

THE ROLE OF THE LEXICON  
IN SYNTACTIC CHANGE

W. S. Allan

Ph.D.

University of Edinburgh

1984



I declare that this thesis is all my own work.

W. S. Allan

October 1984

	<u>Page</u>
CONTENTS	iii
ACKNOWLEDGEMENTS	vi
ABSTRACT	vii
INTRODUCTION	1
CHAPTER 1. The Lexicon: An Overview	5
1.0. The <u>Aspects</u> Lexicon	6
1.1. 'Remarks'	11
1.2. $\bar{X}$ -Syntax: An Excursus	17
1.3. Lexical Grammars	24
1.4. Generalized Phrase Structure Grammar: An Excursus	55
CHAPTER 2. Change in the Lexicon: The Modals - A Paradigm Case	64
2.0. Introduction	65
2.1. The Autonomous View	65
2.2. The Reanalysis	66
2.2.1. The Early Changes	67
2.2.2. The Later Changes	71
2.3. The Data for the Early Changes	77
2.3.1. Preliminaries	77
2.3.2. The Morphological Form of the Third Singular Present Tense	80
2.3.3. Loss of the Direct Object I	84
2.3.4. Loss of the Direct Object II	86
2.3.5. A Semantic Argument for Reanalysis	89
2.3.6. Past Tense Interpretation	93
2.3.7. Rejection of the <u>to</u> -infinitive	94
2.4. The Data for the Later Changes	97
2.4.1. Loss of Modals from Infinitival Constructions	97

	<u>Page</u>
2.4.2. Loss of Modals from Gerunds and Possessives	99
2.4.3. Loss of Double Modal Constructions	102
2.4.4. Loss of Modals from Perfective Constructions	103
2.5. Lexical Account of the Data	105
2.5.1. Lexical Diffusion	107
2.5.2. Syntactic Lexical Diffusion	108
CHAPTER 3. A Dependency Account of Modals and Auxiliaries	116
3.0. Introduction	117
3.1. Modal Opacity	117
3.1.1. Criteria for Auxiliaryhood	117
3.1.2. Need and Dare	119
3.1.3. Ought and Be	121
3.1.4. Have and Non-Modal Be	121
3.1.5. Would Rather	123
3.2. Main Verb - Auxiliary Similarities	124
3.2.1. Affix Hopping	124
3.2.2. Gapping	124
3.3. Syntactic Opacity	126
3.4. A Finite Argument for Main Verb Status	128
3.5. On the Elimination of Aux	131
3.5.1. Affix Hopping Revisited	131
3.5.2. Akmajian, Steele and Wasow	136
3.5.3. Natural Serialization and the Status of Aux	145
3.6. Dependency Syntax: Heads and Modifiers	146
3.6.1. Dependency Syntax: Rules	148

	<u>Page</u>
3.6.2. Dependency Syntax and Natural Serialization	151
3.6.3. Dependency Syntax and Concord	153
3.7. Recursion and Embedding	157
3.7.1. Morphology	173
3.7.2. Subcategorization	175
3.7.3. Passive	178
3.7.4. Question Formation and Negative Placement	184
3.7.5. Lexical Entries	188
3.8. Diachronic Change in the Lexicon	190
CHAPTER 4. The Transparency Principle: A Critical Review	193
4.0. Introduction	194
4.1. The Transparency Principle	194
4.2. Derivational Complexity	197
4.3. Autonomy and Independence	201
4.4. Lightfoot and History	202
4.5. Therapeutic Change and the Individual	203
4.6. Necessary Changes Which Fail	211
4.7. Explanation and Prediction	216
4.8. Markedness	218
4.9. Transparency and Falsifications	220
BIBLIOGRAPHY	225

ACKNOWLEDGEMENTS

A thesis is never, despite the declaration, totally the work of one person but normally involves others. This thesis is no exception and, therefore, I would like to take this opportunity to thank: Audrey Somerville, my typist; Harriet for her help in proof-reading, sarcasm and general abuse; and Roger Lass and Jim Miller, both of whom read and commented on parts of this thesis. Finally, I would like to thank especially my supervisor, John Anderson, for his help, advice and encouragement. Needless to say, none of the above are responsible for any remaining errors.

W.S.A.

October 1984.

ABSTRACT

This thesis is founded on the premise that a more restrictive account of syntactic change may be achieved if it is assumed that change is confined to the lexicon and in particular to the subcategorization frames which govern the availability for various morphological processes. This is, in principle, R. Lakoff's (1968) thesis, except that the proposed model of syntax is more restrictive. The first chapter is essentially an overview of the role of the lexicon in generative grammar from the Aspects model to lexical grammars. Within this overview particular attention is paid to the treatment of morphology. The chapter concludes that morphology should not be conflated with either the syntactic or the lexical components. The second chapter tests this restrictive view of change against the data concerning the development of the N.E. modals and reviews Lightfoot's (1979) position. Chapter three introduces and develops a synchronic dependency grammar for the English verbal system. Problems for the analysis are raised and the evidence for a main verb-auxiliary split and the 'Aux' node are critically surveyed, with both the division and the 'Aux' node being rejected in favour of a verb only non-'Aux' analysis. The final chapter critically reviews the Transparency Principle introduced in Lightfoot (1979).

## INTRODUCTION

One major strand of current syntactic research concerns the role of the lexicon in the grammar. The source of inspiration for these investigations is undoubtedly Chomsky (1970). It was in this paper that he discussed the balance of each component of the grammar, i.e. the base, the lexicon and the transformational component. He decided that the scope of the base and the lexicon should be extended at the expense of the transformational component. Research, at that time, e.g. Ginsburg and Partee (1969), had shown that Aspects type grammars were extremely powerful and, as such, were of a low explanatory value. This excessive power lay in the transformational component and, therefore, Chomsky sought to reduce the power of the grammar by reducing the role of the transformational component. Although Chomsky did not further discuss the role of the lexicon in a generative grammar, a number of other linguists have pursued this line of investigation. The results of this avenue of research were published in a series of papers throughout the late seventies and early eighties by Brame, Bresnan, Jackendoff, Lapointe, Siegel and Wasow amongst others.

Within these discussions of the lexicon, morphology is given prominence. The criticism often levelled at the Aspects model (see e.g. Brown and Miller (1982) and Lapointe (1980)) that the categories which trigger morphological processes are spread throughout the grammar, allowing the double interaction of syntax and morphology, and cutting across the traditional divide between the two, cannot be levelled at the lexical models. In the lexicon (e.g. see Lapointe (1980)) and in some of the more extreme views (e.g. Bresnan (1979))

syntax also becomes a function of the lexicon. These shifts in the role of the lexicon raise issues which we shall discuss in Chapter One and return to in Chapter Three.

Discussions on the enriched role of the lexicon have, so far, concentrated on synchronic issues and have not touched on the role it might play in diachronic change. However, R. Lakoff (1968) has suggested that it might be the case that much 'syntactic change' occurs in the lexicon. While there have been few investigations into what role, if any, the lexicon might play in 'syntactic change', the same cannot be said for phonological change. Chen (1972) and Wang (1977) provide evidence, which suggests that while change might be phonetically abrupt, it is lexically or phonologically gradual. It would appear that all lexical items containing the phone defined by the SD of the rule, occurring in the environment specified by the SC of the rule, are not affected simultaneously, but that individual items, which fulfil the conditions of the rule are affected one at a time. Within this model change is spread over a long time period and it is often the case that not all items available for change are affected. This results in a residue of exceptions. The lexical diffusion model, as it is referred to, guarantees the preservation of communication between generations.

If it can be shown that syntactic change, like phonological change, is gradual then the claim that the lexicon is the locus of syntactic change would gain credence. The evidence I advance for such a view of syntactic change is based on the development of the English modal verbs. The claim that syntactic change is gradual and occurs within the lexicon is counter to the position advanced by Lightfoot (1979), which is discussed, in detail, in the second chapter. I suggest

that the evidence for syntactic change, in the case of the modals, is, at best, marginal and that they may not have undergone any syntactic changes at all.

The view that syntactic change occurs in the lexicon, and, in particular, in the strict subcategorization frames which govern the availability for morphological processes, allows for a more restrictive account of change. If we assume that the P.S. rules do not alter then it is possible to give a more restrictive account of change. Further the lexical view of syntactic change, like the lexical view of phonological change, ensures the preservation of communication between generations.

The main topic of the third chapter is a dependency account of the English verbal system. I argue that the modals are not a distinct class but are members of the class of main verbs, the irregular behaviour of which is best captured, as is their historical development, in the lexicon. I also argue that the AUX node plays no role in the grammar and may be dispensed with. Natural serialization and the head-modifier relationship between constituents are advanced as evidence for the main verb status of modals and auxiliaries. Both of these are best captured by a dependency framework and, therefore, as mentioned above, it is this model which provides the theoretical background for the third chapter.

The discussion of the modal verbs in the second chapter offers no functional or other explanation of why the modals changed. As with the analysis presented in that chapter this lack of functional explanation is in contrast with Lightfoot (1979). In his account he invokes the Transparency Principle. This is essentially a

functional notion which 'explains' why changes occur. In the fourth, and final, chapter I critically review the T.P. as an explanatory device and find it to be vacuous.

CHAPTER ONE

The Lexicon : An Overview

## 1.0. The Aspects Lexicon

In Aspects, Chomsky (1965), the lexicon is a sub component of the base. This lexicon comprises of an unordered list of lexical formatives. Each entry consists of a pair (D,C), where D is a specified phonological distinctive feature matrix and C is a set of specified syntactic features.

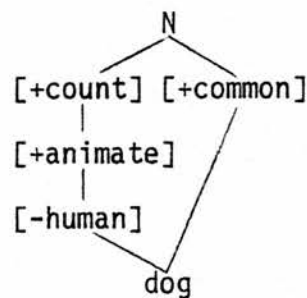
The re-write rules of the base generate derivations terminating in strings which consist of grammatical formatives and complex symbols. These strings are pre-terminal. Terminal strings are derived from pre-terminal strings by inserting a lexical formative in accordance with a lexical rule:

"If Q is a complex symbol of a pre-terminal string and (D,C) is a lexical entry, where C is not distinct from Q, then Q can be replaced by D."

Chomsky (1965:84)

Consider the following example:

1.1. (i)



(ii)  $N \rightarrow [+N \pm common]$

$[+common] \rightarrow [\pm count]$

$[+count] \rightarrow [\pm animate]$

$[+animate] \rightarrow [\pm human]$

(iii) dog [+N, +count, +common, +animate, -human]

The rules in 1.1(ii) generate the complex entry in 1.1(i) and the example entry in 1.1(iii) which matches 1.1(i) can be entered into the structure. However, the lexical entries, such as 1.1(iii), do

not just contain the feature information given in 1.1(iii) above. Information concerning selection restrictions and strict subcategorization must also be included: for example ADMIRE typically selects an animate subject.

- 1.2. (i) The man admired the floral arrangement.  
 (ii) \*The floral arrangement admired the man.

Transitive verbs, for example, are strictly subcategorized for object NP's:

- 1.3. (i) The man admired the picture.  
 (ii) \*The man admired.

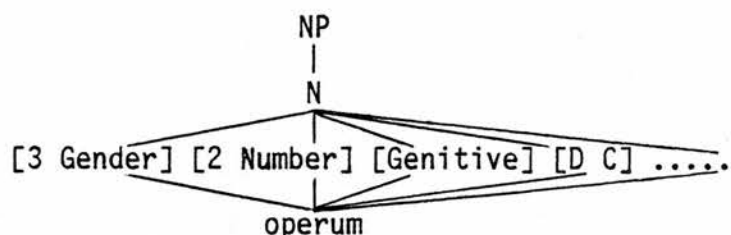
The lexicon will also contain all idiosyncratic properties of lexical items (Chomsky 1965:87). It will include: (i) non-predictable aspects of phonetic structure, (ii) properties relevant to the application of transformational rules, (iii) a dictionary definition necessary for semantic interpretation, (iv) lexical features indicating the positions in which the lexical item may be inserted.

Having outlined the nature of the lexicon in an Aspects grammar let us now turn to morphology, both inflectional and derivational.

Aspects adopts a paradigmatic approach to inflectional morphology (Chomsky:1965:170ff). In a paradigmatic approach, a noun, for example, is described by its place in a paradigm system. The parameters of this description are morphological categories - number, case, gender and declension type. In Latin, for example, operum would be [2 number] (i.e. plural), [genitive], [3 gender] (i.e. neuter) [3 declension]. This would be represented in the following

fashion in an Aspects grammar:

1.4.



However, whereas gender and [D C] are inherent specifications of the lexical entry, number and case are introduced by grammatical rules. Number will be introduced by a context-free rule of the base, e.g.  $N \rightarrow [\Delta, \infty \text{ number}]$  and case is introduced not by a base rule but by a transformational rule. This allows case assignment to be made at shallow surface level after all movement rules have applied.

Chomsky adduces support for this claim from data like the following:

1.5. (i) I shot him.

(ii) He was shot by me.

Sentences 1.5(i) and (ii) are related by the rule of Passive. In 1.5(i) him is the object and takes the oblique form. In 1.5(ii) him of 1.5(i) is in subject position and takes the subject form he. If case were to be assigned at deep structure then the oblique of 1.5(i) would not have the subject form in 1.5(ii).

From the above we can see that morphology, in an Aspects grammar is not concentrated in a single component but spread out throughout the grammar - idiosyncratic in the lexicon, regular in the base and transformational component. This situation, where the categories which trigger morphological rules are spread throughout the grammar, is wholly unsatisfactory.

What is more unsatisfactory is the Aspects treatment of verbal inflection such as tense and the participle affixes. These affixes are treated as syntactic categories and introduced by a context-free base rule: Aux  $\rightarrow$  Tense (M) (Aspect) (Chomsky 1965:107). This treatment of verbal morphology is one to which we shall return at several points below (3.5.1).

In the Aspects model there is, in principle, no difference between morphological and syntactic rule types. This is somewhat akin to the loss of distinction between allophonic and morphophonemic rules which occurs in generative phonology (see Halle 1956 and Chomsky and Halle 1968). Clearly, as with the phonological example, it is possible, and more restrictive, to distinguish the two rule types. Morphological rules are confined to those processes which operate below the level of the word, e.g. inflection, derivation and compounding. Syntactic rules, on the other hand, deal with the combining of words to form phrases, clauses and sentences.

Let us now turn to derivational morphology. This is somewhat more problematical than inflectional morphology for a variety of reasons. Some of the processes are semi-productive while others produce items, the meaning of which cannot be predicted from the sum of the parts. Productive processes such as nominalizations, e.g. destroy-destruction and refuse-refusal are treated transformationally. In the above examples only the first member of each pair is entered in the lexicon together with a feature specification which determines the phonetic form when they appear in nominalized sentences. Similarly, in the case of, for example, sincerity, this is not entered in the lexicon but is the product of a transformation i.e. NP-is-Adj

where NP is empty  $\Rightarrow$  Adj+ity (=N), NP is sincere  $\Rightarrow$  sincer+ity. Semi productive processes e.g. horror - horrid - horrify vs terror - \*terrific - terrify, may simply have to be entered in the lexicon as fully specified lexical items. One alternative considered by Chomsky (1965:187) is to allow such semi-productive processes to occur in the lexicon. To take Chomsky's examples, telegraph, horrify, frighten might be entered in the lexicon as:

- 1.6. (a) ( $\widehat{\text{tele}}\text{stem}_1$  [ $F_1$ ....])  
 (b) ( $\widehat{\text{stem}_2}\text{ify}$  [ $G_1$ ....])  
 (c) ( $\widehat{\text{stem}_3}\text{en}$  [ $H_1$ ....])

and these items are entered into strings by the general lexical rule (see p.1 above). The lexicon would also contain the following:

- 1.7. (a) (graph, [ $+\text{stem}_1$ ....])  
 (b) (horr [ $+\text{stem}_2$ ....])  
 (c) (fright [ $+\text{stem}_3$ ....])

These items being inserted in strings formed by prior insertion in pre-terminal strings of items selected from 1.6 above. In the case of morphologically complex forms there may be several layers<sup>1</sup> of derivation.

Again, as with inflectional morphology, derivational morphology is carried out both transformationally and lexically. We may raise the same objections to this as we raise above in the case of inflectional morphology.

---

<sup>1</sup>Chomsky raises objections to this vis à vis productivity.

### 1.1. Remarks

It would appear that the Aspects type grammar abandoned the traditional distinction between syntax and morphology. This is due to what has been referred to as 'transformational holism' (see Hoekstra et al 1980:1). This transformational holism results from the fact that transformations were considered the only means of expressing relatedness between structures.

The division in syntax between deep and surface structure allows repetition of, for example, the selection restrictions between items in related expressions, to be avoided. Related expressions can be given a single underlying form at the deep structure level, where selection restrictions are handled. Transformations mediate between deep and surface structure. They, alone, allow the repetition of information to be avoided.

As stated above, all distributional similarities between sentences (interrogatives and declaratives) and words (nominalizations) are related by transformations. Derivation and compounding were also handled by transformations (e.g. see Lees (1960) and Botha (1968)). Further, the lack of non-transformational interpretive mechanisms required that anaphora, for example, be given a transformational treatment (e.g. see Lees and Klima (1963)).

If all distributional similarities are treated transformationally and derived from a common deep structure source then this source will be very abstract or remote from the surface. The transformations, due to the variety of relations expressed, will be almost unlimited in power. Investigations in mathematical linguistics

(e.g. Ginsbury and Partee (1969) and Peters and Ritchie (1971, 1973)) had shown that such a powerful system was low on explanatory value. This all points towards a need to reduce the power of the grammar, and also to a question of appropriateness: are lexical derivations structure-changing?

### 1.1.1.

In the late 1960's Chomsky presented a series of lectures at M.I.T., which finally appeared in print as "Remarks on Nominalizations" (Chomsky 1970). Although much of this paper was involved in the Interpretive-Generative Semantics polemic, it marked a major shift in the organization of a generative grammar. This paper, also, was a response to the problems outlined above (1.1).

In "Remarks", Chomsky attempted to reduce the power of the transformational component. This would enhance the learnability of the grammar, increase its explanatory powers and reduce the distance between deep and surface structure (producing a less abstract grammar)<sup>2</sup>. This reduction can be brought about in two ways: (i) enriching the other components of the grammar at the expense of the transformational component or (ii) placing constraints on the transformations themselves. In "Remarks" Chomsky chose the first of these options. One of the ways Chomsky chose to constrain the transformational component was to enrich the lexicon, at the expense of the transformational component. In subsequent works e.g. Chomsky (1973), (1977) he has sought to constrain the transformations.

---

<sup>2</sup>See Bresnan (1978) in Halle, Bresnan and Millar (1978).

## 1.1.2.

Chomsky (1970) claims that a grammar is a tightly organized system and as such modification in one part entails widespread modification in other areas. From this he deduced that enrichment of one component may lead to simplification in another. Chomsky (1970) chose to reduce the power of transformations and enrich the lexicon.

Recall that in the Aspects model the entire burden of relating structures was borne by the transformational component. If it can be shown that certain relations are non-transformational and less powerful than transformations then these non-transformational relations can be filtered out of the transformational component. As this is done the transformational complexity will be reduced, and deep structures will bear a much closer resemblance to their corresponding surface forms. It is also possible that the resulting grammar may be more restrictive (see Wasow (1977) which claims that lexical rules are more restrictive than transformations).

To achieve this we must be able to identify "non-transformational relations". We have already discussed the transformational treatment of morphology (both inflectional and derivational) and this, for a number of reasons, would seem an ideal candidate for a non-transformational treatment. The main reason for choosing morphology is one of power. Morphological processes do not require the full power of transformations. They are essentially local and not long distance processes. The morphological rules of an Aspects type grammar usually involve two adjacent terminal, non-phrasal, nodes, with one being hopped onto the other.

This, clearly, distinguishes them from syntactic processes such as, for example, question formation, passive and extraposition, which involve the movement of phrasal nodes over long distances.

Chomsky (1970) initiated work in this field which has been advanced by Aronoff (1976) and Siegel (1973, 1974) amongst others. Let us consider some examples from Siegel's work.

Siegel (1974) gives the following as examples of word formation rules:

1.8. (i) N or V + -d = Adj

e.g. [<sub>N</sub>salary] + -d = [<sub>A</sub>[<sub>N</sub>salari]ed]

[<sub>V</sub>hurry] + -d = [<sub>A</sub>[<sub>V</sub>hurri]ed]

(ii) un + Adj = Adj

un + [<sub>A</sub>successful] = [<sub>A</sub>un[<sub>A</sub>successful]]

1.8(i) derives adjectives from nouns or verbs, while 1.8(ii) forms adjectives from adjectives<sup>3</sup>. Rule (ii) cannot apply to either nouns or verbs, e.g. \*unsalary or \*unhurry but may apply to adjectives derived from nouns or verbs, e.g. unsalaried or unhurried. Given that the rules involved in the derivation of these examples are local and operate below the level of the word, then the processes involved are better suited to a non-transformational treatment. Siegel (1973) claims that what is of interest in these examples is that the rules allow the expression of generalizations not adequately expressed by transformations. Consider the following example:

---

<sup>3</sup>In 1.8(ii) un- is not the same un- as in unlock where it involves a reversal of activity.



- (ii) \*Fred was suspicious Alice \*Adj NP  
 (iii) Fred was suspicious of Alice Adj PP

In 1.13(i) the verb suspected takes a direct NP complement and the sentence is well formed. 1.13(ii) and (iii) show that the adjectival form requires an intervening preposition.

Further, a small group of passive verbs, like verbs, may also take a direct NP complement. Consider:

- 1.14. (i) Alice was taught cordon-bleu cookery.  
 (ii) \*Alice was untaught cordon-bleu cookery.

In 1.14(i) taught can take a direct noun phrase complement but adjectival untaught cannot. Sentence 1.9 can be analysed as NP be Adj, parallel to 1.13(iii) and derived by rules 1.8(i) and (ii):

- 1.15.  $[_V \text{touch}] + d \Rightarrow [_A [_V \text{touch}] \text{ed}]$   
 $[_A [_V \text{touch}] \text{ed}] \Rightarrow [_A \text{un}[_A [_V \text{touch}] \text{ed}]] \Rightarrow \text{untouched}$

A transformational derivation of these adjectives would force us to give up the morphological generalization expressed in 1.8(ii) and the generalization that passive verbs are verbs.

It would appear that the class of morphological rules, both derivational and inflectional (see 3.4.6 on inflectional morphology), are non-transformational and can be removed from the transformational

---

\*However, consider Fred is like Alice, where like, an adjective, takes a direct object. Such an exception may be treated lexically.

component. This would result in a simplification of the transformational derivations. However, this raises the questions of where these rules belong, and whether they can be distinguished from transformations by non-ad-hoc means. We shall return to this below.

## 1.2. X̄ - Syntax: An Excursus

Having argued for the lexicalist position Chomsky (1970) pursues the consequences of adopting such a position. If the following phrases:

- 1.16. (i) eagerness (for John) to please  
 (ii) refusal of the offer  
 (iii) belief in a supreme being

are all base NP's then pursual of the lexicalist position required that the rules of the categorial component will have to be expanded. This expansion allows the introduction of an extensive range of complements within the NP as in both VP and AP. (Remember that Chomsky was claiming that derived nominals have the internal structure of NP's). Consider the following examples:

- 1.17. (i) ....upset {that she had left  
 {over the news  
 (ii) ....sad {about the failure of the bill  
 {that the bill had failed  
 (iii) his belief that God exists  
 (iv) his belief in the existence of God  
 (v) he believes {in the existence of God  
 {that God exists

In all of these examples we have A, V or N followed by a complement. Based on such examples, Chomsky proposed the following rule schemata:

- 1.18. NP → N Comp  
 VP → V Comp      Comp → S, NP, NP S, PP PP, NP PP etc.  
 AP → A Comp

Apart from the derived nominals, Chomsky (1970:196) claimed the following as independent evidence for the rules in 1.18:

- 1.19. the weather in England      the bottom of the barrel  
 the weather in 1965      the back of the room  
 the story of Bill's exploits      a nation of shopkeepers

Chomsky claims that in these, and other examples, it makes sense to regard the underlined form as a noun of a Det N Comp construction. This construction constitutes a simple base NP. The alternatives according to Chomsky would be to either regard the whole expression as a transform with the underlined element being a nominalized verb or adjective, or to take the complement as a reduced relative. In some cases neither of these analyses seem particularly well motivated.

Given the rules in 1.18 it would appear that a generalization may be captured, that is no matter which category is involved there is always a head and a complement. By using a cover symbol for N, V or A and the rule  $\bar{X} = X + \text{Complement}$  the base rules of the 1.18 may be replaced by a single schema:

- 1.20.  $\bar{X} \rightarrow X$  ( $\bar{X}$  is a phrase with X as its head).

This notation allows us to claim that (i) categories have certain

properties in common and (ii) every construction consists of a head and modifiers.

By using the notation in 1.20 Chomsky claims that the phrases immediately dominating  $\bar{N}$   $\bar{V}$  and  $\bar{A}$  will be designated  $\bar{N}$   $\bar{V}$   $\bar{A}$  respectively. Further the phrase associated with  $\bar{N}$ ,  $\bar{V}$  and  $\bar{A}$  in the base structure will be referred to as the 'specifier' of these elements. The introduction of specifier requires the following schema:

$$1.21. \bar{X} \rightarrow [\text{Spec } \bar{X}] \bar{X}$$

[Spec  $\bar{N}$ ] will be analysed as the determiner, [Spec  $\bar{V}$ ] as the auxiliary, and [Spec  $\bar{A}$ ] as the system of qualifying elements associated with adjective phrases. Having adopted this schema the initial rule of the base will be:

$$1.22. S \rightarrow \bar{N} \bar{V}$$

This system formed the base of the X-Bar system which Jackendoff (1977) attempted to develop into a complete grammatical model.

One of the major faults of P.S.G. was its failure to express the relation between the head of a construction and its modifiers e.g. between NP and N - NP obligatorily includes N (and optional modifiers) and NP's tend to occur in the same syntactic environments as N's. It also fails to express the parallelism between NP, VP AP and PP i.e. they all consist of an obligatory nuclear element and optional modifiers and all phrases occur in the same environments as their obligatory element.

These problems are tied in with one particular aspect in which the

P.S. rules are unconstrained, as was pointed out by Lyons:

Phrase Structure Grammars fail to formalize the fact that NP and VP are not merely mnemonically convenient symbols, but stand for sentence-constituents which are necessarily nominal and verbal, respectively, because they have N and V as an obligatory major constituent. What is required, and what was assumed in traditional grammar, is some way of relating sentence constituents of the form XP to X (where X is any major category: N,V, etc.). It would not only be perverse, but it should be theoretically impossible for any linguist to propose, for example, rules of the following form ....

$$\text{NP} \rightarrow \text{V\_VP}$$

$$\text{VP} \rightarrow \text{DET\_N}$$

Lyons (1968:331)

Despite the fact that PSG uses the symbols NP and N, as far as the rules are concerned NP's are atomic. In PSG only one symbol may be re-written at a time. Therefore, NP is only one symbol and consequently atomic. It is simply a graphic accident that NP and N have N in common. In PSG heads and modifiers display few differences being sisters or equipollent daughters.

Clearly some way must be found to capture the head - modifier relationship, distributional regularities, and to constrain the possible rule types. We could add, for example, \* to the grammar to indicate headhood:

$$1.23. \text{NP} \rightarrow (\text{Det}) \overset{*}{\text{N}}$$

$$\text{PP} \rightarrow \overset{*}{\text{P}} \text{NP}$$

However, this is simply an ad-hoc accretion to the grammar, e.g. one could just as easily write  $\text{P} \rightarrow \text{P} \overset{*}{\text{NP}}$ .

It would appear that what is required is a convention which ensures that e.g. NP must contain a head noun, VP a head V. The rules used to capture these relationships are those of the form given above, e.g. 1.21 which is interpreted as saying that every  $\bar{X}$  i.e. phrasal category, has an  $\bar{X}$  as its head. The question of how successful  $\bar{X}$  is in capturing the head-modifier relation now arises.

In  $\bar{X}$  rules of the form  $\bar{N} \rightarrow [\text{Spec } \bar{N}] \bar{N}$  were introduced and it was stated that  $\bar{N}$  expresses a relation - it is more inclusive than N which it must include as one of its daughters. It is also a function which expresses a relation between phrasal and lexical categories. However, like markedness in phonology (and syntax) it is simply an accretion. The formalism fails to express the regularities it might appear to, e.g., as far as the PSG formalism is concerned  $\bar{N}$  is still atomic, and an interpretive rule is required to express the fact that  $\bar{N}$  is a category which necessarily includes N.

In order to express the parallelism between NPs, VPs and APs a variable X over NP, VP and AP was introduced. Again this is simply an accretion. NP, VP and AP all take modifiers on their left and to capture this a specifier is introduced - [Spec]. Now  $\bar{X} \rightarrow [\text{Spec } X] X$ . This now creates a super category X over N V and A. However problems arise with PPs which take their specifiers to their right. We now have to find some means of relating N V and A, while excluding P. Feature analyses of syntactic categories show that this is not possible. Chomsky (1972) suggested that the four principal lexical categories be considered not as atomic units but as complexes of two syntactic features  $[\pm N]$  and  $[\pm V]$ . This gives us the following:

$$1.24. \begin{bmatrix} +V \\ -N \end{bmatrix} = V \quad \begin{bmatrix} -V \\ +N \end{bmatrix} = N \quad \begin{bmatrix} +V \\ +N \end{bmatrix} = A \quad \begin{bmatrix} -V \\ -N \end{bmatrix} = P$$

Although this feature analysis has some advantages it does not solve the problem of excluding P. Further it does not prevent us from forming a natural class  $\begin{bmatrix} \alpha V \\ \alpha N \end{bmatrix}$  i.e.  $\begin{bmatrix} +V \\ +N \end{bmatrix} = A$  and  $\begin{bmatrix} -V \\ -N \end{bmatrix} = P$ . I can think of no rule which refers exclusively to A and P<sup>5</sup>. This could be considered as the syntactic corollary of the phonological class  $\begin{bmatrix} \alpha \text{VOC} \\ \alpha \text{CONS} \end{bmatrix}$  in Chomsky and Halle (1968). However, feature syntax is peripheral to our discussion and I do not propose to pursue the issue any further at this juncture.

In attempting to adequately characterize what PSG fails to capture,  $\bar{X}$  effectively undoes some of the basic properties of PSG. It denies that labels are atomic and claims that N, V and A belong to a super category, X. These problems are enhanced when we consider  $\bar{X}$  in conjunction with natural serialization<sup>6</sup>.

Bartsch and Vennemann (1972) claim that the operator-operand relationship (head-modifier) tends to be expressed by unidirectional serialization. It is expressed as either:

- 1.25. [operator [operand]] throughout in XV languages  
 [[operand] operator] throughout in ~~VX~~ languages

This captures the parallelism between NP and VP. Let us now consider as an example, the auxiliaries, have and be and main verbs - all

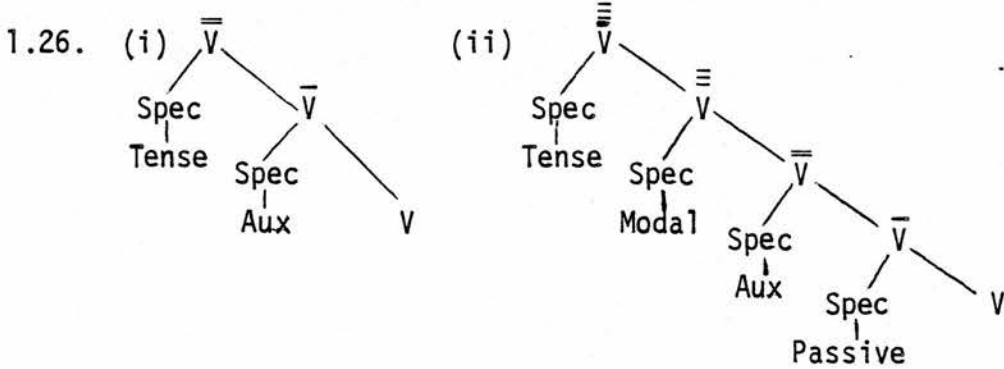
---

<sup>5</sup>Only  $[-\bar{V}]$  appear in focus position in cleft sentences.

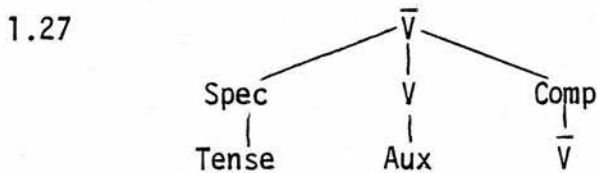
<sup>6</sup>Natural serialization is given expanded discussion in chapter 3.

as members of VP. The question arises whether the auxiliaries are specifiers or verbs.

If we claim they are specifiers then we present ourselves with two problems. Firstly, they are unlike tense in that they are independent. Secondly, we fail to explain the order T(M)(Perf)(Prog) MVb. We could set up the following configurations:



In 1.26(i) Spec of  $\bar{\bar{V}}$  will always be re-written as Tense and in (ii) the Specs immediately dominated by the barred V will always be re-written in the order shown above - Tense, Modal, Aux and Passive<sup>7</sup>. This solves the order problem. However, in both 1.26(i) and (ii) problems arise when optional specifiers, e.g. modal, aux and passive are omitted. We could, however, claim that auxiliaries are main verbs which take complements to their right, e.g.:



<sup>7</sup>Bar levels are a problem in  $\bar{X}$ , e.g. Siegel (1974) has a minimum of four for N, Bresnan (1976) has five for V whereas A.S.W. (1979) have three. See Cann (1984) for a further discussion.

This solution poses no problems if we accept that auxiliaries are, in fact, main verbs (see Ch.3). However, if all auxiliaries are verbs then Spec V will only be Tense, which is an affix and not a word. The occurrence of Tense under Spec is not uncontroversial and raises issues to which we shall return (see 3.4.6.).

Chomsky(1970) dispenses with Comp, claiming that it plays no rôle in transformations. However consider the following sentences:

- 1.28. (i) John sits here and Fred there.  
 (ii) Bert ate the steak and Alan the chicken.  
 (iii) Fred will have to leave and so will Bert.  
 (iv) Fred may have to sit here and Alice there.

