that Early Child English lacks both these positions. If this claim is true, Hyams theory could not possibly be correct, at least not for the early stages of language acquisition.

Another model of parameter setting is known as *Subset Theory*. This was developed by considering binding phenomena. The theory of binding relies on the concept of Governing Category. There seems to be crosslinguistic variation as to the definition of Governing Category. Yang (1983) has surveyed these differences and has come up with the following values:

(3.6) The GC Parameter has five values:

α is a governing category for β if α is the minimal category which contains β , a governor for β and:

- (a) a subject English
- (b) an I Italian *se*
- (c) a TENSE Icelandic *hann*
- (d) an indicative tense Icelandic *sig*
- (e) a root tense Korean *caki*.

The technical details of this are not important to our aims here. What is important is the relationship of the various possible values for the Governing Category parameter to each other. Manzini and Wexler (1987); Wexler and Manzini (1987) (henceforth W&M) have postulated the *Subset Theory* as a way for learners to be able to set the correct parameter for these languages.

There are two strands in this framework. One strand is the Subset Principle itself and the second strand is the Lexical Parameterization Hypothesis which states that parameter values are associated with particular lexical items. I shall describe each of these in turn.

If the assumption that no negative data is available to learners is correct, an English child may assume value (e) for anaphor binding, and would receive no negative data that this is too free. W&M thus conclude that if the languages are subsets of each other in the following pattern

(3.7) $L(a) \subset L(b) \subset L(c) \subset L(d) \subset L(e)$

and the child starts off by assuming the smallest of the grammars, then the other larger grammars will be learnable as the child will encounter positive evidence that the larger grammar is the one used by the language they are learning.

Turning to pronominals the subset relations are exactly the opposite:

(3.8) $L(e) \subset L(d) \subset L(c) \subset L(b) \subset L(a)$

Thus there cannot be an unmarked value for the Governing Category Parameter in general and a parameter can, in principle, be set separately for anaphors and pronominals.

This separation of anaphora and binding into different subset relations is unusual as they are normally in complementary distribution. However, the relationship between them emerges only accidentally in W&M's system.

Based on the fact that two words in Icelandic (*hann* and *sig*) are associated with different values of the GCP, W&M (1987) have proposed the *Lexical Parameterization Hypothesis* which states that parameter values are associated with particular lexical items.

The problem with this hypothesis is that it allows no leeway for the child to generalize, it is more likely to be the case that there are generalizable parameters which are non-subset theoretic and lexically parameterized parameters (not many) which are subset theoretic.

Gibson and Wexler (1994) accept the problems of a rigid theory based purely on subset relations between parameters. They identify instead three ways that parameters can be set.

- General Motor of Learning
- Subset Theory
- Triggers

Under the General Motor of Learning it is assumed that parameters can be set on the basis of primary data by fitting parameter values to the data. Subset parameters are set in the manner given above. The idea of *Triggers* is that for non-subset parameters, there will be a sentence that will be grammatical only under the correct parameter setting and thus this sentence will be the trigger for the setting of that parameter.

Gibson and Wexler provide us with an Algorithm for setting parameters from these triggers and a pair of constraints that will prevent the learner diverging away from the correct grammar.

(3.9) *The Triggering Learning Algorithm*

Given an initial set of values for n binary-valued parameters, the learner attempts to syntactically analyze an incoming sentence S . If S can be successfully analyzed, then the learner's hypothesis regarding the target grammar is left unchanged. If, however, the learner cannot analyze S , then the learner uniformly selects a parameter P , changes the value associated with P , and tries to reprocess S using the new parameter value. If the analysis is now possible, then the parameter value change is adopted. otherwise the original parameter is retained.

(3.10) *The Single Value Constraint*

Assume that the sequence $h_0, h_1, \dots, h_n$ is the successive series of hypotheses proposed by the learner, where h_0 is the initial hypothesis and h_n is the target grammar. then h_i differs from h_{i-1} by the value of at most one parameter for $i > 0$.

(3.11) *The Greediness Constraint*

Upon encountering an input sentence that cannot be analyzed with the current parameter settings, the language learner will adopt a new set of parameter settings only if they allow the unanalyzable input to be syntactically analyzed.

In mathematical terms Gibson and Wexler show that learners may arrive at a situation where one of a set of *local maxima* holds. This is where the learner can find no local triggers to set parameters.

This can be seen as one of the major failings of the theory of syntactic parameters - as parameters are stated in terms of (mathematical) configurations with no notion of language

content or the purposeful use of language, there is a concentration on mathematical arguments regarding these configurations.

Gibson and Wexler suggest two ways that learners might escape such maxima.

- default states
- parameter ordering

If the number of parameters is large it seems unlikely that the parameters could be successfully ordered as they will interact in complex ways. We will see in Chapter 4 an analogous argument for the problem of basing processing theories on PPT theory as the sub-theories cannot be put into an optimal ordering.

On the other hand default states mean that the learner must come to the task with a great deal of innate knowledge of a very specific nature. If so much knowledge is innate, one may expect languages that have more parameters set to the default value to be learned more quickly. There is no evidence that this is the case.

In this section we have seen that each of the theories of how parameters could actually be *set* falls down on various counts. Hyams' theory ultimately collapses as the elements it relies on arguably do not exist in child language at that stage. Subset theory removes useful observations of complementary distribution. The Lexical Parameterization Hypothesis does not allow for generalizations and Gibson and Wexler's triggers rely on too much innate knowledge.

3.4 The Relationship of Ontogeny and Phylogeny

There has been a recent shift in genetic theory (Ho (1995); Jones (1996)) that emphasizes the holistic nature of genetic inheritance. In this view, environment, culture and society interact with the genes at an epigenetic level and the results of the interaction can be passed on to future generations.

In this way, the stability or change of both genes and environment, and specifically for humans, society, can evolve in concert. And following on from this combined evolution, there will be similarities in the historical emergence of language and its ontogeny. The societal forces that led to the emergence of language, perhaps *social grooming* as Dunbar (1993) suggests, are still relevant today and still influence the form and nature of learning and constrain the way language is learned.

And at another level, we may expect parallels to be found in the history of how conceptual relations become grammaticalized into language structure, and the steps children pass through in learning the structure of a language. The relationship between grammaticalization and language acquisition will be discussed in the final section of this chapter.

The case can also be made that phylogenetically and hence ontogenetically early developments will influence the structure of the later-learned language capacity. These starting points for language learning will be discussed in the next section. The hypothesis space is constrained by the phylogenetic history of the language component.

Learnability and learning theory should not be seen in terms of mathematical criteria as in Gold (1967)'s theory, but rather in terms of biological criteria. This emphasis on biological

criteria is yet another link to the phylogenetic approach and should thus be contrasted with the other approach that claims to be based in biology: constructivism.

Constructivism makes the claim that input is enough to determine the structure of the language processor and the order in which the structure of language is acquired. However, most studies of language acquisition have emphasized the poverty of the stimulus.

Others claim that the hypothesis space for parameter setting is set at the beginning of acquisition (Pinker (1984)). This is known as the *Continuity Hypothesis*. However, there are two other possibilities. The child could construct grammars that are in some way at variance with UG and only later converge on a UG compatible grammar. Alternatively, certain grammatical principles may be maturationally determined and only come into effect at later stages of acquisition.

There is also the question as to which extent the ontogeny of a cognitive system corresponds to its phylogeny. Thus, in a strong theory, evolutionarily simpler and prior structures would correlate with what is learned earliest in acquisition. For this claim to be tested we have to determine what form the course of evolution of language took and calculate what the formal linguistic correlates of the stages of evolution are in representational terms. I will discuss this more fully in Chapter 6.

3.5 The Starting Point for Language Acquisition

I believe that early emerging cognitive conceptual abilities will be of the highest salience and will influence the types of conceptual metaphor that thematic-conceptual structure relies on. In this section I shall assess what some of these early emerging conceptualizations are.

Given the arguments of the previous section we will expect these early emerging cognitive abilities to have a long phylogenetic history and to have a particular saliency both for language acquisition and for language processing. This will be explored in the conclusion to this thesis.

There is evidence of early salience of shape, colour, height and width, and in Prototype theory (Keil (1981)) it seems that the ontological trees of children have the same overall form as those of adults but merely have fewer branches and therefore fewer distinctions. Perceptual knowledge, object knowledge, ontological knowledge, primitive understanding of mechanics and number and a rudimentary grasp of causal processes arise early and are all well in place when children make their first two word utterances.

Spelke (1991) suggests that the processes by which humans perceive objects are inseparable from the processes by which humans reason about objects, just as Piaget believed. Physical reasoning and object perception do not, however, undergo revolutionary changes over human development. They develop through a process of enrichment around core principles that are constant.

Although Piaget's writings contain no inventory of the constraints that humans come to appreciate, five constraints figure prominently in his experiments: continuity, solidity, no action at a distance, gravity and inertia.

Studies of infants require tasks within infants' behavioural repertoires. The challenge is to devise tasks that meet this requirement without sacrificing what is essential to Piaget's experiments: The tasks must not be solvable by engaging in habitual actions or by responding

to superficial properties of events.

An example can be found in Leslie (1984): In the causal event Leslie studied, one object (A) caused a second object (B) to begin moving, and B caused A to stop moving, whereas in the reversal of that event, A caused B to stop moving, and B caused A to begin moving. Leslie's experiments provided evidence that infants apprehended causal relations, in accord with the principle of no action at a distance.

This innateness of the understanding of causal events makes possible the clear preference and salience of the endpoint of events in language processing. That is, there is a clear preference to concentrate on the results of an action, or the object acted upon.

The experiments of Baillargeon (1987) provide evidence that infants represent the continued existence of an object that is hidden from view.

In the critical events of the studies, infants are shown an object that moves out of view behind a screen. Then the screen is raised, revealing the object at rest in either of two positions. One resting position is consistent with all physical constraints on object motion; the other resting position is inconsistent with one or more constraints.

Continuity and Solidity

The first experiment carried out by Spelke (1991) provides evidence that 4-month-old infants infer that a hidden object will move on a connected, unobstructed path.

The second set of findings shows that infants infer that no part of a ball will pass through the surface in its path. These findings corroborate Baillargeon's findings that infants represent the size of a hidden object and honour a rigidity constraint, inferring that a hidden object will maintain a constant shape and size.

The studies reveal pre-existing conceptions of object motion and do not 'teach' such conceptions over the course of familiarization. Concerning Piaget's theory, these findings, along with the findings of Leslie and Baillargeon, provide evidence that the capacity to represent and reason about the world develops long before the attainment of major sensorimotor coordinations.

Gravity and Inertia

Spelke also investigated whether infants appreciate that objects move downwards in the absence of support and that objects continue in motion in the absence of obstacles in their path.

Infants looked longer at the outcome of the consistent event test. The experiment therefore provided no evidence that 4-month-old infants were sensitive either to the effect of gravity or to the effect of inertia on the motion of a falling object.

The experiment was repeated with a sample of 6-month-old children with the more familiar pattern of findings. Between the ages of 4 and 6 months, infants evidently began to infer that the hidden, falling object would continue falling to a surface.

These experiments do not reveal what aspects of object motion 6-month-old infants have begun to appreciate. Between 4 and 6 months, infants may develop a general conception that objects require support or a general conception that objects do not stop moving abruptly in

the absence of obstacles. As a third possibility, infants may develop more specific expectations about the behaviour of falling bodies.

The findings of a further experiment provided no evidence that 6-month-old infants have developed a general appreciation that objects require support. Although such infants evidently infer that a falling object will continue falling until it reaches a supporting surface, they do not appear to infer that a stationary object will begin to fall when it loses its support.

Inspection patterns suggested that 6-month-old infants do not appreciate that object motion is subject to inertia.

4-month-old infants looked reliably longer at the superficially novel consistent event, contrary to the inertia constraint. The experiment was repeated with 6 and 8 month-old children. Like the 4-month-old children, the older infants looked reliably longer at the consistent event. The experiment showed no evidence that infants infer that a linearly moving object will continue in linear motion.

A final experiment investigated infants' expectations of linear motion in a situation in which the consistent and inconsistent event outcomes were equally novel. In this situation sensitivity to inertia began to manifest itself between 6 and 8 months of age.

It now seems that knowledge of continuity and solidity is more deeply rooted in human development than is knowledge of gravity and inertia. Also present is a knowledge of causal structure. Knowledge of continuity and solidity may derive from universal, early developing capacities to represent and reason about the physical world. In contrast, knowledge of gravity and inertia may derive from the child's growing acquaintance with particular kinds of events involving physical objects.

3.5.1 Early Capabilities Reflected in Lexical Acquisition

Siskind (1992) has attempted to solve the early bootstrapping problem by proposing a cross-situational learning algorithm which is based on visual perception and the early capacities of children. Siskind takes an approach that is on the right lines, but, because it relies on logical algorithms which cannot be realistically assumed to be in operation in young children, the model is somewhat unrealistic.

He claims that the appropriate semantic representation for the meanings of simple spatial verbs such as *throw*, *pick up*, *put* and *walk* must incorporate the notions of **support**, **contact** and **attachment**.

The question then is how support, contact and attachment relations necessary for these representations can be picked up from the context of utterances. In Siskind's model the putative link between cognition and language can be seen most clearly. He proposes that children learn the meanings of these verbs by using *counter-factual simulation*.

Counter-factual simulation involves imagining the short term future of a potentially modified image under the effects of gravity and other physical forces. Counter-factual simulation in turn relies on an *imagination capacity*. Imagination capacity is normally assumed to be some kind of simulator of kinematic motion. Traditional approaches to kinematic simulation take physical accuracy and the abilities to simulate mechanisms of arbitrary complexity to be primary. This type of kinematic simulator is thus abstracted away from what we need which

is a kinematic simulator that simulates the way human beings perceive the world. Human perception seems to be based on a number of principles which include the very basic principles that Baillargeon, Leslie and Spelke have found to be present in very young children, namely:

- Substantiality: Solid objects do not pass through each other.
- Continuity: Objects follow continuous paths when moving from one location to another. They do not disappear and reappear elsewhere later.
- Gravity: Unsupported objects fall.
- Ground plane: The ground acts as universal support for all objects.

Siskind provides a mechanism by which children can use these features of cognition to construct the semantic representations necessary within the linguistic system from the linguistic input they receive and its context. Given the general argument in this thesis, it would seem hardly surprising that human language mirrors the way we perceive the world and that the representations involved also include elements of human cognition. And it is also hardly surprising that the language learning system employs these aspects of cognition in the learning process if indeed language has evolved from other aspects of human cognition.

It may well be the case that the source of the salience of contact, attachment and so on for the language systems may lie in the gestural stage in the evolution of language (to be discussed in Chapter 6).

We see that the early capacities of children can predict and help explain at least to some extent the learning of a set of verbs. We would also expect these ecological salencies to be operant in language processing. In Chapter 5 I shall show that a similar type of conceptualization (i.e. continuity, solidity, contact and attachment) can explain attachment decisions in NP/PP adjunct attachment ambiguities such as the following.

(3.12) Verb Attachment Preferred:

- a. The spy saw the cop with binoculars.
was preferred to
- b. The spy saw the cop with the revolver.

(3.13) Noun Attachment Preferred:

- a. The woman married the man with money.
was preferred to
- b. The woman married the man with delight.

3.5.2 General Principles of Language Acquisition

It seems to be the case that language goes through maturational stages, but that these stages are not universal across children and the theory has to be flexible enough to capture these differences. We have also seen that children are probably learning language structure in its

basic form by linking a conceptually based language structure to their more general conceptualization of what they cognize. This conceptualization will necessarily be constrained by what they can cognize at this early stage.

In this section of the chapter I shall examine suggestions as to more general principles of language acquisition that have been proposed with specific reference to lexical learning by Clark (1993), but that are, in my opinion, relevant to all aspects of language learning. For each proposal I shall give the original definition as relevant to the lexicon and then discuss what relevance they have for the learning of syntactic structure via thematic-conceptual structure.

1. **MAPPING:** Across languages strong correlations exist between word classes and ontological category (see Croft (1991)). By the age of one children have a good deal of conceptual ontological knowledge. If children assume early on that there must be different word classes on the basis of meaning-form mappings, they can bootstrap themselves into syntactic categories. This is an example of what is known as semantic bootstrapping to syntax. It may be the case that the process could work in the opposite direction. That is, syntactic inflections could be used to predict what class a word is and help to predict its meaning. A lexicon that has been bootstrapped in this conceptual manner would then form the input for the learning of syntax.
2. **CONVENTIONALITY:** What makes an item conventional is that everyone in the speech community agrees with everyone else on which words or expressions denote which kinds. The store of conventional forms can be used by the child to predict production of new words. Such predictive ability would also be the case for the learning of structural aspects of language. That is, whatever is conventional in structural terms could be applied to new cases, and any new forms that the child hypothesizes will be tested against the input to check whether they are conventional. The checking tells the child whether its hypotheses fit into the conventional forms of the language.
3. **CONTRAST:** The major consequence of this principle is that speakers do not tolerate complete synonyms. This makes the following predictions about the major properties of the lexicon:
 - words contrast in meaning - what characterizes lexical domains is that all their members differ in meaning.
 - Established words have priority - if the child finds a word in its conventional lexicon for the expression needed, it will be used.
 - innovative words fill lexical gaps - speakers must make sure the listener can compute the intended meaning.

Contrast offers economy of effort in acquisition. Thus children can immediately set about working out how a word is different from the other words they know. If this idea is carried over to syntax we would expect every sentence structure to be associated with a different meaning. A claim for the principle of contrast in syntax is a strong claim as it has often been assumed that, for example, the passive version of a sentence does not differ in meaning from its active counterpart.

4. **TRANSPARENCY OF MEANING:** Words are transparent when children know the meanings of its morphological elements. Sentences are transparent when children know the meanings of its constituent elements.
5. **SIMPLICITY OF FORM:** words are simple when the elements to be combined require no changes or only minimal changes in form. The principle of simplicity of form also predicts that more simply constructed sentences will be learned before more complex ones.

If we can assume that different word classes are learnt on the basis of mapping, then semantic bootstrapping is possible. Children also show a great sensitivity to what is important in their own language.

Also vital to the future development of my argument is the principle of contrast. This principle means that not only all lexical items, but also all differences in structure are associated with a different meaning. This idea will be central to my reanalysis of the syntactic priming (Bock (1986)) literature in chapter 5.

3.6 Thematic-Conceptual Structure and Language Acquisition

In the previous chapter we saw that one of the motivations for thematic-conceptual structures was to overcome Baker's paradox. There I gave examples of the dative alternation. The child is able to learn the alternation patterns of various verbs from their underlying thematic-conceptual structure. Here I will describe how the dative alternation and its learnability can be captured by thematic-conceptual structures.

The lexical rule for the dative alternation with *to* converts a verb with the semantic structure *X causes Y to go to Z* into a verb with the semantic structure *X causes Z to have Y*. Thus verbs can alternate only if they signify a transfer of an object that can result in its being possessed - note the impossibility of (3.14)

(3.14) She carried the mailbox a letter.

However, negative exceptions do exist (3.15).

(3.15) ?Reach me that book.

To capture negative exceptions, Pinker introduces the notions of broad conflation classes and narrow conflation classes. Membership of a broad conflation class is only a necessary condition for alternation, it is membership of a narrow conflation class which is a necessary condition. The negative exceptions are created by applying the broad range rule without paying heed to whether the verb in question fits into one of the narrow conflation classes. Narrow conflation classes have the following properties: They are verbs that enter into the construction fall into semantically cohesive subclasses involving a narrower range of meanings than that which is directly associated with the argument structure. A common set of elements of meaning, such as contact, motion and effect enter into the definitions of these semantically

cohesive classes. Whether a verb belongs to a class depends not on the characteristic features of the event in the world but on the aspects of the event that its semantic structure constrains.

Pinker proposes the following subclasses of verbs in the dative alternation (N.B. all of these verbs fit into the broad conflation class for the dative alternation):

Subclasses which do alternate:

Verbs of giving.

Verbs of instantaneous imparting of force: *flip kick, toss*.

Verbs of continuous imparting of force with a specified direction: *bring, take*.

Verbs of future having: *offer, promise, bequeath, forward*.

Illocutionary verbs of communication: *tell, show, ask, teach, write, read*.

Instrument of communication verbs: *fax, radio, e-mail*.

Subclasses which do not alternate

Verbs of Continuous imparting of force: *pull, push*.

Verbs of future not having: *cost, spare, envy, bet, refuse*.

Verbs of presentation: *present, credit, reward*.

Manner of speaking verbs: *shout, murmur, grunt*.

The motivation for these subclasses is that if the thematic core of the double object dative involves an actor acting on a recipient in a way that causes the recipient to possess something, then those verb subclasses that suggest that the action inherently involves the beneficiary as patient directly are more likely to dativize.

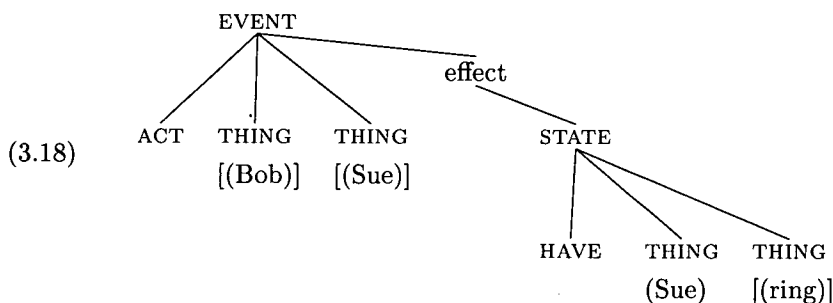
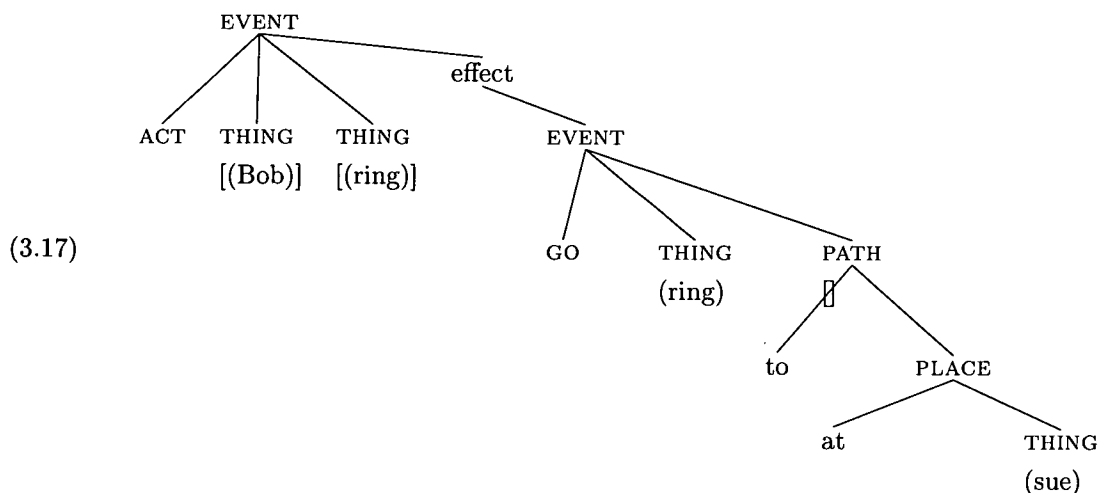
Broad-range rules could not be replaced by narrow-range rules because evidence exists for the on-line application of broad-range rules to produce (in a one-off manner) forms which are not licensed by the narrow-range rules. An example from the causative alternation:

- (3.16) a. The experience grew me up quickly.
b. Sparkle your table with....

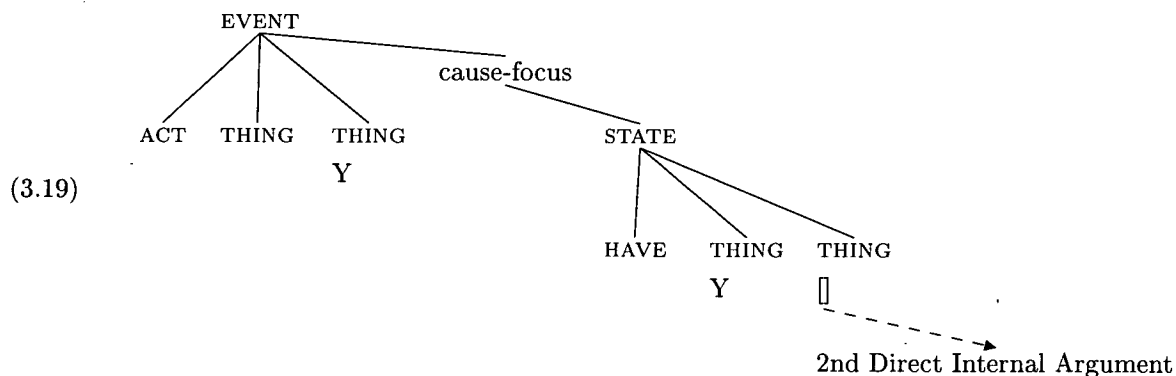
Broad-range rules are thus said to be property-predicting: they dictate what grammatical properties a form would have should it exist. Narrow-range rules on the other hand are existence-predicting: they allow a speaker who possesses one form to add a related form to the grammar automatically.

It now remains to be shown that the restrictions on lexical alternations can be represented within such a system of semantic structures. We have considered the dative alternation with *to*.

The Broad Range rule for the dative alternation can be represented as:



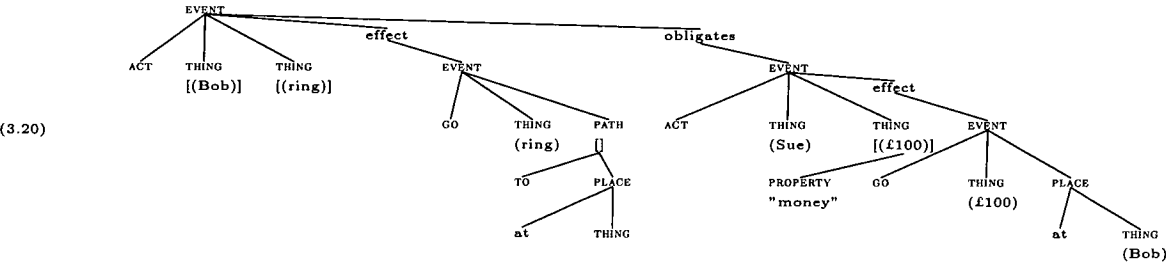
And a possible linking rule for the second direct object in the double object construction is



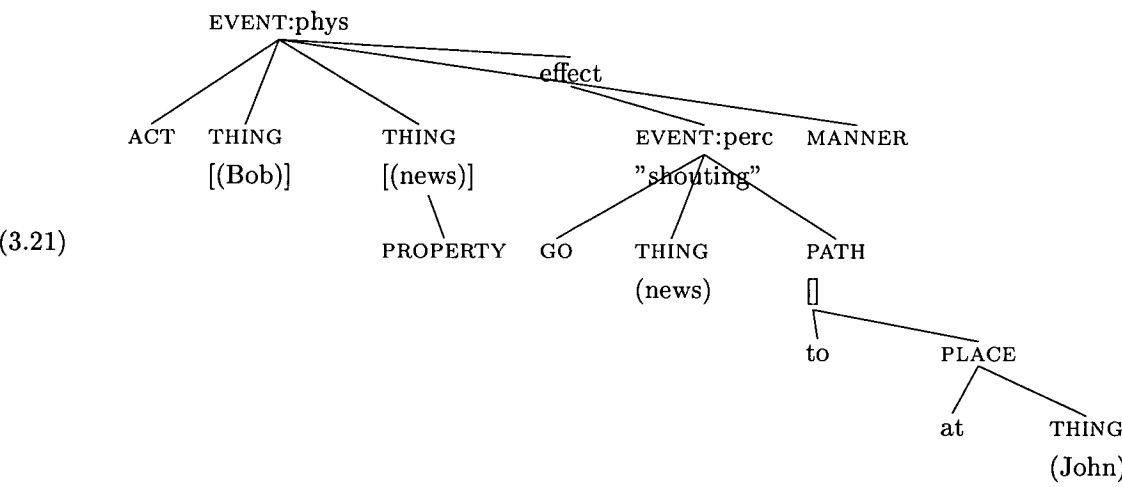
In the prepositional form, the transferred object is the patient, whereas in the double-object form the recipient is the patient. This underlies the pragmatic differences in which argument is construed as 'affected'. In the double-object construction the change of possession is expressed not as an analogue of a motion of the object, but as the direct causation of a state whereby the possessor has (HAVE) the object. This has the consequence that the double-object form is incompatible with the expression of pure physical motion. But as we have seen the Broad Range rule is only a sufficient condition for the alternation to occur, it is the narrow range rules that are a necessary condition and which predict that the alternation will actually occur.

The narrow-range rules are of the following types: First there are the verbs of *Giving*. *Give* would have a representation almost identical to the input of the broad range rule, though the semantic field would be specified as possession. *Pass* would specify possessional: physical custody. *Hand* would introduce an additional MANNER constituent specifying the use of hands.

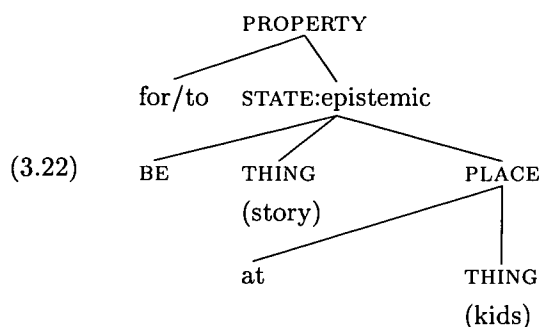
Second, *Sell*, *trade* and *pay*. These verbs have a subordinate counter-transfer event to which the agent is committed. The patient/theme of the counter-transfer (£100) is linked to the oblique argument role because it is embedded under ACT and would be expressed by *for* thanks to a lexical entry for *for* specifically for the counter-transfer.



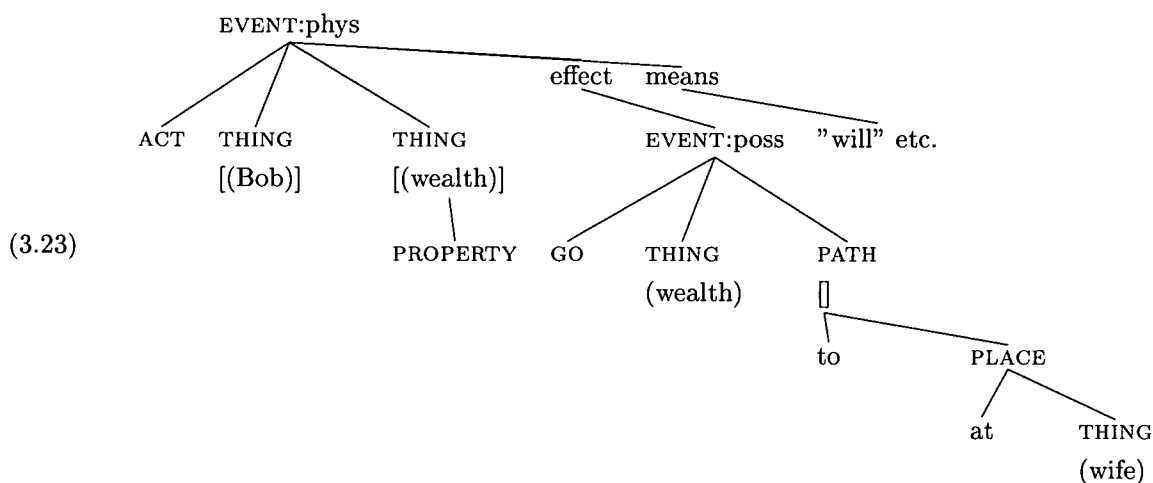
Third, there are the verbs of *Manner of Speaking*. These are non-dativizable. The aspect of the representation that ensures their non-dativizability is that the MANNER refers to a *physical* event, which by its nature cannot be possessed. For example the lexical entry for *shout* is:



Fourth, there are the verbs of *Communication*. These verbs involve a subfield of possession involving the communication of ideas. An example tree for this would be similar to the tree for the verb of manner of speaking, except that the manner substructure is unnecessary and that the property substructure will have the following substructure:



Fifth come the verbs of *future having*, where the main event is not one of giving but of commitment - the possession transfer is an effect at a different point on the time-line. The verb specific information is tentatively put into a means substructure:

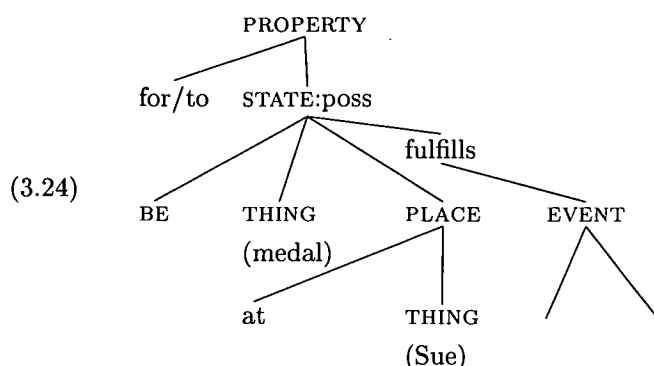


Sixth are the verbs of *Presenting*. These are also non-dativizable. They encode the transferred theme as a *with*-object in the alternative form. The *to*-form is probably derived from the *with*-form. The construction shares features of both the dative verbs (change of possession, leading to use of the preposition *to* rather than *into* and *onto*) and the locative (the change of state and the existence of the *with* form, linked to the entity whose transfer effects the change of state).

The transferred object has a property 'for the recipient to possess it' because of some deontic cause specified idiosyncratically by the individual verbs.

We still have to explain why the dative rule cannot apply to this structure. It could perhaps, yielding a structure which is linked to the double-object but the semantic representation of *with* fused with the linking rule for oblique arguments of ACT blocks or preempts the second object linking rules because it is more specific. Pinker opts for the simpler explanation that the extra complex property structure makes the verbs dissimilar enough for them not to work with the rule.

So for example the property substructure for the sentence *Bob presented the medal to Sue* would be:



Finally, there is the group of verbs of *Causation of Motion*. Here we have to explain *throw* (dativizable), *pull* (non-dativizable) and *bring* (dativizable). The crucial difference between these three is the interaction of aspectual and force -dynamic components - for *throw* the causing is instantaneous and precedes the motion of the object. For *pull* it is a continuous process that is temporally co- extensive with the motion of the object.

Bring doesn't specify MANNER but supplies deictic information concerning the PATH and implying that the Actor moves.

It is obviously important to constrain the semantic features required by the model to a manageable number. The requirement that each feature posited be grammaticalized in some language or other seems a reasonable way of limiting this number. It seems plausible that all the features that various languages grammaticalize are part of the relevant semantics in all languages.