In each of the examples in 1.28 the rule of gapping applies to the second conjunct. This rule optionally deletes the second occurrence of a tensed verb and any dependent non-finite verbs. If the schema of 1.27 is adopted then gapping will apply to complements.

It would appear that  $\bar{X}$  cannot capture successfully the head modifier relationship. Consequently, it fails to capture adequately the principle of natural serialization. If the head-modifier relationship is a dependency relation then it may be that natural serialization can be captured in a dependency grammar. We shall explore both this possibility and its consequences below (esp. Ch.3).

### 1.3. Lexical Grammars

Clearly, from our discussion in 1.1.2, the morphological rules or lexical rules as they are often referred to, should not be handled transformationally. This however, I feel, raises two questions.

Firstly, can we find some non-arbitrary means of distinguishing or separating these lexical rules from transformations? Secondly, where in our grammar do these non-transformational rules belong? Let us tackle these two questions in the order given above.

### 1.3.1.

Wasow (1977) is an attempt to establish an independent set of criteria for distinguishing lexical rules from syntactic rules. These criteria are:

1. Lexical rules do not affect structure (i.e. they are structure preserving), whereas transformations need not be structure preserving.
2. Lexical rules may relate items of different categories but transformations do not change node labels.
3. Lexical rules are local, i.e. they involve only NP's bearing grammatical relations to the items in question, whereas transformations need not be local and are formulated in terms of the structural properties of phrase markers.
4. Lexical rules apply before transformations whereas transformations may be fed by transformations.
5. Lexical rules have idiosyncratic exceptions but transformations have few or no true exceptions.

Wasow argues that these criteria provide an explanation for previously unexplained phenomenon. His strongest example involves the passive in English. It has been suggested that some passive participles are adjectives while others are verbs. He claims that

these two passives differ from each other in certain ways predicted by the criteria given above. The surface subject of an adjectival passive must be the deep object of the corresponding active verb. Such passives may not take the logical subject of a lower clause, the indirect object or an idiom chunk, as their surface subjects e.g.:

- 1.29. (i) \*Margaret is unknown to be a humanitarian.  
 (ii) \*Fred seemed told the plan.  
 (iii) \*Advantage seemed taken of Fred.

Verbal passives, on the other hand, may take, as their surface subject any NP which can immediately follow the corresponding active verb e.g.:

- 1.30. (i) Margaret is not known to be a humanitarian.  
 (ii) Fred was told the plan.  
 (iii) Advantage was taken of Fred.

The differences between 1.29 and 1.30 follow from Wasow's claims. Consequently, the adjectival passives are lexical (see 1.1.2 above) whereas the verbal passives are transformationally derived. Wasow weakens the claim that there are no structure preserving transformations (see Freidin 1974, 1975, Brame 1976, 1978 and Bresnan 1978). He allows for a small class of bound structure-preserving transformations - verbal passives and raising. Lexical treatment of both the verbal passives and raising is excluded by Wasow's interpretation of the boundedness criteria. This states that as the word is the maximal unit of the lexicon then lexical rules only refer to information contained in the subcategorization frame of the items they relate. Wasow tightens this definition by stipulating that

lexical rules are formulated in terms of grammatical functions, i.e. thematic functions. Such a functional interpretation of boundedness limits the scope of lexical rules to the subcategorization domain. The lexical passive, for example, in Wasow's terms applies only to logical objects (themes), whereas dative passives e.g. John was given the money are not lexical as an indirect object is affected. However, under the more general definition of boundedness lexical treatment would not be excluded - the indirect object in a V NP NP structure would be part of the subcategorization frame. For raising cases a V S subcategorization is assumed by Wasow. Consequently, passives of the type: John is expected to win the nomination, cannot be treated lexically as the fronted NP does not bear any grammatical relation to the passivized V. It is also not part of the subcategorization frame. However, as we shall see below, these passives can be given a lexical treatment.<sup>8</sup>

S. Anderson (1977) raises further problems for Wasow's criteria. He points out that the criteria are a consequence of grammars within the E.S.T. framework, the basic assumptions of which are accepted as correct, and do not reflect intrinsic properties of rule types. Anderson also suggests that Wasow is wrong to formulate rules in terms of deep subject and object and what is crucial are thematic relations - theme, agent, source and goal, (see Gruber (1965) and Jackendoff (1972)).

---

<sup>8</sup>Although not central to our discussion, it is perhaps worth mentioning that Wasow makes use of grammatical relations as defined by Perlmutter and Postal (1977). Problems for such relational grammars have been highlighted by, for example, Anderson (1979, 1980). On relational grammar see Johnson (1974, 1977a,b) and Perlmutter and Postal (1977).

Neither of these criticisms is fatal. A theory of thematic relations (or even case relations - see Anderson forthcoming on this) can quite easily be incorporated into Wasow's proposals. The criticism that the criteria follow from the adoption of the E.S.T. model is self-evident and the criteria cannot be shown to fail from outwith the model. It is vacuous to apply these criteria to, for example, McCawley (1968), since lexical redundancy rules do not exist in this model. It would be to the support of the criteria if they could be shown to be model independent, i.e. if they are able to discriminate between the two rule types in other non E.S.T. models which utilize both lexical and transformational rules.

Clearly Wasow attempts to provide criteria to allow transformational rules to be distinguished from lexical ones. However, Wasow has abandoned this position, i.e. that function dependency can be combined with boundedness (Wasow 1978, 1980) and accepts that all bounded structure preserving processes are to be treated lexically. In doing this he has moved closer to the position advanced by Bresnan (1978) to which we shall now turn.

### 1.3.2.

Bresnan (1978) proposes that all lexically governed, bounded, structure preserving processes be treated lexically. Under such a proposal the properties of governance, boundedness and structure preservation - the properties of lexical rules - follow from the fact that lexical rules relate entries associated with finitely specified subcategorization features which must be satisfied by base-generated structures. A transformational approach, it is claimed, will leave these properties unexplained. Bresnan believes that the E.S.T.

theory is loose enough to allow both approaches - lexical and transformational. She proposes to solve this by reducing the role of the transformations to the treatment of long distance processes which fall outwith the local scope of lexical processes.<sup>9</sup>

Bresnan's model also makes use of the VP hypothesis (see Bresnan (1971)) in dealing with control phenomena, which receives a surface treatment. Under this hypothesis, for example, subjectless infinitival complements:

1.31. Ignatius wants to eat.

are represented, syntactically, as VP and not as S. The missing subject is filled in at the semantic representation level associated with V's which select subjectless infinitival complements. The implication of this is that passives of the following type:

1.32. John is expected to win the booby-prize.

which Wasow could not handle lexically because the fronted NP bears no grammatical relation to the passivized V, can be given a lexical treatment. Under the VP hypothesis expect verbs are assigned a V NP  $\overline{VP}$  frame with passivization of the NP being bounded.

As stated above Bresnan decided that the role of the transformational component was to be reduced. Lexical, semantic and pragmatic relations are distinguished from transformational ones and factored out of the transformational component. An instance of lexical relations which are factored out would be the examples above from

---

<sup>9</sup>Bresnan (1979) abandons the transformational approach to long distance processes in favour of an interpretive approach.

Siegel (1974) - see 1.1.2. However, as Chomsky (1970) points out, if the role of one component of the grammar is to be reduced and the grammar is still to achieve its purpose then the other components must be enriched. This assumption holds true for Bresnan's model in which the lexicon is given an expanded rôle. We shall now look at Bresnan's lexicon.

The minimal semantic information stored for verbs in the lexicon is the logical argument structure, i.e. whether the verb is intransitive, transitive or (but not noted by Bresnan) ditransitive. Verbs will be categorized by one, two or three place relations. Other semantic information will be recorded e.g. agent, patient, theme etc. The syntactic contexts in which a verb can appear are not sufficient to represent their argument structure, e.g. two verbs may have different types of argument structure in the same syntactic contexts. Consider wept and teach:

- 1.33. (i) Vladimir wept.  
 (ii) Josef teaches.

Both verbs can be intransitive but in 1.33(ii) there is clearly the implication that Josef taught something. The conclusion being that the verb teach has a logical object even when the grammatical one is absent. In this way the argument structure of weep differs from teach. This raises the major issue of Bresnan's grammar: how are the logical argument structures of verbs related to their syntactic contexts?

Bresnan's solution does not operate upon the syntactic structure of the verb but on the logical argument structure e.g. the argument structure of teach can be converted from a two place relation into a

one place relation. This is carried out by a logical operation: variable binding of quantification e.g.:

- 1.34. (i)  $x$  TEACH  $y$   
 (ii)  $(\exists y)$   $x$  TEACH  $y$ .

1.34(i) is a two place relation whereas 1.34(ii) is a one place relation. According to Bresnan, if we accept 1.34(ii) as the lexical argument structure for the intransitive verb teach, then it is easy to explain how 1.33(ii) differs from 1.33(i) and how 1.33(ii) is related to:

- 1.35. Freda teaches something.

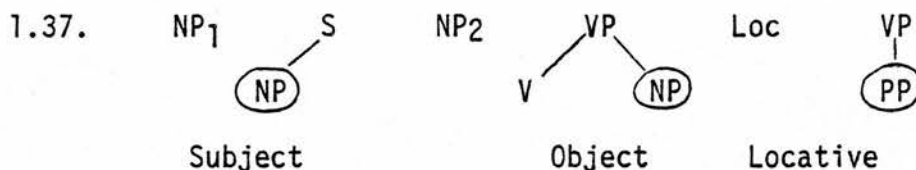
Bresnan (1978:16) claims that providing 1.34(ii) as lexical information seems quite natural because the intransitive use of normally transitive verbs is a property of individual verbs, e.g. teach, eat but not hit and squash. She also claims that entering 1.34(ii) in the lexicon, avoids the counter intuitive conclusion that 1.33(ii) requires more grammatical processing than 1.35. It is not clear why this is counter intuitive. Also 1.33(ii) still requires more processing of some sort in Bresnan's account. The issue is whether that processing is syntactic, e.g. a rule of indefinite object deletion or lexical by a rule of quantification.

To make lexical associations between argument structure and syntactic structures, Bresnan claims that a set of grammatical functions are required. If such a set is not provided then it is impossible to distinguish between:

- 1.36. (i) John teaches something.  
 (ii) Something teaches John.

as both have the argument structure of 1.34(i). However, this may be resolved by appealing to thematic relations, e.g. agent and patient or ergative and absolutive.

Bresnan defines these functional relations configurationally:



Having set up the relation between argument and syntactic structure the following lexical representations are provided by Bresnan:

- 1.38. (i) sleep V, [\_\_ ] NP<sub>1</sub> SLEEP  
 (ii) hit V, [\_\_NP] NP<sub>1</sub> HIT NP<sub>2</sub>  
 (iii) teach V, [\_\_NP] NP<sub>1</sub> TEACH NP<sub>2</sub>  
 [\_\_ ] (∃y) NP<sub>1</sub> TEACH Y

As in the Aspects model, the information in the square brackets provides the syntactic contexts for the insertion of the verb, and the information on the right is referred to as the functional structure. The task of the function structure is to combine grammatical functions with logical argument structures.

This notation can be used to express the function structures of passive verbs. Consider the following:

- 1.39. (i) John was teaching.  
 (ii) John was taught.

In 1.39(ii) the passive verb is syntactically intransitive just as the active verb is, in 1.39(i). The difference between the two structures lies in the logical argument structures. In the former, the logical

object has been eliminated whereas in the latter, the logical subject has been eliminated:

- 1.40. (i)  $(\exists y) x \text{ TEACH } y$  (= 1.39(i))  
 (ii)  $(\exists x) x \text{ TEACH } y$  (= 1.39(ii))

Also in 1.39(i) John plays the rôle of logical subject  $x$ , but in 1.39(ii) it is the logical object. This can be expressed as:

- 1.41. (i)  $(\exists y) \text{NP}_1 \text{ TEACH } y$  (= 1.39(i))  
 (ii)  $(\exists x) x \text{ TEACH } \text{NP}_1$  (= 1.39(ii))

From this it is possible to provide lexical entries for the passive verbs (be) eaten, (be) taught and (be) hit:

- 1.42. (i) eat + en V, [be \_\_ ]  $(\exists x) x \text{ EAT } \text{NP}_1$   
 (ii) teach + Pass V, [be \_\_ ]  $(\exists x) x \text{ TEACH } \text{NP}_1$   
 (iii) hit +  $\emptyset$  V, [be \_\_ ]  $(\exists x) x \text{ HIT } \text{NP}_1$

The agentive by-phrase, which optionally appears with passives can be analysed as an optional prepositional phrase functioning semantically to identify the logical subject of the passive verb:

- 1.43. eat + en: V, [be \_\_ [pp by NP]]  
 $(\exists x) (x \text{ EAT } \text{NP}_1 \text{ and } x = \text{NP}_{\text{by}})$

A comparison between the lexical entries in 1.38 and 1.42 reveals that the syntactic contexts are related by the following simple rule:  $[\text{__}(\text{P}) \text{NP}] \rightarrow [\text{be } \text{__} (\text{P}) \dots]$ . The function structures are related by the operation "Eliminate  $\text{NP}_1$ " and, either "Replace  $\text{NP}_1$  by  $\text{NP}_2$ " or, "Replace  $\text{NP}_p$  by  $\text{NP}_1$ ". Bresnan refers to these operations collectively as the 'Active-Passive Relation'. This relation, as expected, will not apply to non-passivizing verbs, but this is due

more to the way in which the notation is set up, rather than anything intrinsic in the model, i.e. it provides no explanation why some verbs do not passivize.

This problem may be resolved by the same means we appealed to in resolving the difficulties posed by the examples in 1.36. We could invoke a set of thematic relations and make the 'Active-Passive Relation' sensitive to these relations. Following Brown and Miller (1980:325) the logical object of a passive verb would have to be either a [neutral, patient] or a [neutral, result] and the verb itself would be [+active].<sup>10</sup> Only if these conditions were fulfilled would the 'Active-Passive Relation' apply.

The set of lexical relations is not fixed and may change to embody those concepts which become important for communication. This is forced upon Bresnan if she is to allow for historical change. If there are no transformations to be added, lost or reordered then the set of lexical relations must be permitted to change. However, the grammar predicts neither which ones will change nor the rate of change.

It could be argued that the lexical relation is simply a notational variant of the transformation. However, it must be remembered that unlike a transformation the lexical relation is "structure independent"; i.e. according to Bresnan it is function-dependent in the sense that it relates lexical function structures and not grammatical structures (phrase markers). In the case of the active-passive relation - a lexical relation - the deep structure of both active and passive sentences is equal to the surface structure. Bresnan (1978)

---

<sup>10</sup>For a definition of these terms and a further discussion of their rôle in a grammar, see Brown and Miller (1980: Part Three).

does still require transformations in her model to carry out processes which are solely structure dependent - question formation and verb number agreement being two examples.

As with Wasow, Bresnan faces problems concerning her use of functional relations and particularly their configurational definition. These problems are enhanced by Bresnan's claim that:

"Perhaps the active-passive relation belongs to a universal "logic of relations" by which the lexicon of a human language - the repository of meanings - can be organized."

Bresnan (1978:23)

This appears to be a claim for universality of lexical rules.<sup>11</sup> Such universality would have to be based on grammatical functions as the lexical rules operate on function structures. However, can grammatical functions be given a universal definition? Keenan (1976) appears to believe so, but Anderson (1978, 1979a, 1979b, 1980 and 1984) thinks not. At the risk of being diverted from our main discussion let us look briefly at the evidence for a configurational definition of object.

The configurational definition of object, in English, is that NP which is immediately dominated by VP (Chomsky, 1965; Katz, 1972 and Saunders, 1972). The viability of the configurational definition depends crucially upon the universality of VP and the uniqueness of object NP's as daughters of VP. A universal configurational definition of VP is not possible if constituents must be continuous. Both VSO and SOV languages prove problematical (see Schwartz, 1972; Hinds, 1973). McCawley (1970) argues that in a SVO language like English VP is difficult to motivate in base as opposed to derived

---

<sup>11</sup>See also Bresnan (1978:16) "...the set of grammatical functions... can be considered universal".

structures. If VP cannot be motivated configurationally, then Bresnan's function structures have no motivation and subjects are not configurationally distinct from objects. This means that the 'Active-Passive Relation' cannot be formulated to operate on function structures and secondly the following two sentences cannot be distinguished:

- 1.44. (i) John ate something.  
 (ii) Something ate John.

Both sentences use EAT which has the argument structure  $x \text{ EAT } y$  (see above).

If we accept that  $VP^{12}$  is motivated then objects will be well defined configurationally only if they are the sole instances of a particular category dominated by VP. Consider the following example:

- 1.45. Scott gave Harriet the whisky.

This example is problematical unless both NP's dominated by VP are objects and the definition is not required to produce a unitary output.<sup>13</sup> To assign both NP's object status would be disastrous for Bresnan's function dependent lexical relations. Consider the effect of the active passive relation

- 1.46. (i) Harriet was given the whisky by Scott.  
 (ii) \*The whisky was given Harriet by Scott.

Clearly some way must be found to distinguish the post verbal NP's

<sup>12</sup>There are some linguists who are prepared to accept discontinuous VP's, see, e.g. Shieber (1984).

<sup>13</sup>Recall Bresnan does not discuss ditransitive verbs.

from each other. We could label them  $NP_2$  (direct object) and  $NP_3$  (indirect object) or  $NP_2$  and  $NP_{obl}$  (oblique).<sup>14</sup> Consider the following:

1.47. Scott gave the whisky to Harriet.

1.47 is related to 1.45 in transformational accounts by a rule of dative movement.  $NP_2$  in 1.45 is a prepositional object in 1.47 according to Bresnan's framework and  $NP_3$  in 1.45 is a direct object in 1.47 ( $NP_2$ ). From 1.46 we can see that the former prepositional object of 1.47 can undergo the 'Active-Passive Relation' whereas the original direct object cannot. Let us call the lexical rule relating 1.47 and 1.45 the 'Dative Relation'. This relation tells us that  $NP_p$  replaces  $NP_2$ , the preposition is eliminated and  $NP_2$  replaces  $NP_p$ . However,  $NP_p$  must also now become  $NP_2$  to be available for the 'Active-Passive Relation', and  $NP_2$  must become  $NP_3$  or  $NP_{obl}$ . Consequently, grammatical relations are not fixed but some are derived.<sup>15</sup> The implications of this are that grammatical relations cannot be given a universal definition and cannot be primitives in the grammar. If this is the case then Bresnan cannot appeal to configurationally defined function structures and 1.44(i) cannot be distinguished from 1.44(ii) and her formulation of the 'Active-Passive Relation', which is based on configurationally defined function structures, will fail. In the case of 1.45 and 1.47 we could reject the proposal that grammatical relations are not fixed but we are then forced to claim that 1.45 and 1.47 are not related when clearly they are.

---

<sup>14</sup>See Anderson (1978) on the problems of the indirect object.

<sup>15</sup>This is not the usual sense of derived. The point being that the rules which relate the function structures of 1.45 and 1.47 must be allowed to change the functions of the NP's involved.

Although this discussion is short and somewhat inconclusive it does, I feel, show that there are problems involved with configurational definitions of grammatical functions. Furthermore, these difficulties raise problems for Bresnan's model. I feel that this is not an appropriate place to pursue, in detail, these issues, as they are not central to our discussion. For further discussion I refer the reader, once again, to the references of works by Anderson quoted above.

### 1.3.3.

It is clear from the preceding discussion that the rôle of the transformational component can be reduced and that a set of lexical rules can be identified on a non-ad-hoc basis. This answers the first of the two questions posed in 1.3. Both Bresnan (1978) and Wasow (1977) put these rules in the lexicon - whether or not this is appropriate is an issue to which we shall return.

It is also clear from the work by Bresnan and Wasow that lexical grammar is not homogenous but consists of (many) diverse strands. Hoekstra et al (in Hoekstra et al (1980)) attempt to draw on the similarities between lexical models in an effort to provide an overview of lexical grammar. It is to this, more general, overview that I now wish to turn.

### 1.3.4.

It has been argued that the rules which build up complex lexical items are of the same categorial status as those which build up phrase markers. The difference between the two being one of scope of application, e.g. Jackendoff (1975:668) states that the canonical

derivation rules exhibit structure building qualities making them the lexical equivalent of a P.S. rule. Although this may be the case, lexical rules also contain the following types of information: (a) morphophonological, (b) contextual properties associated with a subcategorization frame and (c) semantic representations related to the syntactic frame.

As an example consider the following:

1.48. Adj  $\rightarrow$  V + able.<sup>16</sup>

This rule states that Adjective is expanded into Verb-stem plus an affix - able in this case.

The subcategorization features (or syntactic frame) of the rule serve a double purpose. Firstly, they define the context for lexical insertion. Secondly, when a lexical rule (e.g. 1.48) relates two entries the frames serve as the syntactic input and output of the mapping defined by the rule. All lexical rules are structure preserving.<sup>17</sup> Lexical insertion consists of the replacement of a terminal element  $\Delta$  by a lexical item. The mapping between two syntactic frames of related entries is also structure preserving in that both items must be inserted into base generated structures.

The input to a lexical rule is referred to as the syntactic frame and consists of, as we noted above, a set of subcategorization

---

<sup>16</sup>This is not strictly a P.S. rule, and states that adjectives have the constituents V + able, rather than that in certain circumstances an adjective may be formed by suffixation of -able.

<sup>17</sup>Bresnan (1978) treats long distance process transformationally but in Bresnan (1979) she gives even these processes a lexical treatment.

features. The output, or 'structural change', is referred to as the translation. The relationship between the syntactic frame and its translation is governed by compositionality i.e. subcategorizing an item for a particular frame entails that its meaning is built up in a compositional way as a function with the translation of the phrases mentioned in the frame as arguments. Lexical rules, transformations with subcategorization frames as their S.D.'s, manipulate the function of the frame by re-arranging the order in which the arguments plug into the original function. They may also alter the degree - or valency - of a predicate by either decreasing or increasing the number of arguments it takes. Given this, lexical rules are, in some cases, carrying out permutations, deletions and additions, all of which P.S. rules are forbidden to do. In this respect they bear more resemblance to T-Rules than to P.S. rules.

Let us now consider how these functions may be formalized, taking the able rule (1.48) as an example:

$$\begin{aligned}
 1.49. \quad \text{ABLE}(i) = j \text{ for} \\
 & \langle i, [\Delta]_V, [V \text{ NP}], V'(\text{NP}') \rangle \\
 & \langle j, [[\Delta]_V, \text{able}]_A, [A], \lambda x \exists y [V'(x)(y)] \rangle
 \end{aligned}$$

The input of this rule is a transitive verb and the output is that verb to which able has been suffixed to produce an adjective. The logical object of the input verb becomes the grammatical subject of the output. When the lambda expression is combined with the translation of the grammatical subject it will be understood as the logical object of the verb to which able has been suffixed.

This rule bears a striking similarity to a G.P.S.G. meta-rule (see

below) but neither this nor a meta-rule is, i.e. strictly speaking, a transformation, - although both behave like or have similar powers to a transformation. Furthermore the above discussion allows an interpretation of lexical rules which is quite similar to the lexical building rules of generative semantics, where the same rules for building phrase markers could also be used to build lexical items. One major difference is that generative semantics used the rules to build morphologically non-complex and lexically distinct lexical items e.g. kill < cause to die. Also generative semantics put these rules into the syntax and consequently allowed for abstract lexical items (see Lakoff, 1970 and c.f. Aronoff, 1980).

Hoekstra et al (1980) take the view that word formation rules are not typologically distinct from those dealing with the syntax of words. This is formalized in a system of word structure rules (W.S.R.'s). These rules account for word structure in terms of an  $\bar{X}$  system operating below word level. It is assumed that these context free rules will generate the hierarchical structure of an infinite set of word frames. The insertion of context sensitive affixes is governed by a finite set of insertion conditions associated with each affix.

The W.S.R. component contains a set of ordered pairs of complex symbols consisting of an integer  $i$  and a syntactic feature matrix. These features e.g.  $\{\pm N, \pm V, \pm M, \dots\}$  define the category status of the morpheme while the integer defines the level of the category. Within this model it is either 0 or -1.

Hoekstra et al (1980) utilize an  $\bar{X}$  system which was criticized above - 1.2 - to allow W.S.R.'s to make use of the head-modifier

relationship (which  $\bar{X}$  fails to capture - see 1.2). The W.S.R.'s for suffixation and prefixation are respectively:

- 1.50. (i)  $X^0 \rightarrow Y^0 X^{-1}$  (i)  $\begin{array}{c} A^0 \\ \swarrow \quad \searrow \\ V^0 \quad A^{-1} \end{array}$  (ii)  $\begin{array}{c} V^0 \\ \swarrow \quad \searrow \\ V^{-1} \quad A^0 \end{array}$
- (ii)  $X^0 \rightarrow X^{-1} Y^0$
- drink   able                      en   able

In the above constructions the affixes are considered to be the heads and carry with them a set of syntactic features to determine to which category the resulting structure belongs. This expresses the traditional view that the affix determines the category of the derived word. There are, of course, problems with prefixes which do not usually alter the categorial status of the base and we shall return to these below. Affixes are assigned to level -1, that of bound morphemes and maximally project structures of the 0 level i.e. free morphemes - words. Level 0 is the recursive cycle domain of word formation - the maximal processing unit of W.S.R.'s. The bases or complements selected by an affix are of level 0 and in this the W.S.R. is in keeping with the traditional view of morphology.

As noted above, one of the problems for this system concerns the status of prefixes which do not usually alter the categorial status of the base. Hoekstra et al (1980) claim that amongst prefixes those with head-like properties tend to be the exception and consequently consider 1.50(i) to be the unmarked situation concerning the relationship of head to base in a W.S.R. This forces them to recognize two classes of prefix: those which function as heads and introduced by 1.50(ii), and the class of category neutral prefixes. To handle these prefixes the feature [ $\pm$ major] ([ $\pm$ m]) is introduced. Heads projecting an expansion of the maximal type are [+m] and those which do not are

[-m]. Consequently, category neutral prefixes are [-m] and introduced by the following rule:

$$1.51. \quad X^{\circ} \rightarrow Y^{-1} X^{\circ} \\ \quad \quad \quad [-m]$$

In this case it is the base and not the affix which functions as the head.

This system specifies the following set of rules:

$$1.52. \quad (i) \quad \text{read} \quad \begin{bmatrix} +V \\ -N \\ +m \end{bmatrix}^{\circ} \dots$$

$$(ii) \quad \text{-able} \quad \begin{bmatrix} +V \\ +N \\ +m \end{bmatrix}^{-1} < \begin{bmatrix} +V \\ -N \\ +m \end{bmatrix}^{\circ} \_ >, \dots$$

$$(iii) \quad \text{en-} \quad \begin{bmatrix} +V \\ -N \\ +m \end{bmatrix}^{-1} < \_ \begin{bmatrix} +V \\ +N \\ +m \end{bmatrix}^{\circ} >, \dots$$

$$(iv) \quad \text{pre-} \quad \begin{bmatrix} -V \\ -N \\ -m \end{bmatrix}^{-1} < \_ \begin{bmatrix} +V \\ -N \\ +m \end{bmatrix}^{\circ} >, \dots$$

which, with the apparatus introduced in preceding discussion, constitute the context-free part of the system.

Context-sensitive insertion of bound morphemes is carried out by a series of insertion conditions, which function as S.D.'s for the insertion of these affixes, i.e. subcategorization frames. These conditions factorize the structure created by a W.S.R. and check whether or not it satisfies the S.D. e.g. 1.52(ii)-(iv) give the insertion conditions associated with the respective affixes: they specify where the affix occurs in relation to the base and the categorial status of the base.

Application of the insertion conditions is subject to analysability c.f. transformations, and therefore they are constrained. Consequently, the power they impart into the grammar is reduced. The

conditions are restricted to morphological well-formedness, with all others being accounted for in the appropriate subsection of the lexical rule. Given this consideration, the insertion frames are described by the following (Hoekstra et al, 1980:25):

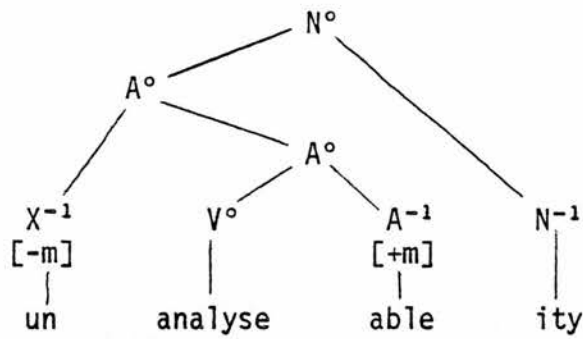
- 1.53. (i) Categorical frames are of the form  $\langle \alpha \_ \beta \rangle$   
 where either  $\alpha$  or  $\beta$  is empty and  $\alpha$  and  $\beta$   
 belong to the set defined by the syntactic  
 features  $\{\pm N, \pm V, \pm m, \dots\}$
- (ii) Morphophonological conditions on a  
 categorical frame are Boolean conditions  
 on (i), referring to:
- (a) the phonological string
  - (b) a specified morpheme
  - (c) a morphological feature.

The constraints on the power of the insertion conditions take the form of restrictions on the values that  $\alpha$  and  $\beta$  in 1.53(i) can take.<sup>18</sup> Williams (1978) has formulated what he calls the Head Condition. This is a locality restriction on the maximal depth of structure analysable by an insertion frame. Such a frame for an affix P can refer to the head of the base Q to which P is attached but not to any complement of that head. Given this, consider the following example:

---

<sup>18</sup>Siegel (1974, 1978) and Aronoff (1976) discuss several restrictions e.g. -al requires that if the stem ends in a consonant then that consonant must be [+ant], -ity can only be attached to a [+linate] base.

1.54.



The affix -ity requires the feature [+lat](inate) to be on the base. When the H.C. factorizes the base for the -ity rule, it can descend to the head of the base i.e. to -able which is [+lat]. It is necessary to go down to -able, as features such as [+lat] do not percolate upwards.

Having established this schema for derivational morphology, Hoekstra et al (1980:26) turn their attention to inflectional morphology. They decide that it is necessary to organize the grammar in such a way that the differences between inflection and derivation are clearly expressed. Inflectional rules are considered to be fully productive and transparent in meaning. They do not discuss either irregular plurals - consider e.g. scissors, which could be singular or plural - or the past tense forms of strong verbs. They would also like to maintain the principle of lexical integrity which treats words, including inflections, as atomic with respect to syntactic rules. This removes the means of inflection marking by generation of inflectional morphemes on abstract base positions with transformations later attaching them to the appropriate stem.

To solve this problem the lexicon is divided into a store and a process component. The store handles all non-inflectional aspects of morphology and is organized as a full entry component together with a set of lexical rules to factorize out all redundant information.

The process component operates on the output of the store. It operates by processing the member of the inflectional paradigms, which can then be inserted in their fully specified inflectional form in a base generated structure.

These rules are formally of the following type and considered as "spell out" rules:

$$1.55. \quad \begin{array}{l} /X/ \rightarrow /X + a/ \\ \left[ \begin{array}{l} +F \\ -G \end{array} \right] \end{array} \quad /VERB/ [+3sg] \rightarrow /VERB + S/$$

Features spelled out by these rules are either morphosyntactic or morpholexical. The latter are local in the sense that they are present only at the  $X^0$  level and do not project up to higher levels e.g. verbal aspectual morphology. The former project up to the phrasal  $X^n$  level and are involved in phrase internal and inter-phrasal agreement.

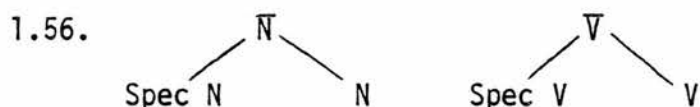
Before moving on I shall pause to discuss some points of the preceding component.

Hoekstra et al (1980) accept without question the  $\bar{X}$  system. This model has already been criticized for its failure to capture the relationship between head and modifier - exactly one of the relationships it was supposed to capture.  $\bar{X}$  is not really a dependency model. Hoekstra et al (1980) weaken this attempt to capture the head-modifier relationship by allowing the head of a complex lexical item to fall on either the left or the right of a modifier.  $\bar{X}$  claims that NP, VP and AP all share a common source i.e. a specifier plus a head.<sup>19</sup> Further, in each case the specifier falls to the left of the head allowing the general schema  $\bar{X} \rightarrow [\text{Spec } X]X$  to be

---

<sup>19</sup>See the following discussion.

posited. This schema gives the following structures:



Consider the rules of 1.50 which produce the following structures:



It would appear that the generalization established at phrasal level does not hold at word level. It is not possible to provide a single rule schema as no generalization over affixes exists.

A second problem arises when we consider the lexical entries. The  $[\pm N]$ ,  $[\pm V]$  feature system was discussed above (1.2) and need not be discussed again at this juncture. Hoekstra et al (1980) make no attempt to justify the  $\begin{bmatrix} -V \\ -N \end{bmatrix}$  feature specification, normally reserved for prepositions (see Jackendoff, 1977), given to pre in 1.52.

Furthermore the addition of  $[\pm m]$  appears ad-hoc and can easily be dispensed with. This alteration to the grammar may also lead to the simplification of the lexical entries for both  $[-m]$  and  $[+m]$  entries. Let us now see how this might be achieved.

Consider the ABLE rule (1.49) which transforms a transitive VP into an intransitive AP. Both the input and the output are specified by the rule. In the case of  $[+m]$  items the output differs from the input, whereas with  $[-m]$  items the subcategorization for the item remains the same (but the semantics would change e.g. kind vs unkind). If heads decide the categorial status of the item then where the input and output of the rule differ in their subcategorization, the affix specified in the rule must be the head and there-

fore [+m]. Here we have two markers of headhood - the form of the rule, which specifies the item and subcategorizes it and the feature [+m]. Consequently, one of these may be disposed of; and as the rule contains all the information required, [ $\pm$ m] can go. Furthermore as the ABLE rule, for example, creates an AP from a transitive VP and adjectives are  $\begin{bmatrix} +V \\ +N \end{bmatrix}$  the whole lexical entry may be disposed of. A similar rule may be posited for adjectives created by -ING suffixation. Rules with [-m] would also be simplified, differing from category changing rules by having the same input and output in the syntactic component. Further simplification could be achieved by treating inflectional morphology as suffixation, where the syntactic subcategorization of the input and output of the rule remains unchanged. This would remove the process component but conflate derivational and inflectional morphology.