This contrasts greatly with the unconstrained features used in connectionist attempts to link semantics and syntax. McClelland and Kawamoto (1986) use 1,052 input features. The training sessions involved require input sentences to be paired with their output representations. A verb could pair the input features [object is food and female], [subject is soft and medium-sized] and [verb is intense and causes chemical change] with the output features [agent of causal event is round], [patient of no-change is compact] and [instrument of shredding is soft]. The result is a massively knowledge driven process of sentence interpretation in which the actual contents of the sentence play a small role, merely activating preexisting associations of what typically happens in the world. Pinker gives examples of how their model assigns arguments to predicates during sentence comprehension.

(3.25) The plate broke. What broke?
A vase or window.

(3.26) The boy broke. What broke?
A piece of furniture.

(3.27) The pillow broke the window. What broke the window?
Something hard.

The approach of semantic structures, where it is assumed that aspects of the representation link to syntactic positions via unchanging linking rules and that the representations explain why only certain verbs enter into various alternations seems far more acceptable. It may also

be the case that semantic structures together with their linking rules play a role in phenomena which are normally explained in terms of purely syntactic representations. (e.g. Minimal Attachment).

3.7 Parameter Setting and Grammaticalization

In this section I shall first describe the sub-discipline of grammaticalization. I will present the principles of grammaticalization that have been proposed. I will argue that grammaticalization is very similar to language acquisition in the tasks it has to perform.

I will then go on to give some examples of the choices languages have made in the way they grammaticalize our conceptualization of the world. These examples present a complexity that a theory of parameter-setting, particularly one based on binary parameters will be hard pushed to capture.

If we see language acquisition as a task that involves associating a linguistic representation with the conceptual relation in the world that humans cognize, then we can see a clear link between language acquisition and grammaticalization. Grammaticalization takes place over time and is the theory of how languages turn particularly salient conceptualizations into a grammatical form. Under this view the form of a language will be a design solution to the task of facilitating communication. But as a design solution it is unlikely that the design decided on will be optimal.

It is interesting to note that, although language acquisition and grammaticalization are very similar in the task that has to be completed, the sorts of arguments that are used in these two sub-disciplines are widely divergent.

Over time languages grammaticalize contentful elements that are particularly important to a particular culture. Thus, in this way, the rules and rigidity of the distinction of synchronic and diachronic linguistics is called into question by the study of grammaticalization.

With regard to the function word/content word distinction it should be noted that function words are very often derived from content words.

The grammatical nature of the element becomes more and more abstract along the following cline:

content item < grammatical word < clitic < inflectional affix

Hopper and Traugott (1993) see grammar as a process for organizing cognitive and communicate content. They adopt a more flexible view of language universals - a universal component *broadly* characterizes the properties of the human constitution and can be modified by outside stimuli and by the functional purposes to which language is put.

Two main processes of grammaticalization are reanalysis and analogy. Reanalysis modifies underlying representations, whether semantic, syntactic or morphological and brings about rule change. Analogy modifies the surface manifestations and does not in itself bring about rule change.

There are three conflicting inputs/motivations to reanalysis. Old information is placed first and new information later. Ideas that are closely connected tend to be placed together and what is uppermost in speakers' minds is first expressed.

These are the processes, but what are the factors that actually enable these processes to come into operation. Here three factors are relevant:

- Language Acquisition
- Role of the types of contact within the speaking community
- The different priorities of the speaker-hearer roles.

Concentrating first on the speaker-hearer roles, hearers process the input in ways that may not match the speaker's intentions. Thus over time, the pragmatic inferences that are made by the hearer with respect to the speaker's utterances may become conventionalized.

The first step in this is semanticization, where the conversational implicature becomes part of the polysemies of the form that these implicatures are associated with. Thus, in English, *since* has come to be understood as signaling causation, whereas it originally was merely temporal.

Within the concept of inferencing two main processes are relevant: Metaphorical processes and metonymic processes.

In *Metaphorical Processes* a more concrete meaning is used to express an abstract meaning. For example *grasping an idea* uses the mind-as-body metaphor. This process is normally seen as a semantic process but Hopper and Traugott (1993) argue that it is in fact pragmatic. Other examples are where *SPACE* is seen in terms of an *OBJECT* as in **behind the cupboard** and where *TIME* is seen in terms of *SPACE* as in *we are behind in paying*. It could also be argued that the semantics of the modal have been based on principles of force dynamics.

Metonymic Processes rely on the conceptual association of one entity with another. For example, the future meaning of *going to* is based on the dual inferences of later time indexed by *go* and the purposive meaning of *to*. From this example we see that the process of metonymy depends on contiguity and reanalysis.

Heine *et al.* (1991) argue that underlying grammaticalization is a principle of *exploitation of old means for new functions*. Using this concrete concepts are used to understand, explain and describe less concrete phenomena. so for example non-physical experiences are understood in terms of physical experiences, time in terms of space, or abstract relations in terms of kinetic processes or spatial relations.

The relevant processes may be described in terms of a few basic categories which can be arranged linearly in the following way:

(3.28) PERSON>OBJECT>PROCESS>SPACE>TIME>QUALITY

Each of these categories can be viewed as representing a domain of conceptualization which is important for structuring experience. The relationship among them is metaphorical i.e. any of them may serve to conceptualize any other category to its right.

The cognitive categories are reflected in various aspects of language structure. Interrogative pronouns are often structured in a way that mirrors these distinctions: Thus all African languages have separate pronouns for the categories PERSON, OBJECT, PROCESS, SPACE, TIME and QUALITY.

There appears to be some correlation between these metaphorical categories and the division of word classes:

Category	Word Type	Constituent Type
PERSON	human noun	NP
OBJECT	non-human noun	NP
PROCESS	verb	VP
SPACE	adverb, adposition	ADVP
TIME	adverb, adposition	ADVP
QUALITY	adjective, adverb	modifier

The task of grammaticalization is to construct an efficient way of communicating our conceptualization of the world. Elements that are particularly salient to a culture's ecological niche are more likely to be grammaticalized. More complex conceptualizations are often grammaticalized using the more central grammaticalized conceptualizations in a metaphorical manner. Thus, once more, central to the form of language is an understanding of what is important to us as humans.

3.7.1 The Complexity of the Grammaticalization Data

I think it is useful to see one example of the complexity of the grammaticalization data in order to realize quite how many small variations need to be captured by parameters. Lehman (1988) looks in detail at the phenomenon of clause linkage. Lehman lists six semantic-syntactic parameters that can be identified in the linking of clauses. The complexity of these 'parameters' does not seem expressible in the form of parameters posited by PPT theory:

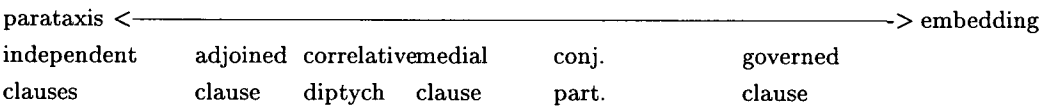
- The hierarchical downgrading of the subordinate clause.
- The main clause syntactic level of the subordinate clause.
- The desententialization of the subordinate clause.
- The grammaticalization of the main verb.
- The interlacing of the two clauses.
- The explicitness of linking.

The link between the two clauses can be seen in terms of a continuum from Autonomy running to Integration.

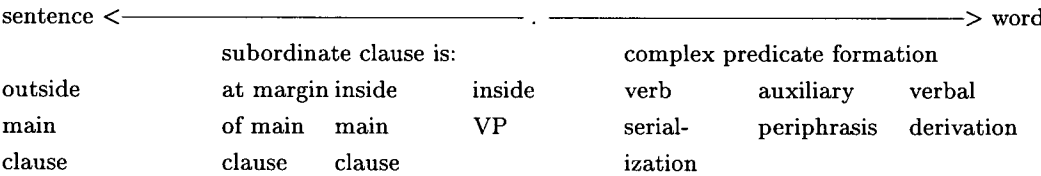
- No hierarchical relation = parataxis.
- Adjoined clauses - one of the two clauses contains a subordinate conjunction - not embedded.
- Correlative diptych - relative clauses subordinate but not embedded. Sometimes the subordinate clause has to appear in the last position as last constituent of the main clause.

- Clause-chaining: verbs in subordinate clauses with loss of tense, aspect and mood are taken to be those of the main verb. Person agreement shows whether the same subject or a different subject of the main clause is the case.
- Embedded clause - object clause governed by main verb.

In line with this continuum there are increasing constraints on the position of the subordinate clause:

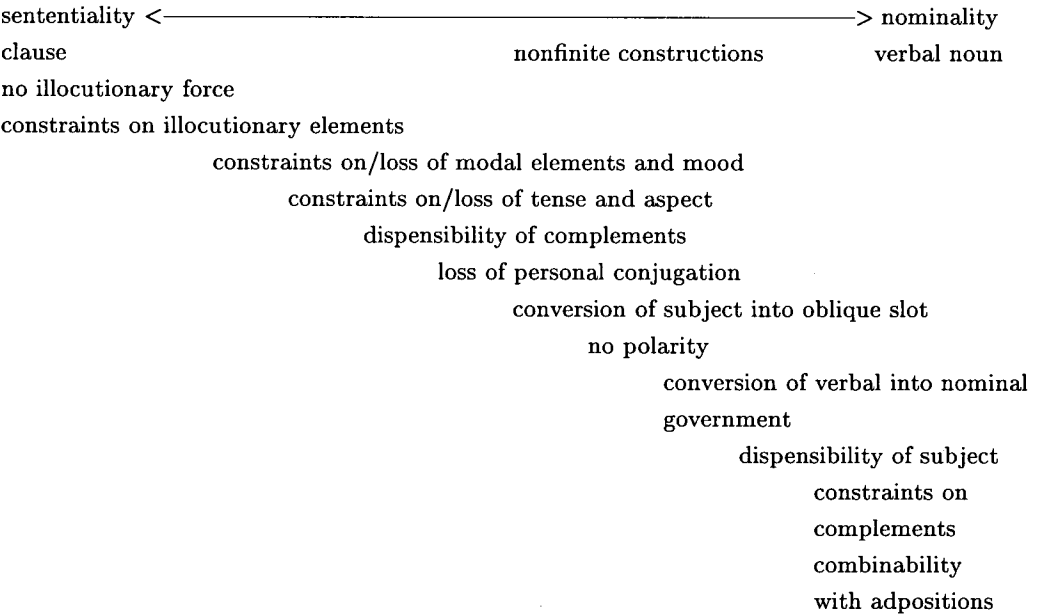


With regard to syntax there is variation according to the main clause syntactic level which the subordinate clause belongs to:

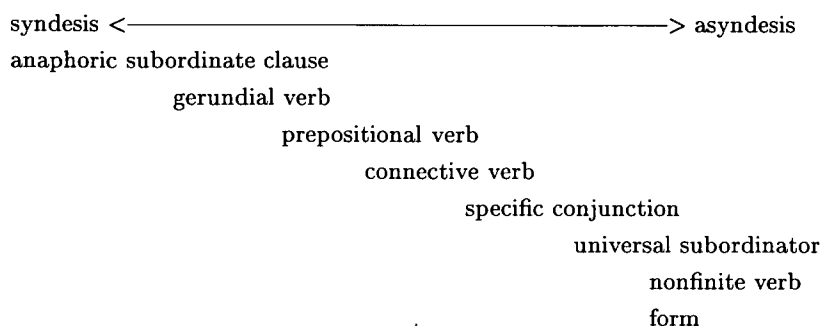


Another aspect is the degree to which the subordinate clause is desententialized - the more desententialized it is the more nominal it becomes. A fully-fledged sentence is characterized by illocutionary force, mood, tense, aspect, actants and circumstants.

The Desententialization hierarchy is:



Languages also vary as to the explicitness of the linking of the two clauses. The absence of a connective device is termed syndesis. The most explicit linking is termed asyndesis.



Lehman has found the following correlations among these various options.

- High integration of subordinates into the main clause correlates positively with its desententialization.
- Strong grammaticalization of the governing predicate presupposes either advanced desententialization or strong interlacing.
- Interlacing of clauses as brought about by raising presupposes downgrading, thus integration of the subordinate clause.
- Interlacing of clauses as brought about by dependent subject control leads to desententialization of the subordinate clause.
- Asyndesis correlates with dependent subject control and thus with interlacing.

The main aim of this section was to show the sheer range of possibilities for joining two clauses. This range of possibilities is not likely to be expressible or described by a small set of binary parameters that pay no attention to content.

3.7.2 A Link between Metaphor and Thematic Conceptual Structure

Sweetser (1990) puts forward a cognitively based theory which takes not the objective 'real world', but human perception and understanding to be the basis for human language.

In our conceptual organization, vision and knowledge are linked and for this reason any theory of semantics has to take this into account. But many theorists have been reluctant to take seriously the idea that language is shaped by cognition. One reason is that linguists have hoped to analyze language relatively independently of the rest of human abilities.

The metaphors manifested in most linguistic systems fall out from a more holistic viewpoint, which takes language as part of our general cognitive system: linguistic structure is then as logical and objective as human cognition, no more and no less.

So for example the following vision metaphors are prevalent in English and other Indo-European Languages:

1. The most common semantic sources for vision verbs are:
 - (a) The physical nature of sight (light, the eyes, facial movement etc)

- (b) Metaphors of vision:
 - i. Vision $\leftarrow$ physical touching manipulation. The probable basis of this is in the channeling and focusing ability connected with our visual sense; vision, far more than the other senses, can pick out and attend to one stimulus amid a multitude of input stimuli.
 - ii. Visual monitoring $\leftrightarrow$ control: The basis for this metaphor is probably the fact that guarding or keeping control often involves visual monitoring of the controlled entity; and the limited domain of physical vision is further analogous to the domain of personal influence or control.
 - (c) Basic Indo-European vision roots: There is a set of basic Indo-European roots which seem to have referred to vision as far back as history can be traced.
2. Target domains for vision verbs. Vision verbs commonly develop abstract senses of mental activity:
- (a) Physical sight $\rightarrow$ knowledge, intellection: This metaphor has its basis in vision's primary status as a source of data; not only does English have expressions like *I saw it with my own eyes* to indicate certainty, but studies of evidentials in many languages show that direct visual data is considered to be the most certain kind of knowledge.
 - (b) Physical Vision $\rightarrow$ mental *vision*: This metaphor is probably based on the strong connection between sight and knowledge, and also on the shared structural properties of the visual and intellectual domains - our ability to focus our mental and visual attention, to monitor stimuli mentally and visually.

This type of metaphorical basis for linguistic expression will form a useful addition to the theory of thematic-conceptual relations I set out in the next chapter. The two in conjunction will be used as part of the psycholinguistic theory I set out in Chapter 5.

3.8 Chapter Summary and Predictions for Psycholinguistics

In the next chapter I will be continuing my survey of the role thematic information has been seen to play in the various linguistic sub-disciplines. In that chapter psycholinguistics will be addressed.

In this chapter I have been considering language acquisition. I have argued for a Classical theory of rules and representations. The representation I have adopted is that of thematic-conceptual structure. We saw in this chapter how this type of representation can explain how it is possible for children to learn which verbs alternate syntactically and which do not via the underlying semantics of the verb as represented by thematic-conceptual structure.

We have seen that it is most likely that children begin with a thematically-based system that then matures into a more syntactic system with functional categories. The putative semantic content of this early thematic system was discussed with such elements as contact,

attachment and support playing a major role. We also saw a proposal from Lebeaux that when a grammatical system acquired later breaks down, the processor will fall back onto a lower (earlier acquired) grammatical system.

In Chapter 2 we saw that the components of thematic-conceptual structure should be limited to those semantic items that are found to be grammaticalized in some language in the world.

In the final section of this chapter, grammaticalization itself was discussed. Language acquisition can be regarded as a synchronic version of the diachronic process of grammaticalization. The child is grammaticalizing the thematic-conceptual structure underlying all language into the particular syntactic structure of the language she is learning. Thus what the child is learning is the linking from thematic-conceptual structure to the surface structure. It is possible that it is at this level that parameters may be able to be defined. The idea of thematic-conceptual structure and linking as a psycholinguistic process will be central to the argument of Chapter 5.

The adoption of thematic-conceptual structure in syntax and language acquisition makes the following predictions for language processing:

1. Thematic-conceptual structures will be a major player in the processing of sentences, as processing is then understood as the linking from thematic-conceptual structure to syntax or vice versa, depending on whether production or comprehension is being considered.
2. The grammaticalization of metaphors that have become central to humans in their ecological niche (See Chapter 6 for further discussion) will play their part in the content that is relevant to processing.
3. The breaking down of a higher processing system will result in the processor falling back on lower processing systems. The lowest level of this processing will be thematic-conceptual structure, as thematic-conceptual structure forms the earliest stages of language acquisition.
4. Linguistic parameters may well be able to be stated and defined at the level of the linking of thematic-conceptual structure to the surface syntactic structure.

In the next chapter I will survey the role thematic information has played in psycholinguistic theory in the past and in Chapter 5 I will address these predictions directly.

Chapter 4

Thematic Information in Psycholinguistic Theory

In Chapter 2 we saw how thematic roles have been built into syntactic theory. We noted that within PPT thematic roles were basically analogous to argument structure. Modifications of the basic formulation of thematic roles incorporated the notion of hierarchies to predict the linking of thematic structure into grammatical functions, or used an extended structure to explain, within thematic structure, transformations that were usually explained with a movement operation. We also saw decompositional representations of thematic information.

In this chapter I shall turn my attention to the part thematic roles have been assumed to play in psycholinguistic theory. *A priori* it may be expected that thematic roles in syntax and thematic roles in psycholinguistics should coincide. However, as we shall see this does not seem to be the case. Thematic roles often seem to be peripheral to the explanations employed by psycholinguists (except in a minority of frameworks).

When thematic roles are used in any coherent explanatory way, there seems to be doubt as to the form of the thematic roles that are assumed to be part of the parsing process. I will show that thematic structure plays an important part in parsing decisions and that the approaches that concede this are explanatorily the most adequate. I shall show in Chapter 5 that thematic-conceptual structure can be incorporated into the thematically based system discussed in this chapter, and that it will also be able to explain phenomena that remain unexplained by that system.

There are four main approaches in psycholinguistics with respect to thematic roles, one of these being not to include any mention of thematic roles within the parsing algorithm. Parsing algorithms based on the statistical properties of language take this approach. A typical paper which takes this stance is Mitchell and Cuetos (1991). In that paper it is assumed that the statistical properties of a lexical item predict the parsing preferences associated with that item. It seems that an approach such as this does not do much to *explain* parsing preferences but serves merely to *reflect* them. The argument seems to be particularly circular.

If the parsing preferences associated with a lexical item are based on the statistical properties of that item, what governs the statistical properties of the item, why does it have the

statistical properties it has? The only answer that seems to be put forward is that it has these statistical properties because they are preferentially produced with one structure associated with them as compared to another structure that could be associated with that lexical item. No explanation is given for why a particular structure is produced more often in association with that lexical item as compared to the less frequent lexical item-structure association.

The three other current approaches in psycholinguistics that do mention thematic roles see thematic roles as (i) a method of bringing latent parallelism into the parsing algorithm, (ii) a filter on what the syntactic processor has produced, (iii) determining parsing decisions. I shall discuss each of these below. Before I do that, however, I think it is worthwhile to list the type of empirical data that psycholinguistics is typically trying to capture.

4.1 What Psycholinguistics is Trying to Explain?

Before we get into the details of the parsing algorithms including thematic roles/structure that have been proposed, it is best to list the phenomena that are the explananda of psycholinguistics in an unbiased way.

It has been argued that since language processing is so fast, it is only when the system breaks down that we will be given clues as to its operation. For this reason most of the research effort has been expended on explaining the parsing preferences of sentences that are ambiguous in some way or other. The breakdown in processing that occurs in ambiguous contexts allows us to make inferences about the processes that operate in normal processing.

The first type of ambiguity is where there is a choice for the processor between constructing a Main Clause or Relative NP. It is generally accepted that the ambiguity is resolved in favour of a main clause analysis. The dispreferred reading results in a *conscious garden-path*. There are certain types of ambiguity that can be resolved with very little effort by the processor and some that are consciously difficult to understand.

- (4.1) a. The boat floated past the house.
 b. The boat floated past the house sank.

The second type of ambiguity is between a preferred Complement Clause reading and the dispreferred and consciously difficult to understand Relative Clause.

- (4.2) a. The farmer persuaded the vet that he was having trouble with his sheep.
 b. The farmer persuaded the vet that he was having trouble with to leave.

A third type of ambiguity involves a noun following a verb which could be analyzed as the object of that verb or the subject of a following verb. The ambiguity is resolved in favour of an object interpretation. The subject interpretation produces conscious difficulty.

- (4.3) a. While Ron was sewing the sock it fell on the floor.
 b. While John was sewing the sock fell on the floor.
- (4.4) a. I suspect John.
 b. I suspect John will come soon.

- (4.5) a. I warned John.
 b. I warned John would come soon.

A similar kind of ambiguity occurs when a noun can be analyzed as being attached to a possessive, pronoun and then proves not to be, as in the second sentence here:

- (4.6) a. Without her it would be impossible.
 b. Without her contributions would be impossible.

The fourth type of ambiguity is the Double-Object Ambiguity. This appears to be resolved in various ways. In some cases there is no conscious difficulty.

- (4.7) a. I gave her presents.
 b. I gave her presents to Ron.

On the other hand, in other cases there is a conscious processing difficulty.

- (4.8) a. Joe put the biscuit in the jar into Jill's mouth.
 b. Fred gave the cat the dog bit a saucer of milk.

There are also ambiguities based on a straightforward ambiguity in the meaning of an individual lexical item. The Lexical Ambiguity at times produces no conscious processing difficulty.

- (4.9) a. The warehouse fires destroyed all the buildings.
 b. The warehouse fires a dozen employees every week.

And at other times there are conscious difficulties, as can be seen in the second of these sentences.

- (4.10) a. The old train chugged down the track.
 b. The old train their dogs.

Similar to the object/subject ambiguity is the case of preposed adverbial clauses with, or without a comma. In the case without the comma a conscious processing difficulty occurs as the NP is attached incorrectly to the preceding verb and then has to be reanalyzed as the subject of the next verb.

- (4.11) After the boy had eaten (,) the remaining food went off.

In reduced relative sentences, the disambiguating region takes longer to process than the disambiguating region in the full relative.

- (4.12) a. The defendant examined by the lawyer turned out to be unreliable.
 b. The defendant that was examined by the lawyer turned out to be unreliable.

There are therefore of the order of only a dozen different ambiguity types in English. There are thus two main questions that the processing literature has to answer. The first is how the ambiguity is handled in the first place. That is, whether ambiguous materials are more difficult to process than unambiguous material. This difficulty can be either in the ambiguous region itself, or the region that disambiguates the ambiguous region.

There are three main options in this regard:

1. Serial Processing: Assemble the structure for just one of the potential meanings (Kimball (1973); Frazier (1979)).
2. Parallel Processing: Construct the structure for both potential analyses (Gorell (1987)).
3. Minimal Commitment Processing: Make a partial analysis and delay decisions on the ambiguity until it is possible to do so successfully (Weinberg (1990)).

The second question concerns the mechanism of initial choice. That is, what decides which of the two potential meanings/structures associated with an ambiguous word will be pursued initially. There are six competing theories here (Mitchell (1994)):

1. Tree-Structure
2. Lexical Frames
3. Thematic Assignment
4. Discourse Driven procedures
5. Exposure Based Procedures
6. General Processing Cost Strategies

In Sections 4.2 to 4.4 three parsing algorithms which include a role for thematic structure are presented.

4.2 Latent Parallelism in Thematic Roles

The approach where thematic roles are seen as providing latent parallelism in a mainly serial parser has been associated with the work of Tanenhaus and his colleagues. The theory was first delineated in Carlson and Tanenhaus (1988).

In this initial paper the authors are unclear what the exact nature of thematic roles is; they could be syntactic, semantic or conceptual in nature. What is most important is that they see thematic roles as a mechanism of interaction between syntax, the discourse model and world knowledge. Thus in terms of the (Fodor (1983)) modularity thesis, where there is a central processor (which includes world knowledge) and input modules (and Fodor argues that language is an input module) thematic roles form a frontier module at the boundary of the language system and the central processor.

4.2.1 The Discourse Strand

With respect to parsing, Carlson and Tanenhaus (1988) argue that all senses of a word are activated in parallel, regardless of context. This seems a reasonable claim given the work of Swinney (1979)¹. They then argue that the most appropriate sense remains active, and that thematic roles are assigned as soon as possible. Any roles that are not assigned remain open in the discourse model and this will make integration of some sentences that follow in the discourse event easier than others. This is one strand of their predictions - let us call it the *Discourse Strand*

Thus in the following sentences (4.13c) is read more quickly after (4.13b) than after (4.13a), they argue, because although *suitcases* are conceptually linked to travel (as in (4.13a)), there is an open unassigned role associated with *unload* that could easily be filled with *the suitcases*, and thus the (4.13c) sentence will be easily integrated into the discourse as the open role will form a conceptual link between the two sentences.

- (4.13) a. Bill hurried to catch his plane.
b. Bill hurried to unload his car.
c. The suitcases were very heavy.

It is predicted then that open roles thus form antecedents for definite NPs. This was tested by constructing a target sentence such as (4.14c) preceded either by a context sentence that introduces an open thematic role that the NP could plausibly fill (4.14a) or a sentence with a plausible context but no open role (4.14b).

- (4.14) a. The miners were drilling a large tunnel. (open Location)
b. The miners were making a large tunnel. (no open role)
c. The rock was hard.

Subjects had to judge whether the final sentence made sense. 97% judged the final sentence to make sense after reading (4.14a) compared to only 84% who judged it to make sense after reading (4.14b). The decision latency was also faster after (4.14a) than after (4.14b); 1628m-sec versus 1847m-sec.

It was also found that after a context such as (4.15 a), sentence (4.15 b) was more easily integrated than sentence (4.15c).

- (4.15) a. Her nephew's birthday was coming up so..
b. Mary sent a book. (open Goal)
c. Mary bought a book. (no open role)

A final test utilized the probe recognition task, where the recognition time to decide that a word occurred in the preceding sentence or short discourse is faster when the word is an antecedent for a subsequent anaphor.

¹ Although this may not be totally true. It seems that the interpretation can be influenced by the context to some extent (Burgess and Simpson (1988)). Burgess and Simpson argue that the two contrasting results can be explained if the two hemispheres of the brain are seen as doing different things. In the right hemisphere the non-dominant meaning is activated for longer. In the left hemisphere, on the other hand, categorical choices are made on the basis of frequency.

- (4.16) I don't know who should sweep the crumbs off the floor.
- a. However, I bet Carol was the one who made the mess (makes sense, but no VP anaphor).
 - b. However, I bet Carol will volunteer to, reluctantly. (VP ellipsis)
 - c. However, I bet Carol will volunteer (to do it) reluctantly. (explicit VP anaphor or unexpressed verbal complement)

The relevant probe was the verb *sweep* here. The recognition time was faster in the VP ellipsis condition (1201m-sec) and the VP anaphora/null complement condition (1176m-sec) than in the non-anaphoric condition (1283m-sec), providing evidence that the priming effects for VP anaphors are similar to those for NPs.

4.2.2 The Thematic Recovery Strand

A second strand is the *Thematic Recovery Strand*. Although Tanenhaus and his colleagues have been convinced that parsing is essentially a serial process, with only one syntactic parse being followed at any one time, Carlson and Tanenhaus argue that recovery from misparses is generally quick (they seem here to be ignoring what seems to be a clear distinction between conscious garden-pathing and non-conscious garden-pathing). The argument is that the parser is able to recover quickly as, although it is only following one syntactic parse, the various thematic grids associated with a lexical item are activated in parallel throughout the parse, and that when a misparse is detected, it is the fact that these thematic grids are all still active that allows for quick recovery. This is reminiscent of the Property of Smooth Degradation as proposed by Lebeaux with respect to language acquisition.

What this argument means with respect to linguistic representation is that thematic grids and subcategorization frames are assumed to be different things, and have different contents and thus must be listed separately for each lexical item. Thus a verb will have a core meaning, a subcategorization frame and a number of thematic grids associated with it.

Thus there are different effects of core meaning ambiguities when compared to thematic ambiguities. Sentences with temporary thematic ambiguities (4.18) show weaker garden paths than sentences with sense ambiguities (4.17), as only the sentences with thematic ambiguities will have the alternative roles still available to the sentence processor, facilitating recovery.

- (4.17) Bill passes the test to his friend.
- (4.18) Bill loaded the truck onto the platform.

4.2.3 Thematic Influence on Initial Decisions

The third strand of prediction in the model can be termed the *Thematic Influence on Initial Decisions*. Thus Tanenhaus *et al* do not see thematic roles as purely enabling recovery from misparses; they also believe that thematic information can influence decisions on parsing on-line.

They argue that animacy is a part of thematic knowledge (if this is the case, they would seem to be arguing for thematic roles that are not primitives, but are built from smaller

conceptual primitives), and that it is the non-animacy of *truck* in the following example from Stowe (1986) (later replicated in a self-paced reading paradigm by Holmes *et al.* (1989)) that makes it an unlikely agent for *stop* and that therefore (4.19b) is easier to parse than (4.19a) as the misparse is less likely to occur.

- (4.19) a. Even before the police stopped the driver was getting nervous.
 b. Even before the truck stopped the driver was getting nervous.

It seems that the form of thematic roles that Tanenhaus and his colleagues are arguing for, in spite of their doubt on this point, is a form where thematic roles are semantic or conceptual rather than purely syntactic. This is very much in line with the more decompositional approach to thematic information we saw in the final sections of Chapter 2.

This argument has its source in Crain and Steedman (1985) who contend that sentence comprehension involves rapidly and optimally integrating lexical, syntactic and contextual information in some form of mental models representation, a position that will also be taken in this thesis. However, Crain and Steedman had not worked out the nature of the representation required nor any specific mechanisms by which input could be integrated into a mental model. I shall argue for thematic-conceptual structure as the relevant representation and linking rules as the mechanism.

Further evidence that thematic information comes to bear early in sentence comprehension can be found in long distance dependencies (Tanenhaus and Carlson (1989)). Filler-gap dependencies allow the properties of different verb frames to come to bear. So, for example, in the following sentences, *which girl* is an implausible object of *hope* and a plausible object of the verbs *persuade* and *hurry*.

- (4.20) a. Which girl did the boy hope --
 b. Which girl did the boy persuade --
 c. Which girl did the boy hurry --

The embedded anomaly technique trades on the difference in plausibility of a filler after different verbs. There is an increased load if a gap is posited at an implausible site. Using this technique it is possible to test for thematic influence on sentence comprehension.

- (4.21) a. The physical therapist wasn't sure which doctor the orderly hurried rapidly towards ---
 b. The physical therapist wasn't sure which bed the orderly hurried rapidly towards ---

Here there is a decoy gap after *hurried*, where *which doctor* would be a plausible filler and *which bed* would be an implausible filler. When sentences with verbs with a transitive preference were compared with sentences with verbs having an intransitive preference, it was found that in sentences with transitive preference verbs, the implausible fillers were read more slowly at the verb. Thus the gap was posited, filled and semantically interpreted at the verb and hence thematic information must have been available and used at this point. In contrast

to this, in sentences with verbs with an intransitive preference, there was no plausibility effect at the verb, there being no available thematic role against which to test the plausibility of the filler.

Boland *et al.* (1990) add a qualification to this. They argue that immediate incongruity is spotted only if the filler is implausible in all potential roles. They use sentences such as the following, where *read* is a verb that is typically transitive and does not allow an infinitival complement, and *remind* is obligatorily transitive and allows a complement clause.

- (4.22) a. Which child did Bill remind to watch the show.
b. Which movie did Bill remind to watch the show.
- (4.23) a. Which book did the child read in bed at night.
b. Which food did the child read in bed at night.

In a self-paced reading stop-making-sense task, where readers read the sentence word by word and push a 'No' button when they feel the sentence has stopped making sense it was found that the *read* sentences produced similar results as in Tanenhaus and Carlson (1989). However, the *remind* type sentences resulted in no difference between the plausible and implausible filler conditions.

It is claimed that the reader is not associating the filler with a role in which it is incongruous until there is no other choice, that is at *to*. This is further evidence that the decision can not be made in terms of syntactic structure alone as it is hard to explain how the processor could judge the congruity of the filler with respect to positions based on the syntactic representation alone. To gauge the congruity requires access to thematic role information.

There is a further issue that has been addressed within this framework. The results up to now with respect to gap filling have been compatible with two accounts:

- The location of the gap must be identified first.
- The filler is initially interpreted using semantic information that was made available as soon as the verb was recognized.

To test which of these was true Tanenhaus *et al.* (1994) created sentences where the implausibility is contingent upon the filler having been assigned a particular role and the implausibility precedes the syntactic location of the empty category. They achieved this by using non-dativizable double-object verbs.

- (4.24) a. Which campus party did John contribute some cheap liquor to -- last week.
b. Which public library did John contribute some cheap liquor to -- last week.

Both of these sentences have an implausible Theme but a plausible Recipient. The thematic filling hypothesis predicts that interpretation can take place at *contribute*. The syntactic hypothesis predicts that interpretation cannot take place until *to* is encountered. In a stop-making-sense experiment subjects were found to be noticing the oddity in the implausible

condition before reaching the syntactic gap and were thus argued to be using the thematic filling strategy².

Trueswell *et al.* (1994) used an eye-tracking experiment to show that thematic information could be used by the processor to lessen the size of a garden-path effect in reduced relative sentences. They were careful to use animate/inanimate NPs in sentences that did not allow plausible past tense completions, thus avoiding the problems of that nature that they identify in Ferreira and Clifton (1986).

- (4.25) a. The man recognized by the spy took off down the street.
 b. The man that was recognized by the spy took off down the street.
 c. The van recognized by the spy took off down the street.
 d. The van that was recognized by the spy took off down the street.

Normal garden-path effects occurred after the animate nouns. However, there was no reliable difference between the reduced/unreduced relative clauses after the inanimate nouns.

This conceptual approach is on the right lines as, arguably, true sentence comprehension centrally involves the computation of the thematic roles that the entities in the sentence play. The pure computation of syntactic structure cannot provide this information. And as Tanenhaus and colleagues argue, the computation of thematic roles is essential to building the discourse model into which following sentences can be integrated. If this claim is true, we would expect thematic roles to have an effect on discourse structure - some of the evidence given above appears to support this claim. In section 4.6 we shall see that it is not merely that thematic roles are implicated in the construction of a discourse model, but that there are differing psychological saliencies associated with the various thematic roles.

4.3 Thematic Roles as a Filter on Syntax

Tanenhaus and Carlson identified two approaches to the immediacy of comprehension in parsing. The first of these involves rapid building of syntactic structure, so rapid that the parser makes hypotheses without using lexical knowledge. This would mean that there would be very many small garden-paths occurring during parsing. The second approach involves the integration of lexical, syntactic and contextual knowledge in some form of mental models representation. The work of Tanenhaus and his colleagues we have seen above is most easily identified with the second approach. The first approach finds its most well-known proponents amongst Lyn Frazier and her colleagues.

Indeed, early work in this framework was based very strongly on mechanisms that took no account of lexical knowledge. A typical paper in this regard is Frazier (1979) where the principles of Minimal Attachment and Late Closure are proposed. These are very similar to a number of the parsing principles to be found in Kimball (1973).

(4.26) *Minimal Attachment*

Attach incoming material into the phrase marker being constructed using the fewest nodes consistent with the well-formedness rules of the language under analysis.

²Although Nicol and Pickering (1993) have found evidence for reactivation at both the verb and the gap.

(4.27) *Late Closure*

When possible attach incoming material into the phrase or clause currently being parsed.

These two principles are defined over tree-structures and are thus susceptible to changes in the theory of syntax. Structures that were assumed to be easier to parse because they have a simpler structure may have been reanalyzed in more recent syntactic theory as having an equally complex or a more complex structure. This change arguably happened when Abney (1987) reanalyzed the noun phrase in terms of more complex determiner phrases.

Over the years the approach has been modified to some extent, so that the effects of other aspects of linguistic representation can be built in. The architectural assumptions that Frazier now makes are made most explicit in her 1991 paper. We will see that Frazier has come remarkably close to the position of Tanenhaus, without wanting to recognize this fact.

In general, what seems to have happened is that the growing part played by thematic roles within syntax has had a knock-on effect in psycholinguistics and, in the same way as in syntax, researchers have tried to build thematic roles into whatever theoretical framework they were advocating before with as little effect on that initial framework as possible. So the differences that seem to be emphasized between various psycholinguistic theories are derived from the starting assumptions of the researchers' early work and not actually by the endpoint.

What needs to be done is an exploration of thematic roles in parsing divorced from other theoretical assumptions. Or, as I argue in this thesis, to install thematic-conceptual structure in a central position within linguistic theory and not a peripheral one.

In the 1991 paper Frazier assesses what bearing psycholinguistics has on the Fodor (1983) modularity thesis. She notes that it is problematic that world knowledge influences the ease of reanalysis of unpreferred structures which have garden-pathed the parser. She concedes that if this only happens on conscious garden-path sentences, it could be argued that the reanalysis takes place outside the language processor. However, she cites data from Frazier and d'Arcais (1989) which shows that in Dutch there are pragmatic effects for automatic reanalysis in non-conscious garden-path sentences and thus it is unlikely that recovery is taking place in the central processor.

Frazier argues that most of the subsystems of PPT are informationally encapsulated modules, having their own vocabulary. However, like Tanenhaus and his colleagues, she regards the thematic processor as being to some extent on the frontier of the language processor and the central processor.

Frazier states that form-driven subsystems of the language module are supplemented by the effects of world-knowledge, once translated *into* the vocabulary of a grammatical module by 'pseudo-encapsulated' linguistic systems.

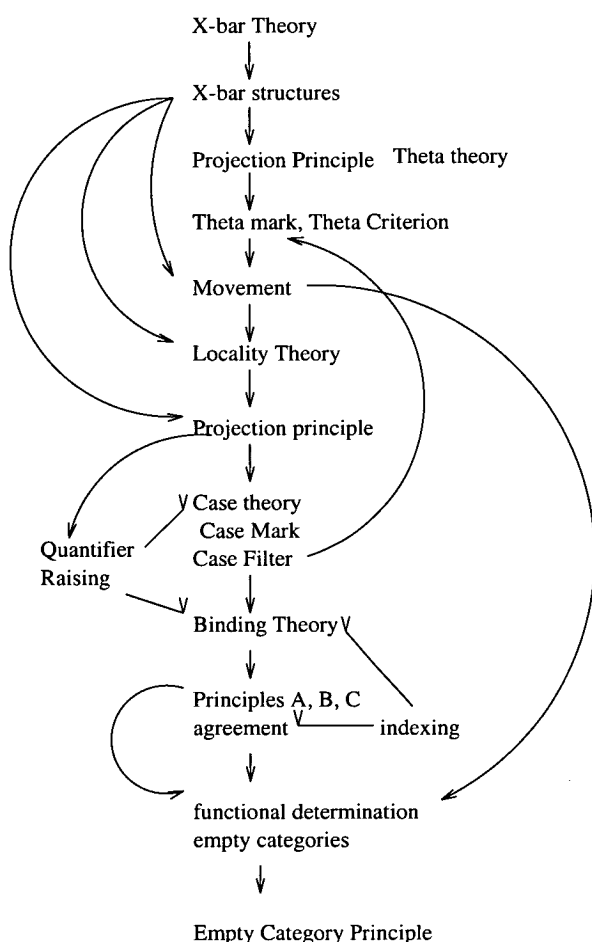
Evidence for such differing processing vocabularies can be found, she claims, in the work of DeVincenzi (1989) on the processing constraints for Binding Theory. The informational encapsulation is there, as there is no guarantee that the constituent structure decisions cannot carry on without checking the information that the binding-processor is providing. Thus the information provided by the Binding theory module and the constituent structure module may be inconsistent and will need to be cleaned up at the end of the sentence.

Frazier follows Rayner *et al.* (1983) in the idea that there is a subsystem of the language processor concerned with thematic and predication assignments which considers alternate analyses in parallel but delays a decision until all the arguments have been received. This arrangement is the latent parallelism of Tanenhaus and his colleagues. The difference is that thematic roles are not assumed to be able to guide the parse, only help to remedy a misparse.

The question is whether this position is coherent. Why should syntax operate in a serial way and thematic structure in a parallel way? As Fodor (1991) points out this is a very tentative suggestion, although it has been made by both Tanenhaus and Frazier. If it could be shown that thematic structure operating in a serial way could better explain the parsing decisions, this parallel/serial dichotomy would be cast seriously into doubt. Elements of the Pritchett (1988)³ approach as delineated in the next section go some way to doing this.

Another problem that is faced by both Frazier and Pritchett and indeed any psycholinguistic theory that bases itself closely on PPT as it stands, is the problem that is also faced by computational modules of PPT grammars: There does not seem to be a globally optimal ordering of the principles - the interrelationship between the principles is too complex.

The interactions involved can be seen in the following diagram from Berwick (1991).



Fong (1991) has considered the issue of principle combination and ordering in some detail. He built a parser where the order of the principles could be easily changed. He found one

³see also Pritchett (1992)

major problem; there does not seem to be a globally optimal ordering. To see this consider (4.28).

- (4.28) a. *John₁ is crucial [_{CP} [_{IP} t₁ to see this]]
 b. *[_{NP} John₁'s mother][_{VP} likes himself₁]
 c. *John₁ seems that he₁ likes t₁.

Example (4.28a) violates the Empty Category Principle. Hence the optimal ordering must invoke the ECP operation before any other operation that it is not dependent on. On the other hand (4.28b) violates Condition A. Hence the optimal ordering must invoke Condition A as soon as possible. In particular given that the two principles are independent, the optimal ordering must order Condition A before the ECP and vice versa. Similarly (4.28c) demands that the Case Condition of Traces' operation must precede the other two operations. Hence a globally optimal ordering is impossible.

Given that no optimal ordering is possible, Fong proposes a heuristic which the parser can use to reduce the amount of unnecessary work. The parser 'predicts' a failing filter. That is it will predict what principle the structure is most likely to violate given certain structural clues. Each structure clue is associated with a list of possible failing filters. Some of these are in the table below.

<i>Structure Cue</i>	<i>Possible failing filters</i>
trace	ECP, and Case Condition on traces
intransitive	Case Filter
passive	Theta Criterion/Case Filter
non-argument	Theta-criterion
+anaphoric	Binding Theory Condition A
+pronominal	Binding Theory Condition B

Although this dynamic ordering principle performs well in many test cases, it is by no means foolproof. There are also many cases where the mechanism triggers an unprofitable reordering of the default order of the operations.

Frazier has to hope that it is when optimal ordering breaks down in this way that garden-paths and, particularly, conscious garden-paths occur. But overall it seems unlikely that processing based on PPT is plausible with our current understanding of the nature of PPT and its computational complexity⁴.

4.4 Thematic Roles Determine Parsing Decisions

Pritchett (1988, 1992) proposes a system where thematic roles are central to the parsing decisions and to predictions as to when a garden-path will be conscious and when it will not.

⁴Although it may be the case that principle ordering could be achieved by using recurrent constraints, where if one principle started to apply the other principles would be relegated. This may be able to decide which principle is most apt at a particular point, but will not be able to calculate the principles that will apply across the whole sentence and the ordering in which they should apply, as the constraints approach will not be able to solve the problems identified by Fong.

He sees the parser as being isomorphous with the grammar. This seems to be the position that Frazier (1991) was (unexpectedly) taking. This is an unusual position as Frazier and most others felt that the parser and the grammar must be distinct following the demise of the Derivational Theory of Complexity (See, for example, Miller and Isard (1963)).

Much of what Pritchett has to say is reminiscent of the work of Tanenhaus and his colleagues. However, the form of thematic roles assumed by Pritchett is somewhat weaker than that assumed by Tanenhaus. Whereas Tanenhaus sees thematic roles as conceptual entities providing a link to mental models, Pritchett sees them as purely structural entities that need not have any content whatsoever.

This view comes as no real surprise as Pritchett's theory has been derived directly from the principles of PPT. Indeed, he expands the thematic principles below to include the on-line application of all of the principles of PPT. We have seen the kinds of problems that such an approach comes up against in the previous section. However, with these provisos in mind, it is still worthwhile investigating Pritchett's mechanisms which do fairly successfully predict when a parsing mistake will become conscious. This aspect of the theory could well be expanded into a richer theory by adding some of the insights of the Tanenhaus framework. What is fairly successful in terms of structural thematic roles may be even more successful when these thematic roles are incorporated into a decompositional representation of conceptual-thematic structure.

The principles Pritchett assumes are:

(4.29) *Theta Attachment*

The theta criterion attempts to be satisfied at every point during processing given the maximal theta-grid.

(4.30) *Theta-Reanalysis Constraint*

Syntactic reanalysis which reinterprets a theta-marked constituent as outside of a current theta-domain is costly.

and the theta domain is defined in the following way:

(4.31) *Theta-Domain*

α is in the theta-domain of β iff α receives the γ theta role from β or α is dominated by a constituent that receives a theta role from β .

If it is accepted that Pritchett's theta roles are really only argument positions, these three principles could be paraphrased as what I will call the Conscious Garden Path Hypothesis.

(4.32) *The Conscious Garden-Path Hypothesis (CGPH)*

In on-line processing, attempt to attach each NP as an argument of a preceding predicate. If an argument, once attached to one argument position of a predicate, has instead to be attached to a different argument position of that predicate, or alternatively to an argument position of a different predicate, a conscious garden-path will occur

Turning now to the data we saw in section 2, we can see how Pritchett's theory accounts for the patterns observed.

(4.33) =(4.1b) The horse raced past the barn fell.

Race has the maximal theta-grid (**Agent**, Location). As soon as *race* is encountered *the horse* is allocated the Agent role. *Past the barn* is then assigned the Theme role. When input then continues *fell* is encountered which has the theta grid (**Agent**). No Agent is available, so reanalysis has to occur, with *the horse* being removed from the theta grid of *race* and reanalyzed as the Agent of *fell*. Or, in terms of the CGPH, *the horse* is no longer an argument of the predicate *race*, rather it is reanalyzed as an argument of the predicate *fell*.

The explanation for the second type of garden path sentence proceeds in an analogous way.

(4.34) = (4.2b) The patient persuaded the doctor that he was having trouble with to leave.

Persuade has the maximal theta-grid (**Agent**, Goal, Proposition), or in terms of the CGPH, three arguments. *The patient* is assigned the Agent/1st argument role when *persuade* is reached. *The doctor* is assigned the Goal/2nd argument role, and the Proposition/3rd argument role is assigned immediately to *that he was having trouble with* as both internal arguments could be satisfied at this stage by doing so. Encountering *to leave* forces reanalysis as a relative clause. Thus *that he was having trouble with* has to be removed from the third argument position and reassigned as part of the second argument/Goal role.

In each of the object/subject ambiguity sentences in (4.3), the object analysis will be preferred as this will mean the relevant NP being attached to a verb that has already appeared with its potential argument position available to the reader. A subject analysis would require waiting for a verb requiring a subject, an event that the reader cannot be sure will happen.

4.5 Thematic Information in Sentence Processing

In general, there seems to be no load difference between ambiguous and non-ambiguous material. This fact favours serial models or resource free parallel models. In terms of the thematic role/structure models this means that each of the Tanenhaus, Frazier and Pritchett models are consistent with this finding.

When it comes to the disambiguating region, however, there do seem to be differences in the processing load. At least, a continuation after the ambiguous region takes longer to process than the same material in a non-ambiguous context. Once more all of the models that include a place for thematic roles are compatible with this observation.

The evidence with regard to the mechanism of initial choice is more confused. Frazier's approach (Section 4.3) was a tree-based approach with a thematic checking mechanism. Both the Tanenhaus account (Section 4.2) and the Pritchett account (Section 4.4) are thematic assignment accounts. The difference between these latter two accounts is that the Tanenhaus account provides latent parallelism and thus predicts ease of recovery, whereas the Pritchett account is strictly serial and predicts conscious difficulty in recovering from misparses.

Problems for the tree-structure account can be found in Taraban and McClelland (1988), where it is shown that in sentences with PPs that could be attached either as modifiers of the NP or as modifiers of the verb, the preferred attachment to the verb that the tree-structure based approach suggests is not always the preferred option. This will be discussed in depth

in the next chapter. There is also the problem of the complexity of PPT in any theory that is to be based on the principles of PPT in totality, as Frazier's most recent suggestions would seem to suggest.

Lexical frame strategies and discourse-driven strategies suffer from the problem that it is difficult to differentiate between the mechanism of initial choice and the mechanism of checking. As mentioned in the introduction of this chapter, exposure-based strategies do not seem to explain processing preferences but merely to reflect them in a circular manner.

The thematic-based accounts suffer from the problem of insufficient coverage of the data. These accounts make no prediction whatsoever for cases where adjuncts are involved (The PP/NP ambiguity in Taraban and McClelland (1988) provides an example of this). However, I will show that a thematic account based on thematic-conceptual structures, as opposed to mere thematic roles) can provide a solution to this problem (See Chapter 5).

4.6 Thematic Structure in Discourse

All things being equal, sentence comprehension can only really be said to have occurred when the reader has calculated the thematic relationships within the sentence. In other words, it is thematic structure that is central to language comprehension. If this is so, it comes as no surprise that thematic structure plays a role in language processing, as I have shown in the previous sections.

When we turn to discourse structure and specifically to the construction of a discourse model in comprehension, there is evidence that thematic structure acts as the interface between syntax and this discourse model. This evidence comes from Stevenson *et al.* (1994). This work shows that thematic roles differ in their saliency in the construction of the discourse model with particular regard to reference resolution.

The Carlson and Tanenhaus model suggests that the nature of described events and the relationship between them may also influence the way that pronouns are understood.

Following this train of thought, we might surmise that the arguments associated with the verb introduce entities into the discourse model. These entities can then serve as the potential antecedent of a pronoun. Thematic roles could then be seen as specifying the relationship of these arguments to the verb.

It may also be the case that some thematic relationships make certain entities more salient than others. This could perhaps be modeled within thematic-conceptual structure. These issues were examined by Stevenson *et al.* They examined preferences for one antecedent over another in sentence continuations such as

(4.35) Ken admired Geoff and he....

Here, the *he* could refer either to *Geoff* or *Ken*. They examined how other factors could interact with the preferred antecedent and the influence of thematic roles. These factors are

- The observed advantage for the first mentioned individual in the sentence.
- The presence or absence of a pronoun (Pronoun/No Pronoun Conditions) in the sentence presented to the subjects.

- (4.36) a. Ken admired Geoff and he... (Pronoun)
 b. Ken admired Geoff and... (No Pronoun)

- The difference between within- and between-sentence references.

- (4.37) a. Ken admired Geoff and he... (within-sentence)
 b. Ken admired Geoff. He... (between-sentence)

To examine the influence of the first mention advantage, the sentence position of the antecedents was varied (Stevenson et al. 1994). A proper evaluation of the relative salience of antecedents with different thematic roles requires that any effect of thematic role is pitted against the first mention effect.

Omitting the pronoun allowed investigation of two aspects of pronoun comprehension:

1. the way that a particular antecedent may be more highly focused. This is termed a *top-down effect* because the focused antecedent can be thought of as being in focus even before a pronoun is encountered.
2. the way that a pronoun may trigger specific search processes. They call this a *bottom-up effect* because it concerns processes that are triggered by linguistic input.

(4.38) **Goal-Source**

Version 1: John seized the comic from Bill. (Goal Theme Source)

Version 2: John passed the comic to Bill. (Source Theme Goal)

(4.39) **Agent-Patient**

Version 1: Joseph hit Patrick. (Agent Patient)

Version 2: Patrick was hit by Joseph. (Patient Agent)

(4.40) **Experiencer-Stimulus**

Version 1: Ken admired Geoff. (Experiencer Stimulus)

Version 2: Ken impressed Geoff. (Stimulus Experiencer)

(4.41) **Agent Goal/Source**

Version 1: Simon ran towards Richard. (Agent Goal)

Version 2: Simon ran away from Richard. (Agent Source)

For all these sentences, the subjects had to produce a continuation, The following pattern was observed.

- In Goal-Source sentences there was a preference for Goal.
- In Agent-Patient sentences there was a preference for Patient.
- In Experiencer-Stimulus sentence there was a preference for the Stimulus.
- In the Agent-Goal/Source sentences there was no preference.

Stevenson *et al.* concluded that the influence of thematic roles is a top-down effect, since the role preferences were observed even in the No Pronoun condition, i.e. where the continuation was elicited merely by *and*. It appears then that the thematic role effect is due to the preferred role becoming more highly focussed and accessible in the mental model.

It now remained for the thematic preference to be characterized. There is an apparent preference for focusing on the consequence of an event. However, this preference can be manipulated to some extent. If the two sentences are conjoined by *and* the Experiencer is preferred to the Stimulus. It is suggested by Stevenson *et al.* that the clausal link between the sentences is nullified by *and*. That is, when there are two sentences there is a focus on the cause of the event and when the sentences are conjoined there was a focus on the consequences of the experience.

This proposed psychological saliency of the consequences of an event was tested by using the connective *so* which focuses on the consequence of the event, and the connective *because* which focuses on the cause of an event. These connectives were expected to reinforce or reduce the focusing due to event structure. Thus the preferences for Goal and Patient are expected to be reduced by *because*.

The lack of focus in the Experiencer-Stimulus sentences leads to the expectation that *so* and *because* will be the sole determinants of the preference, with *so* leading to a preference for the Experience and *because* leading to a preference for the Stimulus. These predictions were all shown to be correct.

We see in the work of Stevenson *et al.* clear evidence for a focusing on the consequences of an event. This psychological saliency with regards to language will be discussed in more general cognitive terms in Chapter 6.

4.7 The Time-Course of Sentence Comprehension

In the course of this discussion, two main themes have emerged: the determinants of parsing decisions and the time-course of parsing decisions. I have discussed the various potential determinants of parsing decision (statistics, phrase-structure, thematic/argument structure) without examining in any principled way when these factors come into play during sentence comprehension.

In this section I will describe what seems to be the contemporary consensus with respect to the time-course of parsing. It is worth pointing out that much of what follows is controversial nonetheless. However, what can be drawn from an overview of psycholinguistic data is that thematic/conceptual structure would seem to be implicated most strongly over the entire act of language comprehension in general and should thus be central rather than peripheral to theories of parsing and language comprehension.

A distinction is often drawn between lexical proposal and lexical filtering. In lexical proposal, combinatory information in the lexicon is used to propose structure in advance of input. In lexical filtering, structure is built, and then filtered, as lexical information is used to reject the structures so generated.

This distinction can be misleading. So for example, *studied* could be either a past-participle or the past form of the verb. If the thematic role of the subject is used to choose between

these options, this could in some sense be seen as filtering. But the choice that has been made could then be said to be proposing further structure.

In general, in sentence comprehension, there are four main elements distributed over time: selection, assembly, checking and revision. Selection involves selecting the lexical information to be used to make initial decisions. On the basis of these decisions, a structure is assembled. This structure is then checked for coherence and if found to be incoherent, the structure is revised.

Any decisions made on psycholinguistic proposals should be made with one thing in mind - psycholinguistics is not merely about structure building, but rather it is about interpretation.

Thus there is the question of whether one interpretation or two is followed over the course of language comprehension. The weight of the evidence suggests that serial, rather than parallel, interpretation takes place, as interpretation seems to be incremental. However, as we have seen Tanenhaus argues for latent parallelism. This parallelism is latent as only one core meaning of the verb in question is followed and the parallelism is in terms only of the possible assignments of thematic roles associated with that core meaning. Another choice would be to suspend decisions on interpretation (Perfetti (1992)), but suspended interpretation decisions seem unlikely given that gap-filling seems to take place very rapidly at the verb rather than at a syntactic gap, or possibly at both (Nicol and Pickering (1993)).

The evidence also seems to favour a model where thematic roles predict structure and where interpretations are assembled on the basis of this information. Revision after checking also seems to have a thematic component (Frazier (1991); Carlson and Tanenhaus (1988)) and in terms of discourse, thematic structure seems to be the crucial link here also.

Assembly must construct linguistic representations. Checking is required to see whether these representations are coherent. The only plausible way for checking to be carried out then is in terms of interpretation. Thus central to both checking and assembly is the process of interpretation and the main aim of psycholinguistics should be to explain the processes of interpretation.

Thus in assembly, models that allow integration of sentences into conceptual or discourse models are to be preferred to assembly procedures that result only in syntactic trees. Whatever the relevant representation turns out to be, it should be able to handle such phenomena as co-reference. Syntactic theories would seem to have the upper hand here, but this is only because no real attempts have been made to explain within conceptual/thematic structure phenomena that have up to now been captured within syntactic theory.

So, for example, the results of experiments on gap-filling would indicate that a conceptual rather than a syntactic theory seems to be most apt as gap-filling seems to be based on a conceptual decision at the verb rather than a syntactic decision at the gap site.

With respect to checking we have seen that Frazier (1991) has proposed a checking device based on thematic information and the Tanenhaus model also provides a recovery procedure based on thematic roles. This finding now opens up the question of whether the checking procedure on discovering a misparse requires the structure to be computed from scratch, or whether the checking procedure itself can provide clues as to the correct structure.

4.8 Language Processing and Language Acquisition

Fodor (1998a) has pointed out that theories of language processing and language acquisition have been to a large part in ignorance of each other. This is especially surprising as models of parameter setting have made it very clear that in order to set parameters from input data, learners need to be able to parse the input data. Indeed, it is difficult to relate the psycholinguistic data discussed above to the discussion of language acquisition in chapter 3.

In this section I will discuss how Fodor attempts to link parsing to learning, and learning to parsing (Fodor (1998a, 1998b)). I will then suggest how I see a way to link the syntactic discussion of Chapter 2 to the language acquisition discussion of Chapter 3, and all this in turn to the psycholinguistic discussion of this chapter. In doing this, I will make some suggestions as to how we would expect these links to shake out in testable psycholinguistic terms. I will explore these suggestions both by referring to previously intransigent psycholinguistic data, as well as by testing experimentally the kinds of predictions I make.

Fodor discusses data from Cuetos and Mitchell (1988), where it was discovered that adjuncts in English tend to attach low, whereas adjuncts in Spanish tend to attach high. With regard to this data, Gibson *et al.* (1996) suggested that there was potentially a parameter for the height of attachment and Mitchell and Cuetos (1991) suggested that the preference could be set as a response to the frequency of a particular construction in the relevant language.

Fodor argues that we have a deeper explanatory theory if what the parser does is not a product of learning. Her evidence for this is that wherever a parser has a choice of construction, there is a systematic preference for one of the other, as was clear from the previous discussion in this chapter. Secondly, the parser does not invariably choose the most frequent construction⁵. Fodor chooses to remain with a purely structural theory (without thematic content) and argues that all other things being equal, the parser prefers the simplest structure.

For Minimal Attachment it thus makes sense that less structure is preferred to more, but for this to be implemented in learning the parser needs to be able to tell what *less* structure is. She proposes that there is a 'race' mechanism in which comparative evaluation of alternative potential attachments takes place. The first attachment that is computed (presumed to be the simplest) wins. Fodor claims that this can occur at no extra cost. This seems an unusual claim, as it clearly seems to be the case, as she herself contends, that one structure is preferred initially, and the evidence overall clearly seems to indicate that only one structure is followed at a time, and if this structure proves to be incorrect, recovery comes at some cost. Thus to argue that all potential attachments being computed at the same time with one winning out is equal to a preferred initial analysis does seem to be having her cake and eating it.

Fodor then turns this around and looks at what the parser has to offer the learning system. Fodor argues that the Human Sentence Processing Mechanism is innate and that as language learning implicates parsing routines, the learner's encounters with language are filtered through the parsing mechanism. There are, however, a number of problems for the setting of parameters by the parser. First, the setting of a parameter is often reliant on facts about a sentence that are not apparent from the surface form. Second, many sentences are parametrically ambiguous, and third as we saw in section 4.4, the interaction of parameters is often very complex.

⁵However, the discussion of Taraban and McClelland (1988) in the next chapter adds even more to this discussion.

Sentences need to be parsed before parameters can be set, but a parser needs a grammar inside it to achieve this. Any sentence a learner can parse successfully is already licensed and thus the learner can learn nothing new from such a sentence. Thus, in order to learn, an unlicensed sentence must be presented. One option for the learner would then be to parse as best it can and then patch up a solution for the remainder of the sentence. This was what the Triggering Learning Algorithm (Gibson and Wexler (1994)) was doing (see discussion in Section 3.3). Thus the child would guess a parameter value, and test whenever an unlicensed sentence occurs. This is not only extremely inefficient, it would also mean that some children would be operating with a language significantly different to the one they are exposed to (for a period of time at least).

Fodor sees the parser helping out in such an impasse by giving clues as to the way the input could be successfully parsed. Specifically, she proposes a set of innate tree-lets, representing parameters. These tree-lets are pieces of a syntactic tree. Tree-lets can be seen as the correlate of the parsing packages Fodor proposed in the companion paper discussed previously in this section. A learner using these tree-lets is termed a *Structural Triggers Learner* (STL) by Fodor. At a point of ambiguity the learner will have access to full set of tree-lets. But parallel testing is unlikely as we have seen that the consensus is that parsers operate in a serial manner.

If the learning system could be told whether or not there is a competing analysis, it would learn parameter values from the unambiguous cases and set aside the ambiguous sentences. This is what Inoue and Fodor (1995) termed *flagged serial parsing*. Thus a learner would engage in no more learning past the ambiguity point. This would be a problem, of course, in languages such as Japanese where most of the ambiguity is in the initial part of the sentence. There are other potential problems with such an approach; some parameters may only occur after points of ambiguity, and ambiguity may be underestimated and parameters may be set on the basis of ambiguous data, where the ambiguity was not detected by the parser. Crucially at the onset of learning, almost all sentences will be ambiguous.

Fodor suggests three possibilities by which such problems could be avoided. First, with *robust parsing* the parser could be allowed to continue past an ambiguity point to see how much of the remaining sentence could be successfully parsed. At ambiguities the alternative parses could be evaluated and preference ratings posted on the flags. It is difficult to know how such preference ratings should be calculated, and how much decision-making the child would have to put 'on hold'.

The second option could be *parametric defaults*. Defaults empower learners by allowing learning to get underway. Ordering of parameters could also be an option, but the discussion in Section 4.4 shows that orderings of principles and parameters is extremely difficult to ascertain. Both of these options may appear stipulative. Fodor argues that they fall out naturally from the notion of learning by parsing as the parser is seen as a least effort device. The default value is whichever value is the quickest for the parser to access, that is, the quickest and simplest in the parsing 'race'. Again this seems to suggest latent parallelism in the parsing mechanism, which goes against the weight of the evidence.

The third option, particularly interesting in light of the discussion on the role of functional categories in language acquisition of the previous chapter, is the idea that some parameters may be subordinate to others. In other words, there is a simpler, set of parameters that need to

be set initially (as well as a simpler set of functional categories) and that additional parameters can only be set when the initial parameters have been set (and/or further functional categories emerge). The younger child would thus have less ambiguity to contend with. This would be very much a maturational theory. The weight of the evidence of Chapter 3 would indeed favour a maturational acquisition theory, with thematic structure forming the all-important initial layer.

4.9 Conclusions

We have seen that thematic roles can explain both initial parsing preferences in sentence processing, and that they are centrally implicated in the construction of a discourse model, and that within this discourse model certain aspects (the consequences) of the event structure are more salient than others.

We certainly need some way of connecting the thematic structure to the arguments of a verb so that Pritchett's parsing principles can be maintained. His relational, plainly syntactic thematic roles cannot be individuated from syntactic arguments in theoretical terms.

We would also like a framework which ensures representation of event structure in a way that allows the psychological saliency of the consequences of an event to be easily derived from the representation. This is an important link to other aspects of the psychology of language that are often ignored. By concentrating on the computational aspects of the processing of language, it is often forgotten that the processing theories must be consistent with the processes and stages involved in the acquisition of language by children. And in addition to this concept, the evolution of language, in the ecological perspective of human beings, will also have shaped certain saliencies that are relevant to processing. These issues will be discussed in the concluding chapter.

In the next chapter I will discuss phenomena from a number of past experiments and evidence from a number of new experiments designed in this thesis that indicate that a thematic-conceptual representation can best capture a large range of psycholinguistic phenomena.

Chapter 5

Psycholinguistics and Thematic-Conceptual Structure

In this chapter I will address some further psycholinguistic issues that do not seem to have been adequately explained in the literature, or to the extent that explanations exist they are erroneous, in my opinion. The topics I will cover here are noun/verb attachment ambiguities, specifically after *with*, and what has come to be known as syntactic priming. All these can be successfully explained by using thematic-conceptual structure.

5.1 Attachment Ambiguities with *with*

Tree-structure based approaches to sentence processing have until recently been arguably the most influential of all psycholinguistic theories. The best known principles in the tree-based framework are Minimal Attachment (MA) and Late Closure (LC), which are identified with the work of Lyn Frazier (See Section 4.3) and her colleagues. This approach has been claimed to have had a number of successes in explanatory terms for a range of ambiguities as can be seen below.

(5.1) *NP/S Ambiguities*

As soon as he had phoned his wife started to prepare for the journey.

As soon as he had phoned his wife she started to prepare for the journey.

(5.2) *Modifier Ambiguities in NP V NP PP Sentences*

The girl hit the boy with a stick.

(5.3) *Reduced Relatives*

The horse raced past the barn fell.

(5.4) *Choice of NP to attach modifier to*

The actor saw the sister of the man with a limp.

However, I argued in Chapter 4 that tree-structural approaches are necessarily sensitive to whatever changes the theory of tree-structures itself undergoes. Many of the structures

assumed in the early work on Minimal Attachment and Late Closure are no longer standard. In addition, in recent years cross-linguistic evidence has started to cast doubt on the validity of claims based purely on tree-structure. This evidence came in the work of Cueto and Mitchell (1988), who compared the attachment preferences in Spanish and English for the types of sentences that had previously been explained in MA/LC terms. They found that in the equivalent English and Spanish sentences (5.5)-(5.6) there was a difference in attachment preference from one language to another.

(5.5) El periodista entrevistó a la hija del coronel que tuvo el accidente.

(5.6) The journalist interviewed the daughter of the colonel who had had the accident.

In Spanish, there is a preference to assign to the first NP, the opposite strategy to Late Closure. In English, Late Closure effects were replicated and the preferred attachment was to the second NP. There are currently two competing explanations for this crosslinguistic difference. One relies on linguistic exposure and the other on memory-based heuristics.

Psycholinguistic parameters were first proposed by Cueto and Mitchell (1988). They propose that the attachment preferences shown by speakers of a particular language are based on the frequency of the preferred attachment to which the speaker of that language has been exposed. Thus it is assumed that attachment to the first NP is most frequent in Spanish and attachment to the second NP is most frequent in English. There is some counter-evidence to this claim, however. In English it would be expected that in NP VP NP PP ambiguities such as (5.7), the verb attachment would be most frequent in the language as this has been claimed to be the preferred attachment in English (we will see that the facts of this case are a good deal more complicated).

(5.7) The man hit the girl with a stick.

However, in one of the very few analyses of the frequencies of contrasting attachments in a large corpus Hindle and Rooth (1993) have shown that it is in fact the noun attachment that is most common in English.

To circumvent this problem Gibson *et al.* (1996) have proposed a theory centred around memory-based parsing heuristics. They provide two principles which the parser uses to give the correct parsing preferences for the two languages Cueto and Mitchell (1988) consider. These principles are

1. *The Property of Recency Preference*

The cost of attaching an item increases as the distance from the attachee increases. This increase in cost is not incremental. Rather it is defined as an increasing exponential decay function.

2. *The Property of Predicate Proximity*

There is a processing cost whenever an element is attached to a site which is further away from the predicate than another position to which the same item could have been attached which is closer to the predicate.

It is then assumed that the Recency Preference dominates in English and that the Predicate Proximity Preference dominates in Spanish. In sum, the memory-based analysis of the crosslinguistic differences seems at this point to be the most successful. But as Gibson et al. (1994) acknowledge, their approach does not explain one of the best-known parsing preferences which we have already touched on: the NP VP NP PP ambiguity. They concede that some form of content-based or plausibility-based approach will probably be necessary.

However, the parsing preferences for this ambiguity are not as clear-cut as was first claimed by Rayner *et al.* (1983). And it is because of the complexity of the data for English alone, that a frequency based account of both Spanish and English data is thrown into doubt.

Taraban and McClelland (1988) have shown that in different contexts the parsing preferences change. This gives further support for a content-based explanation of these parsing preferences. If this linguistic content could be linked in some sensible way to other aspects of human cognition, this would of course be even more interesting from a theoretical standpoint.

I will concentrate on a subset of the sentences that Taraban and McClelland discuss. This subset comprises what is probably the stereotypical NP VP NP PP ambiguity, namely *with*-phrases. The attachment preferences found were the following: (the preferred attachment is the first of the pair in each case)

Verb Attachment Preferred to Attachment to the Noun

- (5.8) a. The spy saw the cop with binoculars.
b. The spy saw the cop with the revolver.
- (5.9) a. The little girl tried to cut the apple with plastic knives.
b. The little girl tried to cut the apple with plastic coating.
- (5.10) a. The landlord painted the walls with a brush.
b. The landlord painted the walls with cracks.
- (5.11) a. Jim played the records with John's needle.
b. Jim played the records with deep scratches.
- (5.12) a. The kid hit the girl with a whip.
b. The kid hit the girl with a wart.
- (5.13) a. The doctor examined the patient with a stethoscope.
b. The doctor examined the patient with a toothache.