Much of the work in the morphological section of lexical grammar appears to have been carried out without historical perspective on the subject. Consider, for example, Hoekstra et al's heading of a discussion of Halle (1973) as "The Rediscovery of Morphology" (see Lipka, 1975 for a discussion of Halle, 1973) or Carlson and Roeper (1980):

"Morphology is a new member among the recognized autonomous systems in linguistic theory. Some of its crucial characteristics are known. It has a vocabulary of primitives (root, morpheme, word, stem, etc.) ...."

Such statements seem odd when one considers the published literature e.g. Brown and Miller (1980:161ff), Hockett (1954, 1958), Lyons (1968:53), Marchand (1969) or Matthews (1965a, 1965b, 1966, 1967, 1972 and 1974) amongst others.

## 1.3.5.

Our discussion of Hoekstra et al has, so far, concentrated only on the word structure aspects of the model. Let us now briefly look at the syntactic aspect of lexical grammar.

Lexical syntax is surface syntax and like Bresnan (1978) all governed processes are given a lexical treatment e.g. causatives, passive and dative. Long distance processes e.g. wh-movement are still treated transformationally.<sup>20</sup>

This is achieved by adopting the translation principle:

Subcategorizing  $X^0$  for a syntactic frame amounts to the claim that its meaning is a complex function with the various set theoretic objects corresponding to the subcategorized phrases playing the role of arguments.

This formulation is based on Dowty (1978:418). Through this translation principle, each syntactic frame is associated with a corresponding expression in the language of semantic representation. The meaning of the semantic representation is built up compositionally from the meaning of the parts mentioned in the syntactic frame.

Adoption of this principle removes the necessity for a syntactic level of deep structure from the grammar. The level of syntactic deep structure is where logical and grammatical functions coincide in a one-to-one fashion.<sup>21</sup> If the lexicon now explicitly associates a translation with each syntactic frame (subcategorization features) then, as the translation principle realizes the mapping between

---

<sup>20</sup> This follows the view that lexical rules are structure preserving - see 1.3.2 above.

<sup>21</sup> This is, in effect, what Bresnan (1978) proposes with her function structures. See 1.3.2 above.

grammatical and logical arguments, the level of deep structure is no longer required and can be dispensed with.

These principles can best be illustrated by looking at a fragment of a lexical grammar, taken from Gazdar (1979). Consider the following syntactic frames and their corresponding translations:

- 1.58. (i)  $\langle 18, {}_{VP}[V NP \overline{VP}]_{\lambda x}[V'(\overline{VP}'(x))](NP') \rangle$   
 $V18^* = \{\text{believe, expect, .....}\}$
- (ii)  $\langle 19, {}_{VP}[V NP \overline{VP}]_{\lambda x}[V'(\overline{VP}'(x))(x)](NP') \rangle$   
 $V19^* = \{\text{persuade, force, ask .....}\}$
- (iii)  $\langle 20, {}_{VP}[V NP \overline{VP}]_{\lambda x}[V'\overline{VP}'(x)](NP')(x) \rangle$   
 $V20^* = \{\text{promise, .....}\}$

Each rule is a triple, of which the first member is a unique arbitrary integer - the number of the rule, the second is the syntactic or sub-categorization frame, and the third is a semantic rule. This semantic rule shows how the semantic representation of the expression in the syntactic frame is built up from the semantic representation of its parts.  $V_n^*$  represents the class of lexical items introduced by rule n.

Categorially, each frame is identical and represents the surface environment for each item without recourse to syntactic abstractness. The differences lie in the translations associated with each frame. Compositionality is not threatened by the one-to-many relation between a single frame and the three distinct translations as each translation is uniquely associated with each verb type by a rule integer.

The verb classes of 1.58 are distinguished by their control properties and argument structures. Constituents count as arguments only when

they function as logical arguments of the predicate in question. Consider, for example, the syntactic object of verbs of the believe type. The translation associated with this class shows that they only have one argument, the propositional object -  $(\overline{VP}'(x))$ , in which the translation of the syntactic object, NP', plays the role of subject. The syntactic object has no logical function to believe itself.

Control properties distinguish between the persuade and promise classes. In the former class the complement is controlled by the matrix object, while in the latter the understood subject of the complement is the matrix subject. This difference is reflected in the translation. In the case of a verb of the persuade class the VP consists of a complex transitive function combined with an NP argument. The argument corresponds to the translation of the syntactic object. The complex function is the expression within the scope of the lambda operator. The function-argument structure is interpreted so that the NP denotation is plugged in for the occurrences of the variable, the lambda operator abstracts on. The translation establishes the fact that the syntactic object NP is the controller of the subject of the complement.

Consider now promise (1.58(iii)). In the case of promise the translation of the promise VP consists only of the complex intransitive function formed by the lambda operator. The translation of the syntactic object of promise is within the scope of this operator. Consequently, within the translation of a promise VP, there is no argument to satisfy the lambda operator. This function can only be satisfied when the intransitive VP translation is combined with the NP denotation corresponding to the subject of promise. By plugging



in the NP denotation of the subject for the occurrences of the variable  $X$  we see that the syntactic subject of promise, in fact, controls the complement.

This interpretive treatment of control is called the VP hypothesis (see 1.3.2 above and Bresnan, 1971).

Within a lexical grammar, passive, for example, would be characterized as an operation defined on the semantic representation of lexical items. It can be defined on structures of the type of 1.59(i) and has 1.59(ii) as its output:

- 1.59. (i)  $\langle m[V\ NP\ X], F(NP') \rangle$   
 (ii)  $\langle n[V\ X], \lambda x[\exists y\ V'(x)(y)] \rangle^{22}$

This rule turns transitive VP's into intransitive ones, where the syntactic subject of the passive construction is mapped onto the logical object of the original function. The variable  $F$  in 1.59(i) ranges over lexical or complex phrasal functions. Rules <18> and <19> (1.58(i)-(iii)) satisfy the input for passive with  $F$  equal to the lambda expression. Rule <20> (1.58(iii)) does not satisfy the input since in the VP translation no NP denotation satisfies the lambda function.

It appears that lexical rules can be used not only to build up word structure by carrying out both derivational and inflectional morphology but they can also be used to perform operations which used to be performed by structure preserving transformations. NP movement processes, for example, can be interpreted in terms of properties of individual lexical items related by lexical rules -

---

<sup>22</sup>See Gazdar (1979:28).

1.59 above - rather than as mappings between phrase markers in which such lexical items may occur.

### 1.3.6.

It is clear from the preceding discussion that many linguists are in agreement with the view expressed in Chomsky (1970); i.e. that the power of the transformational component ought to be reduced. They also agree that it is the lexicon which should be given a greater role. Differences arise over how much work the lexicon ought to do and to what extent the power of the transformational component is to be reduced. Some linguists, for example Gazdar, Pullum and Sag (1981) dispense with transformations totally, while others, notably Bresnan (1978) and Hoekstra et al (1980) retain transformations for long distance unbounded ungoverned processes e.g. Wh-movement.<sup>23</sup>

The view that transformations play too great a role in an Aspects type grammar, where, for example, all morphological processes are treated transformationally, gains support from traditional approaches.

Traditional grammars differentiated between inflection and syntax. Inflection deals with the internal structure of words while syntax account for the way words combine to form sentences.<sup>24</sup>

Inflection, dealing with the internal structure of words, is clearly what we refer to as morphology. Traditionally, it is concerned with the internal structure of words and their relationship to other words within the paradigm. A paradigm is a model based on a member

---

<sup>23</sup>Any discussion of the relative merits of these two approaches lies outwith the scope of this thesis.

<sup>24</sup>See Lyons (1968:4.1.2 and 5.3) and Matthews (1974:2ff and 154ff).

of the declension class (for nouns and adjectives), or conjugation (for verbs), from which the person using the grammar could construct the other members of the type by reference to the paradigm. In this sense morphology deals with paradigmatic relationships.

Syntax, as noted above, deals with the rules for forming phrases and sentences. It is concerned not with the relationship of words to other words within the paradigm but with their external relationship to other words within the sentence - it is concerned with syntagmatic relationships.<sup>25</sup>

This leaves us with the lexicon. It will consist of a list of all the words in the language classified according to their membership of the distributional classes referred to in the grammatical rules. Information concerning selection restrictions and strict subcategorization will be included, as will the meaning of each item. The lexicon will also be the repository of "irregular" forms. These will be items which fail to participate in regular morphological processes e.g. the ablaut series of strong verbs, irregular plurals or the failure of the modals to undergo verb number agreement.

Clearly, none of the models we have reviewed so far, adhere to this traditional position. The Aspects models conflate syntax and morphology by allowing the transformational component to handle both. The lexical grammars are guilty too of such a conflation, but allow the lexicon to carry out both syntactic and morphological processes.

Let us assume that the traditional view is correct and set up a grammatical model in which there is a syntactic and morphological

---

<sup>25</sup>However see work by Halliday and Pike (1963) where this distinction is not maintained.

component demarcated along traditional lines since the basis for the division is, in principle, clear and more restrictive. Furthermore, as our main interest is syntactic change let us assume that the lexicon is the locus of syntactic change; i.e. syntactic change proceeds through changes in the lexical restrictions associated with governed rules. One clear exception to this is word order change which requires a change in the P.S. rules.

It is clearly more restrictive than Aspects or E.S.T. type grammars where change may occur in the P.S. rules, T-rules or lexicon. This is also an improvement over a lexical model where the claim that change occurs in the lexicon is redundant and vacuous as the lexicon is the only component where change can occur.

In the next chapter, I shall test this lexical approach against diachronic data from the history of English to discover whether it, or the syntactic approach, gives a superior account of change.

However, before proceeding to this task, let us look at one more syntactic model to which reference has already been made, i.e. generalized phrase structure grammar, henceforth G.P.S.G.

#### 1.4. Generalized Phrase Structure Grammar: An Excursus

Gazdar, Pullum and Sag (1981), henceforth G. P. and S., investigated the possibilities of using a context free phrase structure grammar as a model of grammatical description, in the wake of restrictions placed upon the expressive power of transformations. These restrictions are:

- (i) Restrictions which decrease the membership of the set of languages that can be defined using grammars permitted by the theory.
- (ii) Restrictions to decrease the membership of the set of grammars permitted by the theory.
- (iii) Restrictions which enhance the learnability of grammars permitted by the theory.

Various linguists have studied the effects of either one or two of these restrictions, but G. P. and S. attempt to apply all three. However, type (iii) restrictions require a sound base of research in language acquisition before being fully pursued, but this is a minor point - at least for our purposes.

G. P. and S. restrict their account by permitting themselves access only to descriptive mechanisms available within the theory of context-free P.S.G. or devices which are equivalent in their effects to such mechanisms. This implies that no transformational, interpretive, deletion or readjustment rules are to be employed. By adhering to this they claim to be upholding type (iii) restrictions.

The notation of G.P.S.G. is developed within the  $\bar{X}$  system (see 1.2 above) which we have already discussed and consequently do not intend to pursue the matter any further.

Within G.P.S.G. the notion 'head' is invoked but G.P.S.G. is not a dependency model (c.f. Robinson, 1970a and Ch.3 below). In place of Robinson's three identifying criteria for headhood (see 3.6), G.P.S.G. invokes four conditions, one of which has three separate conditions attached to it. The notion 'head' is only introduced to ensure that

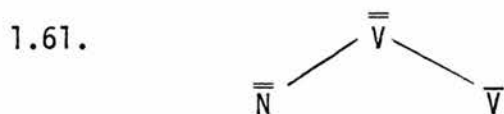
features marked on the  $\bar{N}$  of an  $\bar{N}$  phrase are identical to those on the  $\bar{N}$ . The concept of governance is not utilized at any point, e.g. the object of a transitive verb is not governed by that verb.

The removal of all but the context-free phrase structure rules from the grammar poses a number of problems for G. P. and S. They are still required to handle the same data as T.G. but have no access to the descriptive devices of T.G. One major problem is how to provide for the differing distributions of the members of the same category without selection restrictions or subcategorization frames. Although they would rather handle these facts via semantics and pragmatics, it is possible to adapt the rule schema of G.P.S.G. to handle subcategorization.

Rules in G.P.S.G. take the form of an ordered triple  $(n,R,T)$ , where  $n$  is an arbitrary number associated with the rule - its rule number,  $R$  is the P.S. rule of the form  $[_A X]$ , and  $T$  is a schema used to indicate how the intensional logic representation of the expression created by the rule is to be developed from the intensional logic representations of its immediate constituents e.g. 1.60(i). Within such a schema 1.60(ii) might be the form of the S expression rule in English:

- 1.60. (i)  $\langle \_n [_A X] \quad T \quad \rangle$   
 (ii)  $\langle 1, [_{\bar{V}} \bar{N} \bar{V}] \bar{V}' (\wedge \bar{N}') \rangle$

In this rule  $\underline{n}$  is 1 and  $[_A X]$  is  $[_{\bar{V}} \bar{N} \bar{V}]$  which will be given the tree representation in 1.61.



$\bar{N}$ ' stands for the intensional logic representation of the expression dominated by  $\bar{N}$  e.g. dog' is the expression of intensional logic which translates the word dog in English.

Given this rule form, subcategorization might be handled in the following way: if there were a rule of the grammar  $\alpha$  which introduced a lexical category E but only a proper subset of E might appear in the environment created by the syntactic component of rule  $\alpha$ , then in such a case  $\alpha$  is allowed to be a feature on category E e.g.

1.62.  $\langle \alpha, [\dots E \dots], \dots \rangle$   
                    $[\alpha]$

and by an abbreviation convention this can be reduced to

1.63.  $\langle \alpha, [\dots E \dots], \dots \rangle$

Consider the following example:

1.64.  $\langle \beta[\bar{V} \bar{V} \bar{N} \bar{P}], \dots \rangle$

where  $V[\beta] \rightarrow \{\text{give, bake, } \dots\}$

This rule says that there is a proper subset or class of the category verb which require to be followed by a noun phrase, followed by a prepositional phrase. However, this is simply a notational variant of a subcategorization frame (or context-sensitive lexical insertion). Instead of stating the subcategorization frame in the lexicon along with the class of items which may appear in that frame, the frame is now part of the syntax. Items which may appear in such a frame are listed in the lexicon and linked to the rule via an arbitrary integer. If there were numerous proper subsets of E then each would require its own rule - just as each would require its own subcategorization

frame. At this stage a comparison of a T.G. model and G.P.S.G. with respect to type (iii) restrictions would reveal that there is no significant difference between them. Consequently, one would be incapable of making a principled choice between them. It would be a "matter of faith" whether one found subcategorization frames or syntactic rules easier to learn.

This is not the only problem for G.P.S.G., for within it little is said with respect to the lexicon and morphological operations. G. P. and S. claim that lexical categories may bear morphological features. As an example they state that a tree may contain a node labelled  $V[j+ing]$  immediately dominating, for example, eating. They assume that such forms are listed in the lexicon and not constructed by affixation of -ing to eat by a syntactic rule of affix hopping. This is similar to the lexical grammar approach of Lapointe (1980) (see 3.5.1). G. P. and S. offer no further discussion of verbal morphology, and say nothing about verb number agreement. This can only weaken their case when restriction (iii) is taken into account. How is a grammar in which verbal morphology remains vague and in some instances unspecified more learnable than one in which it is explicitly stated?

In order to generate the vast range of structures and relatedness between structures required by natural language, G.P.S.G. introduced a new piece of descriptive machinery - the meta-rule. The idea behind the meta-rule is one whereby the list of P.S. rules is extended by stating that if  $\alpha$  is a P.S. rule of the form  $\delta$  then  $F(\alpha)$  is also a P.S. rule where  $F(\alpha)$  is a function of  $\alpha$ . G. P. and S. note that this appears somewhat similar to a transformation but whereas a T-

rule maps trees onto trees, a meta-rule maps rules onto rules. If T-rules are added to the grammar then two distinct rule types are being utilized but the addition of meta-rules simply enlarges, in a rule governed way, the set of P.S. rules. The grammar remains essentially P.S.

Formally, this appears sound but the meta-rule simply reintroduces the power of T-rules. Consider the following example:

1.65. 'VP Deletion'

$$\begin{array}{l} \langle [\bar{V} \ V \ \bar{V}], F \rangle \Rightarrow \\ \left[ \begin{array}{l} +AUX \\ -PRP \\ -GER \end{array} \right] \\ \langle [\bar{V} \ V \ \bar{V} \ ] , F \rangle \\ \quad \quad \quad [+NUL] \end{array}$$

This rule looks very much like a T-rule with the top line being the S.D. and the bottom the S.C. These similarities can be extended if we treat T-rules as static as, for example, does Harris (1957). In the static treatment of T-rules the structures exist and the T-rule relates structures by stating that if a structure of the form X exists then it is possible for the structure F('X) to exist, where F is a function on X relating it to 'X. Meta-rules do relate rules to rules but these rule schemas induce trees and, therefore, indirectly they relate trees to trees, just as Harris' transformations do. Therefore, the meta-rule is a notational variant of a transformation which abstracts back one stage.

However, G.P.S.G. is more restrictive despite the use of the meta-rule. The range of structures which can be generated is limited to those which the P.S. rules will generate. By reducing the range of theoretical devices available the grammar is more restrictive and

consequently more learnable.

Finally, G.P.S.G. makes use of numerous features to handle co-occurrence and subcategorization.

In any transformational grammar these features would either be dispensed with or relegated to the lexicon but such a move would take us outwith the G.P.S.G. framework. In G.P.S.G. these features are part of the grammar and occur in the rules of the syntax, to construct a lexicon containing these features for G.P.S.G. would take us outwith the scope and aims of the model. Consider the following from G. P. and S. (1981:11) which is the auxiliary verb schema and associated features:

$$1.66. \langle n[\underset{\alpha}{\nabla} V \underset{[\beta]}{\nabla}], \lambda P[V'(\wedge \nabla'(P))] \rangle$$

+AUX

$n$ ,  $\alpha$  and  $\beta$  given by -

$n$	$\alpha$	$\beta$	$V[n]$ membership
2	+FIN	+BSE	can, may, must, will, could, etc.
3	+FIN	+BSE -AUX	do
4	+ASP	+PSP	have
5	+COP +ASP	+PRP	be
6	+COP	+PAS	be
7	+INF	+BSE	to
8	+FIN $\pm$ COP	+INF	is [+COP], ought [-COP]
9	+COP	+PRO	be

The schema is used, with the above rule, to solve the problem of

ordering of the auxiliaries. However, it is possible to simplify the schema.

The feature [+COP] appears in four different rules - 5, 6, 8 and 9. in 9  $\bar{V}$ [+PRO] is expanded by the following rule:

1.67.  $\langle 10, [\bar{V} \bar{X}], \dots \rangle$  where  $X \in \{A, N, P\}$   
 [+PRD]

In this example I do not believe that the copula ought to be treated as part of the auxiliary schema. If auxiliary verbs are defined as verbs with a right adjacent dependent V and main verbs as those with no right adjacent dependent V then the copula is not an auxiliary and consequently ought not to be in the auxiliary schema.<sup>26</sup>

Rule 8 might be subsumed under rule 2 if ought and be are entered into the lexicon as be to and ought to. This would reduce the number of rules but add complications in the lexicon. The issue cannot be resolved until the trade-off between lexical complexity and rule complexity is fully investigated.

Further reduction could be achieved by having the V[n] member carry the [ $\beta$ ] features. These would be marked on the second or dependent member of a two verb sequence via the head-feature convention or a modified version thereof.<sup>27</sup>

To return to the auxiliary rule schema, consider that if [ $\alpha$ ] is  $\begin{bmatrix} +ASP \\ +COP \end{bmatrix}$  i.e. BE then [ $\beta$ ] is [+PRP] which implies [+ing] via, although it is not entirely clear, [V+ing]. Such processes could be simplified by e.g. stating that if [ $\alpha$ ] is BE then [ $\beta$ ] must be [V+ing] thus bypassing some of the machinery or if  $V_1$  is BE  $V_2$  is Progressive - [V+ing].

---

<sup>26</sup>See Ch.3 for a further discussion of this.

<sup>27</sup>See Ch.3 below.

The main thrust of G.P.S.G. is that context-free P.S.G.'s are more efficient, explanatorily and descriptively adequate than T.G. and consequently more learnable. However, I feel that these claims are open to question. In dispensing with transformations various features for subcategorization and meta-rules are required to provide the full range of structures required. It is then a matter of debate whether transformations and subcategorization frames are more learnable than features and meta-rules. This debate falls outwith the scope of our discussion.

CHAPTER TWO

Change in the Lexicon : The Modals - A Paradigm Case

## 2.0. Introduction

In this chapter I intend to survey, critically, the data for one of the major changes in the history of English syntax. This survey will attempt to ascertain whether the data is compatible with either the autonomous view of syntax and syntactic change or the lexical view advanced in the previous chapter. This assumes an opposition between these two models. The change under consideration is the development of the N.E. modals, with the main question being one of whether they come to constitute a separate class of items, distinct from main verbs or whether they are idiosyncratic main verbs.

The former position, which I shall term the autonomous view, is that advanced by Lightfoot (1974, 1979). The opposing view, i.e. that modals are main verbs, has been put forward by various linguists e.g. Anderson (1976), Huddleston (1974), Ross (1969, 1970), and Pullum and Wilson (1977). Although one view is taken as representing the autonomous viewpoint, no particular theoretical framework will be taken as providing the definitive statement of the opposing view. This is important, given that the autonomous syntax position is compatible with a modals-as-main-verbs analysis, as argued by Pullum and Wilson (1977). The crucial point is that only the autonomous framework permits the reanalysis-of-category-view of the history of the modals.

I propose to first state the autonomous view and then to test the arguments for that position against the data to see to what extent they are compatible.

### 2.1. The Autonomous View

The strongest diachronic case for autonomous syntax is probably that put forward concerning the development of the N.E. modals - can, may, must, shall, and will - (Lightfoot (1974, 1979)). Lightfoot's argument is outlined below.

### 2.1.1.

During the O.E. and M.E. periods verbs such as cunnan, magan, motan, sculan, and willan, usually translated into N.E. by a modal, had all the defining properties of verbs i.e. they had full person number paradigms, conformed to the negative placement and inversion rules, occurred adjacent to each other, and in infinite and gerundial forms. They occur in sentence final position, in subordinate clauses - descriptions of O.E. posit SOV as a basic word order, at least in subordinate clauses. They could occur with normal complementation types, although, as Lightfoot points out, in these cases they cannot be translated into N.E. by a modal. It would appear from this that there would be little justification in positing a class of modal verbs for this period. Between O.E. and E.N.E. however, a series of unrelated changes occurred which, it has been argued, culminate in the reanalysis of the pre-modals as a distinct new class - modal. Such a reanalysis, as will be seen, requires a change in the rules of the base.

### 2.2. The Reanalysis

The reanalysis of the pre-modals as modals occurs in two apparently unrelated stages. The first involves a series of changes which occurred between O.E. and M.E. and during the M.E. period. These changes led to increasing isolation of the pre-modals. The second

series occurs during the sixteenth century and is interpreted as a surface manifestation of the reanalysis and the change in the rules of the base.

### 2.2.1. The Early Changes

The series of changes which led to the identification of the pre-modals as a distinct class are:

(i) The pre-modals lost the ability to take direct objects. According to Lightfoot (1979:101) this change is over by the M.E. period, with the exception of cunnan 'can' which was much more resistant. In his exposition Lightfoot states that mun+D.O. was lost early and the last attested cases for other object taking pre-modals are:

2.1. shall: the leeste ferthyng þat y men shall

c.1425 Hoccleve, Min Poems xxiii 695.

can: yet I can musick too; but such as is beyond  
all voice or touch.

1649 Lovelace, Poems (1659) 120.

may: for all the power thai mocht

1470 Henry, Wallace iii 369

(ii) Most of the pre-modals belonged to a class known traditionally as the preterite-presents. The members of this class were historically strong verbs the preterite forms of which, derived from perfects, had taken on present meaning in pre-Germanic and for which new weak preterite forms had been made. This could be thought of as being analogous to Greek οἶδαι 'I know' and Latin odi 'I hate'

which are perfective in form but present in meaning (see Curtius (1880) on Greek and Palmer (1954) on Latin). These forms are historically the perfect forms of inchoative verbs and are translated into English as present tense forms.

The most interesting thing, in the present context, about the present of this class is that because it is historically a strong preterite it did not have the regular -ep third person singular present ending. Consequently the modals are conservative in failing to undergo modern verb number agreement i.e. they do not have the usual -s ending.

Lightfoot also states that it is even more interesting that all the non-pre-modals of this class were lost. Benugan was lost quite early, unnan c.1320 and purfan c.1485. Witan survives into the nineteenth century but not as a member of the preterite-presents - at least it adopted a new fricative third person singular present tense form. The dating of such a change is unclear but Lightfoot claims it occurs during the M.E. period as he is unable to find any of the unlevelled forms after M.E. Dugan survives in northern dialects until the nineteenth century but as with witan it developed a new third person ending in the fourteenth century. According to Lightfoot dugan survives as both a modal and a main verb in Scottish English.

Lightfoot offers no explanation for the fact that of this inflectional class only the pre-modals survive as such but he merely suggests the development of an identifying criterion in the form of the absence of the third person singular present tense marker. This is a property shared by willan N.E. will which Lightfoot claims

descended, like the pre-modals from the preterite-presents.

(iii) The third change concerns the preterite forms of the pre-modals, which shared with the now productive weak verbs, the feature that there was no phonological distinction between the preterite indicative and the subjunctive except in the second person singular. In early O.E. the plural forms were distinct. The result of this is that, for example, N.E. should corresponds, phonologically to an old preterite subjunctive and indicative. This is true of all weak past tenses. It would appear that the preterites may have been unstable for quite some time, with the result that e.g. N.E. must and ought, historically past tense forms, never carry past tense sense except in indirect speech. The relation between shall/should, can/could, and will/would is rarely based on tense. Consider the following examples:

2.2. (i) When I was young I would play squash.

(ii) I might do this tomorrow.

In 2.2(i) the modal signals not the simple past tense but a habitual event while in 2.2(ii) it signals future possibility. In O.E. the pre-modals carried the time reference appropriate to their morphological form. The breakdown may have begun quite early and been aided by the demise of the subjunctive mood. The subjunctive/indicative was lost for most verbs. Main verbs retain the preterite subjunctive in unreal conditions where it resembles the preterite indicative except with be e.g.:

2.3. Jim acts as if he were a genius.

With the modals the preterite subjunctive is generalised to express unreality outside conditionals e.g.

2.4. As head of department, he would be a disaster.

It would appear that the distinction between e.g. will/would is not based simply on tense or mood.

If we classify words into morphological types - declensions for nouns and adjectives and conjugations for verbs - then for each type we can set up a table or paradigm. This paradigm gives all the forms for one member of the type. The forms of all other members of the type may be constructed by reference to the appropriate paradigm. Given this traditional definition it is possible to claim that would, could and should exist independently of will, can and shall rather than parts of the same paradigm. The relationship between will and would is clearly different from that between close and closed. Also an item like closed, the past tense form of close, is restricted to past time reference. These same restrictions do not apply to a form like could or should. This reinforces the claim that can, shall and will may not be parts of the same paradigm as could, should, and would.

(iv) Another innovation which helped to isolate the pre-modals was the introduction of the to-infinitive. This occurred during the late O.E. - early M.E. period. Lightfoot quotes Visser on the introduction of the to-infinitive, the origin of which is a matter of controversy. Originally the introduction of to was semantically based. It was a true preposition with a sense of direction towards and thus abstractly of purpose. Gradually the directional/purposive force was weakened so that to was used when no purposive or motion

sense was appropriate and the use of the to-infinitive spread quite rapidly during the M.E. period to non-directional/purposive contexts. However, even during the sixteenth century the plain infinitive could still be found. Lightfoot makes the point that the relevant feature here is the absence of the to form immediately after a modal. However, it could occur when separated by other material:

2.5. to do youre bidding, ay we wille

Townley Mystery 266

While to was interpreted as a directional preposition it could be predicted that it would never occur immediately after a pre-modal because they conveyed no sense of motion towards. It is unclear why they resisted to after it had lost its directional/purposive sense. It is possible that the pre-modals had already become identified as a unique class and this may have led to their resistance.

The development of to may be an example of Kronasser's Law, (Traugott: 1977). This law claims that lexical items move from the concrete to abstract and finally to logical connectives. To begins as a full concrete denoting lexical item and gradually moves towards a stage when it is a syntactic function word indicating non-finiteness.

#### 2.2.2. The Later Changes

These then are the early changes which led to the isolation of the pre-modals and their identification as a unique class. These changes were over by the fifteenth century. It is argued that in a grammar in which the pre-modals are main verbs the grammar would count these changes as introducing exception features. However,

once these exceptions reached a certain limit the pre-modals were reanalysed as the category modal - a change which required the introduction of a new base rule into the grammar.

The evidence, Lightfoot claims, for the re-analysis taking place lies in a series of changes which occur during the sixteenth century. These changes can be given a unitary account if we accept, at that time, the introduction of a new category -- modal.

The changes which provide evidence for the second stage of the development of the modals are:

(i) The old pre-modals could no longer appear in infinitival constructions. The last attested case recorded by Visser is:

2.6. "that appeared at the fyrste to mow stande the  
realm in great stede....."

1533 More, Works 885.

(ii) The pre-modals could no longer occur with ing suffixes. The last recorded cases being:

2.7. "....a thyng as mayyng or owing to be done"

c.1512 Linacre, Progymnasmata

"maeyinge suffer no more the loue and deathe of  
Aurelio"

1556 Aurelio and Isab. M ix.

(iii) From the sixteenth century there could only be one modal verb per simple sentence in the standard dialect, the last recorded case being:

- 2.8. "I fear the emperor will depart thence before my letters shall may come into your grace's hands."  
1532 Cranmer, Letters.

However, Lightfoot does note that some Scots dialects still permit double modal constructions e.g. will can, as do some southern U.S. dialects; but he does not know if these are innovating or conservative. These constructions are, in fact, conservative and are discussed in Brown and Millar (1980), Butters (1973) and Šćur (1968).

(iv) The pre-modals could no longer occur with have and an -en affix. The last attested case would appear to be:

- 2.9. "if we had mought conuenient come togyther, ye woulde rather haue chosin to haue harde my minde of mine owne mouthe"  
1528 More, Works 107 H6.

These changes follow automatically from introducing the pre-modals by the following rule

- 2.10. Aux  $\rightarrow$  T(M)

This will entail that there is only one modal per simple sentence, that it cannot occur in either adjectival or participial form, or in a for-to complement where Aux is realised simply as to. Given that (HAVE-EN) and (BE-ING) are in VP and we have a rule of HAVE-BE raising the modal will always be on the left of these constituents.

There are two other changes which occur at this time and although not directly related to this re-analysis they were allegedly a

consequence of it:

(i) It was necessary for the Negative Placement rules to be re-analysed. Previously the negative particle had been placed immediately after the first verb giving the following:

- 2.11. (i) John could not shoot the dog.  
 (ii) John shot not the dog.

When the new category was introduced, if the existing rule had been preserved then the following structures would have been derived:

- 2.12. (i) John could shoot not the dog.  
 (ii) John shot not the dog.

The possible alternatives to this were (a) attach the negative to the right of either the modal or the main verb, which ever came first; i.e. maintaining the surface structures but changing the rules or (b) place the negative after the first main verb i.e. maintain the rules but change the surface structures. The former alternative was adopted and during the sixteenth century two construction types developed for non-modal sentences, one of which involved an extension of the modal pattern using an empty do:

- 2.13. (i) He did not eat.  
 (ii) He ate not.

Eventually, during the seventeenth century, constructions of the latter type declined with neg being restricted to post auxiliary positions, a simplification given the recategorization.

(ii) A similar problem arose with the Subject-Verb inversion rules

which had previously given the following types:

- 2.14. (i) Could John shoot the dog?  
 (ii) Shot John the dog?

Once again the rule could be altered to operate on both modal and verb preserving the surface structures or it could remain unchanged and prepose only either main verbs or auxiliaries. The rules were changed to prepose only auxiliaries. Do-support structures became common in the sixteenth century while those without became rare by the latter half of the seventeenth century.

It would appear from this that the pre-modals were no longer main verbs and no longer to be derived from higher predicates. As the reanalysis took place it would appear that a new set of verbs, some of which are actually adjectives, came into the language. These verbs, the quasi-modals, e.g. be going to, have to, be able to, according to Lightfoot are, despite some subtle nuances (on which he does not expand), almost semantically identical to the modals shall/will, must, and can and differ only in that they have all the syntactic properties of real verbs. It would appear that the re-analysis of the pre-modals as modals created a vacuum which the grammar filled by creating a new set of semantically equivalent verbs. Once again Lightfoot quotes dates from the O.E.D.:

- 2.15. have to he told him he had not to believe, that the  
 couetousness of Virginio.....had moved  
 Ferdiand.  
 1579 Geoffrey Fenton, Guicciard (1618) 6

we have....to strive with a number of heavy prejudices.

1594 Richard Hooker, of the Laws of Ecclesiastical Polity III

be going to thys unhappy soule....was going to be broughte into helle for the synne and onleful lustys of the body.

1482 Monk of Evesham (Ans) 43.

be able to a felde u<sub>p</sub>er lye ....or is able to pasture.

1398 Trevisa Englishing of Bartholomaeus de Proprietatibus Rerum xiv p49.

Here Lightfoot says that in this example be able to means 'suitable' and the earliest convincing example of be able to meaning 'can' is "to be abill to wed hure", 1440 Gesta Romanorum ed Herrtage (1879) p269, but here it could still mean 'suitable'.

According to Lightfoot these data do not prove anything and a different interpretation may be placed on the data. However, a radical re-analysis of the base involving the introduction of the modals does, in his view, account for the simultaneity of the changes which occur in the sixteenth century. He also claims that the relevant dates are remarkably uniform and the data require no tidying up.