NP-Attachment Preferred to Attachment to the Verb

- (5.14) a. The couple admired the house with a garden.
b. The couple admired the house with a friend.
- (5.15) a. The woman married the man with money.
b. The woman married the man with delight

- (5.16) a. The doctor cured the woman with tuberculosis.
- b. The doctor cured the woman with penicillin.
- (5.17) a. The hospital admitted the patient with cancer.
- b. The hospital admitted the patient with urgency.
- (5.18) a. John ordered a pizza with pepperoni.
- b. John ordered a pizza with enthusiasm.

In summary, these results show that for the verbs *see*, *cut*, *paint*, *play*, *hit* and *examine* there is a preference for an instrumental modifier of the verb over a modifier of the noun. And for the verbs *admire*, *marry*, *cure*, *admit* and *order* there is a preference for a noun modifier over either an adverbial modifier of the verb or a comitative modifier of the verb. This difference would then seem to be linked in some way to the semantics of the verb in question.

It is not only the memory-based approach that has trouble explaining the adjunct attachment preferences. Any approach based on verbal subcategorization will, by definition, not have anything to say about adjuncts. These data require an explanation based purely on the semantics of the verb in isolation from their subcategorization properties.

We need then to find some content-based explanation of these results. We want somehow to constrain the theory of content that is relevant to psycholinguistics to as small a subset of all possible meanings as possible. There are a number of plausible strategies that could be applied to this problem.

One such strategy is the one I have taken in this thesis: to look at the range of semantic features that are reflected in the languages of the world. An example of such can be found in Talmy (1985). Alternatively we could try to link the sort of features that influence parsing to those that seem fundamental cognitively - a plausible candidate is the human conception of force dynamics, or cause and effect. It comes as no surprise that there is an overlap between these two approaches. Force-dynamics and cause and effect are grammaticalized in most (if not all) of the world's languages.

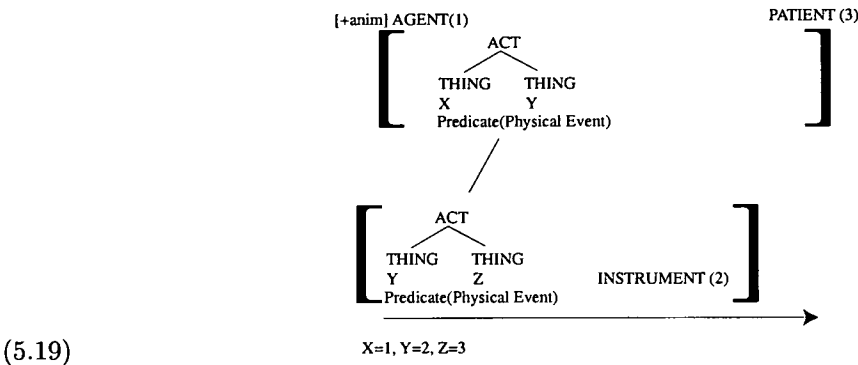
A first approximation to the semantic features involved in these sentences can be found in a subset of the conceptual relationships that are encoded in thematic-conceptual structure: the notion of contact. The set of verbs that result in verbal attachment preference mostly involve some notion of contact. The possible exceptions being *examine* and *see*. If we examine the notion of contact itself we can get a more satisfactory and inclusive feature to explain these differences.

In a contact event an initiator comes into contact with an endpoint. This contact can be mediated when the initiator comes into contact with an item that then in turn comes into contact with the endpoint. So, in effect, what we have in English is an event bounded by a subject and an object, where it is possible for an intermediary item to mediate this event. This mediation is only possible for stereotypical causation-type events. We need then some type of representation for stereotypical causation events which allows this mediation.

In English the initiator of the causal chain is usually associated with the subject and the endpoint of the causal chain is associated with the object. Thus subject and object delimit

the verbal segment of the causal chain.

Using this framework I propose that verbs which license an instrumental adjunct argument will have lexical entries such as (5.19) and verbs that occur only with adverbial adjuncts or a comitative argument will not.



The most important factor in the licensing of an instrumental reading is the simple causation event relating the initiator and the endpoint. This causation event can be mediated. There are further semantic restrictions on the causation event which may well differ from language to language. We have already seen that contact seems to be among these for English. The instrument in English may also not be volitional. The verbs *examine* and *see* seem not to be captured by the restriction of contact - unless of course vision is seen in some way as a metaphorical contact (cf. eye contact, catch sight of). This suggestion would seem to be particularly plausible in the light of the strength of the visual metaphor in the grammaticalization of language (Recall the work of Sweetser in the grammaticalization section of Chapter 3). Alternatively, it may be that vision itself is another of the restrictions.

On reading a verb associated with a lexical entry like that in (5.19), this verb being then followed by *with*, the sentence processor will have the expectation that an instrumental argument will follow the *with* and continuations reflecting this expectation will be processed more quickly than those which violate it. The verbs that do not have a lexical entry similar to (5.19) will never license an instrumental argument, and a noun phrase modifier is the resulting expected continuation.

In all cases there still exists the possibility that the plausibility of the item as either an instrument or as a noun modifier can have an additional effect on the parsing time.

The strength of this content based approach to these parsing preferences comes in the link I have established between the thematic-conceptual structure I have proposed and the properties and relative saliencies of human cognition both in acquisition and as I argue in Chapter 6, in language evolution.

I have proposed above a number of semantic features that seem to have a bearing on psycholinguistic phenomena. It seems that these features are particularly salient to the individual. Evidence for this can be found in experiments performed with very young children. There is evidence that children of a very young age exhibit knowledge of force dynamics as we saw in Chapter 3 in the experiments of Baillargeon (1987); Spelke (1991); Leslie (1984).

Given the argument advanced here, it would seem hardly surprising that human language mirrors the way we perceive the world and that the representations involved also include

elements of human cognition. If language has evolved from other aspects of human cognition, it is hardly surprising that the language learning system, as well as the sentence processor, employs these aspects of cognition in the learning process (See Chapter 6 for a speculative discussion on how this development may have occurred). In addition, if Siskind's hypotheses about language learning prove to be correct it seems that the sentence processor may well use the same principles as the language acquisition device.

The explanatory failings of both tree-based and memory-based approaches with respect to NP-VP-NP-PP ambiguities leads us to accept that some notion of content is essential to psycholinguistic theory.

5.2 Sentence-Level Priming is Not Dependent on Syntactic Structure

Priming is a well-known effect at the lexical level. The phenomenon of word priming has shed a good deal of light on the issues of representation and process in word recognition. Indeed, word-level priming phenomena have been one of the most influential and productive areas in psychology. The basic explanatory principle behind priming is that if the processing of one stimulus affects the processing of another stimulus, then the two are assumed to be related along a dimension relevant to the cognitive system. The phenomenon of sentence priming promises to do the same at a higher level of language processing. Sentence priming is greatly under-researched and is potentially an area of ground-breaking research of even greater importance than lexical priming. Branigan *et al.* (1995b) argue that priming provides a means for tapping in to the underlying commonalities of representation that are basic across language processes underlying both perception and production. Branigan *et al.* (1995a) have explained the relevance of sentence priming to linguistic theories in the following way. Processing theories have been content to assume that theoretical linguistics describes mental representations and that theories of parsing and comprehension draw on such theoretical linguistic constructs in their formulation. Since priming assumes that the primed structure and the prime must be in the same mental category, priming studies are now in the position to suggest the knowledge of language that cognitivist theories seek to describe. Thus priming studies allow for theories of comprehension and production to suggest the content of theoretical linguistic constructs, rather than relying on the output of theoretical linguistic theories to give structures upon which theories of comprehension can be based.

Sentence priming in the sense used here is not to be confused with the related phenomenon on corpus based priming. Schenkein (1980) and Tannen (1984, 1989) found a general tendency for repetition at many levels including the lexical and discourse level and Weiner and Labov (1983) found a tendency for syntactic forms to be repeated in interview situations. Sentence priming in the sense used here is said to occur when the process of reading or producing one sentence results in a facilitation of either the reading or production of another sentence sharing some similarity of form. This phenomenon is particularly exciting as it potentially provides one of the most direct methods available for tapping into the representations that lie behind language use. Also, as I argue below, this is not only at the level of syntactic structure, but

also at the level of thematic-conceptual structure and perhaps at even higher levels of linguistic structure, such as focus. Sentence priming may yield a major step forward in understanding not only the syntactic and conceptual representations behind language, but also the processes involved in their interaction in language use.

The early papers in this field argued for a priming effect based simply on the syntactic structure of the sentence, reflecting the contemporary hegemony of syntax in the psycholinguistic literature. However, there is recent evidence which strongly suggests an alternative interpretation in terms of conceptual structure.

Branigan (1995) concludes from a number of experiments in syntactic priming in comprehension, production and from comprehension to production that the basis of priming is syntactic structure representation and not a procedure that is common to both comprehension and production. She also claims that it is only constituent structure that is relevant to priming and specifically that thematic roles have no influence. Branigan *et al.* (1995a) add further that as syntactic priming takes place between sentences containing different verbs, this indicates that syntactic information is specified over classes of verbs and is not stored separately as part of each verb's lexical entry.

I will show that there is an alternative explanation of the results (See also Evans (1993)), where there is direct implication of thematic roles, or more accurately the extended representational theory of thematic-conceptual structure. In addition, this thematic representation is inextricably linked to a *procedure* which is common to both comprehension and production.

There are then two issues to be addressed. First, can it be shown that thematic-conceptual structure of some kind is involved in the priming of structures where it has been assumed that thematic roles have no influence of any kind. Once this has been shown to be the case it also follows that syntactic information is shared between verbs and is not stored separately in each verb's lexical entry. Commonalities in the syntactic behaviour of verbs based on their underlying thematic-conceptual structure is exactly what we would expect.

The second issue is whether there is a process of some kind which could be relevant to both comprehension and production, but will nonetheless predict the differential results that have been discovered for priming in production versus priming in comprehension. I will argue that this procedure is the linking of constituent structure (among other syntactic elements) to the conceptual structure specifically relevant to language and vice versa. The differential effects found in experiments on comprehension and production will be explained by considering the different strategies that are inherent in being a speaker versus a listener.

I will then describe in more detail how the thematic-conceptual representation and the linking procedure interact to predict the results of Branigan (1995) and Bock and her colleagues (Bock (1986, 1989); Bock and Loebell (1990); Bock *et al.* (1992)). Further predictions will be made to await empirical study. Suggestions as to how this could be achieved are to be found in section 5.4. Some of the experimental programme proposed in section 5.4 has since been carried out by my colleagues and the results are presented at the end of this chapter.

I will also provide further theoretical justification for the *Conceptual Structure and Linking Hypothesis*. We have seen that the justification has been drawn from a wide range of fields as diverse as language acquisition, psycholinguistics, and linguistic theory, all of which conspire to bolster the hypothesis. I will also include more speculative evidence from the fields of

primatology and evolutionary studies in Chapter 6.

5.2.1 Thematic Structure in Syntactic Priming

It has been claimed (Bock (1986, 1989); Bock and Loebell (1990); Branigan (1995)) that syntactic priming occurs at the level of gross constituent structure. The constituent structure is gross as the priming effect occurs between structures that are slightly different in terms of modifiers added to the major constituents having no effect on priming. It is also claimed in these works that thematic roles have no additional effect on the priming mechanism. In this section I will show that, for a number of reasons, it can be claimed that thematic roles or thematic-conceptual structure are involved in priming.

These reasons are that first, in the cases where Bock claims no effect of thematic roles, there are methodological problems which militate against this claim. Second, in the cases where it is claimed that two different structures have the same thematic relations encoded in them, I will show that what on the surface are similar thematic roles are actually quite different in a more detailed, extended version of representation, i.e. at the level of thematic-conceptual structure.

The data specifically of relevance in the argument against the influence of thematic roles in priming is found in Bock (1986); Experiments 2 and 3, Bock (1989), Bock and Loebell (1990), Branigan (1995); Chapter 6: Experiments 1 and 2.

There are two types of construction that are used in these experiments, both of which are alternations¹. The first of these is the passive alternation (5.20) and the second is the Dative alternation with either a Goal second object and the preposition *to* (5.21), or a Beneficiary second and the preposition *for* (5.22).

- (5.20) a. The man kicked the dog.
b. The dog was kicked by the man.
- (5.21) a. The doctor gave the patient a prescription.
b. The doctor gave the prescription to the patient.
- (5.22) a. The cook made the cake for the sailor.
b. The cook made the sailor the cake.

Bock and Branigan use different experimental methods to achieve similar results and make the same explanatory assumptions. Bock uses repetition of a spoken sentence as a prime followed by picture descriptions of the target. On the other hand, Branigan uses a relatively unambiguous written sentence completion as a prime (or a pair of such primes) followed by an ambiguous sentence completion as the target. In all of these cases, it is assumed that producing a certain constituent structure will prime the later production of an analogous structure. For example, repeating or completing a double-object dative will favour the production of a double-object dative when describing a picture or completing an ambiguous sentence compared with producing a prepositional dative structure after the same prime. Alternatively, prepositional

¹One may be a purely lexical alternation and the other a transformational alternation. This issue bears more directly on procedures and will therefore be discussed in the next section.

dative primes should facilitate production of a further prepositional dative compared with producing a double object prime.

Bock 1986 In this paper Bock used sentences such as those in (5.20) and (5.21) to test the syntactic priming hypothesis. These priming sentences were embedded in a series of sentences and pictures in the guise of a sentence and picture memory test. In this experiment, the picture to be described could have been equally well described by either of the alternate forms in (5.20) for depictions of an active/passive, or equally well by either of the alternate forms in (5.21) for dative depictions.

The hypothesis was that the form of the prime sentence would influence the form of the sentence the subject then used to describe a picture. The priming sentence and the relevant picture description did not have open-class words in common, so the priming effect could not be attributed to repetition effects of the open-class words.

However, it was still possible that the closed-class words could be implicated in a repetition effect. Bock (1989) attempted to discount such an effect.

Bock 1989 The results of Bock (1986) could have been explained by a mechanism or representation other than constituent structure. It may have been the case that the priming effect was caused by sentential frames associated with certain prepositions being primed from sentence to sentence. *To* was the preposition in the dative sentences and *by* was the preposition in the active/passive sentences.

Bock (1989) constructed sentences which compared the effect of *to/for*; to see whether changing the preposition in the prime sentence influenced the priming effect in any way.

(5.23) a. The secretary is baking a cake for her boss.

b. The secretary is taking a cake to her boss.

(5.24) The secretary is taking her boss a cake.

She found that both of the prepositions was equally good at priming the subsequent prepositional form of the dative. Bock felt justified in claiming that sentence frames associated with certain prepositions could not be responsible for the priming effect. We shall see, however, that certain methodological problems indicate that this conclusion has been reached somewhat prematurely.

Bock and Loebell 1990 Here the emphasis switched away from sentential frames and towards thematic roles. Bock *et al* tested whether sentences with the same structure but different thematic roles (5.25) were worse primes than sentences with the same thematic roles and the same structure (5.26) for picture descriptions of the form in (5.27).

(5.25) a. The wealthy widow drove her Mercedes to the Church.

b. The 747 was landing by the control tower.

(5.26) a. The wealthy widow gave her Mercedes to the Church.

b. The 747 was landed by the control tower.

- (5.27) a. The girl handed her paintbrush to the boy..
 b. The boy was stung by the bee.

They found that the thematic dimension, as they defined it, had no effect on the strength of the priming effect. However, once again we shall see that methodological problems make this conclusion premature.

Branigan 1995 uses a method where subjects complete relatively unambiguous prime sentences such as:

- (5.28) a. The girl gave the book...
 b. The girl gave the boy...

It was hypothesized that the form of the prime sentence the subject produced would influence the target sentence completion which was ambiguous between the two forms.

- (5.29) The politician showed...

Branigan assumed that the two forms differed only with respect to their constituent structure (although we will have cause to doubt that this is actually the case). Branigan replicated Bock's findings for this construction in this alternative methodology.

In a second experiment Branigan examines whether the priming structure has to be identical to the target or whether it can be substantially different at some lower level, provided that there is a local substructure shared by both prime and target.

- (5.30) a. The teacher gave the book that frightened the child...
 b. The barrister showed....
- (5.31) a. The teacher gave the youth that frightened the child...
 b. The barrister showed....
- (5.32) a. The teacher frightened the child that gave the book...
 b. The barrister showed....
- (5.33) a. The teacher frightened the youth that gave the child...
 b. The barrister showed....

The results were starkly different to what we have seen up to now. Prepositional sentences were produced more often than the double-object in all four conditions. The syntactic structure of the previous sentence had no discernible impact on the proportion of each structure produced for the target sentence. Thus priming does not seem to be based on strictly local structures.

Branigan puts this result down to insufficient activation of the relevant rule. That is, the loss of effect is due to processing factors and not just the representation. In the priming sentences that have no effect there are two VPs and in the sentences where priming is successful only one VP is present. If priming is based on the VP, Branigan argues, there could be

competition between the two VPs in the priming sentence to act as the relevant prime for the target sentence.

We note here that Branigan has stopped talking purely about constituent structure as being behind the priming effect and has started to refer to *rules*. I suggest that these rules are the procedures that are common to both production and comprehension. I shall claim that these procedures will be available to explain not only the failure of priming in this case, but also all of the other findings reported in this section, but in a very different way: it will be proposed that thematic-conceptual structure is indeed centrally implicated in syntactic priming effects.

There are a number of methodological aspects associated with Bock's experiments that cast some doubt on her conclusions.

As we have seen above, her 1989 paper had the aim of showing that the priming effect was not specific to the closed-class lexical item *to/for* that occurred in the prepositional form of the dative construction.

The problem can be put down to Bock's assumption that *to* and *for* will result in different sentence frames being constructed and that if no difference in the priming effect could be found these frames cannot be implicated in the priming procedure.

The problem is that these different prepositions could well be associated with different sentence frames but the methodology of the experiment did not allow these different frames to express themselves in priming.

All the pictures that the subjects had to describe depicted a scene that could only have been described with a prepositional dative with *to* or its associated double object alternation. If the pictures had allowed a description with a *for*-dative, the experiment would have been more interesting. However, it is unlikely that this will be possible with pictures, as producing a *for*-dative requires the inference that the recipient of the action will in some way be pleased by the action described. Thus a mental state has to be inferred. This is not easily depicted.

It may have been the case, then, that the subjects made the closest description possible with respect to the sentential/conceptual frame. Alternatively the conceptual frames associated with *to* and *for* may have been similar enough for there to be no difference expected (as Bock seems to concede p.183). In addition, it seems unwise to base a conclusion on a lack of a result.

This impasse could be solved experimentally if we design an experiment using the written sentence completion paradigm described by Branigan (1995). This paradigm allows the subject to complete the sentence in any way she pleases, she is not constrained by the conceptual relations present in a picture. Thus, if there is a difference in the conceptual frames of a *to*-Dative and a *for*-Dative, we would expect a difference in their suitability as a prime for the complementary structure. So a *to*-Dative should be better than a *for*-Dative prime for another *to*-Dative. We would need material such as the following.

- (5.34) a. **to-Dative Prime** The lecturer gave the book...
 b. **to-Dative Prime** The lecturer loaned the car...
 c. **for-Dative Target** The student baked...

- (5.35) a. **to-Dative Prime** The lecturer gave the book...
- b. **to-Dative Prime** The lecturer loaned the car...
- c. **to-Dative Target** The student brought...
- (5.36) a. **for-Dative Prime** The lecturer baked the cake...
- b. **for-Dative Prime** The lecturer cut the carrot...
- c. **to-Dative Target** The student brought...
- (5.37) a. **for-Dative Prime** The lecturer baked the cake...
- b. **for-Dative Prime** The lecturer cut the carrot...
- c. **for-Dative Target** The student made...

If there is a different strength in the priming effect between contexts with supporting sentential frames (5.35) and (5.37) than in non-supporting sentential frames (5.34) and (5.36) we will have found evidence contra Bock that sentential frames and their underlying conceptual basis do have an effect on priming. If we do not find any differences we have further empirical justification for Bock's position. I leave this as an open topic for further research. It is important for the aims of this thesis to note that the Branigan interpretation is open to doubt.

Bock and Loebell (1990) suffers from the same problems with respect to the pictures used. The pictures allow only one description and they do not allow for any differential effects to appear. The experiment attempted to show that thematic roles had no influence on priming by comparing prepositional dative primes (*The widow drove the old Mercedes to the church*) to prepositional locative primes (*The widow gave the old Mercedes to the church*). Both these sentences are assumed to have the same constituent structure and different thematic roles. A second experiment compared *by*-Agents in passives (*The boy was kicked by the dog*) and *by*-Locatives with a similar constituent structure (*The boy was eating by the lamppost*).

This again could be better decided with a sentence completion experiment with the following materials (Once more I leave this as a topic for future research).

Datives versus Locatives

- (5.38) a. **Prepositional Dative Prime** The widow gave the old Mercedes...
- b. **Prepositional Dative Prime** The lecturer gave the new book...
- c. **Prepositional Locative Target** The widow drove...
- (5.39) a. **Prepositional Dative Prime** The widow gave the old Mercedes...
- b. **Prepositional Dative Prime** The lecturer gave the new book...
- c. **Prepositional Dative Target** The student loaned the car...
- (5.40) a. **Prepositional Locative Prime** The widow drove the old Mercedes...
- b. **Prepositional Locative Prime** The lecturer rode the new cycle...
- c. **Prepositional Dative Target** The pilot gave...

- (5.41) a. **Prepositional Locative Prime** The widow drove the old Mercedes...
- b. **Prepositional Locative Prime** The lecturer rode the new cycle...
- c. **Prepositional Locative Target** The pilot moved...

Agents versus Locatives

- (5.42) a. **Agent Prime** The construction worker was hit...
- b. **Agent Prime** The minister was cut...
- c. **Prepositional Locative Target** The foreigner was loitering...
- (5.43) a. **Agent Prime** The construction worker was hit...
- b. **Agent Prime** The minister was cut...
- c. **Agent Target** The foreigner was confused...
- (5.44) a. **Prepositional Locative Prime** The foreigner was loitering...
- b. **Prepositional Locative Prime** The minister was praying
- c. **Agent Target** The 747 was alerted...
- (5.45) a. **Prepositional Locative Prime** The foreigner was loitering...
- b. **Prepositional Locative Prime** The minister was praying...
- c. **Prepositional Locative Target** The 747 was landing...

Thematically Similar Structures in Parsing?

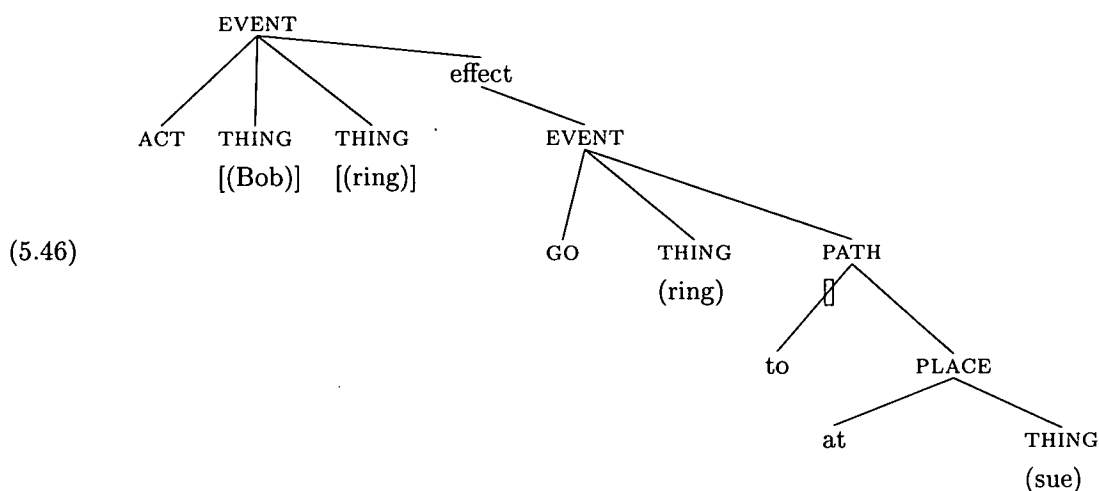
We have seen above that some residual conceptual effects were not easily removed from the priming of the passive/active sentences of Bock *et al.* (1992). It is probably for this reason that Branigan removed the passive/active alternation from her sentence completion experiments. Instead she concentrates on the Double Object/PP Construction which she, following Bock, claims are:

Equivalent in their focus, semantic content and register.

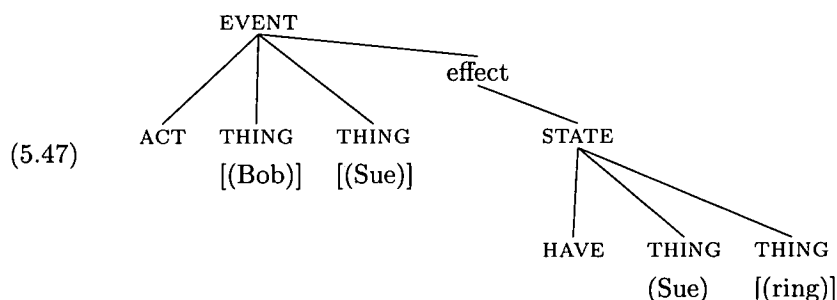
If it could be shown that these two structures do differ in one or more of these aspects, the conclusion that syntactic priming is based purely on constituent structure would become more doubtful and another explanation should be sought.

At the level of what have been termed thematic roles, the structures are indeed similar, but following the discussion of chapters 2 and 3, what are known as thematic roles are merely shorthand for frequently occurring structures within a theory of extended thematic-conceptual structure. And what on the surface seem to be the same are actually different at this deeper level. At a deeper level there are indeed differences between the double-object and the PP form. These differences have been shown by Pinker (1989) to have explanatory potential for the theory of acquisition of lexical alternations by young children. I will now concentrate on the differing conceptual relations that these representations reflect.

The thematic-conceptual core of the prepositional object form is:



and the thematic-conceptual core of the double object form is:



These two conceptual representations reflect two main differences in the thematic-conceptual relations expressed by the propositional and the double-object forms.

First, in the double-object form the transferred object is the patient, in the prepositional object form the recipient is the patient. This causes pragmatic differences in which argument is construed as 'affected' or acted on in the double-object form.

(5.48) What John did to Bill was give him a book.

(5.49) What John did to Bill was give a book to him.

and the entailments are different in terms of to what extent the recipient is affected.

(5.50) John taught English to the students. (unsuccessfully)

(5.51) John taught the students English. (successfully)

The second difference is that in the double-object form the change of possession is expressed not as an analogue of motion (as in the PP) but as a causation-state whereby the possessor has (HAVE) the object.

This is apparent in the way that the double-object form is not available with some verbs, while with other verbs the prepositional object form is not available. If these forms were conceptually identical then every verb which allows for a double-object should have a prepositional object form and vice versa.

- (5.52) a. PP: X causes Y to go to Z
 *They spared the punishment to the policeman.
- b. DO: X causes Z to have Y
 *She carried the mailbox the letter.

We have seen the empirical results of experiments into syntactic priming in production. The materials and methodology of the experiments were also presented. The basic claim in the literature was that it was only the constituent structure of the prime sentence that was relevant to priming. It has been specifically denied in the literature that either thematic roles or sentential frames associated with particular closed-class items are relevant to syntactic priming.

However, a number of problems were found with the materials and the assumptions that the experiments were grounded on. These problems cast doubt on the conclusions reached by Bock and her colleagues and Branigan in two different and complementary ways.

Where no thematic/conceptual difference was assumed to be present, it actually seems to be a clear conceptual difference and this difference predicts the structure that will be used.

On the other hand, where the experiments specifically tried to demonstrate that thematic roles or sentential frames had no relevance to priming, it turns out that the experimental method employed allowed no possible differential priming effect to emerge. A number of experiments were suggested that may further illuminate this issue.

Up to now, I have concentrated only on syntactic priming in production. This concentration was for expository purposes; the results of experiments on syntactic priming in comprehension are quite distinct from those for priming in production. Another reason for concentrating initially on the production experiments is that they were the initial experiments conducted on syntactic priming and thus most of the assumptions concerning syntactic priming in the comprehension literature have been based on the assumptions made in this earlier work.

In the next section I shall propose a procedure that will go some way to explaining the production data. I propose that this procedure is also relevant to syntactic priming in comprehension. I will then proceed to demonstrate that the different communicative strategies used by speakers and listeners acting together with this procedure make the correct predictions with respect to syntactic priming in comprehension, given that a number of other assumptions prove to be correct.

5.2.2 Procedures in Production and Comprehension

To capture the syntactic priming in production data we have seen up to now I propose the Thematic-Conceptual Structure and Linking Hypothesis.

- (5.53) *The Thematic-Conceptual Structure and Linking Hypothesis*(CSLH)
 Any given thematic-conceptual structure is related to its surface representation by a linking rule.

Initially I shall concentrate on the linking from conceptual structure to constituent structure as this is what has up to now been claimed to be relevant to syntactic priming.

We shall see that the CSLH is also relevant to syntactic priming in comprehension. The CSLH is a procedure that can be primed both for comprehension and production. This viewpoint rests on the assumption that the mental representation must be based on a set of knowledge, the units and structures, which provide the speaker with the ability to turn messages into utterances and which underlies the hearer's ability to perform the translation in the opposite direction.

Before I turn my attention to sentence priming in comprehension it remains to be shown how the CSLH can explain the data from production. Turning first to the sentences showing a dative alternation, we have seen that the double object form and the prepositional form are associated with different conceptual structures. The CSLH states that both of these conceptual structures can be potentially associated with a different surface constituent structure. A linking rule maps each conceptual structure to the relevant constituent structure.

Thus the form in all cases will be:

(5.54) Thematic-Conceptual Structure $\rightarrow$ Constituent Structure

The argument is thus that when priming takes place it is not the constituent structure that is primed but rather the linking rule that translates underlying thematic-conceptual structure into surface structure.

As the linking rules specify only the framework of the constituent structure with particular emphasis on arguments and not the details of the constituent structure we have a concise explanation of why it is that the constituent structure of the priming and target clause do not have to match exactly. The linking rules determine the ordering of the constituents and the prepositions necessary to encode the conceptual relations expressed by that structure. As long as this is in place, modifiers can be freely added to this structure without affecting the priming effect.

We also saw though that if two clauses were involved in the priming task, the priming effect disappeared. Branigan (1995) proposed a competition effect between the two clauses as an explanation for this. However, it seems unclear why the local constituent structure should be affected by the clausal status of the contents of that structure. On the other hand, as each clause will by its very clausal nature be associated with a different conceptual structure representing the clausal relations expressed by that clause, it is not at all surprising that the linking rules for each of these clauses will compete in such a way that no priming effect occurs.

In other words, it is much more likely that linking rules will be in competition as opposed to constituent structure.

In some ways this is similar to the approach taken by Bock *et al.* (1992). They identified two ways that the surface structure could be associated with the underlying relations. In one, the interpretation of linguistic relations is in terms of deep-structure subjects and objects and transformations are implicated; this is *mediated mapping*. In the other, syntactic functions are individuated with respect to the verb. There is thus a *direct mapping* to surface syntactic relations and transformations are not involved.

Bock *et al.* examined the passive construction and tested to see whether animacy had any influence on the priming effect. It seems that animacy is a good predictor of subjecthood and this is a challenge to the purely structural interpretations of the subject relation. Thus, in

the encoding of conceptual relations into syntactic structure, animacy may be playing a role. A further reason to suspect this is that in Bock's earlier experiments pictured events with an animate patient and an inanimate agent resulted in a bias towards passive sentences.

Bock *et al* tested the two hypotheses under the following assumptions: The mediated mapping hypothesis predicts that active and passive primes with animate underlying object arguments (the surface objects of actives and the surface subject of passives) should elicit more active targets (all of which had animate objects and inanimate subjects) than active and passive primes with inanimate underlying object arguments. Conversely, the direct mapping hypothesis predicts that active and passive primes with animate object arguments (the surface objects of both) should elicit more active targets than active and passive primes with inanimate object arguments.

Bock *et al* claim that their results show that in language production two procedures are in operation in the realization of syntactic form. These two procedures are:

- The binding of arguments to syntactic relations.
- Building constituent structure.

An alternative explanation is that there is a single procedure and that animacy is but one of the conceptual differences that can result in a preference for an active production or a passive production. Thus sentential focus may be represented as part of the thematic-conceptual structure. If it could be shown that two different structures with similar focusing could prime each other, this hypothesis would be supported. In any case the Bock *et al.* (1992) paper is a significant weakening of the position initially taken by Bock and still taken by Branigan. Bock is conceding that thematic-conceptual structure can have some effect but does not re-examine her earlier work for possible conceptual differences between the alternate structures she has used.

Applying the CSLH to Comprehension

In comprehension the listener/reader is trying to decode the thematic-conceptual structure incrementally and as quickly as possible. In effect the linking rules are being run in reverse. The aim is to derive the full conceptual structure that the sentence expresses as quickly as possible.

There is a theory of syntactic ambiguity resolution that takes a very similar stance. Pritchett (1988, 1992) proposes that the Human Sentence Processing Mechanism (HSPM) tries to assign thematic roles associated with the clause's verb as rapidly as possible in line with that verb's maximal theta-grid.

(5.55) *Theta-Attachment*

The theta criterion attempts to be satisfied at every point during processing given the maximal theta-grid.

If this strategy results in a re-analysis of the input sentence, a second principle accounts for the relative difficulty of reanalysis.

(5.56) *The On-Line Locality Constraint*

The target position (if any) assumed by a constituent must be governed or dominated by its source position (if any), otherwise attachment is impossible for the HSPM.

When this analysis is impossible a conscious garden-path sentence results.

This prediction of conscious garden-path sentences is particularly relevant when we look at the results of experiments on syntactic priming in comprehension. It is precisely the sentences that usually result in conscious garden-path effects that can be primed so that the target sentence no longer seems to result in as serious a garden-path, or possibly no garden-path. Thus it seems that the dispreferred analysis can be primed but not the preferred analysis that would be proposed by Theta-Attachment.

In this way, the dispreferred (b) readings of (5.57) and (5.58) could be primed in comprehension, but sentences similar to those in the production experiments (5.59) and (5.60), sentences including subject/object relatives (5.61), and sentences containing an attachment ambiguity, but not one that results in a conscious garden-path effect (5.62) did not result in any priming effect in comprehension experiments (Evans, 1993, Branigan, 1995).

- (5.57) a. While the woman was eating the soup the pudding went cold.
b. While the woman was eating the pudding went cold.
- (5.58) a. The caretaker reminds the porter that the sleepy boss annoyed the old architect.
b. The caretaker reminds the porter that the sleepy boss annoyed about the dress.
- (5.59) a. The motor-bike injured the walker.
b. The walker was injured by the motor-bike.
- (5.60) a. The hurdler gave the whistle to the judges.
b. The hurdler gave the judges the whistle.
- (5.61) a. The badminton player that assaulted the inspector jogs by the river.
b. The badminton player that the inspector assaulted jogs by the river.
- (5.62) a. The thief opened the safe with a key today.
b. The thief opened the safe with the lock today.

Branigan (1995) appears to concede that a *rule* of some type is being primed in the sentences where priming occurs. She suggests that priming happens due to the level of activation of this rule. The rule associated with the dispreferred structure is boosted after reading a similar dispreferred structure. What Branigan means by rule is unclear as she argues that the priming effect is still based on constituent structure and argues throughout her thesis that priming is based on representation and not on procedures.

However, she does suggest a number of options as to the content of such rules. First it could be that priming of a single rule that specifies a local tree e.g. $VP \rightarrow V NP$ versus $VP \rightarrow V$. In this view, the structures could differ to some extent, as long as this local structure is similar in both prime and target. However, such an account gives us no explanation as to

why one structure is preferred over the other. A similar account would be where the non-local structure has to be the same. This can be criticized for the same reason, why should one of the structures be preferred?

The third possibility is that the only relevant aspect of syntactic structure is the nodes on the right edge of the tree and their daughters.

In sum all three of these possibilities claim that rules (which are most easily identified as Phrase-Structure rules in Branigan's work) are implicated in priming. For some unexplained reason some of these PS-rules are dispreferred and can be primed. This does not explain the lack of effect in comprehension of the sentences that resulted in a priming effect in production as the passive form of the active/passives and the object relative form in the subject/object relatives (Pickering and Shillcock (1992)) are also dispreferred structures. Thus in production, although the form produced could be influenced by priming, this was only a relative effect; relative to a generally preferred production form.

However, if we assume that the reader is trying to derive the thematic-conceptual structure as rapidly as possible from the input we can arrive at a more plausible explanation. The processor uses such principles as Theta-Attachment as a strategy to derive a maximal thematic-conceptual representation. This explains the preferred attachments found for the sentences above at the same time as providing a mechanism for reverse linking. When this strategy fails and reanalysis has to occur, the reanalysis involves invoking a different procedure for reverse linking, or in effect a specific linking rule from the input to the conceptual structure is identified and can thus be primed (this linking rule having proved successful at integrating challenging input).

5.3 Future Research on Sentence-Level Priming

Theoretical Questions

It has been claimed (Bock (1986, 1989); Bock and Loebell (1990); Bock *et al.* (1992); Branigan (1995)) that sentence level priming is dependent on the representational level of syntactic structure. Above I have argued that an alternative explanation can be found at the level of thematic-conceptual structure. In experimental demonstrations of sentence-level priming conducted by Evans (1993) and Branigan (1995), three different factors could have been responsible for the priming effect at sentence level: First, it could result from syntactic processes. That is, they could be related to the sentence's surface representation or a rule implicated in generating a particular syntactic structure, as Branigan (1995) argues. Second, it could arise from the underlying conceptual structure, which in English is canonically realized by a certain syntactic form. For example, focus on the patient rather than the agent of an action is naturally expressed with use of a passive. Or, alternatively, it could be related to the use of a linking rule which is responsible for mapping a particular conceptual structure onto surface syntactic form.

An example of how these possibilities differ as proposed explanations can be seen if we consider one of the structures which seems to provide a robust sentence level priming effect: the Double-Object/to-Dative alternation.

- (5.63) a. Bob gave the ring to Sue.
b. Bob gave Sue the ring.

Branigan (1995, p. 135), following Bock, claims that these alternate forms are 'equivalent in their focus, semantic content and register', and concludes therefore that the locus of any priming effect must be at a syntactic level. If, however, it could be shown that these two structures differ in one or more of these aspects, the conclusion that the locus of sentence priming rests entirely with constituent structure becomes more doubtful and alternative explanations would also need to be considered.

At the level of what has been termed thematic roles, the structures are indeed similar, but following Jackendoff (1983, 1990) and Pinker (1989) it can be argued that what are known as thematic roles are merely shorthand for frequently occurring structures within a theory of extended conceptual structure.

Thus, in this account, the differing surface syntactic structures reflect a different underlying conceptual structure, and it is this which may form the basis of the priming effect. Another possibility is that it is not the underlying conceptual structure that is primed, but rather the process that links the conceptual structure to syntactic structure (and other expressive linguistic devices, such as intonation).

Under this account it is possible that a particular conceptual structure may be associated with more than one linking rule, or alternatively, a single linking rule to syntactic structure may be associated with more than one conceptual structure.

On the basis of current evidence, it appears that both the syntactic explanation and the conceptual explanation (with or without linking rules) are possible and it is very difficult to find empirical data capable of distinguishing the two in a language like English, where a difference in syntactic structure, arguably, goes hand in hand with a difference in conceptual structure.

I have set out a theory of thematic-conceptual structure and linking that can now be applied to the question of whether it is conceptual structure or syntactic structure that is implicated in sentence level priming. The elements of conceptual structure can be expressed in a number of ways, of which syntactic variation is just one.

What is needed, therefore, is a situation where conceptual structure is closely matched in two sentences where the syntactic structure is different.

A solution to the difficulty is to adopt a cross-linguistic methodology. If the same or a similar conceptual structure can be expressed by two different syntactic structures in one language and by only one syntactic structure in another language and, assuming that priming at sentence level is possible across languages, we would expect a different pattern of results if conceptual structure rather than syntactic structure is responsible for the priming effect. Recent evidence suggests this is the case.

In 1996 I suggested a number of experiments to tease apart these possibilities. My suggestions are given in below. In section 5.4 the results of the research following from these suggestions is given. The results thus far seem to show a much greater role for thematic-conceptual structure in sentence priming than has previously been the consensus.

Empirical Background

Heydel and Murray (1996) present work from cross-linguistic priming experiments that suggest it is conceptual structure, rather than syntactic structure, that is centrally implicated in sentence level priming. In their first experiment, bilingual (English/German) subjects were presented with a pair of target pictures; the first picture depicted an intransitive action (e.g. a man swimming) and the second picture depicted an event that could be described by either an active or a passive sentence (e.g. a shark attacking a man). The subject then read a German sentence which was either a transitive Subject-Verb-Object untropicalized active sentence, an Object-Verb-Subject topicalized active sentence, or a passive sentence. The subject was required to decide whether this sentence related to the events shown in the pictures (in the experimental items, it never did) and then to describe in ENGLISH what was happening in the pictures. The syntactic structure of the subject's description was then analyzed in terms of whether it was an active or a passive or some other sentence type.

This task was thus a cross-linguistic priming experiment. No translation was involved, as the picture to be described had no content in common with the sentence that was read (the judgment was included only to ensure processing of the prime sentence). The subject was primed with a German sentence and had to produce an English picture description. The results clearly showed cross-linguistic priming of actives by actives and passives by passives, but, crucially, German topicalizations primed English passives not the structurally similar actives.

Their second experiment employed a similar procedure but involved translation. Here again the subject saw the two target pictures, read a priming sentence (SVO-active, OVS-active, passive) in German, but this time the prime did describe the events in the pictures. The subject was required to produce a translation of this into English which corresponded to the events shown in the pictures. The results from this study again showed reliable priming, but even when the task involved translation, German topicalizations did not prime English actives.

It is striking that the results of both experiments show German topicalizations priming English passives. This is surprising as the topicalization sentence has a syntactic structure most closely resembling an English active rather than an English passive. Thus it appears very likely that the priming effect rests not with syntactic structure per se, but rather with some other (higher level) representation or process, such as conceptual structure.

However, it is unclear from these results whether it is the representational level of conceptual structure itself which is primed, or whether instead it is the procedure of linking the conceptual structure to surface level expressive linguistic devices. The following proposed experiments address this and other important issues necessary for establishing a clear understanding of this process.

The Experiments

The experiments fall into two main groups:

- Monolingual Experiments on Native Speakers of German

- Cross-linguistic Experiments

Each of these groups of experiments sheds light on the questions identified above. The experiments on German speakers examine whether the results of the preliminary experiments on cross-linguistic priming were critically related to the fact that the experiment was cross-linguistic or whether the effect also appears within a single language (as a conceptual structure interpretation would predict). The second set of experiments examine the levels at which conceptual structure can be primed and the strategies speakers adopt when the language they are switching to, or from, is either more restricted or more free in the way it can express a certain conceptual structure with its surface expressive linguistic devices.

The experiments share their methodology with Heydel and Murray (1996). The subjects describe pictures after having read a priming sentence with one of the forms active, passive or topicalization. The subject's productions are then analyzed as to their syntactic structure and, in experiment 5, as to their intonational form.

Experiment 1: German-German Conceptual Priming Subjects: Native German Speakers

It might be suggested that the initial results from the cross-linguistic studies could have been complicated by subjects adopting specialist processing strategies when using their second language. To test this, it is important to establish whether the basic findings can be replicated in a German monolingual study. If the priming is indeed based on conceptual structure then it should be the case that German passives (like c) will prime both passives and topicalizations (b) and, crucially, German topicalizations will prime both passives and topicalizations, but not actives (a). Also, actives should prime neither passives nor topicalizations.

- (5.64)
- a. Active: Ein Hai attackiert den Schwimmer.
 - b. Topicalization: Den Schwimmer attackiert ein Hai.
 - c. Passive: Der Schwimmer wird von einem Hai attackiert.

The first part of the study will establish the baseline for the natural frequency of these structures with a simple picture description task. The second phase, using different subjects, will involve an identical procedure, but with the additional insertion of a priming sentence in the same way as employed by Heydel and Murray. The results from this study will establish both whether conceptual structure effects are found monolingually and, if so, whether these replicate the pattern found in the cross-linguistic studies. They will therefore comment both on the question of whether the previous findings may have involved some specialist processing strategy' and whether the cross-linguistic procedure itself is an appropriate means for investigating effects of conceptual structure.

If the results show priming of the same structures, but no 'transfer' between passives and topicalizations, this would be compatible with an account based solely on identical syntactic structure, but this is not the only possible explanation; the linking rules from conceptual structure to syntactic structure could still form the basis for this apparent syntactic structure effect. At this stage we will have discovered whether it is conceptual structure (as in Heydel and

Murray (1996)) that is responsible for the priming effect, or alternatively that it is syntactic structure that is most likely to be the basis of the effect. In either case, the extent of the involvement of linking rules is unclear. The remaining experiments address this issue and tease apart the contribution of conceptual structure, syntactic constituent structure and linking rules at a more detailed level.

Experiment 2: Main/Subordinate Clause Priming Subjects: Native German Speakers

The importance of linking rules can be established by considering the question of whether a virtually identical conceptual structure represented by different syntactic structures, without any difference in focus, shows sentence priming effects from one of these structures to the other.

In German, main clauses have an SVO word order (a) whereas subordinate clauses have an SOV order (b).

- (5.65) a. Den Schwimmer attackiert ein Hai.
 b. (Ich habe gehoert) dass den Schwimmer ein Hai attackiert.

This will test whether the topicalization structure of a main clause can prime a subordinate clause structure, and vice versa. This approach does not suffer from the problems inherent in previous experiments on sentence priming, since it is difficult to argue for any conceptual difference, at any level, between the two sentence types. Syntactically, however, they are very different. In fact, in syntactic theory in German, (b) is described as scrambling, not a topicalization.

This experiment is important in theoretical terms as it will provide evidence of whether sentences that are as close as possible conceptually (they both focus on the Patient rather than the Agent of an action), but are different syntactically can prime one another. Theories that stress the importance of syntactic structure for sentence priming would predict that no priming can take place between these two structures in spite of their conceptual identity. Also, if rules linking conceptual and syntactic structure are critically involved, we would expect to find no evidence of priming. If priming is found, it clearly indicates the crucial role played by conceptual structure *per se*.

Experiment 3: Cross-linguistic Priming from German to English Subjects: Native German Speakers with Intermediate/Advanced English Language Proficiency.

It could be suggested that native English speakers (as tested in the Heydel and Murray study) find topicalizations to be a more marked structure than passives because of the lack of an English equivalent (there is, for example, evidence that they acquire this structure later and with difficulty (Rogers (1995))). They might therefore either a) assume that there is a stronger focus on the patient in a topicalization than in a passive and therefore feel constrained to produce a marked form in English - a passive, or b) become more form-focused and adopt a processing strategy not normally applied in natural discourse. In contrast, German speakers

will not treat topicalizations as a more marked case since this is a frequently occurring structure. Thus, when producing English they will not be prone to any focus bias. Consequently, if the results with this population mirror those found by Heydel and Murray it will indicate that the previous results cannot be attributed to any such bias.

Further, this study has the potential to test whether linking rules are crucially involved. As shown in the examples below, in German there is a choice of linking rule from the underlying patient-focused conceptual structure to surface structure resulting in either (a), a passive, or (b), a topicalization. Whereas in English there is a single rule linked to (c).

- (5.66)
- a. Der Schwimmer wird von einem Hai attackiert.
 - b. Den Schwimmer attackiert ein Hai.
 - c. The swimmer is being attacked by a shark.
 - d. The swimmer is attacking a shark.

German speakers, with their implicit awareness of this choice, may therefore attempt to identify a linking rule not available in English, resulting in an increase in the frequency of alternative syntactic structures and inappropriate responses, such as (d). This would be the most extreme scenario. Other clues to the interference of linking rules include increased onset time to start speaking, reformulations, or a greater level of pausing. If these occur, it suggests that linking rules, by themselves, may be primed across languages.

Experiment 4: Cross-linguistic Priming from English to German Subjects: Native German Speakers with Intermediate/Advanced English Language Proficiency.

The question of the importance of linking rules can also be tackled from another direction. The conceptual structure of an English passive can be realized in German in two different surface forms. This experiment will test whether Germans primed by English passive sentences are more inclined to produce only passives or both passives and topicalizations. If there is priming from the English linking rule then passives would be expected only to prime other passives. However, if it is the underlying conceptual structure which is most important, then both passives and topicalizations should be primed. (The extent of priming in this experiment can be established by comparison with the baseline results obtained in the first phase of Experiment 1.)

Experiment 5: Intonational Change Subjects: English Speakers with Intermediate/Advanced German Language Proficiency

This experiment will be a preliminary investigation into the way conceptual structure is reflected across languages, not just by syntactic structure, but by a range of expressive linguistic devices. Steedman (1990) has suggested that there is a close organic relationship between syntax and intonation.

The results from Heydel and Murray (1996) show that German topicalization structures prime an English passive. However, as priming is a relative effect, a proportion of the sentences

produced will be in the active form. This experiment will test whether when active structures are produced after a topicalized prime, the subject attempts to express the focus structure inherent in the prime in some other way. The most likely possibility seems to be by the use of intonational cues. Consequently, this experiment will employ a variation on the procedure used in the preceding experiments. Subjects will see pictures and written German primes in the same manner as previous experiments, but will be asked to provide a spoken response, which will be recorded. These responses will be presented (with no indication of their respective primes) to new subjects who will judge whether actives produced following a topicalization prime differ in intonational form from those produced following active primes. The judges will be asked to score each utterance for degree of emphasis on the focused constituent.

Further Implications of the Research

An important distinction between lexical and sentence priming can be drawn. Whereas lexical priming may be argued to be based on discrete atomic entries in the lexicon, the nature of the experiments proposed here argues that sentence-level priming is based on composite entities. In other words, the elements that are being primed would not exist if they were not part of the processing of the sentences that brings them into being. This has implications for the time-scale of both processing and priming. The results of the experiments we propose above will therefore make predictions about the time-scale of processing that will be amenable to further research.

The data obtained in the experiments on cross-linguistic priming would enable us to identify the level of natural language processing at which cross-linguistic transfer takes place and the mechanisms involved in this transfer. This aspect of the research has profound implications for theories of second-language acquisition and their application in language teaching. In general terms, the findings should inform practice in this area. More specifically, if it can be shown that sentence level priming is sensitive to conceptual and/or focus structure, rather than syntactic structure, this would have important implications for teaching practice. It would indicate that second language learners use of a) linking rules that do not exist in the target language, or b) ones that are inappropriate in a given context, needs to be explained as a consequence of teaching practice, rather than as an unavoidable stage in second language acquisition.

5.4 Recent Experiments in Thematic-Conceptual Sentence Priming

5.4.1 Cross-linguistic experiments

Heydel and Murray (2000) have again reiterated in their initial cross-linguistic work that as particular conceptual structures will normally be realized by certain syntactic structures in a particular language, the repeated production of the same syntactic structure does not necessarily imply that it is the syntactic form, or syntactic processes that are primed. Thus it may equally be the case that it is the canonical linking rule from the thematic conceptual structure to the surface syntactic form that may have been primed.

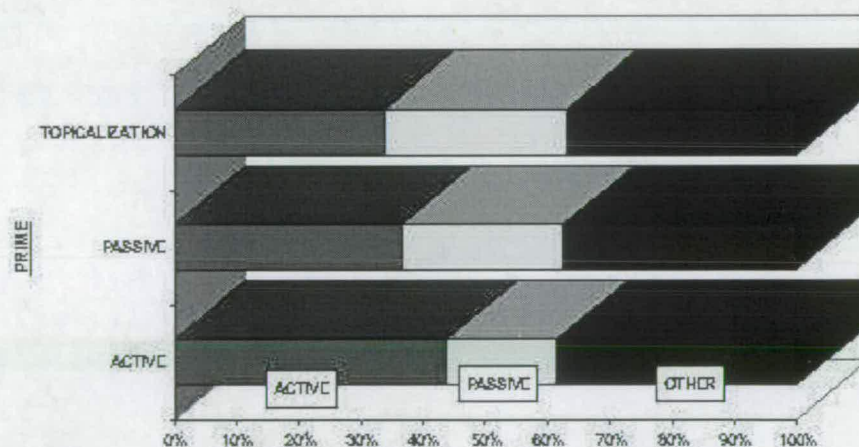
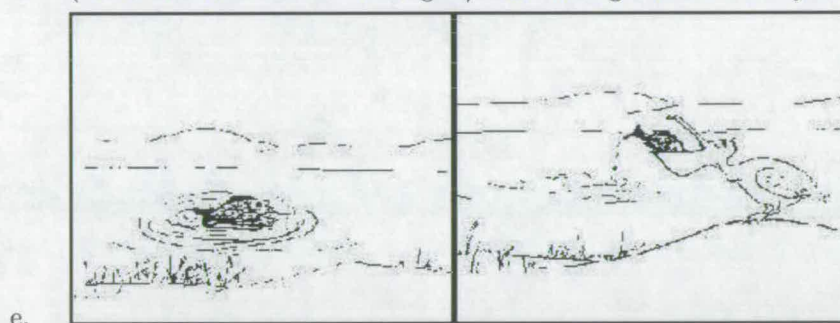
Heydel and Murray have performed at least two studies using crosslinguistic methodology that show that sentence-level priming effects extend cross-linguistically and these studies further suggest that sentence-level priming maybe more closely linked to conceptual rather than to syntactic form.

The studies rely on that fact that in some languages a particular thematic-conceptual structure can be linked to two or more syntactic structures, whereas in another language only one linking rule from thematic-conceptual structure to syntactic structure is available. Thus a focussed patient can be realized as either a passive sentence or a topicalized sentence in German, but only as a passive in English.

Heydel and Murray (1997b) includes two experiments on cross-linguistic sentence-level priming. In the first experiment the task was that participants were presented with picture pairs and German sentences. For each pair of pictures they had to decide whether the German sentence related to the events shown in the picture (this was never the case for experimental items) and to describe each picture in one complete sentence.

(5.67) Prime Types:

- a. (Active) Ein PR-Mann beraet den Manager.
- b. (Passive) Der Manager wird von einem PR-Mann beraten.
- c. (Topicalization) Den Manager beraet ein PR-Mann.
- d. (A PR-man advises the manager / The manager is advised by a PR-man)



It could be argued that the results may have been a side-effect of the 'odd' task in this experiment. The task could be regarded as 'odd' as translation is not required in normal

discourse. This seems unlikely for two reasons: related types of animacy effects have been found in mono-linguistic experiments, e.g. Prat-Sala *et al.* (1996); also, in translation, syntax plays a role which can best be described as 'excessive'. Inexperienced translators, in particular, tend to reproduce the syntactic surface structure of the original utterance, at times regardless of conceptual or grammatical constraints (Krings (1996); Bosshardt and Hagen (1988)).

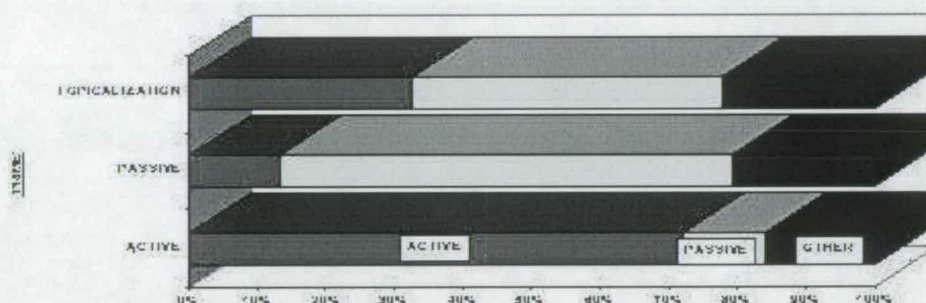
Nevertheless a second experiment was carried out to exclude the pseudo-effect of the translation class from the range of possible explanations:

The same material as for Experiment 1 was used, except that the associated priming sentences always described the events shown in the right hand picture. Participants were asked to provide a description of the left hand picture and translate a German sentence, such as those below, into English.

- (5.68) a. (A) Ein Auto ueberfaehrt den Igel.
 b. (P) Der Igel wird von einem Auto ueberfahren.
 c. (T) Den Igel ueberfaehrt ein Auto.

The responses could be:

- (5.69) a. (A) A car runs over the hedgehog.
 b. (P) The hedgehog is run over by a car.
 c. (A*) *The hedgehog runs over a car.

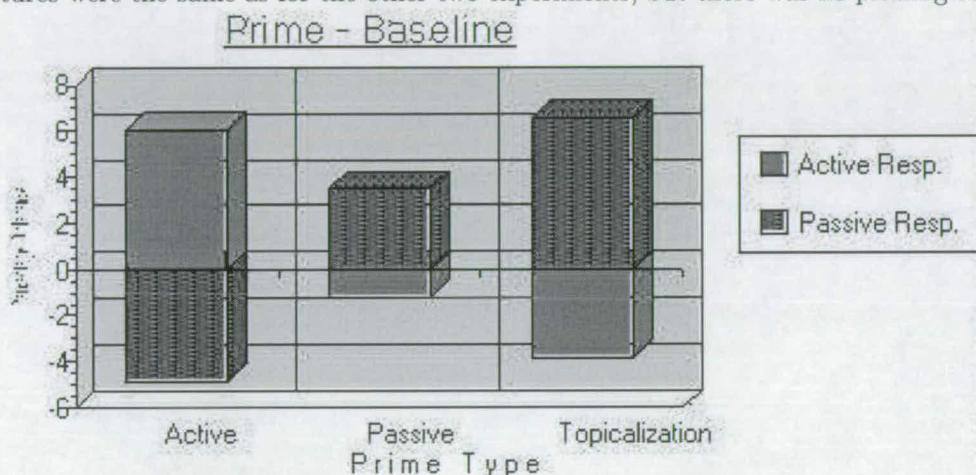


As expected there was a massive effect of syntactic form. Given the magnitude of the effect it might be argued that there is something odd about the translation task. This seems unlikely, given the number of 'Other' responses, and because responses never included any semantically anomalous sentences such as A*. In fact topicalizations were pulled in both directions. But if we take another conceptual feature which would make participants more likely to use passives (such as an inanimate agent), the pattern changes quite dramatically: for animate agents (15 items) there were 56.4% actives, but for inanimate agents (9 items) the percentage of active responses dropped to 13.9%. Heydel and Murray (1997b) conclude that the conceptual effect in Experiment 1 is clearly not due to 'translation'. There is also evidence from lexical priming in language comprehension from bilinguals that suggest that proficient bilinguals do not rely on translation equivalents in order to access the meaning of a word (Frenck-Mestre and Prince (1997)). Thus it would seem perfectly plausible that bilinguals would be able to understand a sentence without translating it into their own language.

There are alternative explanations of the results indicating thematic-conceptual effects in cross-linguistic sentence priming experiments (Heydel and Murray (1997a)).

The first of these explanations is that only one of the structures is primed and the other is not affected. In this scenario, if only the active was primed, then the equivalent results with passives and topicalizations would not in effect be showing the critical role of thematic-conceptual structure in sentence priming. Thus a baseline test is necessary for this counter-explanation to be discounted.

Heydel and Murray (1997b) performed a further study to discount this effect: To establish a baseline rate for the production of actives and passives, participants were asked to provide descriptions of the events shown in the pictures in the same manner as in Experiment 1. The pictures were the same as for the other two experiments, but there was no priming sentence.



Reanalyzing the cross-linguistic experiment against this baseline, we can see very clearly that both forms were primed.

The second counter-explanation is that it is the specific procedure itself that leads to the results shown. This has been discounted by changing the procedure by introducing an intermediate sentence, with no lessening of the effects found.

5.4.2 Mono-lingual experiments

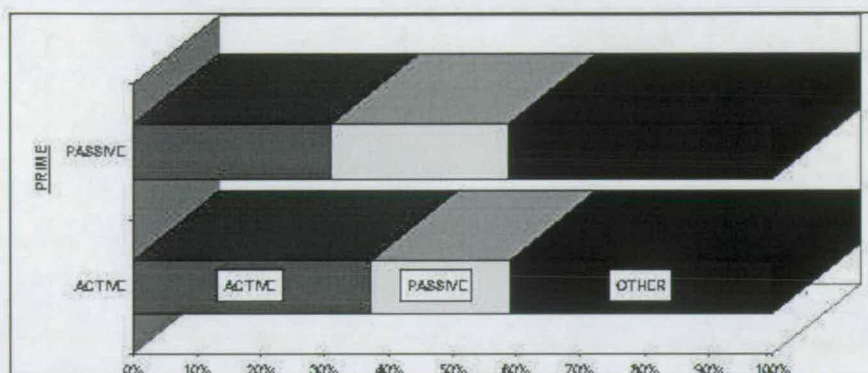
It could perhaps be argued that the results of the cross-linguistic priming experiments were a pseudo result of the cross-linguistic task itself. Heydel *et al.* (1998) carried out an additional monolingual experiments where the result could not be due to translation.

Participants were presented with the same picture pairs used in the original cross-linguistic experiment and with English sentences which were either actives, such as (5.70), or passives, such as (5.71). For each pair of pictures they had to decide whether the sentence related to the events described in the pictures (for experimental items this was never the case) and to describe each picture in one complete sentence.

(5.70) A consultant is advising the manager.

(5.71) The manager is being advised by a consultant.

The results show that mono-lingual priming produces results which are equivalent in magnitude to the ones obtained in the cross-linguistic study (and very similar to those reported by Bock). Crucially, the effects here are not larger - which they would have to be, if syntactic structure and conceptual form had independent additive effects.



It is also not the case that the conceptual effects found in sentence priming are limited to German. Heydel *et al.* (1999) have produced similar results for *monolingual* experiments conducted in Spanish.

Evidence for conceptual effects in cross-linguistic priming has also been found at the lexical level (Ketley *et al.* (1994); Basden *et al.* (1994)). Results from these studies show that cross-linguistic priming takes place if and only if subjects are encouraged to access the underlying conceptual structure of an expression.

5.5 Theoretical and Representational Aspects Of Sentence Level Priming

It has become clear from the discussion above that sentence level priming is a robust phenomenon. The phenomenon is not only fixed to the repetition of the same lexical verb in the prime and the target, but rather acts across classes of verbs. Sentence level priming is also not limited to either production nor comprehension and indeed priming from one modality to the other has also been seen to occur.

There are two competing versions of how this cross-verb sentence level priming should be represented. The first of these is Branigan and Pickering reanalysis of their work in terms of Roelofs (1992, 1993) model. The second is the theory of thematic-conceptual structure expounded in this thesis.

Levelt *et al.* (ress), argued that lexical entries include a *lemma stratum* encoding syntactic information and a *word-form stratum* encoding phonological and morphological information. Pickering and Branigan (1998) identify three types of information that need to be included in the lemma stratum of a verb. These are:

- Category information (N, V, A).
- Featural Information (Number, Person, tense).
- Combinatorial Information (how the verb can combine with other linguistic units).

In Roelofs' model the lemma stratum contains lemma nodes (one for each lexical concept) which are connected to nodes at the conceptual stratum (messages generated) and the word form stratum (where morphology and phonology are specified). Pickering and Branigan (1998) propose an extension to the lemma stratum to incorporate the syntactic aspects of verb representation². They assume the existence of combinatorial nodes which are activated when the verb is used in a particular construction, and are open as to whether this is in terms of subcategorization frames or overall combinatorial potential (including adjuncts)³. The other types of verb information are also activated by the input (Aspect, Number, Tense). The model is then investigated with data from sentence priming. Sentence priming is then explained in syntactic terms based on activation of nodes at the lemma stratum. The crosslinguistic evidence found in the Heydel and Murray experiments already militates against such an account.

The sentence priming experiments in Pickering and Branigan (1998) also show that sentence priming effects occur across verbs, with the priming effect stronger when the verb was the same (expected by both accounts) and that tense, aspect and number had no effect on the magnitude of priming. Pickering and Branigan (1998) conclude that the 'findings argue against approaches to syntactic representation in which combinatorial information is represented separately for each verb (P646)'⁴.

Indeed, Pickering and Branigan (1999) seem to be moving in a direction that incorporates information of a more conceptual nature. They point out that speakers and listeners benefit from priming effects in dialogue and reiterate the findings that interlocutors tend to produce the same kind of forms, and in this way are coordinating their contributions. There is evidence of such coordination at a number of levels. In describing mazes participants converge on the same type of descriptions (Garrod and Anderson (1987); Garrod and Clark (1993); Garrod and Doherty (1994)), and even more interestingly from the thematic-conceptual standpoint, there is semantic coordination in terms of mental models or 'conceptual pacts' in how to refer to an object (Clark and Wilkes-Gibbs (1986); Wilkes-Gibbs and Clark (1992)).

Pickering and Branigan have themselves gone on to show that in picture descriptions with a confederate interlocutor scripted to produce one form or other as a prime (PO, DO), the subject also used the primed form more frequently (75.5% of the time if the verb is the same and 63% of the time if the verb is different)

5.6 Summary

In this chapter it has become clear that thematic-conceptual structure can be used to explain psycholinguistic observations that have otherwise proved difficult to explain.

²It may be worth considering whether the combination of the conceptual stratum and a phonological stratum would be sufficient to explain parsing data. Fodor (1998a) has made the suggestion that prosodic packaging couples with MA and LC would be able to account for many parsing phenomena. In this chapter we have seen that thematic-conceptual structure is required to account for some of the more complex data. A combination of the two without recourse to a separate syntactic extension of the lemma stratum may be sufficient for explanatory purposes.

³This second option comes very close to the Tuning Hypothesis of Mitchell and Cueto (1991).

⁴However in Pickering and Branigan (1999) it is made clear that this representation is only regarded as the least redundant way to have a representation that is drawn on by both production and comprehension. Pickering and Branigan explicitly state that there can be no simple procedural account of priming from comprehension to production as the operation is reversed. I have argued against this saying it is the linking rule that is primed and that the linking rule is a procedure that can run in either direction.

Thematic-conceptual structure was also argued to be the representation type at which sentence-level priming can be explained. This contrasts with the current standard explanation of sentence-level priming purely in terms of syntactic structure.

It was not clear, however, whether priming took place at the level of thematic-conceptual structure itself, or rather at the linking rules that connect the underlying thematic-conceptual structure to the particular surface expressions of a language. I suggested a number of experiments that would clarify this point. The results thus far would certainly seem to give thematic-conceptual structure a much greater role in sentence priming than has been the case up to now.

It should have been quite clear from my approach up to now that I believe that thematic-conceptual structure has a major role to play in explaining psycholinguistic phenomena.

Claims for a distinction between conscious and non-conscious parsing difficulty would seem to be supported by the empirical evidence. The explanation of parsing difficulty from the viewpoint of the thematic-conceptual structure is relatively easy to describe.

If a sentence has been misparsed, recovery will be consciously difficult if:

- An argument has to be relinked as part of a different predicate to the predicate to which it was originally linked; or
- An argument has to be reanalyzed as being linked to an alternative linking point within the same predicate.

The operation of this parsing principle, of course, is dependent on the type of on-line structure building that Pritchett proposed. In the case of thematic-conceptual structure, this on-line structure building is stated in terms of assigning arguments to potentially linked positions as soon as possible, and it is the undoing and rebuilding of this semantically motivated structure that proves difficult in the two cases given above.

This semantic reanalysis is a more plausible source of difficulty than any strictly syntactic explanation. It is unclear why a reanalysis of syntax by itself should be difficult, but it seems a good deal more plausible that once a semantic representation has been built in terms of thematic-conceptual structure, it will be difficult to undo this understanding and to rebuild a new thematic-conceptual structure with a quite different semantics.

The open roles central to the Discourse Strand of the work of Tanenhaus and his colleagues (section 4.2.1) are readily available in a thematic-conceptual structure. The thematic-conceptual structure will always represent all the potential linking points for arguments, whether these are filled or not. The unfilled linking points will be implicit arguments, or open roles as Tanenhaus terms them. The implicit roles will be part of the possible interpretation of a sentence as a result of being part of the semantically and conceptually motivated thematic-conceptual structure. This interpretive element will allow for the easy continuation of discourse based on either filled or implicit arguments.

We also saw in Chapter 4 that the ease of discourse continuation of a sentence is not explained merely in terms of open roles. There also seems to be a difference in the ease of continuation based on the thematic role type that the subject of the continuation plays in the first sentence.

This thesis situates such findings and proposals into a much wider context, both conceptually and linguistically. In the concluding chapter I will draw the relevant sub-disciplines together and explain once more why thematic-conceptual structure should be central to linguistic inquiry. I will also discuss other disciplines and sub-disciplines that I believe also shed some light on the centrality of thematic-conceptual structure to linguistics, but were perhaps a little too speculative to be included in the main body of the thesis.

Chapter 6

Conclusions

The initial chapters of this thesis presented an overview of the role thematic information has been seen to play in linguistic theory. In Chapter 2 we saw how initially thematic information was presented in a decompositional theory known as Generative Semantics. The decomposition of Generative Semantics was unconstrained and thus resulted in very large structures that relied on an equally large number of transformations to convert to surface structure. After some time where thematic information was theoretically regarded as almost equivalent to argument structure, there was a redevelopment of decompositional approaches to thematic structure. The most successful of these in the view of this thesis was what I have termed thematic-conceptual structures.