It would appear then that this is a genuine syntactic change which affects the base rules of the grammar. Over a period of time a number of exception features concerning a group of verbs build up, until at some point in the sixteenth century these exception features reach such a limit that the Transparency Principle (see chapter 4)

intervenes causing a re-analysis of the base to occur. Some diachronic syntacticians would of course disagree with this view and claim that no restructuring took place. Lakoff (1968) attempts to show that much syntactic change is simply a function of changes in the lexical redundancy rules and not even changes in transformations. The implication of this is that syntactic change takes place mostly through changes in the lexical restrictions associated with governed rules. Diachronic syntax becomes much more restrictive if it is assumed that P.S. rules do not change. We shall be looking at the compatibility of the data involved and a lexical account below.

### 2.3. The Data for the Early Changes

In this and the following sections I intend to look at both Lightfoot's data and the analysis based on it. Additional data will be taken into account, leading to an alternative analysis. I intend to divide my approach into two sections. The first will examine some of the more general points made by Lightfoot while the second looks at specific areas, under the same headings as Lightfoot.

#### 2.3.1. Preliminaries

There are three points made by Lightfoot (1979) which must be discussed before moving onto a detailed review of the data:

(i) In his introduction to the historical re-analysis Lightfoot (1979:98) outlines the properties in O.E. of the pre-modals. He claims that they are identical to those of main verbs. Whether or not he is correct is not at issue here but will be examined later. What is of importance is his claim that the verbs which are normally translated into N.E. by a N.E. modal have all the characteristic

properties of main verbs but in the given data they cannot always be translated into N.E. by a N.E. modal. For example:

- 2.16. (i) hwaet þaer forgange, oððe hwaet eftylage,  
 we ne cunnun. what came before, or what comes  
 after we {know not}  
 {\*cannot}  
 Bede.
- (ii) she koude much of wandrynge by the weye.  
 she {knew }  
 {\*could} of wandering by the way.  
 Chaucer.

Lightfoot's view of these examples seems inconsistent. He claims that the pre-modals (the antecedents of the N.E. modals) had all the characteristics of main verbs and could be translated into N.E. by a N.E. modal. Further he adopts the autonomous position where syntactic change proceeds independent of semantic and phonological factors. However, it would appear that, in these examples a semantic shift has taken place as we cannot translate the antecedents of the N.E. modals using a N.E. modal. However, it must be shown that this is semantic rather than, as Lightfoot would probably claim, syntactic.

(ii) Lightfoot includes willan as a member of the class of pre-modals but seems confused over its etymology. He states (1979:103) that willan is usually classed as an anomalous verb and not as a preterite-present but because, like the preterite-presents it had no third person singular fricative ending it descended from the preterite-presents. This is supported by e.g. neither Flasdieck (1937) nor Wright (1910). As we have stated above, the preterite-presents

were originally unreduplicated perfects which acquired a present meaning with respect to which a new weak preterite, an infinitive and a present participle were formed in the primitive Germanic period. They were inflected in the present like the preterite of strong verbs. Willan, however, descends from the class of athematic verbs. The present tense form was originally an optative (subjunctive) form of a verb in -mi, which in primitive Germanic came to be used indicatively. A new infinitive, present participle and weak preterite were added. The endings of the present were originally those of the preterite subjunctive. These differences can be clearly illustrated by the following examples from Gothic:

2.17.	Pres 'will'	Pret Subj 'take'	Pres 'shall'	Pret 'take'
	Singl wiljau	nemjau	skal	nam
	2 wileis	nemeis	skalt	namt
	3 wili	nemi	skal	nam
	Dual2 wileits	nemeits	----	nemuts
	Plurl wileima	nemeima	skulum	nemum
	2 wileiþ	nemeiþ	''	nemuþ
	3 wileina	nemeina	''	nemun

Clearly Lightfoot is mistaken in claiming willan to be descended from the class of preterite-presents. However, it is certainly true that it was drawn into that class.

(ii) Various members of the class of preterite-presents are lost through either falling into disuse or by being re-assigned to the category of main verbs. They then acquire the identifying properties of main verbs i.e. a to infinitive, third person singular present tense marker etc. In these cases there is no syntactic innovation,

with the re-assignment taking place within the lexicon. If this type of change may be given a lexical treatment then suggests that it might be possible to give the behaviour of the pre-modals and modals the same treatment.

### 2.3.2. The Morphological Form of the Third Singular Present Tense

As stated above, Lightfoot (1979:98) claims that during the M.E. period the antecedents of the N.E. modals behaved as other O.E. verbs and that they had full person-number paradigms. However, this is not quite correct. The pre-modals belonged to a class of verbs known as the preterite-presents. These were, as we have already discussed, verbs, the present tense forms of which, through a semantico-morphological shift, were morphologically identical to a strong preterite and for which a new weak preterite had been formed. As these verbs derive their present tense forms from a historically strong preterite it follows that they will not take the usual third person singular present tense ending -eþ. This is the morphological antecedent, although not phonologically, of N.E. -s. The result of this development is that the N.E. modals are conservative in failing to undergo modern verb number agreement.

However, it is not only the lack of eþ which is striking about these paradigms but also the difference between first person singular and the first-third person plural of the preterite-presents and the present indicative of a typical strong verb. These differences can best be illustrated by the relevant parts of the paradigms of: drifan 'drive' a class one vocalic type, cunnan 'know, be able' and magan 'be able' two members of the preterite-presents and pre-modals.

2.18.	drifan	cunnan	magan
	Singl drife	can(n)	maeg
	2 drifst	canst	meaht
	3 drifþ	can(n)	maeg
	Plur1-3 drifaþ	cunnon	magon

A study of the morphology of these three verbs reveals that not only do the preterite-presents differ from the strong verb by lacking the third singular inflection but they also differ in two other aspects. The preterite-presents lack the first person singular inflection and differ from the strong type in their plural ending. These observations are tabulated below.

2.19.	<u>Person/Number</u>	<u>Vocalic</u>	<u>Pret-Pres</u>
	1sg	-e	-∅
	2sg	-st	-st
	3sg	-(e)þ	-∅
	P1	-aþ	-on

Of the four pieces of verbal morphology under examination the preterite-presents share only one with their strong counterparts: the second person singular ending -st. However, not all members of the preterite-present class take this ending: e.g. þū scealt and þū meaht cf. þū ahst; and dugan has no recorded second singular form at all. The second person singular of the preterite-presents differs from the strong past second person singular, which has the same vowel as the plural and ends in -e e.g. draf, drife, drifon the first and third person singular, second person singular, and plural preterite forms of the strong type drifan. The etymology of the preterite-presents would lead us to expect cann, cunne, cunnon for cunnan,

rather than the forms given in 2.19 above. These changes are due not to modifications during the O.E. period but are a result of the historical development of these verbs in Germanic.

The preterite-presents were not the only verbs in O.E. to lack the regular third person singular marker. Consider, for example, the following third person singular forms; rīst 'rise', bītt 'bite', bītt 'wait', býtt 'offer' which are all strong types, and the following weak types; sett 'set', bētt 'make amends'. The partial loss of the third person endings with these verbs is the result of regular phonological processes. In West Saxon the -e- of the second and third person singular was, in most cases, lost. This resulted in the stem final consonant coming into contact with the initial consonant of the inflection. This produced a number of assimilations. These phonological processes can be identified by comparing the paradigm of one of the above verbs with the paradigm of a more regular verb. We could compare, for example, drifan with risan (rist) 'rise'. This anomalous third person singular inflection persists into M.E. Brunner (1963) states that the syncopated forms of the second and third persons singular are common in southern and Midlands M.E., while Kurath and Kuhn (1956) provide examples of syncopated forms alternating with non-syncopated forms into the fifteenth century. Stratmann (1891) records, for biten bite, syncopated forms in Ancren Riwe c.1225, Ayenbite of Inwyt 1310, and in the Old English Homilies and Homiletic Treatises of the 12th and 13th centuries, while non-syncopated forms occur in Ayenbite of Inwyt 1310, Lazamon's Brut c.1205-1275 and The Ormulum c.1200.

Although the partial loss of the third person singular inflection

with these syncopated verbs is the result of regular phonological processes the lack of this regular inflection is clearly a feature they shared with the preterite-presents. Consequently, the preterite-presents are not as isolated by this feature as they might at first appear. The isolation of the preterite-presents would not become fully apparent until the syncopated form adopted the regular person-number inflections. These changes in verbal morphology would, no doubt, be further complicated by the spread of the -es, -s second and third person present indicative endings from the north.

Finally consider the verbal morphology of willan N.E. will, which is traditionally classed along with the anomalous verbs:

2.20. Present Indicative	Singl wille	-e
	2 wilt	-t
	3 will	-e
	Plur 1-3 willað	-að

Willan shares the first singular and plural endings with the strong type drifan, has no third person singular ending and lacks the -s- of the -st second singular ending. It is probably fair to claim that willan has more in common with the strong type than the preterite-presents.

It would appear that the morphology of the preterite-presents differs quite radically from that of the strong type, with which it did share some features, for example the second person singular inflection -st. It, therefore, seems odd that only the third person singular is considered to be an isolating factor. This is even odder when one considers that several weak and strong types also lacked this inflect-

ion and that these morphological differences persist into M.E. One might just as well claim that the lack of an imperative (Quirk and Wrenn (1955) list only witan as having one) to be the distinguishing feature of the class.

### 2.3.3. Loss of the Direct Object I

Another 'isolating factor' in the history of the preterite-presents is their loss of the ability to take direct objects. While Lightfoot is correct to point this out, his account is a gross oversimplification of the facts.

Lightfoot (1979:101) gives the following examples as the last attested cases given by the O.E.D. for object taking pre-modals:

2.21. shall: the leeste ferthyng þat y men shal  
c1425 Hoccleve, Min Poems xxiii 695

can: yet can I musick too; but such as is beyond  
all Voice or Touch.  
1649 Lovelace, Poems (1659) 120

may: for all the power thai mocht  
1470 Henry, Wallace iii 369

He does include further cases of object taking pre-modals:

can: No skill of music can I simple swain  
1710 Philips, Pastorals iv 23

shall: By the feith I shall to God.  
1530 Crt of Love 133

He treats both of these examples as archaisms of doubtful standing. I can see little reason for doing so, particularly as the O.E.D., in its introduction, states that quotations are chosen chronologically so as to give one from each century. Due to lack of space as few quotations as are necessary to illustrate both the signification and development of a word are chosen. This naturally results in gaps in the use of a word of about a century. It would require much painstaking research to discover the uses of an item in the intervening century. Therefore I would prefer to treat these quotations as genuine examples and not as archaisms.

If we include the nonmodal preterite-presents, then the earliest loss of the ability to take a direct object occurs c.1320 with unnan:

2.22. Meriadok was a man þat tristrem trowed ay:

Miche gode he him an.

1320 Sir Tristrem 1928.

The last recorded case is 1710 (see above). This gives a time span of 390 years. If will is included as a pre-modal preterite-present, which Lightfoot claims it is, then the last recorded case is:

2.23. I would that wars should cease, I would the globe

from end to end might sow and reap in peace.

1882 Tennyson, Charge of the Heavy Brigade.

This gives a time span of 562 years. One might argue against the will example by claiming that it is in a poetic register and therefore likely to be archaic. However, this argument may be raised against any of the previous examples, all of which, by necessity, are taken from the literary language, which may well be archaic.

Neither of the time scales, 390 or 562 years, appears to be compatible with an analysis which introduces a new category, modal, in the sixteenth century, by creating a new rule in the base of the grammar. If this analysis were to be the correct one then it would fail to capture the behaviour of both can and will, which would continue to be exceptions for approximately 210 and 382 years respectively. It appears that neither the category modal nor the restrictions are homogeneous. The creation of the category and the introduction of the restrictions cannot be regarded as a unitary event. Consequently, the re-analysis would fail to eliminate the offending opacity (see Ch. 4 below). It may, in fact, be the case that the changes are better suited to a lexical diffusion model of change (see Wang (1977) and, later in this chapter).

#### 2.3.4. Loss of the Direct Object II - Semantics

A study of the semantics of the pre-modals poses further problems for Lightfoot's theory.

Azan N.E. owe, ought to. This verb has a particularly complex history. Originally ought was the past tense form of owe. By c.1160 a new regular third person singular had been formed and a new past tense form had been added by c.1425:

2.24. þat þing þe he ahð

1160 Hatton Gosp, Luke xi 21.

Wheþer owed to loue him bettur þo.

c.1425 Cursor Mundi 14025(Trin).

Ought, as the past tense form of N.E. owe fell into disuse c.1685,

although it occurred with a direct object until c.1694 with the sense of 'owing or having to repay an ill turn or shame' and as the past tense form of N.E. owe in dialects e.g. Scottish and East Anglian English, until c.1896:

2.25. Burton....said he ought him nothing

c.1685 Life. A. Martindale 231 (E.D.D.)

The Devil ought him a Shame and paid him both  
Interest and Principle.

1694 R. L'Estrange Fables cclxviii (1714) 294.

He aught me ten pounds.

c.1825 Forby, Voc E. Anglia

The man as ocht Jerusalem greets becaused the fair  
Circassian winna take him.

1896 Barrie, Tommy 202.

There are several points which can be made concerning the development of azan and the relationship of that development to Lightfoot's claim of re-analysis. The split between owe and ought requires that the semantics of the verbs be taken into account. Both items retain the sense of obligation. In the case of owe a 'pay' component is present. One is obliged to pay or repay money (or the like). Ought, however, does not carry this 'pay' component: one is simply obliged to do something.

Lightfoot claims that the re-analysis of the pre-modals as modals occurs during the mid sixteenth century and that this recategorization is due, in part, to the increasing isolation of the pre-modals as a

distinct class during the late O.E. - early M.E. periods. The split between owe and ought straddles this period, and as we saw in the quotes above, ought continues to be used in its original sense (as the past tense of owe) until c.1896. If, as Lightfoot claims, the pre-modals were isolated from other verbs, then we would not expect the modal form ought, which had been in use since c.1200, to be used as the past tense form of owe. This use of ought also poses problems for the recategorization theory. After recategorisation ought, as a modal, would not be able to take a direct object. It would only be able to take an infinitive to its immediate right. Ought, like can, would require an exception feature to permit its occurrence with a direct object. This suggests, as with can, that the changes might be more adequately captured by lexical diffusion rather than by a change in the rules of the base. It is interesting to note that Lightfoot has nothing to say concerning the development of azan.

Let us now consider the development of cunnan N.E. can. It appears that in its transitive form it conveyed the sense of knowing someone or something. This meaning fell into disuse c.1649-1710, with the modal sense continuing until the present.

The other pre-modals follow a similar pattern. O.E. sculan, N.E. shall, in its transitive form covered the same semantic area as N.E. owe, but this was lost in c.1425. Its use, in the intransitive, as a marker of prediction is first recorded in the O.E.D. as c.900. This change could be interpreted as an occurrence of Kronasser's Law, where the meaning of a lexeme moves from concrete to abstract and space to time to logical connective. Here sculan moves from concrete to time.

It appears that the syntax and semantics of the modals, and, if tenable, of all verbs are irrevocably intertwined. Either the loss of the ability to take a direct object results in the loss of meaning or vice-versa. In each example the modal form develops eventually quite separately, and has, in some cases, quite a different meaning and could be considered to be a completely different form. The semantics of the preterite-presents may be used to distinguish the modal from non-modal significations. In each of the above cases the non-modal signification has a fuller or more concrete meaning, e.g. consider the discussion of owe and ought above. The signification of owe is fuller than that of ought in that it involves not just an obligation but an obligation to pay money. The modal signification is less concrete and involves a move from concrete to abstract or time. This, again, may be interpreted as an instance of Kronasser's Law.

For Lightfoot this poses a dual problem. If he accepts that the transitive form is separate from the pre-modal then he must concede that there has been no change. Consequently, he cannot claim this as an isolating factor. We pursue this question in the following section.

#### 2.3.5. A Semantic Argument for Re-Analysis

It would appear from the above discussion that the semantics of the pre-modals and the loss of the ability to take a direct object are interlinked. This raises the possibility that the re-analysis of the pre-modals might be semantically based. Lightfoot (1979:100) explicitly rejects this possibility. He disagrees with Traugott (1972b) and Stockwell (1976), both of whom adopt the view that there is no such thing as pure syntactic change, where reference is made only to syntactic rules and not at all to semantic or phonological

factors. For Lightfoot the re-analysis of the pre-modals, as modals, in the sixteenth century occurs without reference to the semantics of the verbs. He further claims that changes in the semantics of the (pre-)modals occurs without affecting their syntactic properties.

Our discussion above of the relationship between the semantics of the pre-modals and their loss of the ability to take a direct object throws this latter claim into doubt. We saw that the transitive forms of the pre-modals never carried a modal signification. The loss of the transitive form coincided with the loss of the non-modal signification. Whether the loss of the transitive form induced the loss of the non-modal signification or vice versa is open to debate but what is certain is that the syntax and semantics of these items are closely inter-linked. This interrelation is something to which we shall return; but let us investigate now the possibility of a semantically motivated change.

The semantics of the N.E. modals are quite complex and have provoked much discussion e.g. see Anderson (1971a), Boyd and Thorne (1969), Huddleston (1974, 1979), Palmer (1974, 1977, 1978, 1979) and Perkins (1982). Further evidence for the semantic complexity of these items can be adduced from the difficulties foreign learners have with them. These difficulties existed in O.E. where the range of possible meanings of these items was if anything wider than it is now (see Mitchell, 1975:112ff).

O.E. cunnan, N.E. 'can', signified 'to know or be acquainted with someone', 'to know or have learned a thing', or 'to have practical knowledge of something', in its transitive form, while in its intransitive form it signified 'to know how to do something', 'to have learned how

to do something', and 'to be intellectually able'. This last sense passes into the present usage with no obvious dividing line and gives 'to be able', 'to have the power, ability, or capacity both physically and intellectually to do something'.

Clearly in O.E. cunnan covered a wider range of meanings than N.E. can does. However, there would appear to be a split, and the subsequent development of cunnan supports this, between the transitive and the intransitive forms. The intransitive form moves towards 'ability', while the transitive moves towards, and is eventually subsumed under the area covered by the verb 'to know'. This is illustrated schematically below:

2.26. Early O.E.		Late M.E./E.N.E.
cunnan	}	
to know		<u>NP</u> to know <u>NP</u> → to know
to know how		
to be able		<u>VP</u> to know how <u>VP</u> → N.E. can

The already complex semantics of cunnan are further complicated by the split into Root including deontic and epistemic meanings for the modal signification. Bosworth-Toller (1882-98) records no epistemic sense for cunnan while the O.E.D. records the epistemic meaning from 1250:

2.27. Ic am sonder man, Egipte folc me knowen can  
Gen & Ex 2872.

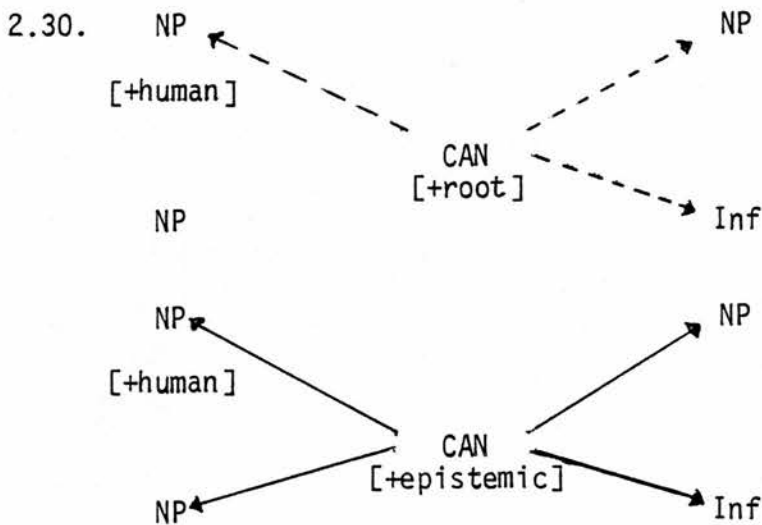
This split has consequences for the subject restrictions that cunnan and all (pre-)modals may impose. Generally the root modals require an animate subject, while the epistemic interpretation does not impose such a restriction on its subject e.g.:

- 2.28. (i) John can swim ten lengths.  
 (ii) The table can go in the corner.

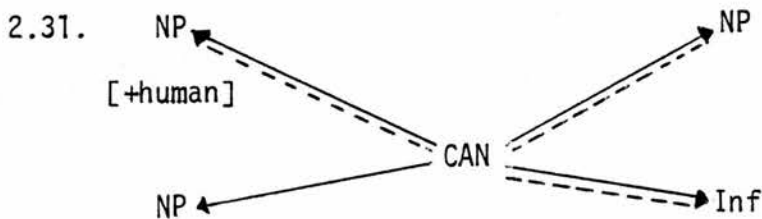
The sentence in 2.28(i) is ambiguous between the following two readings:

- 2.29. (i) John has the ability to swim the lengths.  
 (ii) It is possible for John to swim ten lengths.

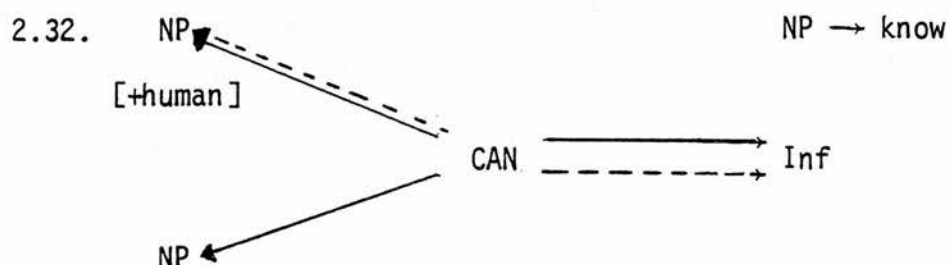
However, the sentence 2.28(ii) has only the interpretation of possibility. It also appears that the epistemic reading of can may only take a VP complement. If the root sense takes both VP and NP complements but is undergoing a split in meaning, then the rather complex semantics of can in the late O.E. early M.E. period could be given the following representation:



Combined this gives:



The introduction of the root and epistemic senses may have placed an intolerable burden on the semantics of the verb and this burden required that a change occur. With can +NP signifying 'to know' rather than 'to be able', one way of resolving the complex semantics would be for can to drop the NP complements, which could easily be handled by know, O.E. cnawan. This would give:



Such a scenario is possible and plausible, although conclusive proof is scarce and the definition of semantic intolerance falls foul of the same problems as the Transparency Principle (see chapter 4 below). However, this shows Lightfoot's analysis to be oversimplistic and inconclusive, as the facts are open to an alternative explanation. The burden of disproving these claims falls on Lightfoot.

### 2.3.6. Past-Tense Interpretation

The opaque relationship between the past tense forms of the pre-modals and their time reference was, according to Lightfoot, another factor which led to their eventual re-analysis. However, this change is also open to an alternative interpretation. All verbs lost the distinction between the indicative and the preterite subjunctive but retained the preterite subjunctive, which looks like the preterite indicative, with non-past reference in unreal conditions. The modals maintained the preterite subjunctive but generalized it to express unreal conditions outside conditionals i.e. in main clauses. Con-

sider the following:

- 2.33. (i) I could play the guitar {five years ago.  
{if we have a concert.}
- (ii) I could go a drink.

In the first example could is ambiguous between past time reference and conditional usage. However, in 2.33(ii) it can only be conditional. When the subjunctive use of the modals is viewed from this angle they have more in common with main verbs than would first appear.

### 2.3.7. Rejection of the to-Infinitive

Throughout their history the modals consistently resisted the introduction of the to-infinitive. It is not clear why this should have been the case but of the reasons proposed the most common is that as the modals conveyed no sense of purpose or motion towards there was no reason for them to take a clearly marked 'dative' form (see Jespersen (1909-49): III 1.3<sub>2</sub> and V 10.2<sub>1</sub> and Visser(1963-73): 947-948). However, this is true of many other verbs, and despite this there are examples of preterite-presents taking to-infinitives:

2.34. a3an: Swilcne laured we a3en to dreden.

c.1175 Lamb, Hom 21

Her me ah to understonden for whi hit sei

[h]alf quic.

ibid 82.

Ought, in N.E. always takes to although Visser (1963-73:1815) claims that the to-less infinitive was originally the more popular. Also to-less infinitive forms are now spreading (see Quirk and Rusiecki

1982). Of the other modals will/would are most commonly followed by the to-infinitive, with shall/should, may/might, must and can in descending order of frequency e.g. Visser (1969-73:183):

2.35. willan: For who so willeth to be contynent, Many a  
luste superflu mote he lete (willeth=wishes)  
c.1420 Hoccleve, Reg Princ 137.

And they wyllen to do that ye welle to do  
(wyllen/welle=wish)

1450-60 Bp Grasseteste's Househ, Stat, in  
Babees BK331.

Thei haue assembled a grete power, and wele  
to conquare this londe be force.

(wele=wish)

c.1450 Merlin, iii, 54.

sculan: How shold he your ful pleasure to accomplice?  
1463 George Ashby, Poems (EETS) 56, 315.

neiper he schal mowe to studie, to preche,  
to speke myche, neiper to singe.

1443 Peacock, Reule Crysten Relig (EETS) 270.

...he shall not mowe longer to occupie the  
same Office.

1455 Rule Prlt 5, 339.

Further to this, a number of verbs in N.E. share the same feature.

Bid does not normally take the to-infinitive and the same may be said for do, hear, let, leave (restricted to certain idioms, e.g. leave it

be and leave go), make and see. It is optional after help.

It would appear that this is another feature which is not peculiar to the modals. However, is there a possible explanation which would account for this?

In O.E. the infinitive occurred in two distinct forms; (i) the substantive form, which appears "when it is the subject or object of the verb". This form occurred with to + anne/enne and is known as the dative infinitive. This form did not normally appear after the preterite-presents. Secondly the verbal or predicative form, which appears "when the verbal or assertive idea is predominant e.g. when it completes the sense of an auxiliary verb" Callaway (1913:3).

These two forms could be given a more formal characterization. The description of the substantive form as the subject or object of the verb permits the characterization of this form as an NP. Whereas the description of the predicative form suggests that it should be characterized as a V or VP. These characterizations suggest a possible explanation for the behaviour of the modals. As was argued above (2.3.4.-2.3.5.) the transitive forms of the preterite-presents did not carry modal signification. This means that the pre-modals never occurred with a direct object. Consequently, as the substantive form is immediately dominated by NP (see Lightfoot 1976 and 1979:4.2.1. for a discussion of the characterization of the substantive form as a NP) it would never appear after a pre-modal. The predicative form, being a VP would always appear after a pre-modal.

This explanation does not extend easily to the non modal verbs listed above. These verbs include various sub-groups; (i) 'permission' -

(ii) 'causatives' - make, (iii) 'perception' - hear, see. It is difficult to tie these groups semantically to the modals, although 'permission' may tie in with can and may, and the 'perception' verbs are subject recipient i.e. the subject is affected by the action of the verb, as are the root modals.

It appears from the preceding discussion that the early isolating factors are more complex and far less uniform than Lightfoot's analysis suggests.

#### 2.4. The data for the Later Changes

I now intend to discuss the changes which occurred during the sixteenth century and constitute the evidence for the reanalysis.

##### 2.4.1. Loss of Modals from Infinitival Constructions

The pre-modals lost the ability to occur in infinitival constructions. Lightfoot (1979:110) gives 1565 as the last occurrence of a pre-modal in an infinitival construction. However, the following are recorded in the O.E.D.:

2.36. The wyse can rule: to can is full as much as though  
he did

1566 Drant Horace's Sat I iii B vij.

In evill, the best condition is not to will the second  
not to can.

1607-12 Bacon, Gt Place, Ess Arb 282.

If from this love thou will thou can'st unbind,  
To will is here to can.

1633 Fletcher Pisc Ed vi xxvi

Man at creation had to wyl or nyl naturally.

1654 Vilvain, Theorm Thoel ii 50.

Two of these examples extend into the seventeenth century and pose problems for the re-analysis by change in the rules of the base theory. If re-analysis proceeds through change in the rules of the base then these counter examples must be accounted for. Once again this suggests that a lexical diffusion model would be better suited to capturing syntactic change. We shall return to this below (2.5.).

It is interesting to note that not all pre-modals appeared in infinitival constructions. The evidence for an infinitival form of shall, at any period, is doubtful, as it is for mun and mot. Further, in the above examples, it is possible that will signifies 'to wish', and can, at least in the second example 'to know'. In both cases neither verb appears in its modal form. This correlates with the fact that the constructions are absolutes i.e. there is an understood direct object which is not present. This may also be the case with other modals appearing in infinitival constructions. It is also possible that these examples are cases of exceptional complementiser insertion but this would still have to be provided for. Again the lexicon would be the ideal location for the treatment of such exceptions.

Instances of pre-modals appearing in to-infinitive constructions are quite rare. It must be borne in mind that in verb phrases the (pre-)modal was usually followed by another verb which was rarely another (pre-)modal. If it were another (pre-)modal then it was quite unlikely that it would be preceded by to - remember the pre-modals never took the to-infinitive after them. This non occurrence of

modals with the to marker may have contributed to the relative infrequent appearance of the modals in infinitive constructions.

#### 2.4.2. Loss of Modals from Gerunds and Progressives

The pre-modals could no longer appear with an -ing suffix. Lightfoot records 1556 as the last case but the O.E.D. also lists:

2.37. In canning the text of the whole New Testament without book.

1563-87 Foxe A&M (1684) II 419

Willing, which still survives in phrases e.g. God willing, is recorded by Visser (1963-73:1964):

2.38. He turned and stared at all of us in that dark prison his big hands held out as if he were willing us to understand.

1965 MacMahon, Post Bellum Blues (Four Sq Books) 143.

Again these examples pose problems for the theory of re-analysis in the base. Also the M-ing forms in these examples are presumably 'main verbs' (for one thing they are all transitive). This is something to which we shall return below.

Visser (1963-73:1963) states that of the modals only four are recorded as having a present participle -- (i) cunning, (ii) mowing/mowende, (iii) owing/awende and (iv) willing. Of cunning, he says that the forms cunnende and the infinitive cunnan(=to know) are not found in O.E. but from the fourteenth century the northern form cunnand and the midlands and southern forms cunning/conning are in regular use. However, as cunning and its related forms signify 'to know', we face

the same question here as we did with the transitive forms of the preterite-presents. Are these forms in any way related to the modal form? If we decide that they are not (see 2.3.4.-2.3.5. for discussion) then we can claim that the pre-modal signification never occurred in an -ing form. Further, Visser (1963-73:1974) states that to be canning ought to be thought of as a Vcop+adjective construction. This view is supported by the O.E.D. which lists canning/cunning as an adjective signifying 'clever, knowing, learned' or 'skilful'.

Mowing/mowende. These forms do not occur, on record, with to be in to be +V-ing constructions.

Owing/awende. Visser (1963-73:1964) states that when the subject of the unit denotes something that has to be paid or rendered then is owing signifies is owed, is due. However, when the subject denotes a human the be V-ing form was formally used but now this has become rare. Ought, being a past tense form, would never appear with an -ing suffix, and therefore we ought to consider owing as a separate form from ought.

Willing/willende. The construction to be willing has been recorded in the language since earliest times. However, a gradual semantic change has taken place whereby willing = 'desirous, wishing', 'disposed to consent or comply, ready to do or to be of use or service'. Although future action is implicit in these cases, will cannot be said to be purely a marker of future tense.

The loss from modal paradigms of this affix results in an inability to appear in the progressive aspect, which, according to the data, was quite rare anyway, as gerundives and as adjectives. Apart from

owing, which is not related to ought synchronically (ought never appeared with -ing anyway), and willing, which has three possible sources none of which was a member of the preterite-presents, the occurrence of be M-ing progressive forms was practically non-existent as to be canning should be considered adjectival (cf knowing in N.E. and cunning which derives historically from can, although it is not thought to be synchronically related to can) and may never appeared with the verb to be. Of the other pre-modals no -ing forms are recorded.

Given that no progressive aspectual forms of the N.E. modals are to be found in their history, we must conclude that the rule  $Aux \rightarrow T(M)$  was already in operation prior to the sixteenth century for some members of the class and had been for quite a time. However, this leads to problems when it comes to explaining the appearance of the same items, excluded from (be-ing) constructions, with (have-en) and with direct objects. We could, of course, claim, as has been argued above (2.3.4.-2.3.5.), that there were two distinct forms -- one signifying the modal and the transitive signifying the other. It might also be possible to make this one of the early isolating changes or to implement the rule and make use of exception features or simply to use exception features without the rule. At any rate the loss of the -ing form does not seem to be a simple manifestation of a sixteenth century restructuring of the type envisaged by Lightfoot.

The inability to appear in progressive, gerundive, or adjectival constructions does not isolate the modals from other verbs. Consider the statives e.g. abhor, doubt, imagine, know, mind, remember, resemble, seem and tend. Statives do not usually occur in the

progressive aspect, do not appear as imperative (neither do modals) and some do not appear in gerundives:

2.39. Fred's sufficing.

However, no-one would invoke a new class distinct from verbs called stative. Perhaps, given this overlap in syntactic distribution, both statives and modals might, in this case, be treated as members of the same class.

2.4.3. Loss of Double Modal Constructions

From the sixteenth century there could only be one modal in any simple sentence, in the standard dialect. Although Lightfoot notes that double modal constructions occur in Scots and southern U.S. dialects, he quotes 1532 as the last recorded occurrence of such a construction in the standard dialect.

However, double modal constructions can be found into the nineteenth century:

2.40. Would I might But ever see that man

1610 Shakespeare, Temp I ii 167.

He'll no can hand doon his head to sneeze.

1816 Scott, Antiq xxiv (Scots)

Lady Macbeth, which I never could and cannot and never shall can act.

1847 Frances Kemble (Mrs Butler) Rec later life (1882)  
III 165.

These examples, once again, present problems for any change by rule. To capture these exceptions we would require some form of marking in the lexicon (see 2.5 below) or perhaps alternative grammars which would co-exist. Both of these suggestions lead to complexity in the grammar, making it difficult to learn -- a crucial point for Lightfoot. These complications would be compounded if these verbs were to behave like modals in certain cases but not in others. The result would be that in certain constructions they would be  $\begin{bmatrix} +V \\ -M \end{bmatrix}$  but in others they would be  $\begin{bmatrix} +V \\ +M \end{bmatrix}$  or  $\begin{bmatrix} -V \\ +M \end{bmatrix}$ . However, these problems arise for any synchronic grammar of English with need and dare.

Double modal constructions are, according to the data, quite rare at any period in the history of the language. This may be due, in part, to the lack of infinitive forms of the (pre-)modals, their rapid decline and eventual loss from the language. Also the dialects with double modal constructions pose problems for the re-analysis by rule theory (remember that these dialects are conservative and not innovating, see 2.2.2(iii) above). Further in these double modal constructions the order would appear to be epistemic-root and not root-epistemic. This can quite easily be stated as a selectional rule so as not to complicate the syntax.

#### 2.4.4. Loss of Modals from Perfective Constructions

The pre-modals lost the ability to appear in perfective aspectual constructions. Lightfoot claims the last recorded case to have appeared in c.1528 but the O.E.D. records the following:

2.41. They had cand their lesson.

1587 Fleming, Contn Holmstead III 1928/1.

If he had would he might easily....occupied the  
monarchy.

1633 Done, Hist Sept 216.

You had'nt ought to do that.

Modern dialect.

He has not could come.

Modern Scots

If I had could find it.

Modern Scots.

In the first of these examples can takes a direct object and carries the sense 'to have learned'. The second example is an absolutive construction with would probably meaning 'wished/desired'. This second example is of further interest as it shows a past participle form occurring over a century after Lightfoot claims it fell into disuse. The final three examples are all dialectal.

Visser (1963-73:2221) makes some interesting observations concerning the occurrence of the modal forms in these constructions. He states that only azan, cunnan, magan and willan are recorded in past participle forms, which he terms the modal pluperfect. Of the data recorded Visser claims that the past participle forms of can and will are not related to their modal signification but to 'to know' and 'to wish/want or desire', respectively (see the examples above). This leaves had might:

2.42. He wild haf venged his fadere if he had myght.

c.1338 Robert of Brunne, Chron 2p15.

Whan the deuyll hath not mowe he can not induce  
the man to goo oute of the fayth.

1490 Caxton, How to Die 7.

Ye haue mought oftentimes, & yet maie desceyue me.

c.1510 More, Picus Wks 7/2.

Visser claims that in the first example had myght is best seen as a noun meaning power or ability. Certainly such an interpretation is possible; however, had myght could also be translated as 'had been permitted', which is also the translation of haue mought in the third example. Visser's claim is perhaps best borne out by the second example where hath not mowe is best interpreted as 'has not the power'. As with both the infinitive and the -ing forms, this leads us to question whether the modals ever appeared except marginally in these constructions (see also 2.3.4.-2.3.5. above) and throws into doubt the claim that this is a major syntactic change.

## 2.5. Lexical Account of the Data

Having surveyed the relevant data it is apparent that Lightfoot's account is historically oversimplistic and inaccurate. There are two angles from which Lightfoot's account may be undermined. Both lead to the same conclusion: the claim that there was a change in the rules of the base of English must be in doubt, as little evidence can be found in its favour.

If the semantics of the preterite-presents are taken into account, then it is possible to claim that the loss of the ability to take direct objects (2.3.4.-2.3.5.) and the late changes (2.4.-2.4.4.) are irrelevant to the development of the modals. As we saw above, the

transitive, infinitival, gerundive, adjectival, progressive and pre-fective forms of the preterite-presents, in some cases never and in others only occasionally, carried modal signification. Consequently the evidence for the appearance of the pre-modals in these constructions and morphological forms is highly marginal. From this we may conclude that the modals have always been present in the language and always lacked those forms Lightfoot claims they lost in the sixteenth century. English has either always had an Aux component (pleasing to Steel et al, 1979) or no changes have occurred and, English lacks both an Aux component and expansion rule. What has happened, is that an orthographically identical and semantically related form of the preterite-presents was lost. This form appeared in a wider range of syntactic structures and morphological forms than its pre-modal counterpart but never carried modal signification.

If we wish to maintain the claim that a syntactic change did occur i.e. Lightfoot's account of the distribution of the pre-modals is correct, then the time scale poses problems for the claim that the change occurred in the mid sixteenth century and was implemented via a sudden wholesale cataclysmic restructuring of the rules of the base. The time scale is less concentrated than it at first appears. The relevant dates are neither as uniform nor as 'clean' as Lightfoot (1979: 113) claims them to be. If a change did occur then the time scale suggests that a lexical diffusion treatment would more adequately reflect the nature of that time scale and offer certain advantages over Lightfoot's account. Let us now investigate the lexical diffusion account in more detail.

### 2.5.1. Lexical Diffusion

Classically much sound change, brought about by articulatory fine tuning, is lexically abrupt but phonetically gradual. All segment types with a particular set of feature specifications in a given set of environments are affected, regardless of lexical or grammatical factors. Such change proceeds by insensible stages until phonologized. However, it might also be possible that change is lexically gradual but phonetically abrupt. Feature values would change directly from + to - or vice-versa without intermediate stages but this change might percolate gradually through the lexicon, infecting one lexical item then another taking, perhaps centuries to become complete. In some instances it might never become complete leaving a 'residue' of unaltered items. Such a theory has, most recently, been proposed by Wang and Chen (see Chen, 1972; Chen and Wang, 1975; and Wang, 1969). However, this concept is not new and had previously been discussed, if not explicitly, by Fries and Pike (1949) and Malkiel (1967).

This theory of change may also be applied to syntactic change. The second set of changes in the development of the pre-modals results in their elimination from all non finite positions. Lightfoot's analysis achieves this by introducing the rule  $Aux \rightarrow T(M)$ . However, the same result may also be affected by lexical diffusion. Elimination of the pre-modals from non finite positions would still be the ultimate result but the locus for change would be the lexicon with the distribution of the pre-modals becoming gradually, but increasingly, more restricted until they may only appear in non finite positions (see 3.8. below on how this might be formalised).

2.5.2. Syntactic Lexical Diffusion

The disparity of the time scale provides the best evidence for the lexical diffusion approach and can best be illustrated by the following tables.

## 2.43.

## Change - Loss of

Verb	Direct Object	Double Modal	Infin	M-ing	Have-en
cunnan	1659-1710	1535-1758	1607-1612	1563-1587	1587
magan	1470	1481	1533-1565	1556	1510
mot	----	----	----	----	----
sculan	1530	c.1600	----	----	----
willan	1882	----	----	see 2.4.2.	1633

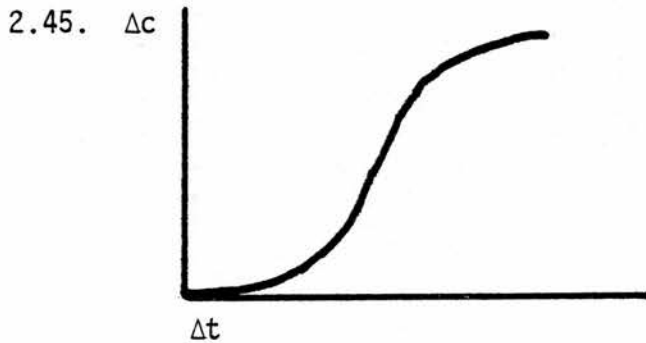
## 2.44.

Date	Verb	Change
1470	magan	Direct object
1481	magan	Double modal
1510	magan	Have-en
c.1530	sculan	Direct object
1533-65	magan	Infinitive
1535-1758	cunnan	Double modal
1556	magan	M-ing
1563-87	cunnan	M-ing
1587	cunnan	Have-en
c.1600	sculan	Double modal
1607-12	cunnan	Infinitive
1633	willan	Have-en
1654	willan	Infinitive
1659-1710	cunnan	Direct object

From these tables we can see that no clear pattern emerges. Fifteen changes occur over a period of 412 years with magan (may) always being first to change and willan (will) last. Mot (must) fails to participate in any change while sculan (shall) is lacking from three.

The spread of dates over which the changes occur makes the dating of the introduction of a rule difficult and the idea of lexical diffusion more appealing (given that lexical diffusion changes often take centuries to complete).

Lexical diffusion changes may be divided into three stages: an initial stage where few items change, a middle stage where the rate of change rapidly accelerates (this is the most productive period), and a late stage where the rate of change falls back and eventually ceases often with a number of items left unaffected by the change. The result of this is that when the number of items changed is plotted against time an S shaped graph is produced.



The shallow incline at the beginning of the graph is produced by the relatively few items which have undergone the change. As the pace of change increases more items change in a shorter period of time and the gradient of the incline increases until finally, when almost all

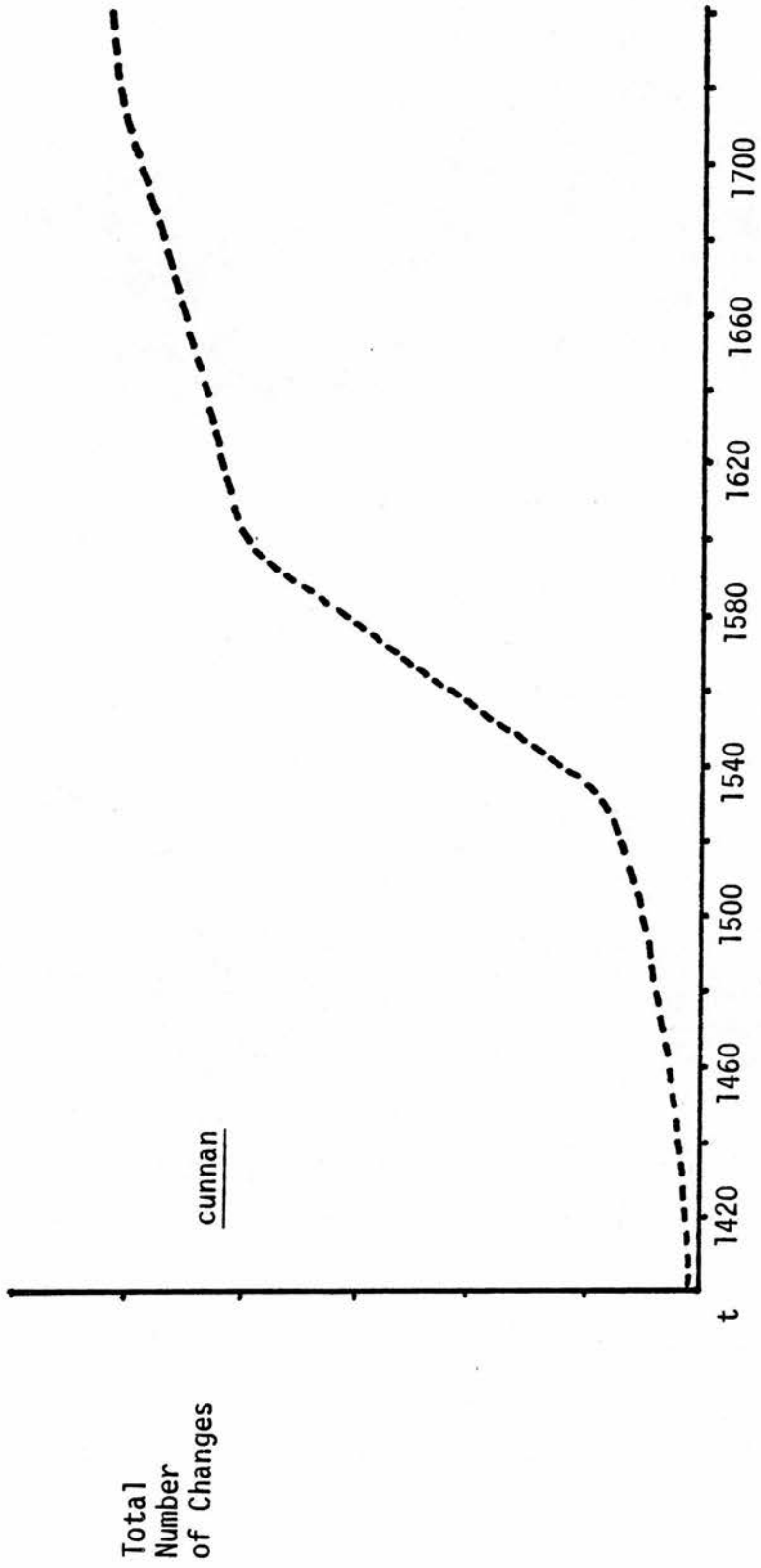
items available have been affected, the pace slows down and the graph flattens out. Hitherto such an analysis has not been applied to syntactic change. The number of items available for change is fewer but consider the following graphs (pp.111-113).

In these three graphs time has been set against the x-axis while the number of changes has been set against the y-axis. The value of the y-axis is the total number of changes which have occurred up to the time indicated by the value on the x-axis, e.g. in 2.48 by 1600 nine changes have occurred. Graph 2.46 plots the development of cunnan while 2.47 records the history of magan and sculan. The total number of changes set against time is recorded in 2.48, which is a graphic representation of 2.44.

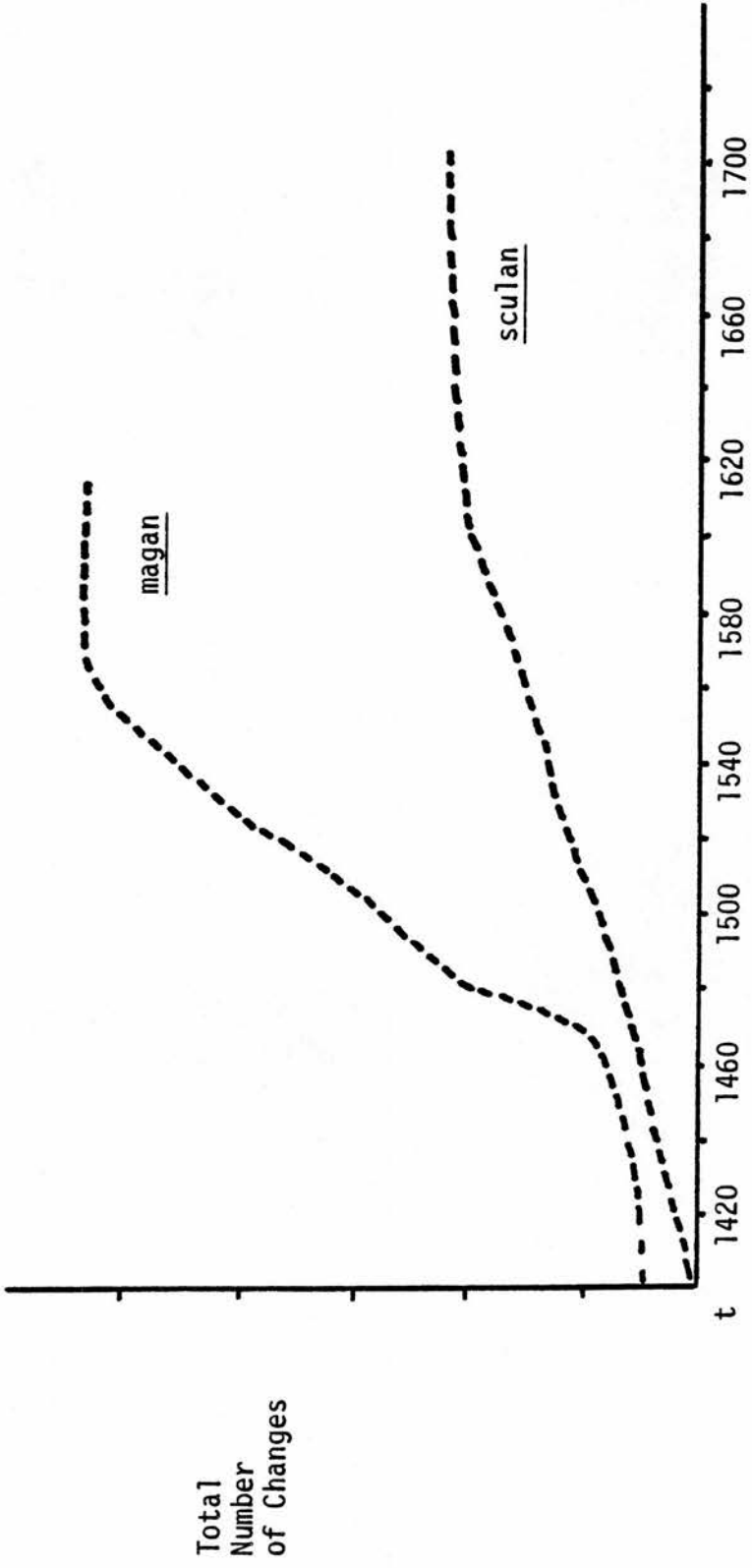
These graphs are not conclusive but their S shape is characteristic of lexical diffusion changes. This lends support to the claim that the development of pre-modals into modals was brought about by lexical diffusion rather than by a re-analysis of the rules of the base. Both the spread of the time scale and the distribution of the changes, illustrated by the graphs, support the lexical diffusion approach. This leaves only the question of what advantages it offers over Lightfoot's account.

Lightfoot dates the introduction of the category modal and the Aux re-write rule as mid sixteenth century. This leaves a number of exceptional constructions to be accounted for (see 2.44 above). These exceptions, for example after the introduction of the rule cunnan retains its infinitive, present and past participle form and may also occur with a direct object, could be handled lexically. The recategorization would still take place in the mid sixteenth

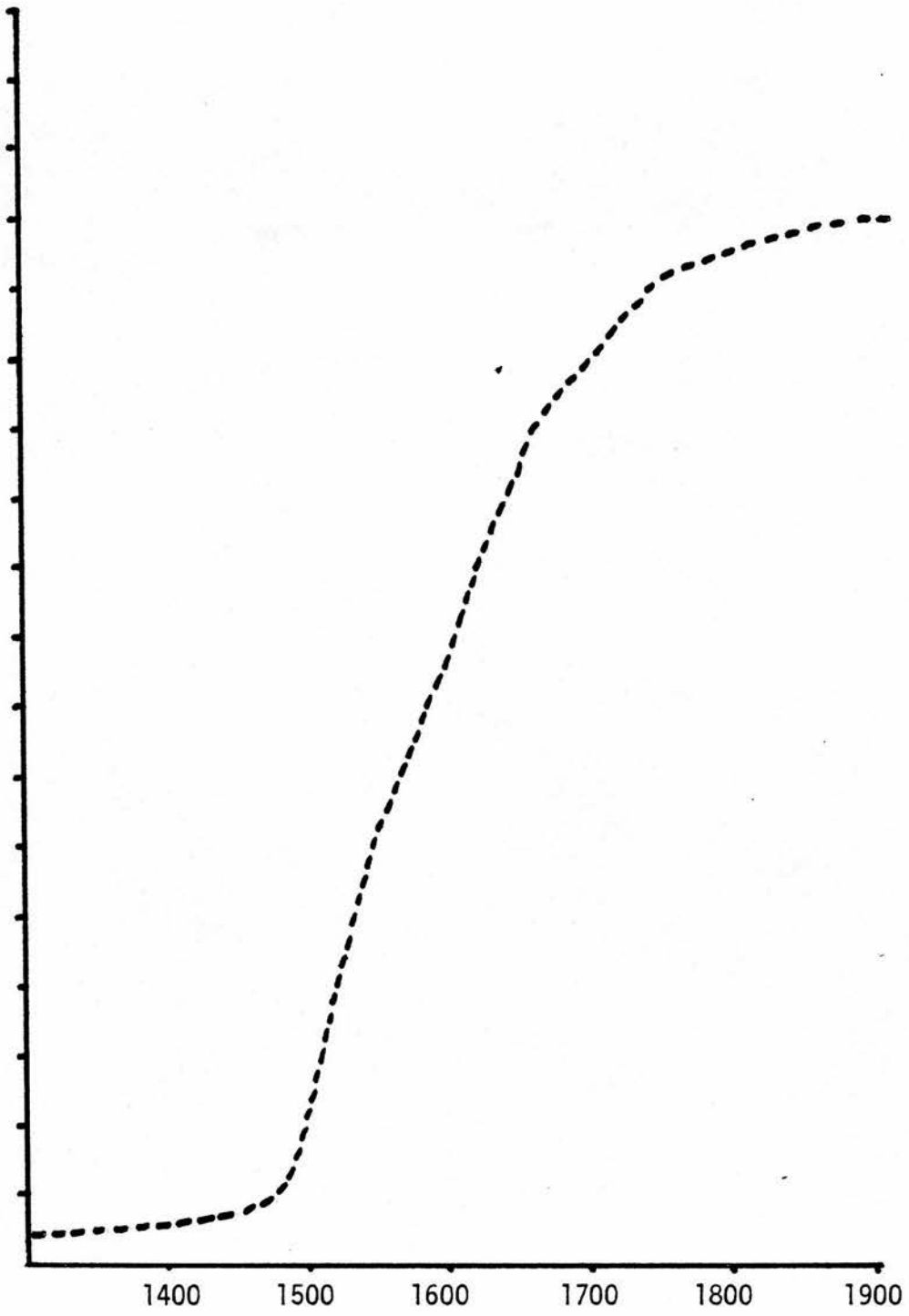
2.46.



2.47.



2.48.



century but those pre-modals which retained the ability to appear in a wider range of morphological forms and syntactic constructions would be marked in the lexicon as exceptions to the Aux rule. Consider, for example, the development of cunnan (the features used here are from Lapointe 1980):

2.49. cunnan

[+M]	(	<u>NP</u>	)	
[+ non part]	(	<u>V</u>	)	
		[+ non part]		
[+ part 1]				
				⇒ [- part 2] → [- part 1] → [- non part]
[+ part 2]				

In this sequence cunnan first loses the ability to appear as a progressive, then as a perfective and then in infinitive constructions. Eventually it would conform to the Aux rule by losing the ability to take a direct object. These features plus the rule would account for the increasingly restricted distribution of cunnan. However, the features themselves reflect this.

The introduction of the features allows us to capture the time spread, over which the changes occur, and to capture the exceptions. We can now dispense with the change in the rules of the base and allow change to proceed via the lexicon. It is still possible to eventually include the rule  $Aux \rightarrow T(M)$ , if we accept that Lightfoot's analysis is the best for N.E. Once all the changes are over we can impose order upon the irregularity by factoring out any generalizations. In the case of the modals this would involve the forming of a new class the behaviour of which may be captured by a syntactic rule. This would occur in the late nineteenth century and not the

mid sixteenth century as Lightfoot claims.

Although the re-analysis by rule theory looks highly unsound there still remains the one problem it was devised to solve, that is: how might we give unitary synchronic syntactic account of the modals without using exception features? This will be discussed more deeply in the next chapter. However, it is interesting to note that the overall effect of the late group of changes was to remove those modals which occurred in non-finite positions from those positions (see Anderson, 1976:40). The result is that the modals are restricted to appearing only in finite positions. This is reflected in their traditional name, the 'anomalous finites'.

CHAPTER THREE

A Dependency Account of Modals and Auxiliaries

### 3.0. Introduction

In this chapter I intend to survey the synchronic distributional evidence for treating the modals as a unique class within the grammar. I shall also discuss the putative role of Aux in English and then propose a dependency account of the English verbal system.

### 3.1. Modal Opacity

It is certainly the case that the modals are morphologically opaque. They lack the third person singular inflection, infinitive, present and past participle forms. However, it must be pointed out that the absence of the last three is explained by the failure of the modals to appear in non-finite positions. They are also morpho-semantically opaque with respect to past time interpretation. Let us now look at the syntactic distribution of the modals and all auxiliaries, of which modals form a proper sub-set, to see if they are syntactically opaque.

#### 3.1.1. Criteria for Auxiliaryhood

Pullum and Wilson (1977), amongst others, list the following as criteria for auxiliaryhood:

(i) Subject-Aux Inversion. An auxiliary permutes with a subject NP to form a Yes-No question; main verbs do not e.g.

3.1. May I leave now?

\*Leave I now?

(ii) Tag Formation. Auxiliary verbs appear in the tag of tag questions; main verbs never do e.g.

3.2. You can go tomorrow, can't you?

\*You go tomorrow, go you?

(iii) Do-Support. Do appears before main verbs in certain constructions but never before auxiliaries:

3.3. You don't leave until Thursday.

\*You don't may leave until Thursday.

(iv) Negative Contraction. The negative particle not may cliticize onto a preceding auxiliary but not onto a main verb e.g.:

3.4. You can't go today.

\*You gon't today.

(v) Auxiliary Reduction. Some auxiliary verbs e.g. will, have, be, when unstressed reduce and cliticize onto the subject NP.

(vi) Quantifier Floating. Quantifiers which follow auxiliaries may be associated semantically with the subject of that sentence e.g.:

3.5. We may all be leaving soon.

\*We leave all soon.

However, some speakers accept the following sentences as grammatical:

3.6. They seem to all like him.

They want to all leave tomorrow.

We expect to all leave tomorrow.

It would appear that with these equi verbs the quantifier may be floated down into the infinitival complement. If this phenomenon were to be more widespread then it could be claimed that, in certain

constructions quantifiers may be floated down to but not past the last verb in the construction. This is, however, peripheral to our discussion and therefore will not be pursued any further at this point.

(vii) Adverb Placement. Certain adverbs normally appear after the first auxiliary, but never after the main verb they are supposed to modify e.g.:

3.7. I can easily swim ten miles.

\*I swim easily ten miles.

This now raises the issue of to what extent these criteria uniquely identify auxiliaries and demarcate them from main verbs or is there any overlap between the two categories and if so where does it occur?

Several items pose problems for these criteria. In particular, need, dare, have, be, is to, used to and ought to are problematical. I now intend to show why the distribution and behaviour of these items is a problem.

### 3.1.2. Need and Dare

The behaviour of these two items is problematical for any analysis which treats modals, and auxiliaries in general, as a distinct class. For most speakers need in positive declarative sentences and all sentences with a nominal object displays the characteristics of a main verb. However, it may behave like a modal in negatives and questions:

3.8. (i) He needs a new theory.

(ii) He needs to find a new theory.

(iii) \*He need a new theory.

- (iv) %He need find a new theory.
- (v) \*Need he a new theory?
- (vi) Need he find a new theory?
- (vii) Does he need to find a new theory?
- (viii) He doesn't need to find a new theory.
- (ix) He needn't find a new theory.
- (x) \*He needn't a new theory.
- (xi) He doesn't need a new theory.

As a main verb need occurs with a third person singular marker ((i) and (ii) above), takes NP complements ((i) and (xi) above), requires Do-Support ((vii), (viii) and (xi)) and takes to complements ((ii), (vii) and (viii)). When it appears in the modal form it loses the third person singular marker ((iv), (v) and (vi)), the ability to take NP complements ((iii) and (x)) and to take a to complementizer ((iv), (vi), and (ix)). It does not require Do-Support in either questions or negatives ((v), (vi), (ix)) and as a modal it permits negative cliticization ((ix) above).

Given an analysis where modals are treated as a separate class the only way of handling the behaviour of need would be to posit two separate entries. One entry would be a main verb which would exhibit main verb behaviour and the other an auxiliary with no complementizer and appearing only in negatives and questions. However, this denies the relatedness of (vi) and (vii), and (viii) and (ix). Similar arguments may be advanced for dare:

- 3.9. (i) He dares to defy my authority.  
 (ii) %He dare defy my authority.  
 (iii) Does he dare to defy my authority?

- (iv) Dare he defy my authority?
- (v) He doesn't dare to defy my authority.
- (vi) He daren't defy my authority.

### 3.1.3. Ought and Be

Ought and modal be pose a different problem. By the criteria listed above, which are uncontroversial, these two items are auxiliaries. However, they obligatorily require a to complementizer. This is normally a feature of main verb status:

- 3.10. (i) Ought he to run for president?
- (ii) He ought to run for president.
- (iii) Is he to arrive on Wednesday?
- (iv) He is to arrive on Wednesday.

If ought is an auxiliary, then, as run is a main verb, to must be introduced into Aux by a mechanism distinct from that which normally introduces complementizers (see Huddleston 1974:219). Alternatively, if ought is a main verb which occurs with tenseless dependent clauses then to can be handled by those rules which cover complementizers. However, the problem of explaining its obligatory auxiliary behaviour remains. Similar problems arise with modal be.

### 3.1.4. Have and Non-Modal Be

If we assume that all sentences rather than utterances contain a main verb then have and be are main verbs in the following:

- 3.11. (i) He is the president.
- (ii) He has the plan.

However, by the criteria listed above, be obligatorily and have optionally behave like auxiliaries:

- 3.12. (i) Is he the president?  
 (ii) Has he the plans?  
 (iii) Does he have the plans?  
 (iv) He is the president, isn't he?  
 (v) He has the plans, hasn't he?  
 (vi) He has the plans, doesn't he?  
 (vii) He isn't the president.  
 (viii) He hasn't the plans.  
 (ix) He doesn't have the plans.  
 (x) He's the president.  
 (xi) \*He doesn't be the president.  
 (xii) I have the plans in my office.  
 (xiii) I've the plans in my office.  
 (xiv) They are all presidents.  
 (xv) They all have copies of the plan.  
 (xvi) They've all copies of the plan.  
 (xvii) He is hardly the president.  
 (xviii) He doesn't really have very much space in  
 his new flat.  
 (xix) He hasn't really very much space in his new  
 flat.

One solution would be to write the auxiliary-choosing rules so that have and be are specifically mentioned in their S.D.'s. However, this can only be, at best, an ad hoc solution to the problem. It seems rather odd to assign disjunctive bracketing to the modals and

only two members of the large class of main verbs. This oddness is obscured by the fact that have and be have aspectual functions. Further, whether or not an item or class of items undergo a transformation is not normally stated in the rule itself but usually in the lexicon. We might conclude from this that the odd behaviour of these verbs ought to be handled in the lexicon rather than in the syntax.

### 3.1.5. Would Rather

Consider the following example:

3.13. I would rather (that) you came tomorrow.

If we insist on would being a modal then we have a very unusual clause structure. The example above contains a modal, a degree adverbial and a dependent sentential object complement but no main verb. If we stick rigidly to the division between auxiliaries and main verbs and to the principle that all sentences contain a main verb then this sentence is clearly problematical. We could permit would to be a main verb in this one construction and a modal in all other constructions or allow rather to be the main verb. Clearly the second of these two alternatives would be completely ad hoc and unworkable. The prospect of treating would as a main verb and then moving it under the Aux node to allow for sentences such as:

3.14. Would you rather I came back tomorrow?

You wouldn't rather I came back tomorrow.

is possible and could also be used for need and dare. However, this is also ad hoc and simply involves a relabelling rule.

### 3.2. Main Verb-Auxiliary Similarities

It would appear that not all auxiliaries adhere to the criteria listed above. There are main verbs which also follow, in some cases obligatorily and in others optionally, these criteria. Further problems for a strict auxiliary-main verb division arise when we consider some regularities which may be couched in the form of transformational rules which refer to both main verbs and auxiliaries.

#### 3.2.1. Affix Hopping

Certain transformational rules refer to both auxiliaries and main verbs. One such rule is Affix Hopping. This rule accounts for the discontinuous morphology of the English verbal system. It permutes affixes over either a following auxiliary or main verb:

3.15. I have+EN be+ING eat.            I have been eating.

Any strict categorial distinction between auxiliaries and main verbs leads to a complication in this rule. It is possible to remove the rule from the syntax and have the process carried out lexically; for example, see Lapointe (1980) and Langendoen (1975) for a discussion. However, if the division between main verbs and auxiliaries is maintained the lexical rule will still be over complex. We would simply have moved the process from the syntax into the lexicon without achieving any overall simplification of the grammar.

#### 3.2.2. Gapping

Gapping is another rule which refers to both auxiliaries and main verbs. In the second of two conjuncts gapping deletes the verb if it is identical to the verb in the first conjunct (Grinder and Elgin,

1973:102-3):

3.16. I ordered Benedictine and Harriet ordered the stomach pump.

I ordered Benedictine and Harriet the stomach pump.

However, this formulation of the rule is simplistic. Consider the following examples:

- 3.17. (i) John may sit and Fred stand.  
 (ii) John has eaten the steak and Fred drunk the wine.  
 (iii) John may have been sitting here and Fred there.  
 (iv) \*I must eat the steak and Fred must the ham-burger.

In (i) it is the modal which is removed and not the main verb. The second example shows that an auxiliary may be gapped while in (iii) the entire sequence of verbs is deleted. If gapping were to be formulated to look for main verbs then (i), for example, which is clearly derived by the application of gapping, would require the application of a rule which performs the same task as gapping but is formulated to operate on auxiliaries. Clearly this misses a generalization. These two processes may be combined, and simplification of the grammar achieved, if we treat auxiliaries as main verbs. Further evidence for the reformulation of gapping arises from (iv) above. If gapping applies only to main verbs then (iv), which is clearly ungrammatical, will be generated. We might reformulate gapping to delete a main verb and, optionally, any dependent verbal complement. If we assume auxiliaries to be main verbs, then this

formulation will generate (i)-(iii) above and exclude (iv). Eat, in (iv) cannot be deleted unless must has first been deleted. This reformulation of gapping requires the rule to apply from left to right in a sequence of verbs. Apart from the leftmost verb any verbal element may only be deleted if the verb form to its immediate left has been deleted.

### 3.3. Syntactic Opacity

Lightfoot claims that the opacity of the modals led to their re-analysis. It is certainly true that they are, in some respects, morphologically and semantically opaque (see 3.1). However, it seems bizarre to consider them to be syntactically opaque. Consider Kiparsky's (1971) definition of phonological opacity:

- 3.18. A rule  $A \rightarrow B/C \_\_\_ D$  is opaque to the extent that there are surface representations of the form:
- (i) A in the environment  $C \_\_\_ D$
  - (ii) B in environments other than  $C \_\_\_ D$ .

Does this when applied to the syntactic distribution of the auxiliaries suggest that they are opaque? The answer to this question must, I think, be no. Like all verbs they can take a subject to their immediate left. This must be the verbal position par-excellence. It is also the case that the non-modal auxiliaries can be dependent on another auxiliary to their immediate left. All main verbs may also be dependent upon an auxiliary which occurs to their immediate left. Therefore this does not distinguish main verbs from non modal auxiliaries. All auxiliaries must take a dependent form to their right and while this is not obligatory for main verbs it is possible for

some of them to take a dependent infinitive form. Consider:

- 3.19. (i) I must go.  
 (ii) I ought to go.  
 (iii) I want to go.  
 (iv) I let go.