Thematic-conceptual structures are constrained to include only those semantic elements that have been grammaticalized in some language of the world. We saw in Chapter 3 how thematic-conceptual structures can explain the learning of a number of verb alternations. Also in Chapter 3 we saw how grammaticalization is not only a most complex operation that may not be easily expressible in terms of binary parameters, but also that grammaticalization seems to operate on metaphors that are particularly salient to humans in their ecological niche. There is more on this link between language structure and phylogeny in the next section.

It seems to be the case that the earliest stages in language acquisition are thematic-conceptual in nature and that such syntactic notions as functional categories develop maturationally. Also, in the earliest stages many of the semantic categories represented in thematic-conceptual structures and in the learning of the earliest alternations and the differentiation of certain word meanings are also seen to be relations in the world that are understood very early by children.

At the end of Chapter 3 this discussion resulted in a number of predictions. These were:

1. Thematic-conceptual structures will be a major player in the processing of sentences, as processing is then understood as the linking from thematic-conceptual structure to syntax or vice versa, depending on whether production or comprehension is being considered.
2. The grammaticalization of metaphors that have become central to humans in their ecological niche (See Chapter 6 for further discussion) will play their part in the content

that is relevant to processing.

3. The breaking down of a higher processing system will result in the processor falling back on lower processing systems. The lowest level of this processing will be thematic-conceptual structure, as thematic-conceptual structure forms the earliest stages of language acquisition.
4. Linguistic parameters may well be able to be stated and defined at the level of the linking of thematic-conceptual structure to the surface syntactic structure.

In Chapter 4, where the role of thematic information in sentence processing and other aspects of psycholinguistics was discussed, the third of these predictions was seen to be the case. When the sentence processor broke down, the parse could be recovered by accessing thematic information. There were also indications that thematic-information may be centrally implicated in the processing of sentences, and that it is thematic information that determines the processing difficulty of the sentence.

The idea that thematic information could be driving sentence processing was examined in detail in Chapter 5, where it was shown that the early indications are that the Thematic-Conceptual Structure and Linking Hypothesis I proposed there was the best predictor for patterns of sentence processing and sentence priming and for the first time represented a process that could be common to both production and comprehension of sentences. We also saw how the metaphorical nature of grammaticalization (specifically the metaphorical link of vision to contact) is implicated in processing patterns.

Provocatively, the idea that processing patterns are determined by the linking or de-linking of thematic-conceptual structure means that it could be argued that linguistic parameters of how a language expresses on the surface the underlying conceptual structure could be expressed at the level of these linking rules. This would mean that syntax would only have one level and the type of movement operations assumed to be behind such aspects as raising and passivization. It is well known that the gaps posited by such operations are hard to pick up psycholinguistically and, encouragingly, the work of Williams discussed in Chapter 2 would seem to suggest that raising at least can be captured at the thematic level without positing gaps. Jackendoff (1990) has also suggested that binding could be captured in a mono-level thematic-conceptual structure. This is a topic ripe for future research and outside the scope of this thesis.

In the remainder of this chapter, I would like to explore some more speculative arguments that I believe have a bearing on the theory of linguistics we should be following. Language is something that has developed evolutionarily and the evolution of language will have a bearing on its manner of operation today. Also language fulfills a role for humans in our ecological niche and will thus be shaped by the elements that are important for humans to cognize in order to survive.

6.1 Thematic-Conceptual Structure and Phylogeny

Thematic-conceptual structure can be seen as the link between our phylogenetically determined conceptualization of the world and the structure of language. It forms the structural

level at which sentence-level priming occurs and the linking rules associated with thematic-conceptual structure can be regarded as a procedure that is common to both production and comprehension.

Linking rules are of central importance in another respect. Linking rules differ from language to language. Crosslinguistic differences occur in the expression of thematic-conceptual structures (which do not differ from one language community to another) in the surface expressions of a language. Linking rules encode in thematic-conceptual terms how a particular language expresses a certain underlying thematic-conceptual structure. The strength of the linking rule over syntactic parameters is that linking rules make reference to the semantic content of the expressions whereas syntactic parameters operate merely in terms of abstract configurational representations. We saw in the discussion of grammaticalization that theories based on a small number of configurational differences are unlikely to be able to capture the wealth of crosslinguistic differences that occur.

There may well be patterns that occur often cross-linguistically in how languages or groups of languages express particular thematic-conceptual relations. These patterns will be expressible in terms of the thematic-conceptual structure of the linking rules. If a language expresses a particular thematic-conceptual relation in a certain way in the surface expression of a language, it may be the case that other patterns of linking rules can be predicted from such a fact, if it is commonly the case that other languages have shown it to be the case that groups of linking rules occur. Thus the number of options that a language will have in expressing sets of thematic-conceptual relations may well be limited in number.

An attractive feature of the proposals laid out in this thesis is that it can be applied to both linguistic structure and language processing. In chapters 2 and 4 we saw that structural and processing theories were formulated in differing notations and with a differing emphasis. There seems to have been very little cross-discipline fertilization with regards to formulating theoretical structures that are relevant to both domains.

Not only has thematic-conceptual structure been shown to be applicable to both structure and processing, it has also been shown to be applicable to both sentence comprehension and sentence production in processing. In production thematic-conceptual structures are converted via linking rules into the structure expressed. In comprehension the linguistic expressions are de-linked to decode the thematic-conceptual structures that allow for the semantic interpretation to take place.

6.2 Grounding Thematic-Conceptual Structure in Evolutionary Terms

We have seen pointers that indicate the explanatory potential of thematic-conceptual structure. That is, we have seen the sorts of things that we would want any theory of thematic-conceptual structure to be able to accomplish. In this chapter I will ground the theory of thematic-conceptual structure in phylogenetic and ontogenetic terms. In so doing, thematic roles will emerge as part of a greater structure that has its basis in conceptual terms.

It would appear that some of the problems that theories of thematic information have

faced up to now stem from the fact that they are embedded in a theoretical framework, where the language system is seen as fundamentally separate from the rest of cognition. I am not denying that this may be the case. However, there is a difference between on the one hand assuming that language is modular now and that other aspects of cognition can have no effect upon its operation, and, on the other hand, the position I take, that whatever the modular status of the language system today the form of the computational procedures taking place within it will have been shaped by the evolution of the language system from its cognitive precursors.

The direct bearing that this approach has on thematic-conceptual structures is that thematic information seems to me to be the clearest reflection of the interface of language and cognition. An understanding of the cognitive and evolutionary precursors of language will illuminate the nature of the thematic relations and their formal properties. These cognitive precursors will determine to a large extent the type of relations that a language can express, and the linguistic structure that is required to express these relations. Thematic-conceptual relations are therefore the bridge between language and cognition.

There are three ways that we can look for the precursors of language. First, we can examine the evolution of language itself. This, of course, is an area where any argument is based on speculation to some extent, but if the speculation is informed I see nothing wrong with this. The difficulty of the task does not preclude it providing information on the nature of language.

Second, and related to the first we can examine computational systems in the brain that probably existed before language and were therefore likely to have been co-opted in the phylogenetic development of language. In addition, it may be the case that ontogeny reflects phylogeny to some extent and that we will see that the formal nature of the early stages of language acquisition reflect more deeply embedded structures of the brain. Brain functions that are in some way more basic may develop earlier in infants and language acquisition will be constrained by the form of the computational processes and knowledge that these earlier brain functions encode.

Third, we can look at our nearest evolutionary relatives; monkeys and apes. Not only do they give us information on the extent to which they can be taught some form of human language, but they can also, and probably more importantly, give us information on capabilities outside of language that seem central to primates and thus are more likely to be deeply embedded in humans as well. Certain of these abilities may have a bearing on the later evolution of language.

The chapter begins with a quick exposition of the most recent view of the evolution of mankind itself. This will form the backdrop for the second section which assesses the evolutionary status of language. The discussion in these two sections will have implications as to the sort of theories of language evolution that will be plausible.

It is likely to be the case that the sorts of computational or cultural adaptations that resulted in the emergence of language will still be of relevance to its structure today. Some of these adaptations will become clear from the discussion in the first three sections.

However, much more can be gained by looking at our primate relatives and examining the language-like abilities that they possess. 'Language-like' should be understood in two different ways. There are the cognitive, computational and structural precursors of language. These

may be used in a domain other than language in other primates. I believe that humans have co-opted into the language system some of these precursors during the evolution of language, and that the structure of modern languages will reflect the structure of these precursors. These possible precursors of language as seen in contemporary primates will be discussed.

The other language-like abilities can be seen in attempts to teach human-type languages to apes. This is addressed in section 4.5. The successes and failures of these attempts will be contrasted. I shall argue that a large part of the success of the bonobo, Kanzi, in learning language is due to the manner in which he learnt language. This manner has been termed cultural learning. I then present the theory of cultural learning and include a discussion of a theory of language evolution based upon cultural change.

6.3 A Brief History of Humankind

Before we look at the evolution of language, it is useful to put its evolution in context of the evolution of Mankind. It is often claimed that it is language (Donald (1991); Edelman (1992); Leakey (1994)) that led directly to the pre-eminence of *homo sapiens* over other species of the *Homo* genus.

Leakey (1994) identifies four key stages in the evolution of Homo Sapiens.

1. The origin of the human family 7 million years ago when bipedal ape-like creatures evolved.
2. The proliferation of bipedal species through adaptive radiation between 7 million years ago and two million years ago.
3. In this proliferation was a species with a significantly larger brain. This species was the first in the genus *Homo*.
4. The origin of modern humans with language, consciousness and technological innovation.

The questions that arise from these stages that are particularly relevant to the task in hand are: What is the shape of the human family tree? When did language first evolve? What caused the dramatic increase in brain size? These questions are particularly important for various reasons. The shape of the family tree will tell us how closely related we are to any other primates we examine for language or language-related abilities. The time when language evolved with respect to other human competences will inform us as to what the language system could have been based on. The cause of the increase in brain size will allow us to examine claims that language evolved only after a certain brain size was achieved (Chomsky (1988)) or whether language itself caused the brain to increase in size, or whether the situation was much more complex than either of these two positions suggests. Later sections will address these questions further and will modify the conclusions we can make from human evolution as presented in this section.

Darwin made two assumptions with respect to human origin. The first was that the origin of humans was in Africa. This assumption is generally agreed with today. The second assumption is what Leakey (1994) terms the *Darwinian Package*. The Darwinian Package is

the belief that bipedalism, technology and an enlarged brain evolved in concert. The reasoning can be put down to the time at which Darwin was forming his theory. It was difficult enough to persuade people that we were in effect a species of ape, but the task was made easier if we could be shown in some way to be a considerable distance apart from the rest of nature, and the co-evolution of the elements of the Darwinian Package provided this distance. The Darwinian Package remained unquestioned until recently and formed the centrepiece of the science of anthropology. It may be that Chomsky's feelings about the evolution of language could have their origin in the Darwinian Package.

The folly of the Darwinian Package was shown in the early 60's when it was claimed on the basis of a fragment of jaw that was found that *Ramapithecus* was the first hominid species around 15 million years ago. The tooth in this jaw had hominid features and thus it was argued that it must have come from a hominid species that was bipedal and had a complex social environment. In other words, because the jaw had one hominid feature, the hominid whose jaw it was must have had all the other hominid features in the Darwinian Package. Sarich and Wilson (1967) working on blood proteins found that they could be used as a molecular clock - the longer a species has been separate from another, the more mutations accumulate in the blood protein. Based on this, they calculated that humans and apes had split only 7 million years ago. Thus *Ramapithecus* was not a hominid but a primitive ape.

There are a number of explanations of human origin, all of which have informed theories of the evolution of language. The theories can be characterized in the following way.

- The Expanding Brain Hypothesis
- Man the Toolmaker
- Man the Killer Ape
- Man the Hunter-Gatherer
- Man the Hunter
- Woman the Gatherer

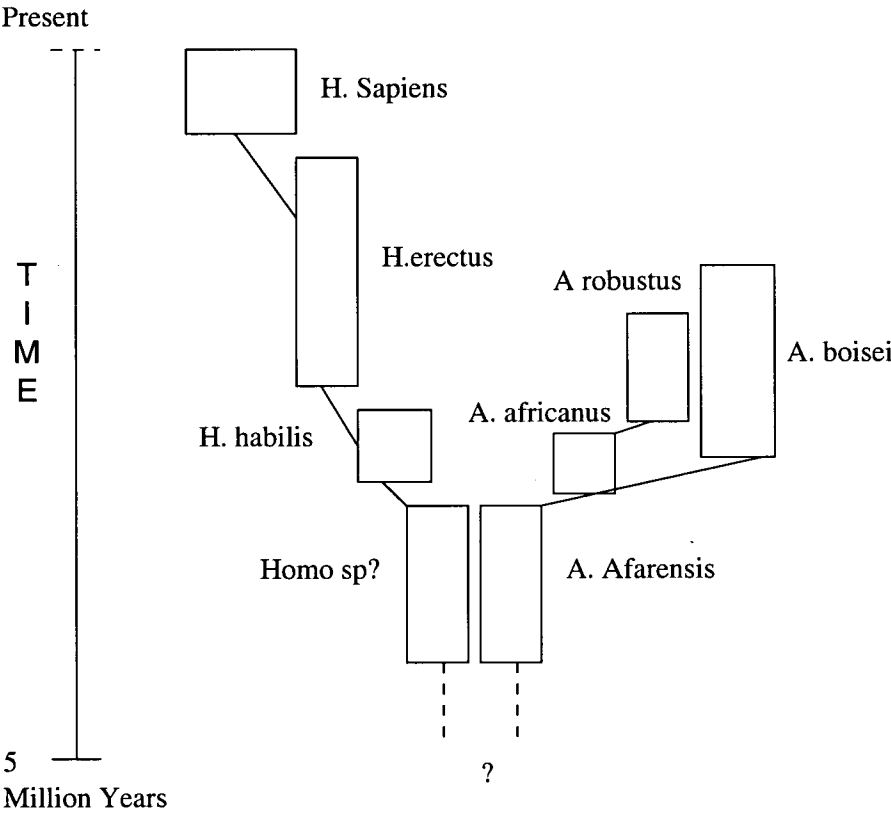
All these have in common the assumption of the Darwinian Package. If it is the case that the Darwinian Package is true, then we would expect evidence in the archaeological record to reflect this truth. This evidence is singularly not there. The earliest tools in the record are from 2.5 million years ago, leaving 5 million years between the evolution of hominids and the first tools. However, it may be the case that there is a link between the advent of technology and the increase in brain size.

The latest hypotheses are couched in biological rather than cultural terms and specifically focus on the origin and selective advantage of bipedalism. It is argued that bipedalism was loaded with evolutionary potential as it freed the upper limbs to manipulate objects¹. This freeing of limbs is a much more likely explanation than the popular but entirely inaccurate 'Out of the trees and onto the savanna' hypothesis. It is inaccurate as the savanna of East Africa only developed 3 million years ago - 4 million years too late.

¹A source of a theory of language evolution (Armstrong *et al.* (1995))(See section 6.5)

Another possible reason for bipedalism was that it was a more efficient means of locomotion (Rodman and McHenry (1980)). This may seem unlikely if we compare bipedalism with dogs and cats, but the comparison is true if we compare like with like and compare bipedalism with the knuckle-walking of chimpanzees.

Moving on from the evolution of hominids to the evolution of *Homo Sapiens* we find our clues as to the family tree. When a new species evolves, a process of adaptive radiation takes place. This is a burgeoning of variants on the species. We thus expect many types of hominid to have evolved. The archaeological record is sparse before 2 million years ago, but we know that 2 million years ago there were at least four species of *Australopithecine* and three species of *Homo*. The hominid family tree is thus probably of the form below.



It was with the evolution of *homo habilis* that the major shift in brain size occurred. There was a jump from 450cc in *homo erectus* to 600cc in early *homo habilis*. The jump in brain size coincided with a difference of hominid lifestyle. When apes are born, their brain is half its adult size. From *homo habilis* onwards, when hominids were born their brain was a third of its adult size. This means that children are born helpless and have a prolonged childhood. There is also an unusual growth spurt in adolescents of 25% compared to a more steady growth in apes of 14%. Bogin (1990) explains this by claiming that children learn better if there is a greater difference in size between adults and children.

This unusual pattern of childhood makes culture possible. Culture could be argued to be the human adaptation. Not only was there a change in child rearing, there also was a change in other aspects of human culture. There was a general change in social context. Males became

closer to females in size, indicating that males were now probably staying in their native groups and were more likely to be cooperating. This cooperation needs to be motivated. This could have been for defence against other tribes. Alternatively, a large brain needs a great deal of energy, which can best be supplied by energy rich foods such as meat. The hunting of meat by a relatively weak, relatively slow hominid would have required cooperation.

In this process we can find a motivation for the evolution of language as a method of more efficient communication: Culture developed and social life became more complex. This complexity led to a need for social and economic coordination. Dunbar (1993) argues that as humans formed larger groups, grooming became inefficient and language replaced hands-on grooming as a more efficient way to integrate groups. Central to both these arguments is the social nature of language, an element which is often ignored in Chomskyan linguistics and theories of language acquisition.

As indicated above there are two main views of language evolution. It is clear that brain size is implicated, but how is not clear. The first view as espoused by Chomsky is that at a certain brain size threshold language appeared rapidly. The other view is that selection acted on various cognitive capacities producing a more complex and enlarged brain. It can also be argued that a large brain had a selectional advantage and that in the same manner as for the peacock's tail, the advantage resulted in an adaptive feedback loop (Wills (1994)).

It is often argued that the abrupt language hypothesis is supported by the sudden difference in human behaviour in the Upper Paleolithic, when complex tools and cave paintings appeared. It can be argued that the Chomskyan view is a more modern version of the Darwinian Package in its motivation; language is so very different to other mental capacities and is largely unrelated to them and thus we are very different from the rest of nature. This may be comforting but it is unlikely to be true. Falk (1983) argues against the Chomskyan view by pointing out that if hominids were not using language in some form or other what else were they doing with their autocatalytically increasing brain size².

Pinker (1994) thinks that Chomsky has it backwards; the brain is more likely to have grown as a result of the evolution of language. The evolutionary evidence favouring an early evolution of language with gradual improvement based on more complex socialization is backed up by fossil evidence on the structure of the brain. Holloway (1983) has found evidence in *homo habilis* of Broca's area and slight brain asymmetry, arguing for some level of laterality and development of areas associated with language, although it is unclear whether these areas were being used for language - or for some other capability which shares some of the computational properties of language. Further evidence for laterality is provided by the research of Toth (1985) into hominid tools. These tools show evidence for right-handed preference.

From the form of the family tree we can safely conclude that the primates should be able to provide data that will be of relevance to our own cognition. We have also seen that evolutionary theories of language that rely on the brain expanding in size before the emergence of language seem dubious. The social aspect of language would seem to be of central importance.

There seems to be a cluster of properties that appear whenever language evolution is discussed and a successful theory of language evolution will have to link them in a satisfying

²Although one could perhaps argue that mountain gorillas use their large brains for keeping track of social interactions.

way. These features are brain size, bipedalism, society and lateralization.

Before I attempt to derive an explanation of the linking of these features (Section 6.5), I will show that an adaptive explanation of language emergence is the only viable theory (Section 6.4).

6.4 The Evolutionary Status of Language

The first task is to show that language is indeed an instinct that has evolved. The argument here will follow closely that of Pinker and Bloom (1990). I agree with much of the framework of their argument, but I reconsider some of the specifics.

Central to the argument that the language system has been formed by the process of evolution is the principle that biological artifacts which show complex design are understood to have been formed by natural selection. Crucially then, it is necessary to show that language has complex design.

The following questions were posed with respect to the human eye by Dawkins (1986). As the first question can be answered negatively and all the others positively for the human eye, Dawkins concludes that the human eye has evolved by natural selection. I hope to show that the same answers will be given for the questions with respect to the human language system.

1. Could language have arisen directly from no language at all, in a single step?
2. Could language have arisen directly from something slightly different from itself, something that we may call X?
3. Is there a continuous set of Xs connecting the modern language system to a state with no language at all?
4. Considering each member of the series of hypothetical Xs connecting language to no language at all, is it plausible that every one of them was made available by random mutation of its predecessor?
5. Considering each member of the series of Xs connecting human language to no language at all, is it plausible that every one of them worked sufficiently well that it assisted the survival and reproduction of the animals concerned?

Language is clearly a complex system. Pinker and Bloom (1990) list the following uncontroversial building blocks of grammar, all of which interact with each other in complex ways.

Grammars are built around symbols for major *lexical categories* that can enter into rules specifying tell-tale surface distributions, inflections and lists of lexical items. Together with *minor categories* that characteristically co-occur with the major ones, the different categories are thus provided with the means of being distinguished in the speech string. These distinctions are exploited to distinguish basic ontological categories such as things, events or states and qualities. This order of development reflects the general process of grammaticalization that we saw discussed in Chapter 3.

Within this general framework a number of other linguistic entities make up the total complex design. Major phrasal categories are constructed with a major lexical item which is allowed to combine with specific kinds of affixes and phrases. The resulting conglomerate is then used to refer to entities in our mental models of the world.

Phrase structure rules force the structure of the string to reflect the semantic connectedness in the underlying proposition and thus provide linear cues of underlying structure. Rules of linear order allow the order of words within these concatenations to distinguish among the argument positions that an entity assumes with respect to a predicate.

Case affixes on nouns and adjectives can take over these functions, marking nouns according to argument roles and linking noun with predicate even when order is scrambled. In addition to case affixes on nouns, verbal affixes signal the temporal distribution of the event that the verb refers to and the time of the event. Also related to the verb, auxiliaries convey relations that have logical scope over the entire proposition such as truth value, modality and illocutionary force.

Mechanisms of complementation and control govern the expression of propositions that are arguments of other propositions, using specific complementizer morphemes. Other features of language design include anaphora, pronouns and wh-movement.

And these aspects of language form only a partial list. The complexities of the interactions of these elements of the language system are only beginning to be understood. Having established the complexity of language in general terms we can then turn to Dawkins's questions.

The first question has already been answered to some extent. It is implausible that a system as complex as language could occur in an individual species in any way other than by natural selection. The alternative seems totally implausible. What possible single adaptation could have resulted in a system of the complexity of the language system?

What X could plausibly have been the basis of language, this X being slightly different to language? To answer this question we need a system making somewhat similar types of computations to the language system. This similar system was then used for the initial stages of language, probably somewhat clumsily. As the language system evolved, the computations required by the language system became more and more specialized and a system devoted to them would have become more and more innate. This is known as the Baldwin Effect (Baldwin (1896)).

One possible brain system that could have provided the initial computational style required by language is the visual system. Language could have arisen and probably did arise by a revamping of primate brain circuits that originally had no role in vocal communication, and by the addition of some new ones. The neuroanatomist Galaburda (1994) has discovered areas in monkey brains that correspond in location, input-output cabling, and cellular composition to the human language areas.

In this respect it is also provocative to note that Rizzolatti and Arbib (1998) propose that there is an observation/execution matching system that provides a bridge from 'doing' to 'communicating', as the link between the actor and observer becomes a link between the sender and receiver of the message. One area of the monkey brain (F5) contains neurons that discharge not only when an action is carried out, but also when it is observed to be carried out. These are termed *mirror neurons*. The area F5 is generally agreed to correspond to

Broca's area in the human brain. It would seem then that Broca's area, before it became more specialized for speech, was endowed with a mechanism for recognizing actions made by others. We see here a link between gesture, action and language that will tie in well with my arguments later in this chapter. The findings of Rizzolatti and Arbib also provide a neural basis for this link.

Another possible brain system would be the system for the hierarchical representation of the social environment, which has been shown to be present in primates. This system may have been co-opted for the representation of syntactic relations.

Is there a continuous set of Xs connecting the modern language system to a state with no language at all? One must be reminded of the fact that tiny selective advantages are sufficient for evolutionary change. According to Haldane (1927), a variant that produces on average 1% more than its alternative allele would increase in frequency from 0.1% to 99.9% of the population in a little more than 4,000 generations.

Is it plausible that the language system could have been gradually modified by small random changes? I believe so. It is hard to separate the cultural evolution of language from its biological evolution. A sensible approach would be to assume that much of Universal Grammar has a biological basis, which in the early stages of language evolution changed in small, biological, brain-based increments resulting either in a language which lost expressive capacity and thus selectional advantage, or in a language of increased expressivity which would have been actively selected for.

Considering each member of the series of Xs connecting human language to no language at all, is it plausible that every one of them worked sufficiently well that it assisted the survival and reproduction of the animals concerned? Once language was in place, we have to assume that each subsequent change had some evolutionary advantage for the species, in the efficiency of either production or comprehension of language, or in its ease of acquisition.

Premack (1990) reviews evidence that pedagogy is a universal and species-specific human trait, and the usefulness of language in pedagogy is not something that can be reasonably doubted. Devices designed for communicating precise information about time, space, predicate-argument relations, restrictive modification and modality are not wasted in such efforts. Recursion, in particular, is extraordinarily useful³.

Humans depend on cooperative efforts for survival. In a group of communicators competing for attention and sympathies, there is a premium on the ability to engage, interest and persuade listeners. This dependence in turn encourages the development of discourse and rhetorical skills and the pragmatically relevant grammatical devices that support them.

Humans fell into a lifestyle that depended on extended cooperation for food, safety, nurture and reproductive opportunities. This lifestyle presents extraordinary opportunities for evolutionary gains and losses. Any adaptation that improved any of these aspects of behaviour would have a selectional advantage for the species.

I believe that the evolutionary status of language has been established. However, there are a number of other possible explanations for the existence of language that have to be discounted. There are two basic arguments that compete with an evolutionary account of the language system. The first is that language is not a separate complex system. This argument

³Hierarchical structure may well have been present in the cognitive system before language.

claims the brain is a general computational device and that language is just one of the abilities that this computational system has (Christiansen (1994)). The important point to note is that the theory claims that the same types of computation can explain many different types of behaviour. Another argument of this same type is that as the brain has grown larger in evolutionary time, its sheer size has made more complex computations such as those involved in language possible.

The second type of argument concerns the theory of evolution itself. Gould and Lewontin (1979) have argued that many adaptations present in creatures first evolved to perform some other task than that which they are used for today, and these changes were then found to be useful for their current purpose at some later time. Their contention is that language is one of these evolutionary 'spandrels', as Gould and Lewontin term them. Another evolutionary argument of this second type is that changes in the structure of a being during its lifetime can be passed on to its offspring. This approach is known as Larmarckism after its most famous proponent.

Common to the arguments against natural selection as the origin of the language system is the idea that language is a mere side effect of other forms of cognitive development. There are reasons why the brain should not be seen as a general multipurpose learning device. It should be borne in mind that language learning is not direct programming. The language learner is provided only with sentences of the language not the rules of the language. The rules have to be derived by some learning device. The evolutionary argument posits that it is natural selection that is the programmer (or has provided the program).

The opposing claim to this is that the mind is like a machine equipped with a single learning device that could learn from examples. However, Pinker and Bloom claim that there is no psychologically realistic multipurpose learning program that can acquire language as a special case, because the kinds of generalizations that need to be made to acquire a grammar do not align themselves with those that are useful in acquiring other systems of knowledge from examples (Chomsky (1986)). This is a large claim which is not necessarily substantiated in the text of Pinker and Bloom.

There is, however, another strong argument that motivates against the multipurpose learning device position. This argument is the Poverty of the Stimulus: The grammar of the language encapsulates the constraints and functions as a set of specifications that allow for the production of all the grammatical sentences of a language. The language learner has to extract this set of specifications from a finite set of instances. This is quite impossible, it is claimed. There is an infinite set of grammars that could explain a finite set of utterances.

If a language learner hypothesizes a plausible rule from the data and the learner has not yet identified the constraints on language constructions, the learner's language will probably include constructions outside the restricted set of the target language. We then have to assume that there is a *language informant* of some kind or other that allows the learner to obtain evidence about the validity of the rule. Negative feedback would provide linguistic information on inappropriate sentences.

But there is little negative information available to the learner (Pinker (1989)). Most people speak accurately most of the time. Non-instances are therefore a statistical rarity. However, negative evidence could be argued to exist in the form of the non-occurrence of particular

forms. Adults rarely correct the grammatical mistakes their children make (Hirsch-Pasek *et al.* (1984)), focusing much more on the factual accuracy of what they say.

Children do not get systematic negative feedback on what they say, and the feedback they do get under-determines the correctness of their productions. Given that constraints on language are negative statements and that there is little negative data available to the learner, how can the constraints be learned? This learning of constraints despite insufficient negative feedback is the paradox of language learning.

Thus the evidence does not seem to support the multipurpose learning device position.

The second argument of this type to repudiate is that changes in brain quantity result in changes of brain quality. Mere largeness of brain is neither a necessary nor a sufficient condition for language, nor is there any reason to believe that merely more and more neurons in a circuit will result in computationally interesting abilities. There is also no good evolutionary reason why a creature should evolve a metabolically very expensive large brain in the hope of some useful ability evolving at a later date.

The second type of competing explanation of the appearance of language phylogenetically sees language as a side-effect of other evolutionary events. It is perhaps instructive to discuss the status of the side-effects of other evolutionary changes within the theory of evolution. Gould and Lewontin (1979) speak of natural selection being replaced by exaptation. This is the process where a side-effect of some other evolutionary event proves to be useful and thus has reproductive advantage and is then maintained. This type of situation is an example of an unmodified side-effect. Natural selection would play no role here.

An uninteresting case would be where side-effects are not involved in any aspect of behaviour - such as redness of blood. Much more interesting is the case Gould is arguing for here; whereby side effects are modified and put to use. Natural selection plays a central role in these cases. A designer is needed and that designer is natural selection.

It is hard to see how language could be an exaptation. How could one side-effect of another evolutionary event result in a system as complex as language?⁴ The complexity of design in the language system lends far more credence to an account based squarely on the more central process of evolution, namely, natural selection⁵.

The only elements adopted by Neo-Lamarckians are basically two: the inheritance of acquired characteristics, and the principle of use and disuse. The principle of use and disuse states that those parts of an organism's body that are used grow larger. Those parts that are not used tend to wither away. All evidence suggests that the idea that acquired characteristics are inherited is false.

Having discounted these alternative explanations of the phylogenetic development of language, one may think that the development of language by natural selection has been established. However, there have been arguments against a natural selection account and these also have to be shown to be spurious.

⁴ Although this discounting of language as an exaptation, of course, depends on how complex the language system as a whole may be. It is perfectly plausible in my view that parts of the language system may be exaptations of other parts of our cognitive apparatus. For example, later in this chapter, we will see that vervet monkeys have a hierarchical representation of social relations. This representation may well have been exapted to represent linguistic syntax.

⁵ Although, here again, there is another interpretation where natural selection acts on a series of exaptations to produce the language system.

The first objection is that grammatical devices and expressive functions do not pair up in a one-to-one fashion. For example, some languages use word order to convey who did what to whom; others use case or agreement for this purpose and reserve word order to distinguish topic from comment. How can one say that the mental devices governing word order evolved under selection pressure from expressing grammatical relations if many languages do not use them for that purpose?

The answer to this objection is that the evolution of structures that serve not one, but a small number of definite functions, perhaps to different extents in different environments, is common to biology. Indeed, though grammatical devices are put to different uses in different languages, the possible pairings are very circumscribed. No language uses noun affixes to express tense. Equally, other aspects of cognition may explain the circumscriptions in the language system.

Even a rudimentary grammatical analysis reveals that surface diversity is often a manifestation of minor differences in the underlying mental grammars. When one looks at more abstract linguistic analyses, the underlying unity of natural language is even more apparent. Chomsky has suggested that anything you find in one language can also be found in every other language, perhaps at a more abstract level of representation.

Piattelli-Palmarani (1989) presents a different kind of argument: Grammar is not completely predictable as an adaption to communication, therefore, he claims, it lacks design and did not evolve by selection. This argument takes two forms:

- Language could be better than it is.
- Language could be different to what it is.

However, even if it could be shown that one part of a language had no function and the structure of language was therefore not optimal, that would not mean that all parts of language had no function. Indeed, many constraints that at the moment seem to have no real basis could later be shown to be part of more underlying principles.

Further, the idea that natural selection aspires towards perfection has long been discarded in evolutionary theory. Trade-offs among conflicting adaptive goals are a ubiquitous limitation on optimality in the design of organisms. The eye is a classic example. The blind-spot is caused by much of the wiring of the eye being on the front instead of the back. Wiring on the back would clearly be a more perfect design, but the evolutionary history of the eye has resulted in the extant structure.

Trade-offs of utility within language are also unavoidable. There is a conflict of interest between speaker and hearer. Speakers want to minimize articulatory effort and hence tend toward brevity and phonological reduction. Hearers want to minimize the effort of understanding and hence desire explicitness and clarity. Clearly any shared system of communication is going to have to adopt a code that is a compromise among these demands, and so will appear to be arbitrary from the point of view of any one criterion.

The second form of this type of argument was that the system of language could be different to the way that it actually is. The fact that we can conceive of a biological system being different from the way it actually is says nothing about whether it is an adaptation.

With respect to language, any communicative system requires a coding protocol that can be arbitrary as long as it is shared. This is the requirement of parity. In the evolution of the language faculty, many arbitrary constraints may have been selected simply because they defined parts of a standardized communicative code in the brains of some critical mass of speakers. This does not mean that the language system is optimum to any extent, as Chomsky (1994) seems to imply.

To be sure, some combination of historical accidents, epiphenomena of other cognitive processes, and neurodevelopmental constraints must have played a large role in the breaking of the symmetry that was needed to get the grammar fixation process running away in one direction or another. But it still must have been selection that resulted in the convention then becoming deeply entrenched.

In this section, following Pinker and Bloom, I have presented arguments for the phylogenetic development of language based on natural selection. These arguments show that the language system with its most general properties shows the same kind of complex design as the eye. The arguments against natural selection have been shown to be spurious. Alternative theories for the emergence of language have also been shown to be unlikely. Although I agree with the general thrust of Pinker and Bloom's arguments, I do not agree that what has evolved is the theory of Principles and Parameters as they seem to suggest, rather it is the general properties of language that have evolved, what language expresses, to what uses it is put, and to some extent, the devices language uses to express the conceptualizations we have as humans. Pinker would like the evolved state of language to mirror PPT exactly as he believes in the Continuity Hypothesis (Pinker (1984)) of the language system throughout the child's development. In other words, all the principles of PPT are available to the child innately. The Continuity Hypothesis has been shown to be dubious in Chapter 3 so it is not necessary to make the extreme assumptions that Pinker and Bloom have made.