It is clear that these sentences share surface structure similarities. This is particularly obvious with (ii) and (iii) where both the main verb and the modal auxiliary take a to complement. Most modal auxiliaries do not take a to complement, but as (iv) shows this is not a feature unique to modal auxiliaries. Within a model of grammar which retains the Aux node (ii) and (iii) would differ in their derivations. Their underlying representations would be:

- 3.20. [I    [[ought to]    go]]  
       S    VP AUX
- [I    [want    [I    [go]]]]  
       S    VP        S    VP

The underlying structure of (iii) contains an embedded sentence, the subject of which is deleted by equi. The underlying structure of (ii) consists only of a simple sentence. Within a theory which employs traces and  $\alpha$ -movement (see, for example Chomsky, 1980 or Radford, 1981) (ii) and (iii) would derive from similar underlying structures:

- 3.21.    [[e<sub>i</sub>]<sub>S</sub> ought [ I<sub>i</sub> [to go]]]  
       ⇒ [ I<sub>i</sub>    ought [ [e<sub>i</sub>]<sub>NP</sub> [to go]]]  
       [[e<sub>i</sub>]<sub>S</sub> want [ I<sub>i</sub> [to go]]]

$$\Rightarrow \begin{array}{c} [I_i \\ S \end{array} \quad \text{want} \quad \begin{array}{c} [ [e_i] \\ S \quad \text{NP} \end{array} \quad [\text{to go}]]$$

The rule of NP movement applies to both underlying structures, leaving a coindexed trace  $[e_i]_{\text{NP}}$  behind.

Within a lexicalist grammar, as with the trace theory derivation, (ii) and (iii) would derive from similar deep structures but no rule of NP movement would apply (see 1.3.6):

3.22.  $\begin{array}{c} [I \\ S \end{array} \quad \text{ought} \quad \begin{array}{c} [to go] \\ VP \end{array}$   
 $\begin{array}{c} [I \\ S \end{array} \quad \text{want} \quad \begin{array}{c} [to go] \\ VP \end{array}$

In these examples the lexicalist treatment of ought and, presumably, of all auxiliaries, including modals, is equivalent to that given to any verb which takes a verbal complement. VP is introduced by the rule  $VP \rightarrow V VP$  and  $VP \rightarrow \text{Comp } VP$ . This parallels Bresnan (1970) where S was introduced and expanded as  $\text{Comp } S$  (see also Bresnan, 1978). The distributional and morphological differences between e.g. ought and want would be handled by the lexicon, allowing the auxiliaries to be members of the syntactic class to which they really belong - verbs.

It would appear that there is no justification for considering the modals and all auxiliaries to be syntactically opaque.

#### 3.4. A Finite Argument for Main Verb Status

A further argument for main verb status concerns the ability to take tense. This may appear trivial and is certainly self-evident but the ability to occur in a tensed form is the criterion which uniquely

identifies a verb. This ability is, however, dependent upon another equally important distinction: the distinction between finite and non-finite verb forms. These forms may be distinguished as follows:

- (i) Finite verb forms have tense distinction i.e. they distinguish between present and past tense:

3.23. He  $\left. \begin{array}{l} \text{(reads)} \\ \text{(read)} \end{array} \right\}$  Hardy.

- (ii) Finite verb forms occur as the verb element of a clause and take a subject. There is person number concord between the subject and the finite verb. This is particularly overt with be but with most verbs it is restricted to a contrast between the third and non-third person singular present tense. With the modal auxiliaries there is no person number concord.
- (iii) Finite verb forms have mood distinctions. It is possible to distinguish between the indicative, imperative and subjunctive moods.
- (iv) Non-finite verb forms are the infinitive, present and past participles. They do not occur with an unmarked subject.

As with the criteria in 3.1.1 not all verbs fulfil these criteria. Statives do not occur in either the imperative mood or a present participle form. Modals do not occur in non-finite forms, the imperative mood and fail to undergo person number concord. Valency cannot be invoked as an identifying criterion - some verbs and adjectives have a valency of one.

It would appear that the criterion which identifies all auxiliaries and main verbs as members of the same class - verb - is the ability

to take tense. This is linked to the subject taking requirement and we shall return to this below. In any sentence in which they appear, the modal auxiliaries may only occur in a tensed form. This restricts them to finite forms. No other verb has this restriction placed upon it. We have made reference to this restriction above (see 2.5-2.5.2). It is from this one simple fact that all other restrictions on the distribution of the modals follow. If there can only be one finite verb per simple sentence then we would not expect to find double modals. Given that the modals have no non-finite forms we do not expect to find them embedded under either the perfective or progressive - the form of the verb embedded must be non-finite.

This traditional view regards the modals as verbs which are paradigmatically defective in lacking a non-finite paradigm. It is a view echoing Anderson (1976), McCawley (1975) and Pullum and Wilson (1977). Any analysis which fails to reflect this major schism between the two sets misses a significant generalization. Analyses of English which contain the Aux node and re-write rule fail to capture, directly, the restriction of the modals to finite positions. Further, these analyses treat the ability of modals to take a subject and infinitival object complements almost as an accident.

Lightfoot would claim that his analysis, which introduces the Aux node and re-write rule, does capture these restrictions concerning the behaviour of the modals. The Aux re-write rule,  $Aux \rightarrow T(M)$ , ensures that there will only be one modal per simple sentence, that, by occurring to the immediate right of T(ense), it will be finite and that, by occurring to the left of perfective HAVE and progressive BE, it will never appear as either a past or present participle. How-

ever, the introduction of this node and rule simply ensures that the modals will fail to undergo processes which ought not to be in the syntax but in the morphology component as they deal with the internal structure of lexical items (see 1.3.6). This would lead us to remove both the rule and these processes from the syntax. Let us look in more detail at the arguments advanced in favour of eliminating the Aux node from our grammar.

### 3.5. On the Elimination of Aux

It would appear that there is a substantial amount of evidence in favour of treating the modals and all auxiliaries as main verbs. This evidence is distributional and rule based. If accepted it would lead us to conclude that the Aux node may be eliminated, thus simplifying the grammar.

#### 3.5.1. Affix Hopping Revisited

Brown and Miller (1982) advance a number of arguments for the elimination of the Aux node in English. They begin by examining the rule of Affix Hopping, which we claimed above (3.2.1), following Pullum and Wilson (1977), was a rule which provided support for a main verb analysis of auxiliaries. This rule requires in its structural description the identification of several categories. It permutes constituents and creates extra structure via Chomsky adjunction. Further, the rule makes reference to the internal structure of words. This is due to the desire to show that, for example, singing consists of a base plus suffix and realises a lexeme and a grammatical morpheme.

Both Brown and Miller (1982) and Lapointe (1980) state that the rule

cuts across the traditional division between syntax and morphology - a distinction we discussed in (1.3.6) above. Lapointe goes further in his criticism by listing the following:

(i) T, -EN, -ING are the only affix like objects to have an independent status in both syntactic rules and structures. Morphemes which actually occur as parts of surface words e.g. the past tense marker -ed do not have this independence.

(ii) The functions performed by the re-adjustment rules, which turn the base+affix sequences into surface forms are very varied e.g. the replacement of an abstract suffix by an occurring one - PLAY+EN → PLAYED, deletion MAY+Present → MAY, internal phonetic change RING+PAST → RANG and finally suppletion BE+PAST → WAS. Lapointe (1980:221) claims that these processes suggest that the readjustment rules are actually morphological word structures mechanisms which apply, for him, in the lexicon, rather than syntactic readjustment rules whose application allow phonological rules to apply. (See Langendoen 1975 for such rules.)

(iii) This analysis allows the morphology to interact twice with the syntactic structures -- once during lexical insertion and again when the readjustment rules apply. This permits a wider range of descriptive possibilities than a more restricted account which would allow such interaction only once.

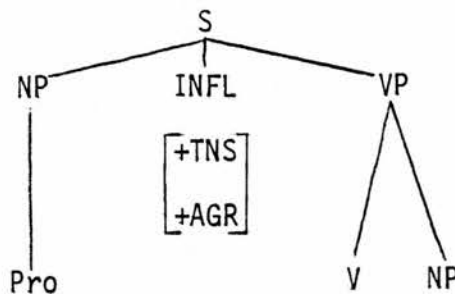
(iv) The post-syntactic readjustment rules allow the use of abstract cover symbols and transformations in accounting for the morphological distribution of auxiliary verbs.

Brown and Miller and Lapointe eliminate Affix Hopping by allotting a greater role to the lexicon. However, Brown and Miller remove the Aux node while Lapointe retains it. Despite this difference, both are clearly correct in arguing that inclusion of Aux and Affix Hopping permits syntax and morphology to be mixed. This is particularly obvious in the case of (iv) above, where it was stated that the inclusion of the Aux node allows transformational rules to be used to account for the morphological distribution of the auxiliaries. These morphological factors ought not to be handled by the syntax but by a separate morphological component. This would provide a more restrictive account, as the syntax and morphology would not be able to interact twice (see (iii) above). Interaction would only occur once, during lexical insertion. We shall return to this below, but it offers the advantage that it maintains the traditional distinction between syntax and morphology which we discussed in 1.3.6.

Similar arguments can be raised against INFL(ECTION) (see Chomsky (1981)). INFL is the category of the tense and subject agreement morpheme. Further it is a type 0 major category. This means that it is a lexical rather than a phrasal or minor category.

Consider the following phrase marker:

3.24.

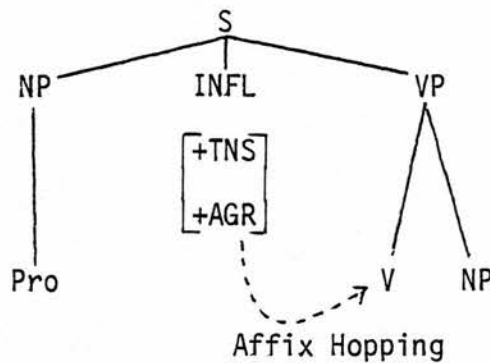


From this diagram it is clear that INFL is a lexical node on the same

level as N, V, A, or P. However, unlike other lexical or type 0 nodes it dominates, not a lexical item but a finite set of features which are ultimately realised by a small group of bound morphemes. (This is somewhat simplistic as deletion, suppletion and internal phonetic change are all involved.) The INFL node gives the bound morphemes of tense and agreement an independent status in the grammar which no other bound morpheme has. No bound morphemes, including tense and subject agreement, may be considered to be of the same status as, for example, man, dog, write, and blue. This parallels (i) above.

As with the Aux node a rule of Affix Hopping or some type of realizational rule is required to map the features dominated by the INFL node onto the V (see Bresnan (1982) for a discussion):

3.25.



This raises the same problems as Affix Hopping and Aux did concerning the interaction between syntax and morphology -- see (iii) and (iv) above. Although the INFL Hopping rule maps the tense and agreement features onto the verb a number of readjustment rules will still be required to ensure that the correct surface form surfaces. Further, as Bresnan (1982) points out, this hopping rule might just as easily hop the features dominated by INFL onto any other node.

Clearly, as with the Aux node and Affix Hopping rule, INFL and the

rules associated with it cut across the traditional divide between syntax and morphology. For this reason and those discussed above a more restrictive analysis may be given if we remove the INFL node from the grammar.

To return to Aux, Brown and Miller (1982:134) state that generative grammar has taken over the term Auxiliary for the node which groups together those constituents in the verb group other than the main verb. These constituents do not occur in a random order and this is used as a reason for giving the auxiliary constituents different labels. This makes it easy to write a PS rule which produces the correct order. They remark that not one of these labels mentions the term VERB. Modal relates to the meaning of the class of items it labels. Perfective and progressive refer to aspectual distinctions and the particular construction types which grammaticalise these concepts. Nowhere is it indicated that the constituents involved are verbs. This is equivalent to setting up categories of COUNT and MASS rather than using these as features on nouns. It is certainly the case that the elimination of the terms modal, perfective and progressive might lead to ordering problems but these might be resolved by relegating these terms to features on verbs.

It is clear from the evidence discussed above that auxiliaries should be treated as verbs. Following such evidence, Brown and Miller (1982) argue that the Aux node be eliminated from the grammar and the VP be expanded as a series of verbs, with 'strict subcategorization frames' in the lexicon handling the ordering problem - Brown and Miller (1982:138).

Lapointe (1980), although allowing the lexicon to play a greater rôle

in his grammar, retains the Aux node. In doing this he accepts Akmajian, Steele and Wasow's (1979) analysis of Aux as a universal category. We shall now turn to A.S.W.'s arguments to see if they provide any evidence for retaining Aux in English. If we refute the existence of Aux in English then it cannot be a true universal.

### 3.5.2. Akmajian, Steele and Wasow

A.S.W. adopt a base schema for English which is based on Emonds (1976) and Culicover (1976). In this base schema Aux contains tense plus either a modal or do in deep structure. Have and be are generated as verbs taking VP complements and if they are present they will obligatorily replace do in Aux. Further do deletes if immediately followed by a verb. This definition redefines VP in an interesting way. In this analysis there are three distinct VP types which stack up one on top of the other.  $V^0$  is the label for main verbs,  $V^1$  immediately dominates  $V^0$  with an optional passive be to its left and whatever complements it may have to its right.  $V^2$  immediately dominates  $V^1$  with an optional progressive to its left.  $V^3$  immediately dominates  $V^2$  and takes an optional perfective to its left. This can be shown schematically by the following rules:

3.26.  $S \rightarrow NP \text{ Aux } V$

$$\text{Aux} \rightarrow \left\{ \begin{array}{l} \text{Tense do} \\ \text{Modal} \end{array} \right\}$$

$$V^n \rightarrow \left( \left[ \begin{array}{l} +V \\ +\text{Aux} \end{array} \right] \right) V^{n-1}$$

Lapointe (1980) uses a slightly revised version of these rules:

3.27.  $S \rightarrow N^3 (\text{Aux}_1) V^3$

$$V^3 \rightarrow (\text{Aux}_2) V^2 \dots$$

$v^2 \rightarrow (\text{Aux}_3) v^1 \dots$

$v^1 \rightarrow v^0$

Aux<sub>1</sub>: modals, tensed forms of have and be

Aux<sub>2</sub>: infinitival have

Aux<sub>3</sub>: be, been

The differences between these two rule schemas need not trouble us here. The distributional behaviour of ought, need, dare and used pose problems for the above analysis. Let us first consider ought, which behaves morphologically and syntactically like a modal but takes a to infinitive complement. Pullum and Wilson (1977) decide to treat ought as a main verb, as they do all auxiliaries, which takes infinitival complements, in the same way as, for example, want does. However, for Pullum and Wilson items like ought, having been generated under VP, are then subject to an obligatory hopping rule. This hopping rule moves, for example, ought from under the VP node and places it in Aux, removing do at the same time. It would perhaps be more accurate to call this hopping rule a relabelling rule as its main function is to relabel a specified class of items as Aux to ensure that they are picked up by the rules which refer to Aux (see 3.1.5 and Pullum (1976) on what might be considered the phonological equivalent of this).

In A.S.W.'s analysis ought would appear under  $v^0$  in  $v^1$  but would be marked [+Aux] lexically. This allows a hopping or Do-Replacement rule to pick up ought and transfer it to Aux. However, there still remains the problem of preventing other auxiliaries from appearing before it. Remembering that ought, in common with all modals, has no non-finite paradigm, we must prevent it from appearing in non-

finite positions. It would be simple enough to block ought, and all modals, from non-finite positions lexically, utilizing the distinction between modals and other verbs.

This solution is not adopted by A.S.W. Pullum (1981) points out that if A.S.W. were to adopt this solution they would be undermining their own analysis:

"if ought, with its lack of third person singular -s, could be treated as a morphologically aberrant verb with the special property of being subject to Do-Replacement, as the copula is, there would be no point in setting up a category 'Modal', nor any criteria for membership in such a category. All the 'modals' could be treated like ought, except that they would take the bare infinitive instead of to+infinitive. And once the category 'Modal' was gone 'Aux' would never dominate anything but Tense and do in deep structure, so it would be essentially a slot for introducing a finite auxiliary verb. Do has the paradigm of a verb, and all the elements that can replace it in 'Aux' are actually labelled V. The case that 'Aux' exists independently of V would just crumble away."

However, to maintain the category modal as a homogeneous category A.S.W. would assign ought to modal. This may be justified on the grounds that it lacks a third person singular present tense marker, does not permit do-support and behaves like an auxiliary with respect to the criteria in 3.1.1. If this is the case then to remains a problem for A.S.W. It could be generated under the Aux node as an optional element following Modal. However, the Aux-choosing rules would then produce the wrong output:

3.28. \*Ought to you do that?

\*You ought ton't do that.

These sentences will be the output of such rules given that they re-

fer to Aux. A.S.W. could re-formulate the rules to refer to Modal, Tense + Do or Tense + V. However this destroys the Do-Replacement analysis and poses problems to which we shall return.

Pullum (1981) suggests that to could be generated outside 'Aux'. This might be done by creating another VP node,  $V^4$ , with a P.S. expansion (to)  $\_\_\_V^3$ . Alternatively to could be generated under S between 'Aux' and  $V^3$ . However, Pullum raises a number of problems for these analyses. The problems stem from the fact that ought must be able to subcategorize the adjacent to while other members of the modals do not require it.

In an Aspects type theory strict subcategorization requires that a V node is replaced by a complex symbol containing the category membership feature [+V] and a set of contextual features that code the properties of its syntactic context. The part of the tree to which this applies is limited to a string  $\alpha$  such that  $V\alpha$  is a VP. This gives the following rule:

3.29.  $V \rightarrow CS/\_\_\_\alpha$ , where  $\alpha$  is a string such that  $V\alpha$  is a VP

Pullum (1982) suggests that this be re-interpreted as the following to allow N, A, and P to be subcategorized for certain deep structure environments:

3.30.  $X \rightarrow CS/\_\_\_\alpha$ , where  $X \in \{N, V, A, P\}$  and  $\alpha$  is a string such that  $X\alpha$  is an  $\bar{X}$ .

Both of these rules state that a lexical category  $\psi$  cannot be subcategorized for a complement C unless  $\psi$  is immediately dominated by a phrasal category  $\bar{\psi}$  of which  $\psi$  is the head.



then optionally hopped into Aux for negatives and interrogatives. In each of these cases the choice is between generating the items in the VP, thus categorizing all auxiliaries as main verbs, and consequently rendering the Aux node redundant or generating them under Aux posing problems for the theory of strict subcategorization. We could, of course, weaken it but in doing so we would be rendering it less restrictive than many of the more recent proposals would permit.

Further problems for A.S.W. arise when we consider their definition of Aux. For example consider Steele et al (1981:21):

"the elements of the definition comprise a small set: constituent, tense, and modality."

or A.S.W.

"Aux is a category - i.e. distinct in its syntactic behaviour from the behaviour of other syntactic categories - labelling a constituent that includes elements expressing the notional categories of Tense and or Modality."

These definitions are produced because A.S.W. (1979:3) claim, without specifying where, that there has been confusion between the category Aux and auxiliary verbs. It must be considered unfortunate that the definitions do not clarify the position at all. They simply give a vague impression of what is to be expected in Aux. What is worse than this is that they assign Aux status to items which A.S.W. would not desire to be given such a description. In Italian, for example, a category, distinct from others, can be found, which includes elements for the expression of tense (credo 'I believe', credero 'will believe', credevo 'was believing', credei 'believed') and modality (dovere 'must', potere 'can' and volere 'will'). Consequently, by the defin-

ition given above it must be labelled Aux. However, were this to be the case then Italian would not have the category V because these items are usually regarded as verbs. Nevertheless, by the definition above they are clearly Aux. Italian is not unique in this as the same situation arises in French.

Clearly we cannot create Aux at the expense of V. Aux in Italian is presumably, in Steele et al's terms the Tense and Mood inflections on the verb (see Steele et al 1981:114-116). The Aux node would not dominate a constituent in the way, for example, N, V, A, and P do but would dominate a number of bound grammatical morphemes. This mixing of syntax and morphology is, of course, open to those criticisms levelled at Aux and Affix Hopping and INFL above (see p.131).

In most languages there are many different categories which meet the definition of Aux. Pullum (1981) points out that in English S, V, A, and Adv can all meet the criteria and therefore will be defined as Aux. For example, V may express modality - 'needs' - and regularly expresses tense on surface structure, A can express ability (able), necessity (necessary), and possibility (possible) all of which are partially synonymous with can, must, and may respectively. It would appear from this that A.S.W.'s definition is not restrictive enough as it defines too many items.

Steel (1978) provides a more restrictive definition of Aux stating that it must (i) be non-verbal (ii) occur in the same clause as the main verb it modifies; that is, it must not determine subordinate morphology on the verb (iii) have a fixed position in the sentence (this rules out adverbs) (iv) not be able to be interrupted by any major constituents in the sentence and (v) the elements of Aux must

appear in a fixed order.

These criteria certainly constrain what might possibly constitute an Aux. However, they are too restrictive. This can be seen by simply applying them to the surface structure of English.

Criterion (i) would rule out the verbs in the following examples as Aux:

3.32. Was it fine today?

Have you a machine gun?

Would you rather I didn't submit?

The second criterion runs into difficulties when the status of -en, -ing, and to is considered. Traditionally these items are regarded as markers of subordination and this is certainly the case with examples like: Fred wants to go and Fred intends leaving. Consequently have, be and ought in the following examples cannot be instances of Aux:

3.33. Has he eaten?

Is he leaving?

Is he to leave now?

He ought to leave.

Let us now consider the third criterion. This stated that Aux be in a fixed position. However, consider these examples:

3.34. He has of course been posing problems.

Of course he has been posing problems.

He has been of course posing problems.

The finite element has, in these examples, not been rigidly restricted to initial, second or any other position in the sentence. Therefore, it fails to adhere to the third criterion. It is, of course, possible to relate this to the variability of of course. However, if we were to adhere rigidly to Steele's criteria, then the finite element would have to be restricted to a fixed sentential position.

It must be concluded, on the basis of the data above, that A.S.W. fail to provide a coherent definition of the category Aux. This, however, is not the end of their problems.

In no analysis of English is Aux only the category label for a terminal element. Aux is not the category to which the traditional auxiliaries belong, nor does it exhaustively dominate them. No auxiliary bears the "is a" relation to Aux in the same way that a verb bears the "is a" relation to the node labelled V. As Huddleston (1976: 125) points out the feature [+V] in the lexical entry of a verb is interpreted as compatibility with a V node for lexical insertion, but the feature [+Aux] is used to mark certain lexical items as being eligible for inversion in questions and is not directly connected to the node Aux.

Finally, if Aux is to be a category then what is to be its head? If the head is obligatory, a lexical item and characteristic of the construction of which it is the head, then what is the head of Aux? As tense is not a lexical item and neither modal nor do are obligatory Aux is left without a head (see 3.6 below on the criteria for the identification of headhood).

Given these considerations it would appear that there is no motiva-

tion for an Aux constituent. Therefore, we must conclude that Brown and Miller (1982) are fully justified in abandoning Aux, although they do not give as many supporting arguments as we do above. Lapointe's retention of Aux seems rather odd considering that he extends the role of the lexicon. The use of Lexical Restriction Frames (Lapointe 1980) would allow ordering of the auxiliaries to be achieved without the presence of an Aux node in the grammar.

### 3.5.3. Natural Serialization and the Status of Aux

Further evidence for the main verb status of auxiliaries can be aduced from a consideration of natural serialization. The concept of natural serialization arises from work by Bartsch and Vennemann (1971). They define, firstly, a principle referred to as natural constituent structure:

"...elements belonging together in the hierarchy of semantic representation tend to be lexicalised and serialized in the surface representation in such a way that hierarchial dependencies are directly reflected in categorial operator-operand relationships and closeness of constituents to each other in the surface string."

This leads on to a definition of natural serialization:

"The natural serialization principle comprises the natural constituent structure but says that furthermore, the operator-operand relationship tends to be expressed by unidirectional serialisation: {Operator{Operand}} tends to be serialised as either [Operator[Operand]] throughout in XV languages or as [[Operand]Operator] throughout in VX languages."

Vennemann (1973:49) defines operator and operand both semantically and syntactically:

"...semantically, that in a constituent structure

AB the specified element is the operand and the specifying element the operator and syntactically, that element which determines the syntactic category of AB is the operand, the other the operator."

This correlates with Greenberg's universals (Greenberg (1963)). Consider, for example, Universal 24, which states that if the relative expression precedes the noun as either the only construction or as an alternative construction, then either the language is postpositional or the adjective precedes the noun or both. In this example the constructions correlate in such a way that the operator precedes the operand. Anderson (1976 and 1979c) interprets the operator-operand relationship as a dependency relationship. I fully agree with this and will now illustrate how this relationship might be captured in a formal grammar.

### 3.6. Dependency Syntax: Heads and Modifiers

The basic notion underlying dependency syntax have already been noted in this thesis (see 1.2 above). This approach to syntax was outlined by Hays (1964), Robinson (1967, 1970a and 1970b) and developed by, amongst others, John Anderson (see Anderson 1971b, 1976, 1977, and 1979c). The main theoretical claim behind this model is that phrases in natural language consist of a head and (optional) modifiers. In a phrasal construction the head is associated with Vennemann's operand and the modifiers with the operator. This is not circular because independent criteria for the identification of the head/operand have been posited. Robinson (1970a) suggests that heads be identified by the following criteria, which were introduced above (see 3.5.2):

- 3.35. (i) it is obligatory  
 (ii) it is atomic i.e. a lexical category  
 (iii) it is characteristic i.e. it uniquely identifies the phrase.

Consider now a prepositional phrase. It consists of a preposition and a noun phrase. Both are obligatory but only the preposition is atomic - the noun phrase is not a lexical category. Therefore, we would conclude that the preposition is the head. This conclusion is confirmed by the third criterion.

Given that we can independently identify the head of a construction we are now in a position to make use of the dependency relationship. This relationship is considered to be binary, asymmetric, and intransitive. Ordered pairs of elements in any string over a finite vocabulary are considered marked by the relation. Any string structured by this relation is said to be simple if and only if two specific conditions are met:

- 3.36. (i) there is a unique element which depends on no other  
 (ii) all other elements depend on only and at least one element.

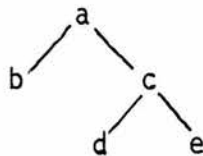
The unique element specified by (i) is the central or main element in the string. The binary relation that arises from the transitive closure of dependency is termed subordination. Informally, one element is subordinate to another if they are linked by a path of dependencies terminating in one and initiating from the other. Consequently all elements in a simple structured string are sub-

ordinate to the central element. Such strings may be given a graph theoretic representation. If a simple string is associated with a proper tree - a directed graph such that the elements of the string are paired with vertices or nodes and the dependencies with arcs arriving in the dependent node and initiating from the governor - then the central element will be identified with the root of the tree. Proper trees meet two conditions equivalent to those given above:

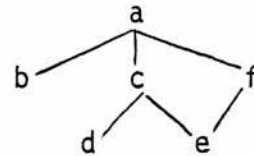
- 3.37. (i) there is a unique vertex which terminates no arc  
 (ii) all other vertices terminate at least one arc.

Given these conditions then (a) below constitutes a proper tree whereas (b), (c), and (d) do not.

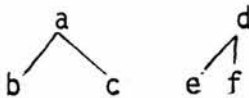
3.38. (a)



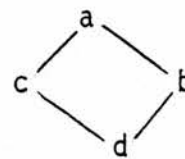
(b)



(c)



(d)



In (b) e terminates two arcs, (c) has two roots and (d) has no root - it is a circuit.

### 3.6.1. Dependency Syntax: Rules

Having shown that the notion head can be independently identified and that the governor/head - modifier/dependent relationship can be captured via a directed graph we are now in a position to provide a set of rules for the derivation of simple sentences.

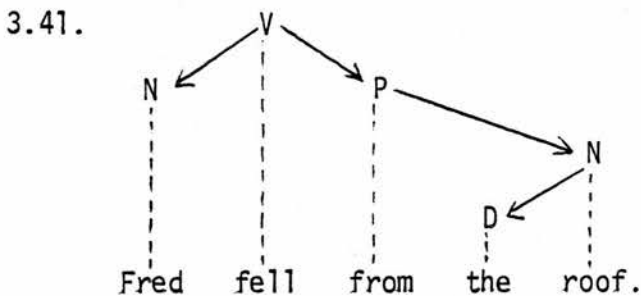
Consider the following sentence:

3.39. [Fred [fell [from [the roof.]]]]

The bracketing of this sentence into its constituent parts (both lexical and phrasal) does not lead to immediate identification of the head of each construction. Therefore, dependency grammar uses \* for this purpose, where headhood is assigned in accordance with the criteria listed above:

3.40. [[Fred\*] [fell\*] [from\* [the\* roof.\*]]]

However, once \* has been added to the grammar there is no longer any need to identify pre-lexical categories. If \* identifies, for example, N then it also identifies NP. Pre-lexical categories can be predicted from lexical categories, giving a structure based on not pre-lexical and lexical but heads and non-heads. Ignoring, for the present, all morphology, the phrase marker associated with the above sentence would be:



Before proceeding to produce a set of phrase structure dependency rules we must consider whether we are correct to assume that V is the head of the sentence as we have assumed in the above phrase marker. Robinson (1970a) proposes T(ype), an auxiliary constituent, as the head of the sentence. Type carries tense, mood and intonation.

This makes nonsense of the claim that the head must be a lexical category. Further, if we were to accept T as the head of the sentence then we would be open to the same criticisms as we have levelled at Aux in both an Aspects type grammar and Steele et al's analysis and INFL. I would prefer to claim that the main verb is the head of the sentence. This does not seem at all unreasonable. The verb dictates the range of complements, both subject and objects e.g. frighten requires an animate object, while eat normally requires an animate subject. Further the verb dictates the choice of complementizer in complement sentences e.g. want takes the infinitive complementizer:

3.42. I want to leave.

I want you to leave.

There would appear to be evidence in favour of taking the main verb as the head of the sentence. However, we would expect the main verb to be the head of the VP. This is problematical because the head is supposed to be characteristic of the construction. Consequently, the main verb cannot be the head of both constructions. We have already discussed the status of VP (see 1.4.2) and decided that it was possible to claim that it is not to be accorded any non-derived status within the grammar. It is also possible to claim that VP and S are basically the same construction, with one being simply more inclusive than the other and also differing in the direction of modification. S would be a left and right modified V and VP a right modified V. This may be captured by including a precedence limitation in our set of rules.

Taking the main verb to be the head of the sentence and eliminating VP gives the following set of rules for our example above:



auxiliary always precedes the main verb. In languages with dominant order SOV an inflected auxiliary always follows the main verb.

This universal poses problems for the principle of natural serialization. If auxiliaries are taken as modifying main verbs (as e.g. by Matthews (1981)) then this universal appears to be in conflict with the principle of natural serialization. Compare it with the following universals:

Universal 7. If in a language with dominant SOV order, there is no alternative basic order, or only OSV as the alternative then all adverbial modifiers of the verb likewise precede the verb.

Universal 13. If the nominal object always precedes the verb, then verb forms subordinate to the main verb also precede it.

Universal 7 suggests that verbal modifiers in SOV languages precede the verb, and Universal 13 suggests that a subordinate verbal form precedes the verb it modifies. Both these universals are in conflict with 16. Clearly SOV languages prefer to place their heads to the right (and VO languages prefer their heads to the left) if 7 and 13 are to be followed. Only auxiliaries seem to contradict this claim - Universal 16. Clearly something is amiss. (We have been arguing throughout this chapter that auxiliaries are, in fact, main verbs.)

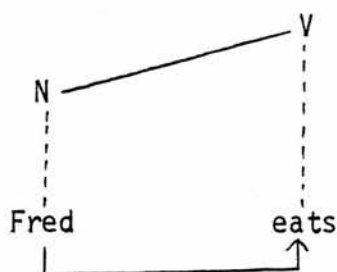
It would perhaps be appropriate to suggest that this apparent unnatural serialization might be resolved if auxiliaries were to be regarded as main verbs. Then Universal 16 could be abandoned and

replaced or subsumed under 13. We have already used distributional arguments to suggest this but let us now consider the role of the auxiliary within a dependency grammar to see if it provides evidence in support of our view that auxiliaries are, in fact, main verbs.

### 3.6.3. Dependency Syntax and Concord

Let us first consider, in support of our claims, the notion of concord. Person-number concord in English is marked on the first verb within the verb group. This will be the left most auxiliary, if one is present. If no auxiliary is present then it will be the main verb. It appears that the concord rules involve the copying of a specification from the noun onto the predicate on which the noun involved is (most immediately) dependent. This is illustrated in the following diagram:

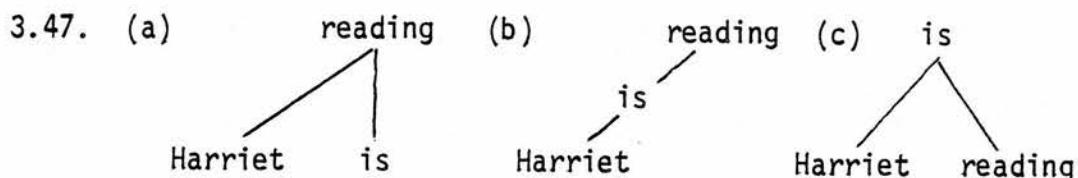
3.45.



In this example both the noun and the verb are inflected for the category of number. As noted above, in English, if an auxiliary verb precedes the main verb and is the first verb in the group then it is the auxiliary and not the main verb which is inflected for number e.g.:

3.46. Harriet is reading.

This sentence could be assigned any one of the following three graphs



If it is correct that concord is counter dependency, as in 3.45, then only (b) and (c) can be appropriate graph representations for 3.46 at the point when agreement is determined. (a) is ruled out because it involves concord between elements which do not enter into a dependency relationship.

In 2.2.1 and 2.3.2 we discussed the failure of the modals to undergo modern verb number agreement. Synchronically this is correct and may be used as an identifying characteristic of the class. However, historically the modals did not always fail to undergo number agreement. When English displayed a wider range of person-number inflections, the modals likewise displayed overt signs of having undergone concord e.g. some modals had the regular second person singular present indicative st ending, which they lost in the late sixteenth early seventeenth century when all verbs lost this inflection. Further, when a modal is the first verb in a sequence of verbs although it shows no sign of having undergone number agreement none of the following verb forms show any sign of having undergone agreement either. From this we may conclude that although the modals display no formal sign of having participated in verb number agreement they are to be accorded the same status as any other verb which appears in the initial position in a sequence of verbs. This failure to mark agreement overtly with a suffix is paralleled by certain nouns in English, which do not distinguish between singular and plural e.g. sheep, fish, and deer.

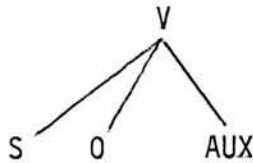
If concord is counter dependency then government, it would appear, follows dependency i.e. the head governs the form of the modifier. Lyons (1968:241) defines government:

"...under government the principle and the dependent member of a syntactic construction do not exhibit the same category : instead the dependent member is determined with respect to the relevant category (e.g. case) by the principle member."

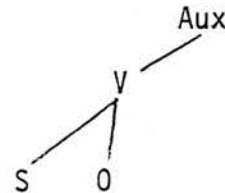
Clearly, in English, the form of the main verb is governed by the auxiliary - modals govern a bare infinitive, perfectives a past-participle and gerundives are governed by the progressive. It appears that the auxiliary and the main verb enter into a relationship where the choice of one dictates the form of the other. In English, in a sequence of verbs each verb, except the initial, is governed by the verb to its immediate left and in turn governs the verb to its immediate right, if there should be one. Therefore, we must conclude that the auxiliary governs the main verb. This supports the claim that modals, and all initial auxiliary verbs, are the head of the sentence in which they occur. Of the three graphs in 3.47 only (c) reflects this view.

If we now recall Universal 16, it is apparent that the choice of the first auxiliary as the head of the sentence would then be in agreement with natural serialization:

3.48. (a)



(b)



However, the defining characteristic of 'head' which stated that they

must be obligatory raises a problem for our claim that auxiliaries are the heads of sentences. The auxiliary is optional and therefore, fails to fulfill this requirement. This is not really a problem if we make, as we have been arguing on other grounds, all auxiliaries, including modals, members of the class of main verbs. Every sentence will now contain a main verb, which is the head of the sentence, followed by an optional series of dependent verbs. Further, the verb which is the head of the sentence is always finite in form. This means that in any group of verbs the one which takes a dependent subject and on which tense, mood and number distinctions are marked (and not all of the last three may be marked) will be the head of the sentence.

This applies only to finite sentences, which we shall take to be the unmarked case (non-finite sentences being the marked case). There are numerous reasons for adopting this position, some of which we discussed in 3.4. Finite sentences take a subject while non-finite ones do not require, and most often occur without, a subject (see Quirk and Greenbaum 1973:310). Moreover, the finite sentence may stand on its own as a simple sentence i.e. it is not dependent on any other construction. However, non-finite sentences are not independent and usually, in a complex sentence, they constitute the dependent or subordinate clause. These dependent clauses depend on the final verb of the verb group of the finite or superordinate clause. It is the final verb of this clause which governs the form of the complementizer, which in turn governs the form of the verb in the dependent or subordinate clause. We shall return to this below.

If the finite verb (tensed verb) is the head of the sentence then

this entails that the modals will always be the head of any sentence in which they appear. Modals, unlike other verbs, are, as we have discussed, restricted to the tensed position by lacking a non-finite sub-part to their paradigm. This results in the modals, within our dependency framework, being restricted to the head position. They will also fail to appear in any non-finite sentence. Further, the changes which have occurred, in so far as there have been any (recall that in 2.5 I claimed that the evidence for syntactic change in the modals was almost non-existent), restrict the modals to the head position: i.e. as the dependent verb forms are all non-finite, the modals, by lacking non-finite forms, are restricted to the head (tensed or finite) position. This restriction can be achieved by changes in the lexical specifications of the items concerned rather than by a change in the rules of the base. Before discussing the form these lexical entries might take and the changes which are claimed to have taken place, I shall outline the set of rules required to generate the base structures.

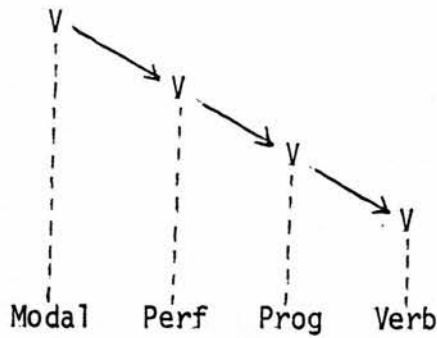
### 3.7. Recursion and Embedding

The set of rules for sentence structure based on the notation of dependency grammar is of the following form, as we discussed above:

- 3.49. R1 \* [V]  
 R2 V [N \* (N) (P)]  
 R3 N [(D) \*]  
 R4 P [\* N]

However, these rules require amending to generate the following sequence of verbs:

3.50.



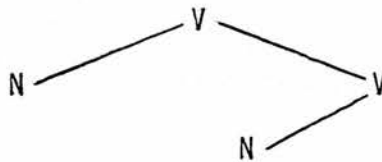
and to introduce the tensed verb as head of the construction.

To generate the verb sequence above in English, it is necessary to amend R2 in order that it may apply recursively to V. This would give:

3.51.  $R2' \quad V [N * \left( \left( \left( V \right) (N) (P) \right) \right) ]$

This rule now applies recursively. There is, however, a problem in that the rule can now create the following structures, which is, of course, ill formed.

3.52.



It appears that R2' requires further amendment. There are a number of ways that the grammar might be amended but first we shall revise R2' to prevent structures of the type illustrated above in 3.52 from being generated:

3.53.  $R2'' \quad V [* \left( \left( \left( V \right) (N) (P) \right) \right) ]$

We now require an additional rule to introduce the left dependent N:

3.54.  $R1a \quad V [N *]$

This now gives the following set of rules:

- 3.55.      R1    \* [V]  
              R2    V [N \*]  
              R3    V [ $\left\{ \left\{ \begin{matrix} V \\ (N) \end{matrix} \right\} (P) \right\}$  ]  
              R4    N [(D) \*]  
              R5    P [\* N]

These rules will form the base of our grammar but we require a further amendment in order to prevent structures of the type illustrated in 3.52 above being generated, due to the possible recursive application of R2 in 3.55.

We could consider extrinsically ordering the rules to ensure that R2 does not apply after R3. However, there are strong arguments against allowing extrinsic ordering in grammars (see, for example, Lass (1984) on the results of extrinsic ordering in phonology). By permitting extrinsic ordering we are relaxing the constraints on the grammar and going against our attempts to formulate a more constrained model, which does not allow abstract analyses and severely limits the range of possible grammars for any given set of data.

A second possibility would be to make R3 context dependent giving:

- 3.56.      V [ $\left\{ \left\{ \begin{matrix} V \\ (N) \end{matrix} \right\} (P) \right\}$  ] /X\_\_\_

This would achieve the desired effect. Context-free phrase structure grammars are, of course, inherently more restrictive than context-sensitive ones. However, to amend one rule, making it context dependent, is more constrained than, and therefore preferable to,

extrinsic ordering.

The rules in 3.55 together with the amended form of R3 - 3.56 above will now form the base of our grammar. We can also provide a rule based distinction between S and VP. A sentence is a construction with a left modified head - R1 and R2, whereas a VP is a construction with a right modified head - R3. These rules produce the following structures:

3.57.



I suggested above that the tensed verb should be taken as the head of the sentence. This is based on the grounds that it is obligatory, lexical, unitary (ignoring morphology), and characteristic. We must now distinguish this tensed verb from other verbs in any sequence of verbs. It would be possible to immediately introduce the tensed verb by amending R1 to

3.58. R1 \* [V<sub>t</sub>] (t = Tense)

However, this solution requires that the other rules be altered, and in some cases duplicated:

3.59. R1 \* [V<sub>t</sub>]  
 R2 V<sub>t</sub>[N \*]  
 R3 V<sub>t</sub>[\*  $\left\{ \begin{array}{l} (V) \\ (N) \end{array} \right\} (P) \} ] / X \_ \_ \_$   
 R4 V [\*  $\left\{ \begin{array}{l} (V) \\ (N) \end{array} \right\} (P) \} ] / X \_ \_ \_$   
 R5 N [(D) \*]  
 R6 P [\* N]

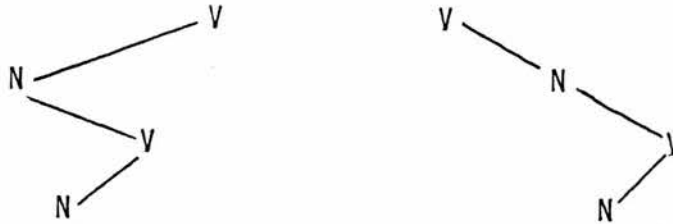
Another solution follows from the fact that the first  $V$  in any sequence, in a simple sentence, is always the tensed (or finite) one and all subsequent or dependent verbs are non-finite i.e. unmarked for person, number, and tense. We may now simply interpret the first verb in a sequence as being  $V_t$  (or  $V_{Finite}$ ).

So far we have been concentrating on simplex sentences. Let us now consider how we might handle complex sentences within our dependency model.

To judge by the literature (see for example; Bresnan (1970, 1971), Kiparsky and Kiparsky (1968), Lakoff (1968) and Rosenbaum (1967)), the treatment of complementizers and complement types has provoked much discussion. Therefore, I do not intend, in this section, to attempt an exhaustive investigation but rather to show that the dependency model is capable of generating these more complex structures. The discussion will be divided into three parts. The first will deal with relative clauses, the second with that complements and the third with infinitive complements.

In phrase structure grammars relative clauses, in English, are embedded within noun phrases. This translates quite easily into our dependency model. Traditionally, relative clauses modify the noun phrase in which they are embedded. In dependency terms the relative clause will be governed by the noun to its immediate left. This gives the following structures:

3.60.

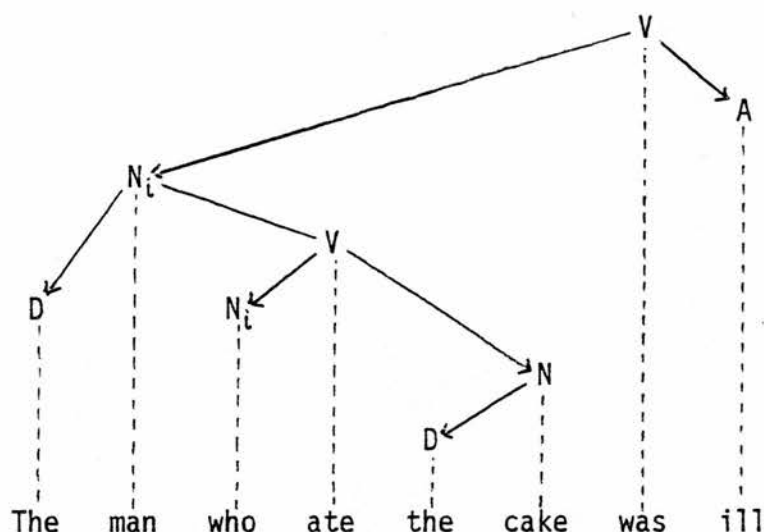


The relative clause will have the internal structure of a sentence leaving us to decide on the status of the relative pronoun. The relative pronoun must have the status of a noun. This follows quite simply from the fact that it represents a noun phrase - either subject or object. Further evidence for noun status arises from a study of the subject relative pronoun. As we discussed above, one of the requirements of a finite verb form is that it occurs with an overt subject, except in commands and ellipsis. Non-finite verb forms need not, and frequently do not, occur with a subject. (We shall return to this below.) The relative pronoun may, under certain circumstances be deleted (see Huddleston 1976:103-104 for a discussion). If the relative pronoun is deleted then the resultant subjectless clause must, like all subjectless clauses, be non-finite. This may be illustrated by the following examples:

- 3.61. (i) The man who lives upstairs is ill  
 (ii) The man living upstairs is ill.

In (i) the relative clause contains a subject who and is finite i.e. the verb lives is marked for number and is tensed. However, in (ii) the subject of the relative clause has been deleted and the now subjectless clause is non-finite. Within our dependency model relative clauses are given the following representation:

3.62.



We have not discussed the structure or treatment of that relatives, within our dependency model. I do not intend to discuss them as they raise issues and problems which fall outside the scope of our present discussion, for example a that relative may not be governed by a preposition in a relative clause:

3.63. The room  $\begin{pmatrix} \text{that} \\ \text{which} \end{pmatrix}$  I live in is small.

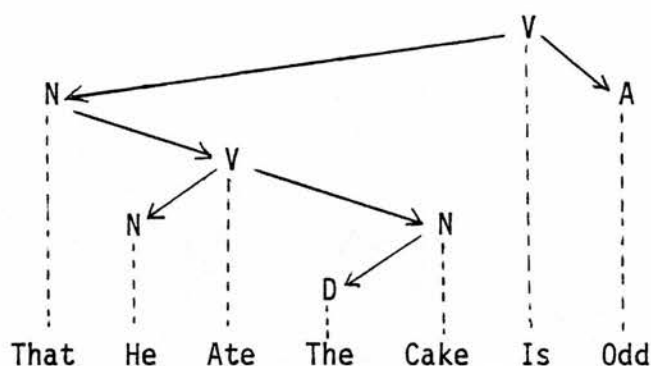
The room in  $\begin{pmatrix} * \text{that} \\ \text{which} \end{pmatrix}$  I live is small.

This suggests that the that relative behaves more like a complementizer than a relative pronoun.

Before amending our base rules to handle relative clauses let us look at the treatment of that complement clauses within our dependency model. Within generative theory that complements, like relative clauses, are embedded in noun phrases. Again I do not intend to dispute this view. The internal structure of that complements, like that of relative clauses, is related to a simple sentence. It has a subject, a finite verb and may contain the full range of verbal complements.

In dependency terms that complements would be assigned the following representation:

3.64.



Our present treatment of relative clauses and that complements raises at least one problem. The structures in 3.62 and 3.64 are similar in that they both contain a  $N \rightarrow V \rightarrow N$ ,  $N \langle N \langle N$ , sequence - man who ate in 3.62 and that he ate in 3.64. Consequently, we cannot structurally distinguish the relative clause from the that complement, which is problematical when we consider the behaviour of relative clauses and that complements in inversion:

- 3.65. (i) The man who ate the cake was ill.  
 (ii) Was the man who ate the cake ill?  
 (iii) That he ate the cake is odd.  
 (iv) \*Is that he ate the cake odd?

Clearly, that complements do not undergo inversion, yet, structurally we cannot distinguish between the two. One possible solution would be to introduce case nodes or thematic roles and make inversion sensitive to these (see Anderson 1977 for such an approach to this problem). Once again to pursue this would take us too far from our main purpose and therefore, I do not intend to further this discussion.

The Introduction of both relative clauses and that complements requires us to amend our dependency rules. The modification required is quite simple and involves only the noun re-write rule, as both complement structures are embedded within a noun phrase. The noun re-write rule is expanded to allow an optional right dependent verb. This gives:

3.66. R2 N [(D) \* (V)]

which will generate the structures in 3.62 and 3.64.

The treatment of the infinitive complement raises more problems than either the relative or that complement. Recent discussion of these complement structures (see for example; Brown and Miller (1982) and Schachter (1976)) have tended to treat some occurrences of them as prepositional phrases. The evidence for this position is varied but, nevertheless, a case can be made.

Firstly, not all occurrences of for-to go together as a complementizer. Consider the following sentence:

3.67. For the president to fail in the negotiations would  
be a disaster.

In this example for the president can be regarded as a prepositional phrase, independent of the infinitival phrase. This analysis is borne out by the possibility of the following example, where the prepositional phrase occurs at the end of the sentence:

3.68. To fail in the negotiations would be a disaster for  
the president.

However, there are instances where for-to can be treated as a complementizer. This is illustrated by the following example:

- 3.69. For the president to fail in the negotiations would be a disaster for the department.

In this example it is still possible to treat the initial for phrase as a prepositional phrase. In parsing 3.69 above for, a preposition, governs a noun, giving  $P \rightarrow N$ , which is the structure of a prepositional phrase. Also when the noun governed by for is a pronoun it is always in the accusative form, which suggests that it is a prepositional object, and consequently that the for phrase is a prepositional phrase. The evidence for the prepositional status of to is more convincing. Consider the following sentences:

- 3.70. (i) This led to him being released.  
 (ii) He knew what inspired him to paint.  
 (iii) They set off to climb the mountain.  
 (iv) He has grown to be quite troublesome.

If to were to be treated as a preposition in these sentences it would throw light on difficulties which arise in the discussion of pseudo-cleft constructions. This construction is of the form WHAT NP BE V NP: What Harriet likes is a novel by Thomas Hardy. This construction used to be used as a test for NP status. It was thought that only NP's could occur after BE. However, consider the following example:

- 3.71. What he is, is extremely foolish.

If this sentence is considered acceptable then we cannot use the

pseudo-cleft construction as a test for NP status alone unless instead of an AP, extremely foolish can be analysed as an NP. A more serious drawback for any linguist who wishes to treat infinitives as NP's is the failure of some infinitives to appear in pseudo-cleft constructions:

3.72. (i) The professor condescended to see the student.

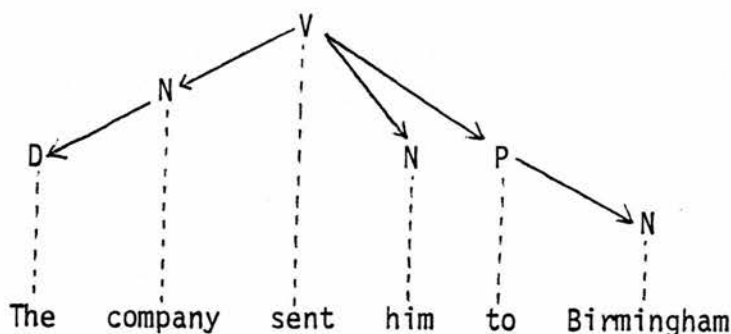
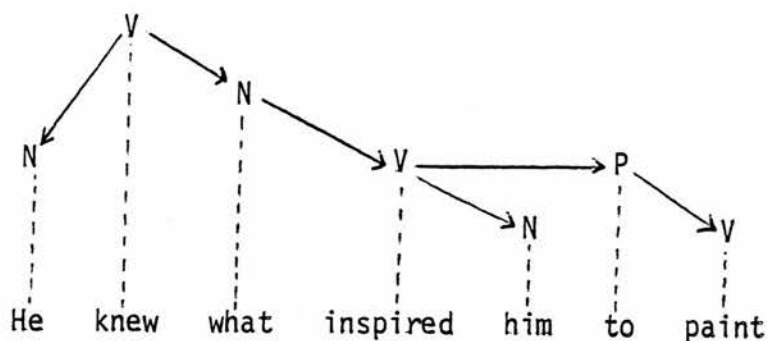
\*What the professor condescended to was to see the student.

(ii) The guard tends to be suspicious of new cleaners.

\*What the guard tends is to be suspicious of new cleaners.

If the ability to appear in pseudo-cleft construction is a test for NP and AP status then the failure of these infinitives to appear in such a construction suggests, strongly, that they are neither NP's nor AP's. The unacceptability of the pseudo-cleft examples above can be explained if to is treated as a preposition. If condescend and tend are taken as verbs of movement and analysed by analogy with, e.g., The discovery of gum tempted thousands of people to Coromandel, then the structure now proposed for the sentences in 3.70 is a verb followed by a prepositional phrase, except for 3.70(ii) which has a verb followed by a noun phrase and then by a prepositional phrase. The structure of 3.70(ii) is parallel to that assigned to The company sent him to Birmingham. These parallels are reflected in the following diagrams:

3.73.



This analysis of infinitive complements as prepositional phrase will be extended to all occurrences of such complements. In adopting such a view, we are extending the analysis presented in Brown and Miller (1981:154ff). However, we do succeed in eliminating the comp node from our grammar. We must now face the question of how we might amend our rules to allow for the generation of these structures.

We have already claimed there to be structural similarities between, for example:

3.74. The company sent him to Birmingham.

The company sent him to negotiate.

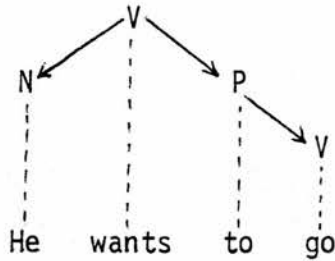
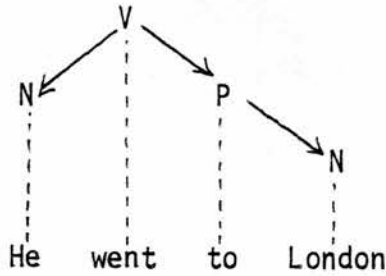
Both sentences contain a verb followed by a noun phrase which is in turn followed by a prepositional phrase. The difference between the two structures being that in the second one the preposition governs a verb while in the first it governs a noun. A simple amendment to the prepositional re-write rule will allow it to dominate either a

noun or a verb. This is illustrated below:

3.75.  $P [ * \begin{matrix} (N) \\ (V) \end{matrix} ]$

This together with our other rules allows the following structures to be generated:

3.76.



This brings us to object infinitive complements which contain a subject, e.g.

3.77. He appealed for witnesses to come forward.

These structures can be considered as parallel to, e.g.

3.78. He appealed for witnesses to the accident.

Both structures involve a verb followed by two prepositional phrases but in first the second prepositional phrase dominates a verb rather than a noun. Clearly we must first amend our previous R3 to allow a verb to be succeeded by two prepositional phrases. This gives:

3.79.

$$R3 \quad V \left[ * \left( \left\{ \left\{ \begin{array}{c} V \\ \{ \{ P \} \} \\ \{ \{ N \} \} \end{array} \right\} (P) \right\} \right) \right] / X \underline{\quad}$$

The amendment to the prepositional re-write rule given above (3.75) will now generate the required structures. However, these adjustments will also generate a number of ungrammatical structures, e.g., it is possible to have both prepositions dominating verbs or the first one dominating a verb while the second dominates a noun. It is, however, easy to impose a general constraint on the grammar to prevent these structures from being generated. If a verb governs two prepositional phrases but there is no direct dependency relationship between these two prepositional phrases (neither governs the other nor is there a relationship of mutual dependency) then if the second governs a verb the first must govern a noun, and if the second governs a noun then so must the first. This can be stated formally using dependency and precedence conditions:

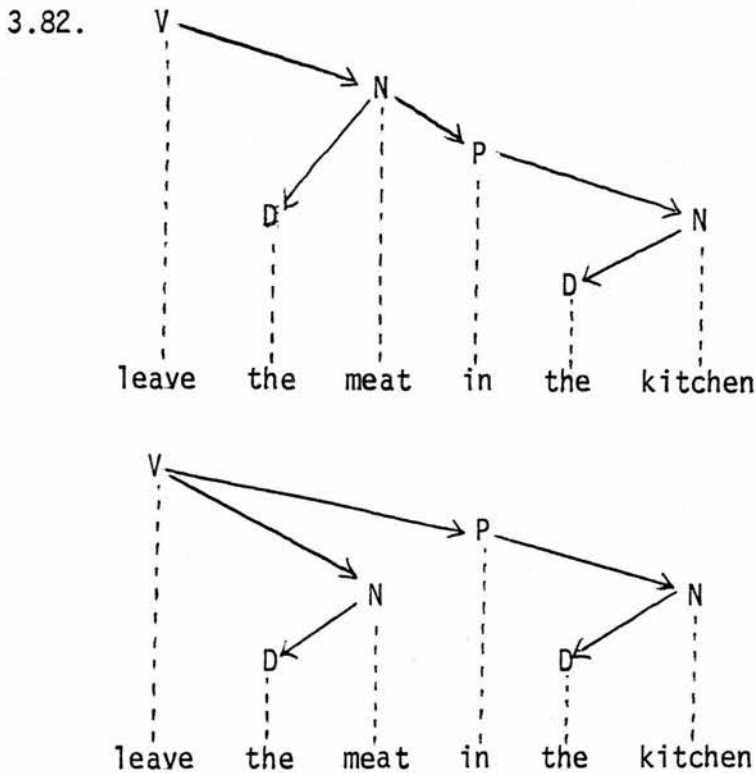
3.80. If  $V_n \dot{\neq} P_i$  and  $V_n \dot{\neq} P_j$ ,  $V_n < P_i < P_j$  then if  $P_j \dot{\neq} V_x$   
 $P_i \dot{\neq} N$ , or if  $P_j \dot{\neq} N$   $P_i \dot{\neq} N$

This leaves us to deal with subject infinitive complements. These will be embedded under a noun which is dominated by the finite verb - recall that finite verb forms require a subject. This requires us to amend the noun re-write rule to allow it to take a dependent preposition. Such an adjustment is not wholly unmotivated. Consider the following example:

3.81. Leave the meat in the kitchen.

Matthews (1981:80-81) suggests that this structure is ambiguous between two possible readings. One reading may be paraphrased by

saying that there is some meat in the kitchen and this meat is to be left untouched. The second possibility is that there is some unspecified meat which is to be left in the kitchen as opposed to anywhere else in the house. The ambiguity of the construction may be reflected in the dependency structure assigned to it:



The first of these two diagrams reflects the first interpretation while the second reflects the alternative interpretation. The first diagram requires a noun re-write rule which allows nouns to take right dependent prepositions. The amended preposition re-write rule now allows for both noun and verb modifiers.

It could be argued that the first of the above structures derives from an underlying relative clause, e.g., Leave the meat which is in the kitchen. However, there are a number of arguments against this position. Firstly, the hearer has access only to the surface struc-

ture and in these terms prepositional phrases are quite different from relative clauses. Secondly, Chomsky (1970:196) advances an argument against the underlying relative clause analysis based on the following data:

- 3.83. the weather in England.      the bottom of the barrel.  
       the weather in 1965.          a nation of shopkeepers.

Chomsky claims that in these, and other examples, it makes sense to regard the underlined form as the noun of a determiner-noun-complement construction, which constitutes a simple base noun phrase. The alternatives, according to Chomsky, would be to either regard the whole expression as a transform with the underlined element being a nominalized verb or adjective or to take the complement as a reduced relative. In some of these cases neither of these analyses seem particularly well motivated, especially given that the source relatives are all pretty peculiar. Therefore, it is certainly the case that our analysis is not without motivation.

This discussion of recursion is, as stated above, not exhaustive and is open to criticism and revision. However, it shows that recursion is not outwith the scope of our model but further discussion at this juncture would most likely take us too far from our main topic - a dependency account of the English verbal system.

With all these revisions, our grammar now contains the following rules:

- 3.84.      R1    \* [V]  
             R2    V [N \*]

$$R3 \quad V \left[ * \left( \left\{ \left\{ \begin{matrix} V \\ (P) \end{matrix} \right\} \right\} \right) \right] / X \underline{\quad}$$

$$R4 \quad N \left[ (D) * \left( \left\{ \begin{matrix} V \\ P \end{matrix} \right\} \right) \right]$$

$$R5 \quad P \left[ * \left\{ \begin{matrix} V \\ N \end{matrix} \right\} \right]$$

The third rule allows for infinite recursion, however, the maximum number of verbs in English which may occur in a sequence is four - modal, perfective, progressive, and passive or the main verb. If the rule generates anything beyond this then any empty nodes will be deleted. How the correct sequencing of verbs is achieved shall be discussed below. It may also be possible to eliminate the context in R3 by invoking a general syntactic constraint. Just as we can state that for English the maximum number of verbs in a sequence is four, we can also state that only the first verb in a sequence, i.e. the finite verb, can take a left dependent noun. This would rule out the following structure:

3.85.



Having proposed a rather sketchy fragment of the base of a dependency grammar we are left with the problems of lexical insertion, the form of the lexicon and morphology. Let us now turn to these problems.

### 3.7.1. Morphology

Recall that previously we decided that the syntactic, morphological, and lexical components are separate (see Ch.1). The syntactic component builds syntactic structure but does not process either deri-

vational or inflectional morphology. Consequently, there is no rule of affix hopping. All regular morphological processes form the morphological component. The lexicon stores the lexical items and morphological irregularities. It states the form classes of items, their strict subcategorization, and selection restrictions. It also lists all irregular forms which do not undergo regular morphological processes e.g., strong past tenses, umlaut plurals and irregular participles. Regular morphology is processed in the morphology component, which contains a number of rules, e.g.

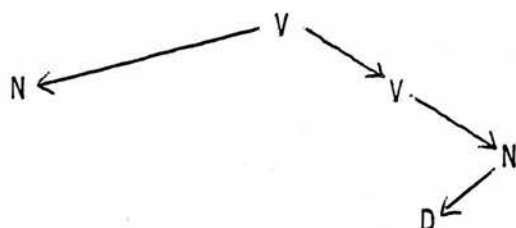
3.86.	Inflectional Rules	Derivational Rules
	R1 V $\Rightarrow$ V+ <u>ed</u>	V+ABLE $\Rightarrow$ Adj
	R2 V $\Rightarrow$ V+ <u>ing</u>	V+ER $\Rightarrow$ N
	R3 V $\Rightarrow$ V+ <u>s</u>	N+IC $\Rightarrow$ Adj
	R4 N $\Rightarrow$ N+ <u>s</u>	N+HOOD $\Rightarrow$ N
	R5 V $\Rightarrow$ V+ <u>en</u>	

These inflectional rules are backed up by a set of redundancy rules, e.g.

3.87.	$\begin{array}{l} \text{N} \\ [-\text{sg}] \end{array} \Rightarrow \text{R4}$
	$\begin{array}{l} \text{V} \\ [+ \text{prog}] \end{array} \Rightarrow \text{R2}$
	$\begin{bmatrix} \text{V} \\ -\text{past} \\ +\text{sg} \\ +3 \end{bmatrix} \Rightarrow \text{R3}$

Exceptions to these rules are stored in the lexicon. If no exceptional form is included in the lexical entry of an item then the appropriate rule applies. Consider, as an illustration, the following derivations:

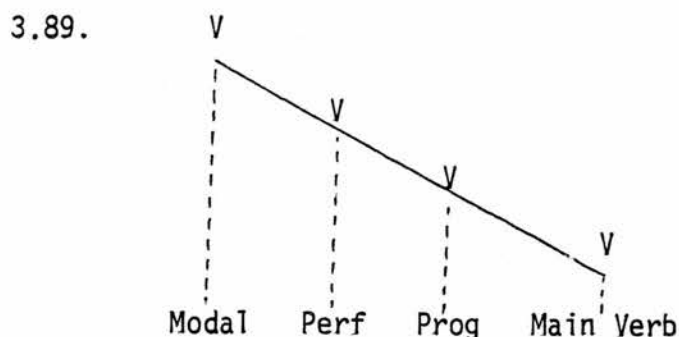
3.88.



First the head of the sentence is chosen, e.g. HAVE, which in this case will be the perfective auxiliary. Next the left dependent N, e.g. I, and then the right dependent V, the form of which is governed by the perfective auxiliary and, therefore, must be a perfective participle. The correct form of HAVE arises by having the person number features copied from the subject noun onto the first verb (head of the sentence) and then the correct form is sought in the lexicon. In this example the correct form is have, which governs its right dependent verb. This right dependent verb receives the feature [+perf] from its governor. Let us say that in this example CALL is the item in question. To obtain the correct morphological form CALL, [+perf] is looked up in the lexicon, where called will be found. If we had chosen, e.g., EAT then, in the lexical entry we would find [+perf] telling us that EAT has a perfective form. The redundancy rules would state that [+perf]  $\Rightarrow$  R5 in the morphology component. This rule would apply giving eaten, which is not included in the lexical entry.

### 3.7.2. Subcategorization

Let us now consider how we might subcategorize the modals, aspectual auxiliaries and main verbs to obtain the correct order:



and identify the tensed verb.

Auxiliaries are verbs which require an adjacent right dependent verb, whereas main verbs do not take a right adjacent dependent verb:

3.90.  $V \dot{\neq} V$  vs  $V \sim \dot{\neq} V$

Modals are members of the class of auxiliaries in that they require a right adjacent dependent verb but unlike the aspectual auxiliaries are restricted to finite positions. Perfective HAVE can be dependent on a modal, while progressive BE may be a modifier of either a modal or the perfective auxiliary. This distribution requires that modals be given a more restrictive subcategorization.

In English, as traditional grammars point out (see, e.g. Quirk and Greenbaum, 1973: §11.3), finiteness is linked to subjects - we have, in fact, discussed this at several points above 3.4. A finite clause always contains a subject except in commands and ellipsis. However, consider the following sentences:

3.91. It's John that Mary thinks will come.

Who does Mary think will come?

If modals are restricted to finite positions then will, in both these examples must be finite. However, neither occurrence of will is in a

clause with an overt subject. Further, neither sentence involves either ellipsis or imperative formation. The first involves extraposition and the second question formation. The requirement that a finite clause contain a subject applies to the initial dependency marker. Subsequent rules may either delete the subject or move it outwith the domain of the governor, in these instances the subject leaves a trace on the verb. This trace is not an abstract marker but an overt marker of finiteness. The subjectless verb will be marked for tense, person and number. This is illustrated by the following:

3.92. It's John that Mary thinks is coming.

It's the Smiths that Mary thinks are coming.

It was John that Mary thought was coming.

It was the Smiths that Mary thought were coming.

In these examples the progressive auxiliary receives tense and number distinctions from the extraposed subject before it is moved. These tense and number distinctions can be used in semantic interpretation to associate the extraposed subject with the correct verb in the sequence.

Non-finite clauses need not contain a subject and most frequently do not contain one. Modals, apart from those cases above, cannot appear in subjectless clauses. This allows us to make the following statements. Firstly the tensed verb is that verb which takes the adjacent left dependent noun, i.e.

3.93.  $N \neq V_i \supset V_i [+Tense]$



not intend this discussion of passive to be exhaustive. To judge by the literature (see for example: Awbery (1976), Baker (1978), Chomsky (1957), Freidin (1975), Keenan and Faltz (1978, 1979) and Wasow (1977)) the passive has provoked much debate. Any extensive discussion of the passive would, inevitably, lead us too far from our main issue. Therefore, it is sufficient to show that our model is capable of handling the construction.

There are strong arguments for dispensing with the passive transformation. It is a complex rule, which permutes constituents, like affix hopping, and introduces new constituent structure. In many transformational accounts of English the prepositional phrase introduced by the passive is the only one not introduced by the phrase structure rules.

A further defect of the passive rule is that it ignores the distribution of forms like eaten, to which traditional grammars assigned the label participle. This term claims that these forms are both verbs and adjectives. However, transformational accounts, as Brown and Miller (1982: 149) point out, treat participles as verbs, to which affixes are added.