6.5 Scenarios of Language Evolution

In the section on human evolution we saw that socialization required a more efficient form of communication. Bickerton (1991) argues that language cannot be seen purely as a system of communication but also has to be seen as a system of representation and that we cannot study the function of language without also studying the structure that underlies that function. Language he argues can be nothing but an evolutionary adaptation and as such it must have antecedents of some kind, both social and structural.

Bickerton sees the difference between human communication and animal communication as a freeing up of referential abilities. The adaptive function of communication for other species is a specific set of referential abilities that rely on calls for things for which sensory evidence exists, whereas for humans the adaptive advantage comes with a system of reference as a whole.

However, we cannot have a representation without a medium in which to represent it. The representations are thus constrained by the nature of the medium that represents them i.e the human brain. Representations for any creature are not veridical. Thus the nature of the representations of dogs (with a large range of olfactory perception) and bats (with a large range

of auditory perception) and humans is constrained by the evolutionary needs and ecological niche of each species and is reflected in the form and nature of their input devices and neural structure. In other words the categories we can distinguish and represent are determined not by the nature of reality but by the nature of our nervous system and what we need to survive and reproduce.

This has often been found difficult to accept but if it seems easy to accept that bats or insects have adapted to fit their environment, why should it not be equally easy to accept this for humans? We may be a lot more complex, but we have our evolutionarily determined limitations as well.

Bickerton sees language as a representational system *par excellence*, but a representational system that has two sides: An atlas of reality which is reflected in semantics, and a method of traversing this atlas which is reflected in syntax. The nature of our mapping of reality is not arbitrary and relies on three mappings:

- Reality → Sensory Perception.
- Sensory Perception → Categorization.
- Categorization → Language.

Bickerton argues that we learn concepts on the basis of their functionality and that these concepts are encoded in semantic space which is constrained by what *we* think exists as required by what we, as a creature, want to be there. That is, there may well be any number of concepts in the world that we do not encode as they are not relevant to us as humans in our ecological niche. We only encode, and perhaps only perceive, the concepts that are relevant to us as humans.

We encode what we perceive in the world as a thing and that thing's behaviour. This distinction is unexpected as it corresponds to nothing in nature⁶. We could just as well have a holistic representation of such a situation. But the subject and predicate distinction is fundamental to us. This fundamental status could be because, phylogenetically, concepts of entities preceded concepts of behaviours. The functional utility of the later development of verbs could be that they represent chunks of behaviour that are functionally significant for our species⁷.

The ordered learning of concepts based on their phylogenetic age can be seen in the learning of colour terms. Berlin and Kay (1969) showed that colour concepts are learned in a fixed order. This fixed order reflects the stages of colour discrimination in all primates and specifically reflects their phylogenetic age.

This is a first example of the link between phylogeny and ontogeny. With respect to language it may be the case that some aspects of language's computational structure emerged earlier than others and will be of greater salience when it comes to language processing.

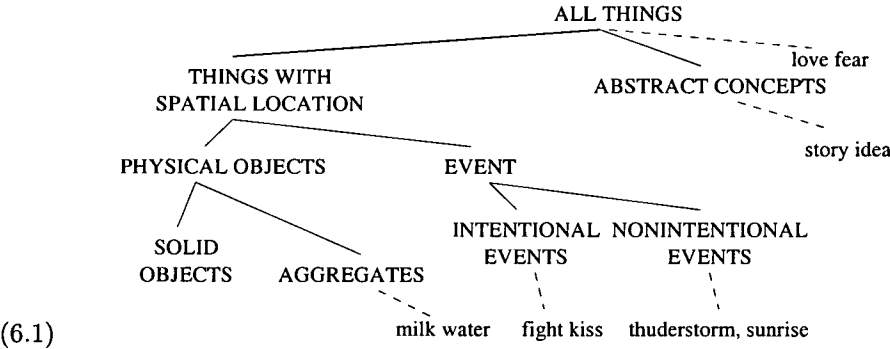
⁶Armstrong *et al.* (1995) propose a method by which we could have separated holistic events into subject and predicate. This method will be explained below.

⁷It could also reflect the fact that different brain structures are responsible for recognizing things and carrying out actions. If our representation of actions of others is parasitic upon an internal representation of an action in terms of its productions and its putative effects, then this representation would make the two distinct.

The semantic space in which we encode concepts is organized as a hierarchical structure and shows contiguity constraints. So, for example, only contiguous sectors of the colour continuum can be encoded. And in some languages there is a single verb for *existence*, *location*, *possession* and *ownership* but our conceptual system automatically links certain concepts together so that where languages use one verb to encode one or more of these relations we find no language which has a single verb for *location* and *possession* and another verb for *existence* and *ownership*.

There would seem, then, to be a limitation on how we as humans categorize the world. Cultural limitations on our conceptualization of the world will be reflected in the language we use. This would seem to be the Sapir-Whorf⁸ hypothesis in reverse.

So the conceptual system is constrained by the way we perceive the world as determined by our ecological niche. In addition to this, the syntax that allows us to cross this conceptual map in forming sentences is itself constrained by the form of the conceptual map. So for example we can produce predicability trees (Keil (1981)) that represent what can be predicated by what. Thus a quality at the top of the tree can be predicated of any class beneath it.



These trees reflect the contiguity constraints in the conceptual map, thus if a property can be predicated of humans and plants, it must also be predicable of mammals. These trees of concepts also appear to be acquired in a certain order reflecting innate discriminatory saliences.

Half the words in a language as it is produced are grammatical items. These grammatical items do not mirror anything specific in experience but they allow things to be plotted relative to the observer in space, time, number and so on. There seems to be a short list of grammatical items but an infinite number of relations in the world.

Turning to syntax itself, we can see a sentence as a play with roles being filled, the number of roles is small. The roles are particularly important as, if anything is consistent across languages, it is the roles that languages are capable of expressing. The same cannot be said of tense, formal syntactic structure or the semantic range of words. And to understand a sentence and to integrate it into discourse, it is essential to work out the thematic relations that the sentence encodes.

The less contentful notions of syntactic structure such as Case and Government can be regarded as a back-up system for the efficient processing of the products of the argument

⁸The Sapir-Whorf hypothesis claims that our view of the world is influenced and shaped by the particular language we speak.

structure system and should thus be seen as subservient to the thematic relations and not vice versa.

Bickerton gives the following definition of Representation:

Responding or having permanent propensity to respond to x, an entity or event in the external world, in terms of y, a particular pattern of neural activity.

Simple representations evolved so that creatures could respond to the environment. Initial unrepresented responses can be seen in plants where the same cells that receive the stimulation are the ones that respond. The first step to representation comes when it is different cells that receive and respond. Thus in insectivorous plants, one cell notices contact with the plant and a different cell causes the leaves to trap the insect. However, there is no element of learning at this stage.

The next stage is the freeing up of stimulus and response. If two types of sensory cell are linked to one intermediate cell the creature can now give a number of responses.

To categorize these perceptions a process of selection takes place determined by species membership. It is possible that some categories are innate but their number will be limited as they will be linked to a single response and will therefore be inflexible. There are thus three stages in concept formation.

- Object in External World.
- Patterns of cell activity triggered by that object.
- The observer's internal and external responses to these patterns.

Bickerton proposes that the first step to language was probably a proto-language with a form analogous to the language of the under-twos, feral-children and pidgin languages. He claims all of these have features in common. The differences between proto-language and real language are that there is a low incidence of grammatical items, no attention is paid to subcategorization requirements, any recursion that occurs has been learnt in a rote fashion, word order changes are not tied to emphasis or presupposition and the use of null elements, or more accurately, implicit arguments, is unconstrained.

Bickerton is thus making a claim about the nature of language acquisition in that if children's early production is based on proto-language it may be that they have no true language grammar and no neural structures to support such a grammar until they are around two years of age. Combinations before this age are an exercise in semantics and not syntax. If this is the case it may be expected that the combinations are constrained in some way, although Bickerton does not address this.

The nature of the first step to language and the motivations for it have to be explained. There are three possibilities as to the nature of the change: syntax before lexical, lexical before syntax or simultaneous. Simultaneous evolution is unlikely according to Bickerton as it is clear that proto-language can operate largely by ignoring syntax. Syntax can only have evolved first if it performed some other established function. Bickerton discounts this argument as he feels syntax needs a lexicon to operate. This may well be true for language but

it does not preclude the possibility that before language and lexica there was some syntax-like computational process taking place in the brain that was later co-opted for language.

With respect to human evolution and the time course of the first step to language, Bickerton's position seems rather confused. He does not believe that language is connected with tool use as linguistic structure is hierarchical and tool use is serial, although he does concede this connection would be enough for proto-language. As he is here explaining the emergence of proto-language, his claim is odd.

He disputes any claim that sign-language had a role as it complicates linguistic history, it does not answer how syntax developed or how elements acquired reference and children do not go through a signing stage. All of these facts are hotly disputed by Armstrong *et al.* (1995) as we shall see later in this section. Indeed, when Bickerton turns his attention to our ecological readiness for language he uses a similar argument to Armstrong *et al.* in saying that the freeing of the hands allowed for an enrichment of neural connections with cross-modal connections being made between the visual the auditory and the tactile - precisely the kinds of interconnections he argued made categorization possible. Other factors Bickerton considers relevant are the formation of social groups with varied eating habits over long ranges, requiring complex mapping abilities (found in many creatures without language) and long pursuit of prey requiring a planning ability.

The most contentious part of the claims of Bickerton (1991) about the evolution of language is his view that the change from proto-language to true language was a catastrophic, all-at-once event, and not an event that took place gradually over time. Curiously he uses arguments from the fully formed language system to back this up. Thus, he claims, in the course of a child's acquisition of language, the steps from the proto-language stage to the true language stage will show the clustered appearance of the systematic expansion of structure, the obligatory expression of subcategorized elements and the automatic identification of null elements. It suffices now to say they are dubious. In any case it is unclear whether the emergence of language is informed by such 'facts'.

A second curious claim against gradualism that Bickerton uses is that there is not gradualism in the transition from pidgins to creoles. The first generation of children of immigrants grammaticalize the pidgin into a full language, a *créole*. Thus he identifies a species-specific biological characteristic - the ability to create a language in the absence of a specific model. This is well and good, but it explains a competence of the *evolved* linguistic system and says nothing about its evolution.

When he turns to phylogeny itself, he points out that evolutionary abruptness is different from everyday abruptness - although his previous two arguments were based on everyday abruptness. In other words evolutionary abruptness is over a relatively long time - gradualness in disguise, in other words. However, he points to the sudden change in the paleontological record in the Upper Pleistocene as evidence for the sudden emergence of language. The sudden increase in artifacts may however have come well after language evolved, and why should syntactic language and not proto-language be the relevant trigger?

In other words Bickerton is looking for a crucial single mutation, that could impose a hierarchical structure on language. This hierarchical structure being the three level X-bar theory. The mutation he argues for is a linking of the lexicon brain area to the brain area

where the structure of events is analyzed.

The structure of events is indeed important and will form a major part of the representation I have adopted for thematic-conceptual structure.

There are a number of comments that can be made about this scenario. First, it assumes that the structure that evolution moved towards is the structure that contemporary linguistic theory claims to be correct. This seems to be argumentation in reverse. We should be looking for what evolution can tell us about the structure of language and not what a highly abstract theory of linguistic structure can tell us about evolution. Second, it is unclear what kind of single *genetic* mutation can have resulted in such a sudden linking of two supposedly separate functional areas of the brain.

There are a number of ideas that we can take from Bickerton's work. He has stressed, and it is important to realize, that cognition is species-specific and constrained by its ecological niche. We should not be afraid to apply this insight to humans. Thus the constraints on our conceptualizations are reflected in the structure of language.

Bickerton sees the linking of event structure to the lexicon as being of vital importance in the evolution of language and this linking of event structure and lexicon would appear to be much more plausible than the other element of his approach, a catastrophic shift from proto-language to true language.

There is also the expectation that language will have societal and structural antecedents. To find evidence for these I will discuss evidence from research into primates. Also of relevance are models of human society as can be found in the work of Dunbar (1993).

Armstrong *et al.* (1995) take an almost diametrically opposite position to Bickerton. Not only do they see language evolution as a gradual process, but they see gesture as *the* major step towards language. For them, sign language is a biologically related homologue to spoken language as they share the same evolutionary history.

They state a number of reasons why gesture and language cannot be separated. Kendon (1991) and McNeill (1985) have shown that gestures that accompany speech are very closely integrated into the timing of the speech act and are therefore probably being generated by the same processes. Intonation, also, seems to straddle gesture and language.

The outline of their theory and general approach is the same as the line that I am taking in this dissertation. That is, they are suspicious of taking the formal categories proposed by contemporary linguistics as the starting point and arguing back from these to possible precursors and alleged brain mechanisms. Instead we start with the raw materials of communication in humans and non-humans.

Humans working in groups led to a selective advantage and communication within the group led to a further advantage. The specific emphasis on gesture that is made in their book is justified by their view of humans as primarily visual creatures. Thus visual language (gesture) is more likely to have been a viable first step to language than invisible language (speech).

They propose that the key to building syntax incrementally is the discovery of relationships within symbols and that the germ of this discovery is inherent in the form of simple visible gestures. They claim that a basic sign has a structure where something does something to something else, or in other words SVO, the basis of syntax. Getting from a sign like this

to syntax is an exercise in analysis and not synthesis. Syntax is formed, then, by taking a gesture apart rather than stumbling on a large number of logical categories that combine spoken words.

The fact that a gesture involves a hand and its movement allows for early humans to see that what acts, and the action it performs, are distinct events as it is learning internal to the creature. In some ways, it could be argued that the ordering of signs and their form are the natural order of things - indeed hominids may have dimly perceived this from the external world. However, it is another step to symbolize them, and this symbolization is more likely to have come about by neural connections based on internal bodily functions rather than by an analysis of the external world.

The next step of course was to equate the relations in a gestural 'sentence' with relations in the outer world. This could be achieved by pairing visible events in the world with similarly structured visible gestures.

Up to this point of the argument, the syntax that the human being has achieved is very basic. There has been no explanation of how adverbs, adjectives and subordination could have entered into the system. In contemporary sign language such elements can be added by simultaneous expression, thus the possibility of modification was also inherent in the structure of visible gestures.

By taking this approach to language evolution, it is argued, we can see a much more obvious link between handedness and left-hemisphere language preferences. They are both forms of praxis, where complex motor actions are planned. It has been shown that there is a clear preference for handedness in gesture and that in deaf children the preference for handedness in gesture precedes the handedness preference for object manipulation (Bonvillian and Richards (1993)).

The view that the basis of language is in its bodily realization has consequences for the conception of the form of grammar. The body is located in time and space and grammar will reflect this. The Spatialization of Form Hypothesis (Lakoff (1987)) argues that grammar is ultimately spatial and makes a number of predictions.

- The acquisition of grammar takes place when linguistic information is routed to, and processed by, the spatial centres of the brain.
- Linguistic expressions are processed in terms of image schemas.
- Image schemas are embodied schemas which represent cognitive models of the body and its interaction with the environment.

It is easy to see how this can be related to language having its first step in gesture. It also has implications for the role of the body in perceptual categorization. Armstrong *et al* adopt the learning theory of Edelman (1987), which argues that categorization requires at least two separate sensory channels which each supplies signals to neural maps. These maps interact with each other in a process termed *reentry* by Edelman.

These reentrant maps interact in turn with non-mapped regions (frontal lobes, basal ganglia and cerebellum) to form global maps which result in categorizations.

Thus, in learning, embodied action is involved. Cognition depends on the kinds of experience that come from having bodies with various sensorimotor capacities. Perception consists of perceptually guided action and the cognitive structures that are formed from this action emerge from recurrent sensorimotor patterns that enable action to be perceptually guided (Varela *et al.* (1991)). If this is true, the cognitive structures underlying language should also have emerged from perceptually guided action. This action is argued to be gesture.

Thus the global mappings that can be made given the nature of our perceptual systems will be the basis of the form of language. The global mappings that are formed by correlating objects, actions and spatial and temporal relations will be the best conception of thematic-conceptual relations we can arrive at. These thematic-conceptual relations will be phylogenetically and functionally determined and therefore will be language universal. Functional theories of linguistics such as Cognitive Grammar, Cognitive Semantics and Functional Grammar address this issue. However, they do not address the full form of language that involves great linguistic differences in the encoding of these thematic relations in *structure*.

Armstrong *et al.* also give us indications that humans are basically visual creatures. In addition to this visual basis, there is a tactile side to sign language that may well have left a vestige as an increased saliency of tactile metaphors in language content and structure. Such tactile relations would include contact, support, and so on. Concepts that we have seen in Chapter 3 appear early in cognitive development.

In general, what comes to be internalized and available for incorporation into the language system is constrained by our ecologically determined perception of the world.

6.6 Primate Behaviour and Primate Cognition

An analysis of primate behaviour and the clues it gives us about primate cognition is important as it provides evidence as to the cognitive abilities and processing and representation systems of our closest evolutionary relatives. We will be able to get an idea of the sorts of cognition that will also be particularly salient for human beings and of possible language-like computational processes that could have been co-opted for language processing and language representations.

Indeed, if it can be shown that primates have representations of some form or another and we know the sorts of behaviours these representations govern, we will have clues as to the types of behaviour and cognitive abilities that can lead to representations being formed. This in turn will allow us to reassess the various theories of categorization and concept formation we have encountered above.

We will not be confusing the account in this section with the literature on the specific *linguistic* capabilities of primates. We will address that issue in the next section as it will be able to give us an idea as to what *extra* is required over and above the normal abilities of primates to result in a full linguistic system. This section, however, as indicated above, seeks only to find nonlinguistic capabilities that had a computational form that could have later been useful for language.

Cheney and Seyfarth (1990) have looked at the behaviour of vervet monkeys in two main areas, social behaviour and vocal communication. They have found that vervets do not interact at random but base their actions on five features of vervet social interaction: kinship, dom-

inance, reciprocity, sexual attraction and defence. They identify three motives that underlie this behaviour: attraction to kin, deference to those of high rank and a desire to increase their own status. They argue that these motives make evolutionary sense as there are evolutionary benefits of kin selection, and deferring to those of higher rank as well as trying to increase one's own rank reflects a mixed strategy common in groups where high rank is correlated with reproductive success.

Vervet monkeys act as if they understand such relations as kinship and dominance, but it could be that we are imposing our categories on them. Cheney and Seyfarth go on to investigate whether there is evidence that monkeys represent this knowledge mentally.

Overall they found that a number of different species appear to observe interactions in which they are not themselves involved, and to recognize relationships that exist among others. For example, if females hear a juvenile's cry for help, they expect a response from that juvenile's mother, as evidenced by their looking at the juvenile's mother when they hear its call. Monkeys' classification of social relationships seems to be based on abstractions rather than on physical features or the number and type of interactions. The abstraction seems to be based on all these elements. They seem, in some way, to be using a kind of analogical reasoning that results in them representing hierarchies of individuals.

Cheney and Seyfarth argue that this kind of analogical reasoning seems similar to the sort of analogical reasoning that can be carried out by language-trained chimpanzees. What is important at the moment is that vervet monkeys show evidence that they can represent hierarchical relations. Hierarchical relations were what Bickerton saw as being the vital mutation in hominids that allowed for the development of a full syntactic language. If what Cheney and Seyfarth claim is accurate it would seem, however, that the hierarchical relations of language could have been based on a pre-existing system of hierarchical representation of society.

Turning to the vocal communication of vervet monkeys, Cheney and Seyfarth point out that it is impossible to tell whether a monkey deliberately intends to communicate or whether calls simply come out as part of their ongoing behaviour.

Vervet monkeys give different sounding alarm calls in response to at least three different predators: leopards, eagles and snakes. Each alarm elicits very different escape responses from other vervets nearby. It seemed then that vervets were using calls to denote different external referents. It was suggested that the difference in the calls was perhaps only due to the level of panic in the vervet making the call. Seyfarth *et al.* (1980) tested this by playing recordings of the three alarm calls adjusted for length and amplitude to vervets in the wild. They found that these alterations made no difference to the escape strategies that each of the calls elicited.

Vervet monkeys also grunt in at least four different social situations:

- When approaching a more dominant individual.
- When approaching a subordinate.
- Watching another animal or initiating a movement across an open plain.
- When a monkey has spotted members of another group.

In contrast to the strikingly different alarm calls, the grunts all sound similar. However, the experiments of Cheney and Seyfarth (1982) have shown that the vervets seem able to

distinguish the grunts. The grunts were played to the monkeys in different contexts and consistent differences in the responses to the grunts were observed, except in the case of grunts to a dominant male and grunts to a dominant female. Marler (1976) has suggested that dividing a graded stream of sounds into discrete categories is the most difficult problem to surmount on the course to language. However, the fact that vervets have so few grunts suggests that other constraints are at work and therefore that humans have an additional capacity to monkeys.

It now remains to examine to what extent the vocalizations of monkeys mean something to them. We have seen how vervets' vocalizations *function* in their society, but it is unclear up to now whether vervets recognize that a vocalization X *means* Y which would be based on an abstraction away from the function and the vocalization's acoustic properties.

Cheney and Seyfarth utilize the theory of intentional stance (Dennett (1987)), which identifies a number of levels of intentionality, to assess the meaning of vervets' vocalizations. The levels of intentionality are:

- Level 0 - no beliefs or desires. In other words, calls are given due to fear.
- Level 1 - Beliefs and desires but no beliefs *about* beliefs. The caller has no conception of the audience's state of mind. A vervet at this level gives a call because he believes a predator is near and wants others to react.
- Level 2 and beyond - The caller has some conception about both his own and other individuals' state of mind. Here a vervet would give a call because he wants others to believe there is a certain predator nearby.

It seems that vervets surpass Level 0 as they have some degree of voluntary control over their vocalizations, and can modify their calling rate dependent on their audience (higher rates with kin, males producing higher rates to a female than to another male).

With respect to the abstraction of meaning, research on chimpanzees (Premack (1976)) has shown that they can associate an abstract chip with its referent in the real world. Thus Premack's chimpanzee, Sarah, was taught that the symbol of a blue triangle referred to an apple and when asked to describe the properties of the symbol she described it as red and round. Sarah could also reverse this and describe the symbol for the apple as being triangular and blue. Based on their behaviour in the wild, it does not seem that vervets have this particular ability.

Armstrong *et al.* have contended that language originated in the mental links that could be made through representational visible gestures. Another line of research that is particularly interesting with respect to this contention is found in Perrett *et al.* (1989). They found that cerebral cortical cells in the superior temporal sulcus of macaque monkeys responded to visual displays of goal-centred activities of the type where human subjects would attribute causal and intentional relationships. Some cells respond selectively when viewing seven different actions: *reach for*, *retrieve*, *manipulate*, *pick*, *tear*, *present* and *hold*. Other cells differentiate agents of an action, thus there is a greater sensitivity for hand-object interactions compared to object-object interactions.

The conclusions we can reach in this section are diverse, but I believe they do shed light on the evolutionary theories of language we have previously considered. It seems that vervet monkeys can represent social relations in terms of hierarchies and that this was not therefore a mutation that was needed specifically for language. It was probably latent within the hominid line, as our social relationships were most likely to have been generally similar to other primates. This latent ability needed only to be built into the language system as it developed, that is, as other cognitive precursors were also built in. The interactions of these cognitive precursors in addition to their own computational style will have constrained the nature of the fully-formed linguistic system. It seems then, that Bickerton's idea of a crucial, catastrophic mutation is misguided⁹.

Some of what Bickerton has claimed is supported by the data from apes. Apes do seem able to identify concepts and to represent them and these conceptualizations and representations do indeed seem to be based on the creature's internal and external responses to activations in response to an object's presence.

Perrett's research would seem to bolster the claims of Armstrong *et al.* It also seems that certain goal-directed activities that are evolutionarily significant for a species can be preferentially neurally represented and differentiated. It may therefore be the case that certain evolutionarily determined goals of humans are more deeply embedded in their linguistic representation and that they will in turn be more easily encoded and decoded from the speech stream. This would be a more teleological approach to language than is fashionable.

6.7 Teaching Human Language to Apes

There have been six notable attempts to teach language to apes¹⁰. Gardner and Gardner (1969) attempted to teach Washoe, a chimpanzee. Premack (1976) worked with the chimpanzee Sarah. Rumbaugh (1977) worked with another chimp, Lana. An unusual study involved a gorilla named Koko (Patterson (1978)). Terrace (1979), who was sceptical about ape-language programs, worked with a chimpanzee, Nim Chimsky. Finally, Savage-Rumbaugh (1984) worked with a bonobo named Kanzi.

In all cases, except for the case of Kanzi, the results were disappointing. Much of this can be put down to the methodology of the experiments. It was recognized early on that the shape of an ape's vocal tract is not conducive to the articulation of language, and thus, other forms of linguistic communication were sought. Two main strategies were used. A form of American Sign Language (ASL) was taught to Washoe, Koko and Nim. Sarah and Lana used lexigrams on plastic chips or a keyboard to communicate. Each of these apes was taught intensively with much moulding and shaping of the ASL signs. Although none of the apes could be shown conclusively to have syntactic knowledge, they did have a number of language-like properties in common.

Washoe learnt 132 signs by 51 months and extended some of these signs to other referents. 78% of Washoe's two-word utterances fitted into seven of Brown's (1973) categories¹¹. Sarah

⁹As he now seems to believe himself (Bickerton (1996)).

¹⁰The references for each of these six attempts is the first relevant reference of the many produced for each study.

¹¹These categories were very similar to the semantically motivated thematic roles that Jackendoff proposed.

and Lana were able to learn categories other than basic-level categories. They could thus sort items out according to colour, or could learn the class-name for a set of objects, such as tools. Koko had learnt 250 items by 48 months and 75% of her multiword utterances could be classified into 12 of Brown's semantic relationships. She also coined new modifier-noun combinations. By the age of three Nim had learnt 125 signs, would overgeneralize and undergeneralize terms to referents and 84% of his multiword utterances could be classed into 8 of Brown's semantic categories.

On the face of it, these results seem impressive and, indeed, in a limited way they are useful to the aims of this thesis. Of particular interest is the fact that the utterances of chimps fit into a limited number of semantic relations in a way very similar to young children at the two-word stage. In general terms, however, these early studies into ape language have been regarded as being unsuccessful. To see why this is so, we will have to look into how the data have been reassessed by other researchers.

For example, an analysis of Lana's multiword strings reveals that all of them were variants of only six sequences that could have been learnt by conditioned discrimination coupled with paired-associate learning. A further problem was that the chips Lana used were coloured according to semantic category and the strings produced by Lana could have equally well been guided by a patterning of colours rather than any syntactic or even semantic understanding.

Terrace (1979) argued that the language taught to Sarah, or more accurately, the manner in which it had been taught made it not much different to what can be learnt by rats and pigeons. Thus strings were taught by getting Sarah to repeat strings and then replacing just one of the words in that string. There was thus no semanticity in the string, it was just a matter of learning orderings. The proposal that the language learnt by Sarah was not in any true sense language-like is bolstered by the finding that people with profound language-learning deficits could learn the 'language' that Sarah had learnt.

It has also been argued (Wallman (1992)) that the vocabulary counts attested for the apes have been inflated due to over-interpretation of natural gestures that apes make. It may also be the case that apes are not referring but are merely making signs that have become associated with rewards¹². Thus, except in the case of Kanzi, there is a virtual absence of displaced reference, i.e. reference to an object that is not present at the time of utterance.

Turning to syntactic ability in general it has been argued (Wallman (1992)) that much of what seems to be syntactic ability has been over-interpretation of the data. So, for example, many 'two-word' utterances are in fact a combination of a taught sign/symbol and a natural ape gesture. Also, although many of the utterances could be categorized into a small number of semantic relations, there seemed to be no syntactic patterning of these relations, so that differing word orders were used (by the same chimp) to express the same semantic relation. What word order is shown by chimps in these early studies can be argued to be a result of assigning words to a stable sentence position across utterances, and not in any way due to its semantic category.

Wallman concludes with respect to these studies and earlier diary studies of Kanzi that:

See Chapter 2.

¹²This interpretation does not seem to be accurate when we consider the representational abilities of vervet monkeys, or the results of a study on Kanzi that is described below.

Given that Kanzi knows the meaning of these words or at least has forged appropriate associations between these words and their corresponding things or actions, and given he is able to 'select the critical words' and 'respond appropriately', there is only one way we should expect Kanzi to respond to the coupling of *grab* with *Jeannine* or to the combination of *give*, *trash* and *Jeannine*.

Given the results of a more carefully controlled experiment (as can be seen below), this conclusion seems to have been hasty.

Savage-Rumbaugh *et al.* (1993) attempted to counter criticisms such as Wallman's by addressing the issue of syntax head-on in a controlled setting. The main difference in the approach is that it examines what the ape can comprehend and not what he can produce. The study in question also compared the comprehension abilities of Kanzi at age 8 with the comprehension abilities of a human child at age 2.

A vital difference can be found in the way Kanzi acquired language. In each of the earlier attempts to teach language to an ape, explicit instruction was used. The same was true for Kanzi's mother, Matata. And in each of the cases where explicit instruction was used, the results were disappointing. Kanzi differs from this in that he learnt language in the way that a child learns language, without explicit instruction. Kanzi observed his mother being taught language from the age of 6 months to the age of 2 years 6 months. Kanzi was thus younger than most of the apes used in earlier studies. This combination of starting young and non-explicit learning of language appears to have resulted in much better language learning in Kanzi. When Kanzi's mother was taken away from Kanzi for breeding purposes, Kanzi showed a good knowledge of the lexigrams the researchers had been attempting to teach Matata.

Kanzi was soon producing multiword utterances. Most of these were requests, as was usually the case for the other apes, but some were comments on events. What Kanzi appeared to have done was to learn through a combination of observation and social interaction. At age 6 when Kanzi was tested on his comprehension of 310 multiword utterances, he was correct on 298 of them. However, this is the experiment that produced the sceptical quote from Wallman.

In the study in question here, both the human and the ape subject were tested in surroundings that were familiar to them. There was an initial pretest to see that each of the individual words to be used in the studies was known to the subjects. Sentences were spoken to the subjects in both blind and non-blind trials. In the blind trials, the sentences were read by a researcher hidden behind a one way-mirror. The aim of the blind trials was to prevent cuing from the researchers that would guide the subject's actions¹³. However, the subjects were initially ill at ease with the blind trials, so non-blind trials were used to overcome this unease. Objects were present in arrays in front of the subjects, or in other rooms¹⁴, as relevant to the sentence being tested. Sentences were generated by randomly selecting objects, locations and agents. The structure of the sentences was varied systematically. Multiple sentences were used for each of the possible arrays of objects. Unusual nontypical actions were often expressed by the sentences e.g. *Wash the hot-dogs*.

¹³Terrace had argued that much of the behaviour that Nim demonstrated was due to the humans around him inadvertently cuing Nim towards the correct behaviour. This is known as a Clever Hans effect.

¹⁴The experimental design included two rooms. In each room there was a selection of objects. Some objects were only in one room. Other objects were duplicated in each of the rooms.

In sum Alia (the child) took part in 180 non-blind trials and 410 blind trials and Kanzi took part in 244 non-blind trials and 416 blind trials.

The sentence constructions used were (with examples):

- (6.2) a. Put object X in/on transportable object Y (*Put the ball on the pine needles*)
b. Put object X in non-transportable object Y (*Put the ice water in the potty*)
- (6.3) a. Give (or show) object X to animate A (*Give the lighter to Rose*)/pretend animates were also used, e.g. toy dog.
b. Give object X and object Y to animate A (*Give the peas and the potatoes to Kelly*)
c. (Do) action A on animate A (*Give Rose a hug*)
d. Do action A on animate A with object X (*Get Rose with the snake*)
- (6.4) Do Action A on object X (with object Y) (*Knife the sweet potato*)
- (6.5) Announce information (*The surprise is hiding in the dishwasher*)
- (6.6) a. Take object X to location Y (*Take the snake outdoors*).
b. Go to location Y and get object X (*Go to the fridge and get a banana*)
c. Go get object X that's in location Y (*Go get the carrot that's in the microwave*)
- (6.7) Make pretend animate A do action on recipient Y (*Make the doggie bite the snake*).
- (6.8) All other sentence types.

Overall Kanzi was correct on 74% of all trials and 74% of the blind trials. Alia was correct on 66% of all trials and 65% of the blind trials. These overall results break down by sentence type in the following way. Each of the figures refers to the percentage of correct responses to each sentence type.

	Kanzi		Alia	
	<i>All</i>	<i>Blind</i>	<i>All</i>	<i>Blind</i>
4.1A	64	63	72	73
4.1B	74	77	72	71
4.2A	81	78	81	84
4.2B	33	37	57	57
4.2C	89	91	87	91
4.2D	71	76	60	61
4.3	70	82	65	63
4.4	63	67	85	83
4.5A	75	78	69	71
4.5B	85	82	51	45
4.5C	77	77	52	52
4.6	70	67	60	56
4.7	64	78	30	33

We see that the ape and the child are very similar in their abilities. Indeed Kanzi does better than Alia for sentence types 4.5B and 4.5C. The sentences were varied enough in structure and sequential ordering to discount any claims that apparent syntactic ability was a mere side-effect of learning recurrent positions for particular lexical items, or that particular items were associated with a single action. Overall, there were three word-order manipulations. (A) verb and word order changes and responses differ (*Could you take the pine needles outside/Go outside and get the pine needles*), (B) word order constant and response differs (*Take the rock outdoors/Go get the rock that's outside*), (C) word order changes and response changes (*Put the juice on the egg/Put the egg in the juice*).

The results with respect to these types of sentences were:

	Kanzi	Alia
Type A		
Sentences	83	59
Pairs	74	38
Type B		
Sentences	79	67
Pairs	57	38
Type C		
Sentences	79	69
Pairs	57	39

Although errors on comparable pairs of sentence were higher, most of these errors were not inversion errors (reversing the direction of causality of the action), but were rather semantic errors (selecting the wrong item).

The problems of many of the early investigations of ape language parallel the problems of neural net approaches in many ways. Both approaches are basically behaviourist in outlook in that they assume a bare minimum of cognitive apparatus which comprises perceptual mechanisms, a capacity to represent in durable format the results of perception, and the ability to form associations between those representations. They see language as an example of verbal behaviour.

These approaches in their minimality make the assumption that there are no species specific mechanisms involved, resulting in cross-species homogeneity. Both approaches also rely on intensive input and training, which is not present for the learning of language in humans and seems to have been counterproductive in attempts to teach language to apes.

6.8 Cultural Learning and the Cultural Evolution of Language

We have been attempting to constrain both the computational form of the language processor and to constrain the content of the conceptual representation that the processor acts on by considering possible evolutionary precursors of language. We have done this in two ways. We

have examined the history of human evolution and we have investigated the cognitive and language-like abilities of the primates.

Recent work on Cultural Learning (Tomasello *et al.* (1993)) sheds further light on both these issues and identifies an alternative source of data on what is required for the language system to operate.

If we look at the evidence we have seen thus far we might like to conclude that the primary evolutionary context in which primate cognition evolved was a social one, we may be able to isolate an element of that social context that is specific to humans and made language possible, and in so doing added its own particular flavour to the nature of language.

Tomasello *et al* propose that this vital element was cultural learning. They note that other creatures may live in social groups but only humans have culture. They identify three types of learning that culture makes possible and claim that in ontogeny these three types of cultural learning appear in a particular order because they are dependent on the socio-cognitive bases that underlie them.

Basic to all forms of cultural learning is that the learner learns not from another person but *through* them, i.e. they try to see the relevant situation in the way the person they are learning through sees it. The three types of cultural learning are imitative, instructed and collaborative.

In *imitative learning* the infant reproduces the adult's actual behavioural strategies in an appropriate functional context. This ability emerges towards the end of the infant's first year in the domains of object-directed actions and the use of communicative symbols. The socio-cognitive bases of this ability can be isolated if we examine the other abilities a child shows at this age. At this age infants are both following the gaze of adults and trying to focus other people's attention by alternating gaze between a person and the object of interest. Thus they are showing understanding of people as intentional agents.

In *instinctive learning* the child learns about the adult's understanding of the task and how that compares to the child's own understanding. Children re-enact the adult's instructions in regulating their own behaviour in a similar task. The socio-cognitive bases of such learning is *inter-subjectivity*. The child shows an understanding that other people are mental beings with their own individual beliefs. This awareness emerges at around four years of age.

At six or seven years of age children understand complex, second-order mental states. This provides the socio-cognitive basis for *collaborative learning*. This manifests the same inter-subjectivity as instructed learning, but neither interactant is an authority or expert; inter-subjectivity is symmetric. Thus learning at this stage is often the result of socio-cognitive conflict - a classic example being learning in the conservation task.

The contention is thus that cultural learning is absolutely crucial for the acquisition of many of the most important cognitive skills including, importantly for our aims, language. It is therefore predicted that a loss or diminution of these skills will result in language deficit. By examining this deficit we can gauge what the cultural element contributes to the language system.

Here we have discovered another source of evidence as regards the cognitive precursors or contributors to language. If we look at the language learning abilities of children with particular cognitive or sensory deficits we will be able to see what each of the functioning

cognitive or sensory systems contributes to language content and language processing. An example specific to cultural learning is the case of autistic children (Fay (1993)). The common denominator in the diagnoses of autistic children is their inability to relate to others. Given this it is not surprising that they show little evidence of cultural learning and correspondingly the diminution of capacities that it has been argued depend on cultural learning.

Thus we find that autistic children do not collaborate in collaborative learning, nor instructed learning, and even with respect to imitative learning, autistic children show difficulty in reproducing any novel behaviour. Half never learn language at all, and those that do, do not use it normally. By looking at the differences between the language of autistic children and normal children we can see what cultural learning contributes to normal language use.

With the cultural perspective in place we can examine another theory of the evolution of language - in some ways similar to Armstrong *et al* - that relies on a cultural basis.

Donald (1991) has provided us with one model for the transition from ape-like hominids to *homo sapiens* with language. His theory is based on a series of major adaptations, each of which results in a new representational system. However, each major adaptation does not totally wipe out the properties of the earlier adaptations. As we have seen evolutionary theory results in what may be described as an 'if it works, don't fix it' phenomenon. Thus each earlier representational system will be left intact and cognitive vestiges will still be detectable today.

In this classification of the development of cognition with respect to representational strategy the starting point is *Episodic Culture*. Within this system the apes' lives are lived entirely in the present and the highest element in the system of memory representation is at the level of event representation.

The first major adaptation is the transition from episodic culture to *Mimetic Culture*. Mimetic skill rests on the ability to produce conscious, self-initiated acts that are intentional but not linguistic. These acts are defined with respect to their representational function. Reflexive and instinctual locomotor acts are thus excluded. Mimetic skill is prior to language but shares many characteristics that are commonly held to be essential for language, namely: Intentionality, generativity, communicativity, reference and autocuing (self-initiation).

These mimetic skills resulted in the sharing of knowledge and permitted such evolutionary advantageous behaviours as the modeling of social structure, reciprocal mimetic games, conformity and communication within the society, group mimetic acts and pedagogy.

If the theory is correct the mimetic adaptation which characterized *erectus* should have left vestiges in its descendants and the mimetic culture which allowed for the advantageous behaviours listed above would continue to be useful, even after the adoption of language. And indeed the observations of Eibl-Eibesfeld (1989) would seem to back up the claim. He has documented the commonalities of human non-verbal expression and has shown that many patterns of communication recur in all human societies, from hunter-gatherer to modern.

The previous representational system (episodic) is however encapsulated by the newer (mimetic) and the mimetic controller thus becomes the unencapsulated central system. In the absence of language, the mimetic controller remains the dominant representational device.

The second adaptation involves the transition from mimetic culture to *Mythic Culture*, where language appears. Donald sees speech and language as fulfilling the role of the sort of adaptation that could explain the explosion of tools, artifacts and inventions of all sorts and

the eventual creation and maintenance of tribal, political and social structures which regulated the society. Donald claims that all of these features rely on language to some extent.

It seems that in surviving Stone Age societies, language greatly outstrips technology. Technology in these societies is primitive, while language in social contexts soars to great heights. The most elevated use of language in these societies is in the area of mythic invention, that is, in the construction of models of the human universe.

Perhaps then language's initial function was tied to the development of integrative thought. Time-bounded snippets of information could be synthesized by the use of language. The invention of symbols characteristic of the mimetic phase can be linked to the mythic phase by considering the case of linguistic gesticulation.(cf. Armstrong *et al.* (1995))

The use of this type of gesture is sometimes lost in aphasia, whereas the use of mime and emblematic gesture are seldom lost. There are three subtypes of linguistic gesture. The first of these is iconic gesturing where images are traced in the air which relate to the meaning of the utterance. McNeill and Levy (1982) studied this type of gesturing in six narrators, producing a complex that had 44 movement features and 38 meaning features. Thus, the gestures were not completely arbitrary and had a predictable relation to the meaning of the utterance.

The second type of linguistic gesture are *metaphoric gestures* where the link to meaning is more abstract, and the third type are *beats*. Beats have no content of their own but indicate an extra narrative content. A beat might show for example that the material it accompanies is not part of the main narrative.

The loss of gesture in aphasia reinforces the independence of the mimetic and linguistic systems. Broca's aphasics continue to produce numerous and elaborate iconic gestures, but no beats. Wernicke's aphasics seem to lose iconix and metaphorix but retain beats.

Language could not have developed without changes in the mental models held by individuals; the processor defining the world is bringing mental models under symbolic control. In other words symbolic thought is primary; it is the driving force behind word use. So from an evolutionary perspective, the language system brought not only a new vocal apparatus but a wholly new system for representing reality. The propositional storage system and verbal semantic memory were products of language and form part of one adaptation, unified under a linguistic controller which encapsulates the mimetic controller.

Donald has given us an explanation of how language was a useful ability from an evolutionary standpoint. He has also outlined how the evolution of language could have occurred in steps each of which had a selectional advantage. A vital point to understand is that language goes hand in hand with a new way to represent reality. The most salient idea with respect to the approach of this thesis is that of cognitive vestiges. Earlier representational systems will still play a part in performance, and many aspects of those earlier systems may well form the backbone of the current language system.

It would seem that mythic culture is the vital step that allows cultural learning. Apes do not appear to have mythic culture, and, in the wild, do not learn language. When they are in an environment where cultural learning is available, however, limited language learning does appear to be possible, as we have witnessed with Kanzi above.

6.9 Chapter Summary

In this chapter I have started to outline the philosophical background to thematic-conceptual structure. I presented the latest theory of human evolution and a number of evolutionary scenarios for language that would appear to be impossible. However, out of this discussion a number of elements that would seem to be implicated in any explanation of language evolution emerged. These were society, brain size, bipedalism and laterality.

I then explained why an adaptationist evolutionary explanation of language emergence is justified before turning to promising theories of language evolution.

A vital link between the hierarchical structure, which we saw was present in vervet monkeys as the representation of their society, and language was argued to be vital by Bickerton. It seems that the analysis of events made possible by the internalization of causal structure in gesture may be the explanation of this linking of representational types.

The central role of gesture also allows us to explain a cluster of properties that seem to have emerged concurrently with language: bipedalism and laterality. Bipedalism freed up the hands to make gesture possible. Language and gesture both involve praxis and this praxis is associated with laterality; the same laterality for language and handedness being a natural consequence of this link¹⁵.

Gesture involves a representation of visual events and their conversion into tactile terms. Thus, visual and tactile metaphors and the way that hands can represent them may have determined the structure of language in its early stages and are likely to be important metaphors for linguistic expression today.

The fact that the hierarchical representation co-opted for language structure originated in social representation is provocative when we consider that we have found evidence that the development of this society into a culture was vital for the emergence of language. In this vein, it is also provocative that when a bonobo ape (with a society such as ours) is exposed to cultural learning, it is able to learn language to a limited extent.

We have also seen that constraints on our cognition, or how we understand the world, as determined by the structure of our brain, its computational properties, and the nature of our representations, are likely to have constrained the structure of language.

I propose a pair of hypotheses that can be drawn from the discussion in this chapter. The first is that much of the structure of language will be determined by its evolution. That is, there will be cognitive vestiges of earlier stages that will still be present today. These vestiges may be in computational, content, structural, or cultural/societal terms.

The second hypothesis is that ontogeny in many respects will reflect phylogeny. We have seen that this is the case for colour perception. I see no reason that this should not be the case for language. In the discussion of language acquisition by children, it should be examined whether the order that children learn particular aspects of language reflects the course of evolution of language.

¹⁵Brain lateralization and other non-cognitive lateralizations go way down evolutionary time to simple organisms.

Bibliography

- Abney, S. (1987). *The Noun Phrase in Its Sentential Aspect*. Ph.D. thesis, MIT.
- Armstrong, D., Stokoe, W., and Wilcox, S. (1995). *Gesture and the Nature of Language*. Cambridge University Press, Cambridge.
- Atkinson, M. (1996). Now, hang on a minute: some reflections on emerging orthodoxies. In H. Clahsen, editor, *Generative Perspectives on Language Acquisition*, pages 451–485. Benjamins, Amsterdam.
- Baillargeon, R. (1987). Young infants' reasoning about the physical and spatial properties of a hidden object. *Cognitive Development*, **2**, 179–200.
- Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.
- Baldwin, J. (1896). A new factor in evolution. *American Naturalist*, **30**, 441–451.
- Basden, B., Bonilla-Meeks, J., and Basden, D. (1994). Cross-language priming in word fragment completion. *Journal of Memory and Language*, **33**, 69–82.
- Berlin, B. and Kay, P. (1969). *Basic Colour Terms: Their Universality and Evolution*. University of California Press, Berkeley.
- Berwick, R. (1991). Principle-based parsing. In P. Sells, S. Schieber, and T. Wasow, editors, *Foundational Issues in Natural Language Processing*, pages 115–226. MIT Press, Cambridge, Mass.
- Bickerton, D. (1991). *Language and Species*. University of Chicago Press, Chicago.
- Bickerton, D. (1996). Paper presented at Evolution of Language Conference, University of Edinburgh.
- Bock, J. (1986). Syntactic persistence in language production. *Cognitive Psychology*, **18**, 355–387.
- Bock, K. (1989). Closed-class immanence in sentence production. *Cognition*, **31**, 163–186.
- Bock, K. and Loebell, H. (1990). Framing sentences. *Cognition*, **35**, 1–39.
- Bock, K., Loebell, H., and Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, **99**, 150–171.

- Bogin, B. (1990). The evolution of human childhood. *Bioscience*, **40**, 16–25.
- Boland, J., Tanenhaus, M., and Garnsey, S. (1990). Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory and Language*, **29**, 413–432.
- Bonvillian, J. and Richards, H. (1993). The development of hand preference in children's early signing. *Sign Language Studies*, **78**, 1–14.
- Bosshardt, H. and Hagen, B. (1988). Einfluss der englischsprachigen Textoberflaeche auf ihre deutsche Uebersetzung. *Archiv der Psychologie*, **140**, 127–135.
- Branigan, H. (1995). *Language Processing and the Mental Representation of Syntactic Structure*. Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh.
- Branigan, H., Pickering, M., Liversedge, S., Stewart, A., and Urbach, T. (1995a). Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research*, **24**, 489–506.
- Branigan, H., Pickering, M., Liversedge, S., Stewart, A., and Urbach, T. (1995b). Syntactic priming: Investing the mental representation of language. *Journal of Psycholinguistic Research*.
- Burgess, C. and Simpson, G. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language*, **33**, 86–103.
- Carlson, G. and Tanenhaus, M. (1988). Thematic roles and language comprehension. In W. Wilkens, editor, *Syntax and Semantics*, volume 21, pages 263–288. Academic Press, New York.
- Chalmers, D. (1994). Why Fodor and Pylyshyn were wrong: The simplest refutation. In *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, pages 340–347.
- Chater, N. and Oaksford, M. (1990). Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn. *Cognition*, **34**, 93–107.
- Cheney, D. and Seyfarth, R. (1982). How vervet monkeys perceive their grunts: Field playback experiments. *Animal Behavior*, **30**, 739–751.
- Cheney, D. and Seyfarth, R. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. University of Chicago Press, Chicago.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs and P. Rosenbaum, editors, *Readings in English Transformational Grammar*. Ginn & Co., Waltham, Mass.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht.
- Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge, Mass.

- Chomsky, N. (1988). *Language and the Problems of Knowledge: the Managua Lectures*. MIT Press, Cambridge, Mass.
- Chomsky, N. (1994). Bare phrase structure. MIT Occasional Papers in Linguistics 5, MIT.
- Christiansen, M. (1994). *Infinite Language, Finite Minds: Connectionism, Learning and Linguistic Structure*. Ph.D. thesis, Centre for Cognitive Science, University of Edinburgh.
- Clahsen, H., Penke, M., and Parodi, T. (1993). Functional categories in early child German. *Language Acquisition*, **3**, 395–429.
- Clark, E. (1993). *The Lexicon in Acquisition*. Cambridge University Press, Cambridge.
- Clark, H. and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, **22**, 1–39.
- Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, **14**, 597–612.
- Crain, S. and Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, and A. Zwicky, editors, *Natural Language Parsing*, pages 320–358. Cambridge University Press, Cambridge.
- Croft, W. (1991). *Syntactic Categories and Grammatical Relations*. University of Chicago Press, Chicago.
- Cuetos, F. and Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of late closure in Spanish. *Cognition*, **30**, 73–105.
- Dawkins, R. (1986). *The Blind Watchmaker*. Longman, London.
- Dennett, D. (1987). *The Intentional Stance*. Bradford Books, Cambridge, Mass.
- Deprez, V. and Pierce, A. (1993). Negation and functional projections in early grammar. *Linguistic Inquiry*, **24**, 25–57.
- Deprez, V. and Pierce, A. (1994). Crosslinguistic evidence for functional categories in early child grammar. In T. Hoekstra and B. Schwartz, editors, *Language Acquisition Studies in Generative Grammar*, pages 57–80. Benjamins, Amsterdam.
- DeVincenzi, M. (1989). *Syntactic Parsing Strategies in a Null Subject Language*. Ph.D. thesis, University of Massachusetts.
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press, Cambridge, Mass.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, **67**, 547–619.
- Dunbar, R. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, **16**, 681–735.

- Edelman, G. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books, New York.
- Edelman, G. (1992). *Bright Air, Brilliant Fire: On the Matter of Mind*. Basic Books, New York.
- Eibl-Eibesfeld, I. (1989). *Human Ethology*. de Gruyter, New York.
- Evans, G. (1993). *Syntactic Priming in Sentence Comprehension*. Master's thesis, Centre for Cognitive Science, University of Edinburgh.
- Falk, D. (1983). Cerebral cortices of East African early hominids. *Science*, **221**, 1072–74.
- Fay, W. (1993). Infantile autism. In D. Bishop and F. Mogford, editors, *Language Development in Exceptional Circumstances*, pages 190–202. Lawrence Erlbaum Associates, Hove.
- Ferreira, F. and Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, **25**, 348–368.
- Fillmore, C. (1968). The case for case. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*, pages 1–90. Holt, Rinehart and Winston, New York.
- Fodor, J. (1970). Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry*, **1**, 429–438.
- Fodor, J. (1983). *The Modularity of Mind*. Bradford Books, Cambridge, Mass.
- Fodor, J. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, **4**, 155–210.
- Fodor, J. (1991). Thematic roles and modularity: Comments on chapters by Frazier and Tanenhaus et al. In G. Altman, editor, *Cognitive Models of Speech Processing*. MIT Press, Cambridge, Mass.
- Fodor, J. (1998a). Learning to parse? *Journal of Psycholinguistic Research*, **27**(2), 285–319.
- Fodor, J. (1998b). Parsing to learn. *Journal of Psycholinguistic Research*, **27**(3), 339–374.
- Fodor, J. and Pylyshyn, Z. (1988). Connectionism and cognitive architecture. *Cognition*, **28**, 3–71.
- Fong, S. (1991). The computational implementation of principle-based parsers. In R. Berwick, S. Abney, and C. Tenny, editors, *Principle-Based Parsing: Computation and Psycholinguistics*, pages 65–82. Kluwer, Dordrecht.
- Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*. Ph.D. thesis, University of Connecticut.
- Frazier, L. (1991). Exploring the architecture of the language system. In G. Altman, editor, *Cognitive Models of Speech Processing*. MIT Press, Cambridge, Mass.

- Frazier, L. and d'Arcais, G. F. (1989). Filler-driven parsing: A study of gap-filling in Dutch. *Journal of Memory and Language*, **28**, 331–344.
- French-Mestre, C. and Prince, P. (1997). Second language autonomy. *Journal of Memory and Language*, **37**, 481–501.
- Fukui, N. (1986). *A theory of category projection and its application*. Ph.D. thesis, MIT.
- Galaburda, A. (1994). Language areas, lateralization and the innateness of language. *Discussions in Neuroscience*, **10**, 118–130.
- Gardner, R. and Gardner, B. (1969). Teaching sign language to a chimpanzee. *Science*, **165**, 664–72.
- Garrod, S. and Anderson, A. (1987). Saying what you mean in dialogue: a study of conceptual and semantic coordination. *Cognition*, **27**, 181–218.
- Garrod, S. and Clark, A. (1993). The development of dialogue coordination skills in schoolchildren. *Language and Cognitive Processes*, **8**, 101–126.
- Garrod, S. and Doherty, G. (1994). Conversation, coordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, **53**, 181–215.
- Gibson, E. and Wexler, K. (1994). Triggers. *Linguistic Inquiry*, **25**, 407–454.
- Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., and Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, **59**, 23–59.
- Gold, E. (1967). Language identification in the limit. *Information and Control*, **16**, 447–474.
- Gorell, P. (1987). *Studies of Human Sentence Processing: Ranked Parallel versus Serial Models*. Ph.D. thesis, University of Connecticut.
- Gould, S. and Lewontin, R. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptionist programme. *Proceeding of the Royal Society of London*, **205**, 281–288.
- Grimshaw, J. (1988). Form, function and the language acquisition device. In C. Baker and J. McCarthy, editors, *The Logical Problem of Language Acquisition*. MIT Press, Cambridge, Mass.
- Grimshaw, J. (1990). *Argument Structure*. MIT Press, Cambridge, Mass.
- Gruber, J. (1965). *Studies in Lexical Relations*. Ph.D. thesis, MIT.
- Guilfoyle, E. and Noonan, M. (1992). Functional categories and language acquisition. *Canadian Journal of Linguistics*, **37**, 241–272.
- Haegeman, L. (1994). *An Introduction to Government and Binding Theory*. Blackwell, Oxford.
- Haldane, J. (1927). A mathematical theory of natural and artificial selection, part v: Selection and mutation. *Proceedings of the Cambridge Philosophical Society*, **23**, 838–44.

- Hale, K. and Keyser, S. (1992). The syntactic character of thematic structure. In I. Roca, editor, *Thematic Structure: Its Role in Grammar*, pages 107–143. Foris, Dordrecht.
- Hale, K. and Keyser, S. (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale and S. Keyser, editors, *The View from Building 20*, pages 53–109. MIT Press.
- Halliday, M. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, and R. H. Robins, editors, *In Memory of J. R. Firth*, pages 148–162. Longmans, London.
- Harris, T. and Wexler, K. (1996). The optional-infinitive stage in child English. In H. Clahsen, editor, *Generative Perspectives on Language Acquisition*, pages 1–42. Benjamins, Amsterdam.
- Hawkins, J. (1995). *A Performance Theory of Word Order and Constituency*. Cambridge University Press, Cambridge.
- Heine, B., Claudi, U., and Huennemeyer, F. (1991). From cognition to grammar: Evidence from African languages. In E. Traugott and B. Heine, editors, *Approaches to Grammaticalization*. Benjamins, Amsterdam.
- Heydel, M. and Murray, W. (1996). Conceptual structure in sentence production: Evidence from crosslinguistic priming. University of Dundee and Abertay University.
- Heydel, M. and Murray, W. (1997a). The conceptual basis of cross-linguistic sentence priming. In *Paper Presented at 3rd AMLAP*, Edinburgh.
- Heydel, M. and Murray, W. (1997b). Conceptual form and the basis of sentence priming: Cross-linguistic evidence. Poster Presented to 10th CUNY Sentence Processing Conference, Santa Monica, Cal.
- Heydel, M. and Murray, W. (2000). Conceptual effects in sentence priming: A cross-linguistic perspective. In *Cross-linguistic Perspectives on Language Processing*, volume 25 of *Studies in Theoretical Linguistics*, pages 227–254. Kluwer, Dordrecht.
- Heydel, M., Murray, W., and Oria-Merino, L. (1998). Syntactic priming? maybe not! Poster Presented to 11th CUNY Sentence Processing Conference, New Brunswick, N.J.
- Heydel, M., Oria-Merino, L., and Murray, W. (1999). Why sentence priming is not based on phrase structure. Poster presented at 5th AMLAP, Edinburgh.
- Hindle, D. and Rooth, M. (1993). Structural ambiguity and lexical relations. *Computational Linguistics*, **19**, 103–120.
- Hirsch-Pasek, K., Treiman, R., and Schneiderman, M. (1984). Brown and Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, **11**, 81–88.
- Ho, M.-W. (1995). Unravelling gene biotechnology. *Soundings*, **1**, 77–98.

- Hoekstra, T. and Hyams, N. (1998). Agreement and the finiteness of V2: Evidence from child language. In *Proceedings of Boston University Conference on Language Development*, volume 22, pages 360–373.
- Holloway, R. (1983). Human paleontological evidence relevant to language behaviour. *Human Neurobiology*, **2**, 105–14.
- Holmes, V., Stowe, L., and Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language*, **28**, 668–689.
- Hopper, P. and Traugott, E. (1993). *Grammaticalization*. Cambridge University Press, Cambridge.
- Hyams, N. (1986). *Language Acquisition and the Theory of Parameters*. Reidel, Dordrecht.
- Hyams, N. (1987). The theory of parameters and syntactic development. In Roeper and Williams (1987).
- Hyams, N. (1994). VP, null arguments and COMP projections. In T. Hoekstra and B. Schwartz, editors, *Language Acquisition Studies in Generative Grammar*, pages 21–55. John Benjamins.
- Inoue, A. and Fodor, J. (1995). Information-paced parsing of Japanese. In R. Mazuka and N. Nagai, editors, *Japanese Sentence Processing*. Erlbaum, Hillsdale, NJ.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, Mass.
- Jackendoff, R. (1977). *X-Bar Syntax: A Study of Phrase Structure*. MIT Press, Cambridge, Mass.
- Jackendoff, R. (1983). *Semantics and Cognition*. MIT Press, Cambridge, Mass.
- Jackendoff, R. (1990). *Semantic Structures*. MIT Press, Cambridge, Mass.
- Jones, S. (1996). The rise of neurogenetic determinism. *Soundings*, **2**, 53–70.
- Keil, F. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, **88**, 197–227.
- Kendon, A. (1991). Some considerations for a theory of language origins. *Man*, **26**, 199–221.
- Ketley, C., Spinks, J., and de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory and Cognition*, **22**, 70–84.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, **2**, 15–47.
- Koopman, H. (1984). *The Syntax of Verbs*. Foris, Dordrecht.
- Krings, H. (1996). *Was in den Koepfen von Uebersetzern vorgeht: Eine empirische Untersuchung zur Struktur des Uebersetzungsprozess an fortgeschrittenen Franzoesischlernern*. Narr, Tuebingen.

- Lakoff, G. (1987). *Women, Fire and Dangerous Things*. University of Chicago Press, Chicago.
- Leakey, R. (1994). *The Origin of Humankind*. Weidenfeld and Nicolson, London.
- Lebeaux, D. (1988). *Language Acquisition and the Form of the Grammar*. Ph.D. thesis, University of Massachusetts at Amherst.
- Lehman, C. (1988). Towards a typology of clause linkage. In J. Haiman and S. Thompson, editors, *Clause Combining in Grammar and Discourse*. Benjamins, Amsterdam.
- Leslie, A. (1984). Spatiotemporal continuity and the perception of causality in children. *Perception*, **13**, 387–305.
- Levelt, W., Roelofs, A., and Meyer, A. (in press). A theory of lexical access in speech production. *Behavioral and Brain Sciences*.
- Levin, B. and Rappaport, M. (1995). *Unaccusativity: At the Syntax-Lexical Semantics Interface*. MIT Press, Cambridge, Mass.
- Lightfoot, D. (1991). *How to Set Parameters*. MIT Press, Cambridge, Mass.
- Manzini, R. and Wexler, K. (1987). Parameters, binding theory and learnability. *Linguistic Inquiry*, **18**, 413–44.
- Marler, P. (1976). An ethological theory of the origin of vocal learning. *Annals of the New York Academic Society*, **280**, 708–17.
- Marslen-Wilson, W. (1989). *Lexical Representation and Process*. MIT Press, Cambridge, Mass.
- McCawley, J. (1968). The role of semantics in grammar. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*, pages 125–70. Holt, Rinehart and Winston, New York.
- McNeill, D. (1985). So you think gestures are nonverbal. *Psychological Review*, **92**, 350–371.
- McNeill, D. and Levy, E. (1982). Conceptual representations in language activity and gesture. In R. Jarvella and W. Klein, editors, *Speech, Place and Action: Studies in Deixis and Related Topics*. Wiley, Chichester.
- Miller, G. and Isard, S. (1963). Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, **2**, 217–28.
- Mitchell, D. (1994). Sentence parsing. In M. Gernsbacher, editor, *Handbook of Psycholinguistics*. Academic Press, New York.
- Mitchell, D. and Cuetos, F. (1991). The origins of parsing strategies. In C. Smith, editor, *Current Issues in Natural Language Processing*. Center for Cognitive Science, University of Texas.
- Nicol, J. and Pickering, M. (1993). Processing syntactically ambiguous sentences: Evidence from semantic priming. *Journal of Psycholinguistic Research*, **22**, 207–237.

- Patterson, F. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, **5**, 72–97.
- Perfetti, C. (1992). The cooperative language processors: Semantic influences in an autonomous syntax. In D. Balota, G. F. d'Arcais, and K. Rayner, editors, *Comprehension Processes in Reading*, pages 285–302. Erlbaum, Hillsdale.
- Perrett, D., Harries, M., Bevan, R., Thomans, S., Benson, P., Mistlin, A., Chitty, A., Hietanen, J., and Ortega, J. (1989). Frameworks of analysis for the neural representation of animate objects and action. *Journal of Experimental Biology*, **146**, 87–113.
- Piattelli-Palmarani, M. (1989). Evolution, selection and cognition: From 'learning' to parameter setting in biology and the study of language. *Cognition*, **31**, 1–44.
- Pickering, M. and Branigan, H. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, **39**, 633–651.
- Pickering, M. and Branigan, H. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, **3**, 136–141.
- Pickering, M. and Shillcock, R. (1992). Processing subject extractions. In H. Goodluck and M. Rochemont, editors, *Island Constraints*, pages 295–320. Kluwer, Dordrecht.
- Pinker, S. (1984). *Language Learnability and Language Development*. Harvard University Press, Cambridge, Mass.
- Pinker, S. (1989). *Learnability and Cognition: the Acquisition of Argument Structure*. MIT Press, Cambridge, Mass.
- Pinker, S. (1994). *The Language Instinct: The New Science of Language and Mind*. Allen Lane, London.
- Pinker, S. and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, **13**, 707–84.
- Poeppel, D. and Wexler, K. (1993). The full competence hypothesis of clause structure in Early German. *Language*, **69**, 1–33.
- Prat-Sala, M., Branigan, H., Pickering, M., and Shillcock, R. (1996). Is word order influenced by conceptual accessibility? Poster Presented at 9th CUNY Sentence Processing Conference.
- Premack, D. (1976). *Intelligence in Ape and Man*. Lawrence Erlbaum Associates, Hillsdale, N.J.
- Premack, D. (1990). The aesthetic basis of pedagogy. In R. Hoffman and D. Palermo, editors, *Festschrift for J.J. Jenkins*. Lawrence Erlbaum Associates, Hillsdale, N.J.
- Pritchett, B. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, **64**, 539–576.

- Pritchett, B. (1992). *Linguistic Competence and Parsing Performance*. University of Chicago Press, Chicago.
- Quartz, S. and Sejnowski, T. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioural and Brain Sciences*, **20**, 537–596.
- Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax: the Nature of Early Child Grammars in English*. Blackwell, Oxford.
- Rayner, K., Carlson, M., and Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye-movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, **19**, 358–374.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Foris, Dordrecht.
- Rizzi, L. (1994). Early null subjects and root null subjects. In T. Hoekstra and B. Schwartz, editors, *Language Acquisition Studies in Generative Grammar*, pages 151–176. Benjamins, Amsterdam.
- Rizzolatti, G. and Arbib, M. (1998). Language within our grasp. *Trends in Neuroscience*, **21**, 188–194.
- Rodman, P. and McHenry, H. (1980). Bioenergetics of hominid bipedalism. *American Journal of Physical Anthropology*, **52**, 103–106.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, **42**, 107–142.
- Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: Retrieval of verbs. *Cognition*, **47**, 59–87.
- Roeper, T. and Williams, E., editors (1987). *Parameter Setting*. Reidel, Dordrecht.
- Rogers, M. (1995). Interpreting interlanguage data: the example of German word order. *Linguistische Berichte*, **157**, 186–215.
- Rumbaugh, D. (1977). *Language Learning by a Chimpanzee: The Lana Project*. Academic Press, New York.
- Sarich, V. and Wilson, A. (1967). Immunological time scale for hominid evolution. *Science*, **158**, 1200–1203.
- Savage-Rumbaugh, E. (1984). Acquisition of functional symbol usage in apes and children. In H. Roitblat, T. Bever, and H. Terrace, editors, *Animal Cognition*, pages 291–310. Lawrence Erlbaum Associates, Hillsdale, N.J.
- Savage-Rumbaugh, E., Murphy, J., and Sevcik, R. (1993). *Language Comprehension in Ape and Child*. University of Chicago Press, Chicago.
- Schenkein, J. (1980). A taxonomy for repeating action sequences in natural conversation. In B. Butterworth, editor, *Language Production*, volume 1, pages 21–47. Academic Press, London.

- Seyfarth, R., Cheney, D., and Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, **28**, 1070–94.
- Siskind, J. (1992). *Naive Physics, Event Perception, Lexical Semantics, and Language Acquisition*. Ph.D. thesis, MIT.
- Spelke, E. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey and R. Gelman, editors, *The Epigenesis of Mind: Essays on Biology and Cognition*. Lawrence Erlbaum Associates, Hove.
- Steedman, M. (1990). Syntax and intonational structure in a combinatorial grammar. In G. Altman, editor, *Cognitive Models of Speech Processing*. MIT Press, Cambridge, MIT Press.
- Stevenson, R., Crawley, R., and Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, **9**, 519–548.
- Stowe, L. (1986). Evidence for on-line gap location. *Language and Cognitive Processes*, **1**, 227–245.
- Sweetser, E. (1990). *From Etymology to Pragmatics*. Cambridge University Press, Cambridge.
- Swinney, D. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Memory and Language*, **18**, 545–569.
- Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In C. Shopen, editor, *Language Typology and Syntactic Description*, volume 3, pages 57–149. Cambridge University Press, Cambridge.
- Tanenhaus, M. and Carlson, G. (1989). Lexical structure and language comprehension. In Marslen-Wilson (1989).
- Tanenhaus, M., Boland, J., Mauner, G., and Carlson, G. (1994). More on combinatorial lexical information: thematic structure in parsing and interpretation. Manuscript.
- Tannen, D. (1984). *Conversational Style: Analyzing Talk among friends*. Ablex, Norwood, NJ.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue and imagery in conversational discourse*. Cambridge University Press, Cambridge.
- Taraban, R. and McClelland, J. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based explanations. *Journal of Memory and Language*, **27**, 597–632.
- Terrace, H. (1979). *Nim*. Knopf, New York.
- Tomasello, M., Kruger, A., and Ratner, H. (1993). Cultural learning. *Behavioral and Brain Science*, **16**, 495–552.

- Toth, N. (1985). Archaeological evidence for preferential right-handedness in the lower and middle pleistocene, and its possible implications. *Journal of Human Evolution*, **14**, 607–14.
- Travis, L. (1984). *Parameters and Effects of Word Order Variation*. Ph.D. thesis, MIT.
- Travis, L. (1989). Parameters of phrase structure. In M. Baltin and A. Kroch, editors, *Alternative Conceptions of Phrase Structure*. University of Chicago Press, Chicago.
- Trueswell, J., Tanenhaus, M., and Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, **33**, 285–318.
- Varela, F., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, Cambridge, Mass.
- Wallman, J. (1992). *Aping Language*. Cambridge University Press, Cambridge.
- Weinberg, A. (1990). A parsing theory for the nineties: Minimal Commitment. *Manuscript, University of Maryland*.
- Weiner, E. and Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, **19**, 29–58.
- Wexler, K. (1994). Optional infinitives, head movement, and the economy of derivations. In D. Lightfoot and N. Hornstein, editors, *Verb Movement*, pages 305–350. Cambridge University Press, Cambridge.
- Wexler, K. and Manzini, M. (1987). Parameters and learnability in binding theory. In Roeper and Williams (1987).
- Wilkes-Gibbs, D. and Clark, H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, **31**, 183–194.
- Williams, E. (1984). Grammatical relations. *Linguistic Inquiry*, **15**, 639–673.
- Williams, E. (1994). *Thematic Structure*. MIT Press, Cambridge, Mass.
- Wills, C. (1994). *The Runaway Brain*. Harper Collins, London.
- Yang, D. (1983). The extended binding theory of anaphors. *Language Research*, **19**, 169–192.