The evidence for treating the passive participles as adjectives is quite convincing. Firstly many participles and adjectives are identical in form:

3.97. The paper was folded by John/the folded paper

The door was locked/the locked door

Secondly, if passive participles are treated as adjectives then the

passive construction would not be alone in consisting of the sequence BE A PP. This sequence occurs in sentences like:

3.98. He was reticent about my idea.

The plan was unintelligible to most of us.

However, the distribution of passive participles and adjectives does not co-incide completely. Some, but not all participles occur after SEEM:

3.99. He seemed ill.

The car seemed dented.

\*The record seems played.

Alternatively, not all adjectives occur in attributive position:

3.100. \*The asleep cat

and some do not appear in predicative position:

3.101. \*The series is major.

\*The chance is main.

A further problem is raised by verbs like melt, which have separate adjective and passive participle forms:

3.102. The wax was melted by the heat of the flames.

The molten wax trickled down the candle.

\*The wax was molten by the heat of the flames.

%The melted wax trickled down the candle.

Molten denotes a state, which results from a process, whereas melted relates to the process itself. The first example above describes a

process, which is one reason why molten does not appear in it. However, melted is coming to denote the resultant state:

3.103. The melted wax clung to the candle.

Molten, which was originally a passive participle, came to occur only in those frames available to adjectives denoting states. This close connection between adjectives and passive participles in the historical development of the language can be seen as supporting evidence for our analysis. The difference between molten and melted - lexical versus syntactic passive - is reflected not only in their derivations (the former being listed in the lexicon while the latter is derived via the morphology component) but also in the semantics of the two items. The semantics can also be used to capture the ambiguity of, e.g. surprised in He was surprised.

One final objection to this analysis is based on constructions like the following:

3.104. He was taught a lesson by his master.

He was promised limitless power by the generals.

While other NP BE A PP constructions occur in the language, there are no NP BE A NP PP constructions in the language, other than those illustrated above. However, these constructions do not pose problems for us. It is quite simple to ensure that only some of the forms we want to label adjectives occur in this construction. We can use the lexicon to account for the distribution of such adjectives. Exactly how this is to be achieved shall be discussed below.

A non-transformational account of passive is, in essence, forced upon

us by the nature of the model. We argued above that affix hopping should be disposed of because it was a syntactic rule which operated on the internal structure of word forms. In this it cut across the traditional distinction between syntax and morphology. To avoid this and to delimit the range of possible analyses, we argued that syntax and morphology ought to be separated out, with lexical items being inserted into the syntactic structure in their fully inflected form and if we were to retain a passive movement rule then we could not maintain this position. The morphology component would interact twice with the syntax - once in the generation of the active and then again in the derivation of the passive from the active. Having removed affix hopping for mixing syntax and morphology, it would be inconsistent if we were to permit passive to be retained, when it is guilty of the same offence. Therefore, we must generate the passive structures via our base rules and avoid the double application of morphology.

Having decided that the passive participle is an adjective, we must now decide how our rules are to be amended and what information must be included in the lexicon.

If passive participles are adjectives, then, as noted above, the maximum number of verbs which may occur in a sequence is still four. Consequently, the ordering constraint in 3.101 need not be revised. If V2 is the passive BE then it will govern an adjective and not a verb form.

In English most transitive verbs may appear in the passive construction. Given this it is quite easy to formulate a lexical redundancy rule to express this. This rule would state that a verb which

governs a right dependent noun, is related to an adjective which occurs in a structure where the right dependent noun is a left modifier of a verb, which is not identical to the original verb. Further, this adjective may be derived by a rule of the morphology component. This redundancy rule and any attendant morphological rule is called up by the feature [+pass(ive)], which is included in the lexical entry of the verb. This feature may seem somewhat ad-hoc but given that those transitive verbs which fail to appear in the passive seem to have nothing in common syntactically, although semantically we could relate them by introducing case roles or thematic relations which would be part of a case frame assigned to verbs and making passive sensitive to these case frames, it is justified. The lexical redundancy rule may be formalised as follows:

3.105. If  $V_i \rightarrow N_j$ ,  $V_i < N_j$  then  $V_i \rightarrow A_i$ ,  $V_k \rightarrow N_j$ ,  $N_j < V_k$   
           [+pass]  
            $i \neq k$

The means by which the adjective is derived from the verb would be indicated in the lexical entry. The [+pass] feature would either imply that a particular morphological rule applies, or, in the case of irregular passive participles the form would be listed in the entry with the feature passive implying that the form is adjectival. In those cases where the adjectival form is produced by a morphological rule the inclusion of the [+pass] feature in the lexical entry would imply that the output of the relevant morphological rule was adjectival. All that remains is to amend our base rules to allow an adjective to appear after a verb. The subcategorization frames will dictate which verbs adjectives may occur after. The amended form of R3, in 3.84 is:

3.106.

$$V \left[ * \left( \left\{ \begin{array}{c} V \\ \left\{ \begin{array}{c} N \\ P \\ A \end{array} \right\} \end{array} \right\} (P) \right) \right]$$

and the rule for A is

3.107. A [\* (N) (P)]

This rule allows an adjective to govern either a following noun or preposition or both, e.g.

3.108. He was taught by me.

He was taught a lesson.

He was taught a lesson by me.

In each of the above examples the noun and preposition do not enter into a dependency relationship directly but are dependent upon the adjective. Deletion of the adjective requires that the following modifiers be deleted.

This discussion, although not exhaustive, shows that our model is capable of handling passives. The above analysis is, of course, an initial solution to an issue which has provoked much debate, and as such is open to criticism and revision.

Before turning to the nature of the lexical entries let us look briefly at how our model might handle question formation, negative placement and account for the auxiliary behaviour of main verb have and be.

#### 3.7.4. Question Formation and Negative Placement

Question formation and negative placement may be treated together, as

both constructions display certain similarities. Consider the following examples:

- 3.109. (i) May we leave now?  
 (ii) Shall I shoot him?  
 (iii) Was he eating at the time?  
 (iv) Have you seen my chequebook?  
 (v) I can't cope with this.  
 (vi) I shalln't eat this.  
 (vii) He isn't eating the poison apple.  
 (viii) They haven't read the book.

All of these examples involve auxiliaries. The interrogatives can be derived by preposing the first (finite) verb to the front of the sentence. In the negative sentences the negative particle is attached to the right of the first (finite) verb. In both cases the same set of verbs is involved - the auxiliaries. Main verbs require do in both interrogatives and negatives:

- 3.110. Did they go to Belgium?  
 She didn't finish the wine.

The auxiliary do is restricted, like the modals, to finite positions. However, unlike the modals, do may only take a main verb to its right. Both question formation and negative placement operate on the first finite verb in a sequence of verbs. If this is the structure involved then we need not mention the term auxiliary in our formulation of the rules.

Let us now formalise the rules. First question formation:

3.111.  $V_i \not\rightarrow N \& V_i \not\rightarrow V_j \& N < V_i < V_j \Rightarrow V_i < N < N_j$

and now negative placement:

3.112.  $V_i \not\rightarrow N \& V_i \not\rightarrow V_j \& N < V_i < V_j \Rightarrow V_i \not\rightarrow \text{NEG} \& N < V_i < \text{NEG} < V_j$

Both these rules or alternative serializations are optional and the negative placement rule or serialization applies before lexical insertion. This allows us to keep syntax and lexical insertion and morphology apart. Further evidence for keeping the syntax and morphology of the negative constructions apart comes from a consideration of negative contraction - see examples (v)-(viii) in 3.109.

It has been argued (see for example Lapointe (1980)) that the contracted negative form n't is an inflectional suffix attached by morphological rules to finite auxiliaries. If we accept this, then negative contraction cannot be carried out by a transformational rule but will be attached by a morphological rule allowing the sequence of finite auxiliary plus n't to be inserted directly into deep structure.

The morphological rule which adds the contracted negative to the verb will be triggered by the feature [+neg] in the lexical entry. This feature allows the optional application of the following rule -  $V \Rightarrow V + \text{n't}$ . There are, of course, exceptions to this rule, e.g. won't, shan't in place of willn't and shalln't and mayn't (only in its epistemic sense) and amn't do not occur in standard N.E. These irregularities are listed in the lexical entry of the items concerned. Main verbs will not have the feature [+neg] in their lexical entries but will be [-neg]. This now allows us to distinguish between do and the other auxiliaries.



do, in turn, may be a modifier of a higher predicate. In this case do will be deleted and the higher predicate, e.g. have, will require that the dependent verb, originally a base form when dependent upon do, be a perfective participle. As do cannot be deleted until question formation and negative placement have applied, the surface form of the verbal modifier of do is not decided at the point of lexical insertion which occurs after the syntactic rules have applied with the lexical items being inserted into the syntactic structure in their fully inflected form. The consequence of this is that the syntax and morphology interact in exactly the fashion we have been at pains to avoid (see 3.7.3 above).

As with our discussions of complements and passive, this discussion, while recognizing that a problem exists, does not attempt to be exhaustive and is, at best, an initial solution to the problem.

Having discussed, if briefly the issues outlined in the initial paragraph of this section, let us now formulate the lexical entries.

### 3.7.5. Lexical Entries

We have outlined the syntactic and morphological components of our grammar, leaving only the lexical entries to be formulated. Consider the following lexical entries:

3.114.

CAN  
[+V]

OUGHT  
[+V]

N  $\Leftarrow$  \_\_\_\_\_  $\rightarrow$  V

N  $\Leftarrow$  \_\_\_\_\_  $\rightarrow$  P

[+NEG]  
[+PAST]  $\Rightarrow$  COULD

[+NEG]  
[+PAST]

The entry for can contains form class membership [+V] and strict

subcategorization,  $N \leftarrow \text{---} \rightarrow V$ , i.e. CAN must occur with a left dependent N and a right dependent V. Further, the dependent verb is in the base form, which is equivalent to the head of the lexical entry. The past tense form of CAN is COULD. The entry for OUGHT is similar to the for CAN except that it has no past tense and takes a dependent infinitive. The dependent infinitive is, following our discussion of complement structures above, a preposition followed by the base form of the verb. Given this, OUGHT takes a right dependent P and its lexical entry would specify that P governs a V and that P must be to. In fact a lexical redundancy rule would state that any P governing V must be to.

NEED and DARE require larger lexical entries to account for their dual status as main verbs and auxiliaries:

3.115.

NEED

[+V]

$$\left. \begin{array}{l} \text{---} \rightarrow N \\ \text{---} \rightarrow P \end{array} \right\} \Rightarrow R3 \left[ \begin{array}{l} +sg \\ +3 \\ -part \end{array} \right]$$

$$\text{---} \rightarrow V \Rightarrow \left\{ \begin{array}{l} [+NEG] \\ [+INTEROG] \end{array} \right\}$$

[+perf]  $\Rightarrow$  R1  
 [+past]  $\Rightarrow$  R1

This now gives us the basis of a synchronic grammar of English, which has advantages over A.S.W., Lapointe, P&W, and G.P.S.G. Both A.S.W. and Lapointe retain the AUX node, against which we have already advanced arguments. These arguments need not be repeated here. By generating auxiliaries in the same sentence as they appear in on the surface, we avoid the criticisms of Ross (1969) - see Lightfoot (1979). Our analysis also avoids the necessity of a 'hopping' rule to account

for constructions like would rather, where an auxiliary takes an object but otherwise behaves like an auxiliary in questions and negatives. In this case it would be lexically specified that would may take a sentential complement under specified circumstances. As with A.S.W.'s account, we have already discussed G.P.S.G. (see 1.4) and need not repeat the discussion here. However, it must be noted that our account both deals with and specifies lexical entries and morphology.

### 3.8. Diachronic Change in the Lexicon

Having illustrated a synchronic dependency grammar of English we must look at our model's attempt to capture the diachronic changes. We claimed above that, excepting word order changes, the rules of the base do not change. This imposes restrictions on the types of change which may occur. The time spread of the changes in the history of the modals, for which, as we saw in chapter two, there is only marginal evidence, suggests that the method of implementation of the change was lexical diffusion. This type of change proceeds via the lexicon. Let us now take an example to illustrate the changes and how the lexicon handles them. Consider, magan, N.E. may:

#### 3.116. Change

Loss of direct object.	1470
Loss of double modal.	1481
Loss of perfective.	1510
Loss of infinitive.	1533(-65)
Loss of progressive.	1556

Time scale 1470-1556(-65) 86(-95) years.

3.117.	Pre 1470	MAGAN	>	MAGAN	>	MAGAN	>	MAY
		[+V]		[+V]		[+V]		[+V]
		[-R3]		[-R3]		[-R3]		[-R3]
		— † N		— † V		— † V		— † V
		— † V				V † —		V † —
						[-perf]		[-perf]
						[-inf]		[-prog]
								[-inf]

The addition of the features [-perfective], [-progressive] and [-infinitive] exclude magan from all non-finite positions. It is eventually possible to make a generalization over these features, replacing them with  $N \not\approx \text{—} \not\approx V$  which is the distribution of all finite forms. This restricts may to this position and excludes it from all non-finite positions. As may cannot be governed by any other verb it will never be in a position to be inserted into a slot which requires a perfective or progressive form. In this model it is the distribution or subcategorization frame which captures the generalization that the modals are restricted to non-finite positions.

We can see that what changes is not the rules of the base, but rather the nature of the lexical entry and, in particular, the strict subcategorization frames. In the case of the modals, they are gradually restricted in their distribution. The increasing restrictions on their distribution results in their unavailability for morphological processes in the morphology component. If a verb cannot, for example, be governed by BE then it will never appear in a progressive or present participle form. The exclusion of modals from gerunds is due to the requirement that modals occur with an unmarked subject.

From our discussion of the modals in chapter two, it is obvious that the lexical entries of these verbs would already have been complex.

None of the modals ever underwent number agreement, and only cunnan and magan had full lexical entries permitting them to occur in the total range of structures. Sculan would have been marked as not appearing in progressive, perfective, or infinitive constructions, i.e. it was restricted to non-finite positions. Mōt participated in no changes. Its lexical entry was already  $N \neq \_ \neq V, \sim[+past]$ . In fact the evidence surveyed in chapter two was such as to suggest that no syntactic changes ever occurred.

We must conclude that it is preferable to capture the behaviour of the N.E. modals through changes in the strict subcategorization of these items in the lexicon. This is in keeping with the lexical diffusion model and is supported by the data. Finally, we restrict the types of change which may occur. It is not the rules of the base of transformations which change but only the lexical entries - in this case the strict subcategorization. This is, in fact, a reformulation of R. Lakoff's thesis within a much more tightly constrained grammar.

CHAPTER FOUR

The Transparency Principle : A Critical Review

#### 4.0. Introduction

In the discussion of the data concerning the development of the N.E. modals several references were made to the Transparency Principle (henceforth T.P.) but no further explanation was offered. I now intend to critically discuss the T.P.

#### 4.1. The Transparency Principle

It is crucial that the ontological status of the T.P. within a theory of grammar, and its role in stimulating diachronic change be fully understood prior to critically evaluating it. Therefore, I intend to begin with a detailed exposition based on Lightfoot (1979:121-166). Such an exposition will, doubtless, raise many issues, some of which will be dealt with immediately and others given a more detailed discussion later.

The T.P. is conceived as part of a theory of grammar and not as a component of a theory of linguistic change. It requires that derivations be minimally complex and that initial, underlying structures be 'close' to their respective surface structures. It is considered to be autonomous but may be subsumed under a general perceptual principle.

A certain amount of derivational opacity is tolerated by the grammar and the limits on this are characterized by the T.P. Once the limit is reached, or exceeded, the principle intervenes and prompts a re-analysis to remove the offending opacity, i.e. it predicts only when change is necessary but not what form it will take. These re-analyses are characterized by some form of the catastrophe theory, i.e. according to Lightfoot (1979:122):

"The second part of the story is a sudden, cataclysmic, wholesale re-structuring of the grammar whereby the exceptionality is, in a sense, institutionalised and the derivational complexity is eliminated at a stroke."

However, this use of 'catastrophe theory' should be considered metaphorical rather than literal. The catastrophe theory is an attempt to reconcile gradualness and change. Woodcock and Davis (1980:42) define a catastrophe as being a discontinuous change that occurs when a system can have more than one stable state. It can be thought of as a flow of water from one channel to another. In this case the transition is discontinuous not because there are no intervening states but because none of them is stable. The passage from the initial state to the final is likely to be short in comparison to the time spent in the stable states.

The re-structuring which occurs is therapeutic and not prophylactic, i.e. the change solves only a local problem and may create opacity elsewhere in the grammar. This is the case in the development of the modals, the creation of which causes problems for question formation and negative placement.

The level of tolerable opacity is to be gauged by some form of markedness theory (Chomsky and Halle (1968) and Lightfoot (1979:§1.5)).

Various changes in the grammar will have an effect on the level of opacity and once this exceeds the level set by the T.P. re-analysis will occur. Lightfoot gives no indication of how the level is to be set and how changes are to be evaluated vis à vis opacity. Do all changes increase complexity and therefore, increase opacity, or might some not actually reduce complexity (i.e. other than those provoked

by the T.P.)? Do those changes which increase opacity do so to exactly the same degree or might it not be the case that some changes create more opacity than others? If this latter position is correct, then how do we decide the relative opacity-creating value of the changes?

These questions, I believe, arise from the fact that Lightfoot fails to formalize the T.P.:-

"I shall make no attempt to formalize the  
 Transparency Principle or to give a precise account  
of the permitted degree of derivational opacity."

Lightfoot (1979:344)

(Emphasis mine W.S.A.)

We shall return to the issues raised here below, but if Lightfoot were to be correct, then the T.P. would allow us to predict the point at which re-analysis occurs, but only if formalised.

As stated above, the changes triggered by this principle are therapeutic; however, further to this, these therapeutic changes occur only and always when necessary and not at random.

Finally, the T.P. is, in principle, according to Lightfoot, falsifiable within a broadly Popperian-Lakatosian framework (see Lakatos (1970) and Popper (1959, 1963, and 1973)), but the possibility of such falsification, to which we shall return, does not worry Lightfoot because the T.P. does not constitute the 'hard core' of his research programme (see Lakatos (1970:133)). I find this claim rather puzzling given that the T.P. is invoked to account for a number of changes discussed by Lightfoot in subsequent chapters, e.g. the introduction of the N.E. modals and quantifiers, loss of NP status of the to-infinitive, loss of the English impersonal constructions, development of the transformational passive and the development of

serial verbs in Kwa. Surely, falsification of the T.P. would pose immense problems for Lightfoot's 'explanation' of the above changes.

The T.P. is by no means an innovation but is another attempt to formulate a functional explanation of change (see Lightfoot (1979:129)). The proposal that marked or complex sections of the grammar are amenable to therapeutic change, in order to resolve dysfunction, is not novel. Many linguists have, in one form or another, proposed that language change is functional or teleological, i.e. goal directed. This functional aspect is one to which we shall return below.

The discussion of Lightfoot's position vis à vis the development of the N.E. modals illustrates the operation of the T.P. without actually discussing the form and function of the principle. Having already discussed the evidence for the development of the N.E. modals and presented arguments against Lightfoot's analysis, I do not intend to re-iterate those arguments here but to refer back to them and the data.

Having outlined the ontological status of the T.P., I shall now return to some of the issues raised in the above exposition and discuss them in greater detail.

#### 4.2. Derivational Complexity

As stated above, the T.P. requires that derivations be minimally complex and that initial, underlying structures be close to their respective surface structures. Therefore, the T.P. can be seen as a constraint on the distance between underlying and surface structures. If the underlying structures become less like their corresponding surface structures, the degree of opacity increases correspondingly and it becomes more difficult for a child to construct a grammar.

This gives an opportunity to make more precise the T.P. Let us say that the degree of difference between two structures is called relative opacity, so that given structures A and B, where  $A \neq B$  and A and B are related by B being derived from A. An example of this would be:

- 4.1. (i) John ate the cake.  
 (ii) The cake was eaten by John.  
 (iii) The cake was eaten by whom?  
 (iv) It was John who ate the cake.  
 (v) Who was it who ate the cake?

In these examples (i) is less opaque than (ii), i.e. (i) is closer to the underlying structure than (ii), which is, in turn, less opaque than (iii). The status of the T.P. is to reduce relative opacity, usually by reducing the difference between initial and surface structures.

The main problems for such a proposal are, in fact, exactly those raised above in the exposition of the T.P. In this case all we are measuring is the number of transformations which apply in any derivation. This is not at all illuminating, and should be compared to the derivational theory of complexity (Miller (1962)). Further, within the model of E.S.T., to which Lightfoot adheres, there are only two cyclic transformations; NP pre-posing and Wh-movement. Within the above proposal, we would have to talk about numbers of applications rather than the number of transformations which have applied. However, Lightfoot, in the development of the modals, shifts the locus of change from the transformational component, where earlier studies, e.g. King (1969) and Traugott (1972), had concentrated, to the phrase structure rules. Such a restricted transformational component leaves

little scope, if any, for a reduction in opacity.

Bennett (1979) argues that the development of the modals, emergence of the quantifiers and loss of the impersonal constructions fail to support Lightfoot's characterization of the T.P. He argues that the development of the pre-modals, i.e. their exceptional behaviour, does not lead to increased opacity in the grammar. Given this he claims that there is no justification for Lightfoot's view that the introduction of a new category - modal - leads to a reduction in opacity by simplifying the transformational component and bringing the initial structure closer to the surface structure. Bennett presents a number of arguments against a potential role in the evolution of the modals for the T.P. and it is these that we shall now review:

- (i) Loss of direct object. This is simply a change in the strict subcategorization of the pre-modals, i.e. +NP is simply lost from their lexical entries. This separates pre-modals from transitive verbs but nothing else. After all the same feature is lacking from the lexical entries of all intransitives.
- (ii) Lack of third person singular present tense affix. Again this, misnamed as a change by Lightfoot, can be handled by a lexical entry, e.g. 3sg pres MAY is may and not mays. Further, this can be predicted from class membership, i.e. it is a distinguishing mark of the class of preterite-presents. This isolates the pre-modals as a distinct sub-class but does not entail any complexity in the grammar. Consequently, it cannot be claimed that this feature increases the distance

between deep and surface structure.

- (iii) Past tense interpretation. The result of this change is that past time interpretation is blocked in the case of a pre-modal with a past tense affix. This may be handled by a lexical exception to a rule of semantic interpretation. Therefore, this cannot affect the relationship between deep and surface structure.
- (iv) Refusal to take the to-infinitive. If the to - infinitive was originally an NP (Lightfoot (1979:122) and Callaway (1913:3)) and the pre-modals lost the ability to take NP objects then the modals would not co-occur with the to-infinitive. (This point would be strengthened if it could be shown that the to-infinitive in its NP form never occurred with the pre-modals (see 2.3.7)). Again subcategorization in the lexicon would handle this, without increasing the opacity of the system.

None of these changes increases opacity. The complexity of the transformational component does not increase nor does the distance between deep and surface structure. Therefore, it is difficult to see how the development of the pre-modals and their re-analysis as modals supports the T.P. as formulated by Lightfoot. Re-analysis does not eliminate the features which cause the opacity, but regularizes them. The irregularities of the pre-modals are regularized by becoming the defining characteristics of the new class. Further, the re-analysis leads to (i) a new phrase structure rule, (ii) a new category and (iii) a re-analysis of question formation and negative

placement. Clearly the re-analysis creates opacity in other areas. This 'new' opacity might be greater than that which has been eliminated. However, this is not problematical for Lightfoot. He states (1979:123) that grammars practice therapy and not prophylaxis. That an opacity eliminating change should produce opacity elsewhere in the system does not, if therapy is the aim, falsify the T.P. However, this local therapeutic application is simply a blocking device (see Botha (1971)) to forestall criticism and falsification. We shall return to this issue later.

#### 4.3. Autonomy and Independence

The T.P. is, according to Lightfoot (1979:114, 239), part of a theory of grammar and not a component of a theory of change. As such it places restrictions on what constitutes a possible syntactic component. This is entirely in harmony with the autonomous view of syntax. This thesis requires that syntactic rules operate independently of considerations of meaning and use, and have no access to semantic information whatsoever. However, Lightfoot argues at other points (e.g. 1979:121) that the T.P. might be subsumed under more general perceptual principles, as they play a role in defining what constitutes a possible grammar. Further, Lightfoot (1979:150) quite explicitly ties together the theory of grammar and a set of general perceptual mechanisms, which define less highly valued grammars. This is surely paradoxical, for while he claims to be adopting a strict autonomy hypothesis he is also claiming that at some level the T.P. interacts with perceptual processes.

Romaine (1981a), discussing this point, claims that this conflict between the status of the T.P. as an independent metagrammatical

principle and its functional therapeutic role, whereby it renders languages more learnable by speakers through the elimination of opacity, renders the T.P. vacuous as an explanation of change. Further, if syntactic change is autonomous and meaning plays no role in change, then why should it be therapeutic and adaptive to the speakers needs? For Romaine the T.P. has no explanatory force.

#### 4.4. Lightfoot and History

As discussed above, syntactic re-analyses are characterized by some form of the catastrophe theory (see Lightfoot (1979:122)). The change is sudden and cataclysmic. However,

"Thus the possibilities for re-analyses are severely limited, but they are not limited formally except as dictated by the theory of grammar. Although there is an element of chance in re-analysis like that involving the English modals, the grammar's room for manoeuvre is limited. This becomes clear when one thinks of the task of each generation of language learners, who must hypothesize or 'abduce' a grammar permitting communication with their models and must do so within the limits which grammars must meet "

Lightfoot (1979:407)

From this it appears that although change is sudden, cataclysmic and eliminates derivational complexity at a stroke, communication between generations must be preserved. Again we have a paradoxical situation, i.e. change is sudden but, in order to preserve communication between generations, minimal. Further, although I have claimed that no re-analysis ever occurred, my expanded discussion of the data concerning the re-analysis of the modals in chapter two, shows that even if a change did occur then it was gradual. This is the formal nature of a catastrophe. The view of change as being gradual gains credence from a number of other studies, e.g. see Aitchison (1980), Bennett

(1981) and Bickerton (1977) on syntactic change and Labov et al (1972), Lass (1978), and Wang (1977) on phonological change. However, Lightfoot (1979:377) dismisses the gradualness issue as uninteresting. This is an extremely weak response to what is a strong counter-argument to his claim that change is sudden and cataclysmic. In his discussion of the development of the English passive construction he states (1979: 280) that the apparent gradual spread of some of the English passives does not follow from E.S.T. but is an artifact of restricted data. Here we have what must be another blocking device to neutralize apparent counter-examples to the claim that change is sudden and cataclysmic, rather than gradual. This tactic seems inexcusable, when some of the cataclysmic changes have not yet been carried out in many dialects of English. Consider, for example, Brown and Millar (1980), who discuss the occurrence of double modals in Scottish English. These constructions are conservative and not innovative (see Scur<sup>v</sup> (1968) and Grant and Dixon (1921)).

Given the nature of the data, we must reject Lightfoot's cataclysmic approach to change, and regard his attempts to block criticism of it as untenable and serving, ultimately, only to weaken his arguments.

#### 4.5. Therapeutic Change and the Individual

According to Lightfoot, language change is therapeutic and not prophylactic, (1979:123). As we have seen, this effectively forestalls the criticism that some changes create opacity within the grammar. Further, this claim of therapeutic change requires that we take speakers into account. This point is, in fact, explicitly stated by Lightfoot, who asks how smart language learners are and

what the limits are to the abstractions that they can postulate, (1979: 129). Once again, we are obviously not treating the language system as an autonomous object but actually taking the speaker's cognitive capacity into account. We are claiming that speakers have the ability to compare their grammar against other grammars to discover the limits of abstraction.

If the restructuring of a system has a psychological base and children, as language learners, are the main agents for change, then it necessarily follows that they have a lower tolerance of opacity than adults. This may be the case but no evidence has been forwarded to substantiate it. Further, such a position denies any major role in language change to adults. This must be compared with Sankoff and Brown (1976), who observed that in Tok Pisin the adult speakers in the pidgin-speaking community were the first to use a new relativization strategy (see also Labov et al (1972)).

Making the speaker's cognitive capacity explicitly the locus of change, while maintaining autonomy, is not only paradoxical but also problematical. Given the time scale of some of the changes, together with the failure of some dialects to implement a change which occurs in other dialects, it appears that different speakers, or groups of speakers, vary in the amount of opacity they are willing to tolerate. It appears that Lightfoot is claiming that while the locus of change is a language spoken by the collective members of a speech community, the motivation for change is described in terms of the psychology of individual speakers - see Andersen's notion of abductive change (1973, 1974), which Lightfoot (1979:148, 151) is in agreement with. This might be referred to as the 'individualist' approach (Lass (1980b:267).

O'Neill (1973), quoted in Lass (1980b), pre-supposes that:

- (i) A speech community is nothing more than the sum of its members, i.e. it has no emergent properties or properties any different from those of any individual speaker.
- (ii) Therefore, the locus of 'language' is the individual speaker; language is to be defined individualistically, as a property of his 'monadic' cognitive make up. This means that language does not exist in the collectivity in any mode different from that of its existence in the individual; more importantly, of its existence in any particular individual speaker.

Such principles underlie all functionalist theories, which is what the T.P. certainly is. The psychological substrate of these theories is individualist and proponents of them must assume the applicability of any given explanatory strategy for any arbitrary individual. Given a competence performance division, and the claim or fact that language resides wholly in the individual, then the 'competence' of all individuals is equal. It follows from this that any functional motivation, i.e. therapeutic change, will be experienced by all members of the community simultaneously and unidirectionally. However, the data discussed in chapter two show this not to be the case. Correspondingly, we could, as Lass (1980b:267) suggests, re-interpret the functional motivation so that it 'is binding on just those individuals who feel like being bound' but in doing so we render it vacuous as a predictive and explanatory principle. (I have not yet discussed the status of the T.P. vis a vis explanation and prediction

but Lightfoot (1979:114) claims that it both predicts and explains change.)

Further problems for the individual-based motivation for change arise when the nature of statistical variation in speech communities is taken into consideration, e.g. the problem of the group score for any given linguistic variable in a variationist study. It is known, in such studies, that sets of interacting conditions appear to define a range of permitted values for the index score of a particular variable. Specific individuals will never register exactly the same score, i.e. there will always be some variation but the average group score will remain constant. Individual based change must assume that the individual speaker knows how to adjust his score so that he knows when to produce the proper variant, which prevents him from producing the statistical pattern belonging to, e.g. an other age group. But how does the individual achieve this, and how does he manage to keep his score within the range, which maintains the group average? This last question is particularly pertinent as there appears to be little doubt that variability serves as a method of implementing change.

Bickerton (1971:461) formulates the problem by asking how can the typical variation patterns we observe come about unless something somewhere is counting environments and keeping a running score of percentages? (Emphasis in the original.) Any individualistic interpretation, he argues, leads to the following:

".....since the group figure is the crucial one, and since.....individual scores will vary round it, each individual must.....keep track, not merely of his own environments and percentages but also of those produced by all other members of the group; in other words speaker B must be continually saying to himself

things like: 'Good Lord! A's percentage of contraction in the environment +V +NP has fallen to 77%! I'll have to step mine up to - let's see A's production of this environment-type stands to mine in the ratio of 65:35 over the last 100 token occurrences, so I'd better compensate by stepping mine upto.....what? About 86%' And to crown it all, he must not only be able to perform all these highly sophisticated calculations - he must also (since the rules apply to 'single' as well as 'group' styles) somehow continue to do so even in the physical absence of all other group members!"

Is it possible to attribute such computational ability to all individuals? Bickerton doesn't think so and I agree. However, he states that 'something' must be adjusting individual behaviour in accordance with group norms, and if the mind of the individual is ruled out then we are left with some supra-individual entity, e.g. a group or collective mind (see Durkheim (1895)). On a more linguistic level we may not wish to argue for a group mind but to consider language as an ontologically real entity of a Popperian 'third world'.

Popper (1973) in discussing the evolution of life and the emergence of man and the development of civilization makes use of the notion not only of an objective world of material things - World 1 - and a subjective world of minds - World 2 - but also of a third world. This 'third world' is the world of objective structures which are the products, not always intentional, of the minds of living creatures. These objects, once produced, exist independently of the mind(s) that produced them. World 3 is the world of, e.g. art, ideas, language, science. These entities, once created, may pose and create their own autonomous problems. An example of this would be the decimal number system, which is certainly the product of the human mind but prime numbers are an inherent property of that system. Prime numbers are not created independently of the number system but are an

unintentional by-product of the number system, which we must discover.

This highly abstracted supra-individual 'third world' view of language comes very close, as Linell (1979:18) points out, to Chomsky's (1965: 3-9) view of competence. If we view language as a 'third world' phenomenon we could claim that change occurs cataclysmically in a speaker's competence and then filters gradually through to his performance. This would allow a resolution of the cataclysmic / gradual conflict (see 4.4).

However, the removal of language or the locus of change from the individual, and consequently, making it an autonomous entity does not lead to a resolution of the problems of the T.P. Instead of asking how much opacity do speakers tolerate, we ask how much opacity do languages tolerate? The problem of what we measure and how we measure it still remains unresolved.

If we admit therapeutic change then we must also admit the possibility of dysfunction, which poses problems for the T.P. Lightfoot (1979: 124) claims that not only is change therapeutic but that therapeutic changes occur only when necessary and not at random. A consequence of this is that all change must be in the direction of reducing opacity, i.e. increasing ease of learning. Therefore, there must have been a time when language was, in general, more difficult than it is now, e.g. O.E. must have been more difficult than M.E., which, in turn, must have been more difficult than E.N.E. Lightfoot (1979: 152) explicitly recognizes this problem, but not, explicitly, its consequences, when he states that it is not possible for N.E. to undergo all those changes 'backwards' and arrive back at O.E. Many of the O.E.-to-N.E. changes were a function of the T.P. Therefore,

to undergo these changes backwards would clearly entail massive increases in opacity. (Emphasis D.W.L.) However, he has not considered the development of the English vowel system from O.E. to M.E. (see Lass and Anderson (1975:60-61)).

The use of 'many' (many of the O.E.-to-M.E. changes) in the text could be used as a blocking device, i.e. not all changes were a function of the T.P. Therefore, some changes, not provoked by the T.P., could create opacity for later changes to mop up. This, together with the view that therapeutic change solves only local problems and may create opacity elsewhere, provides an escape route from this criticism.

A strict interpretation of functionalism forces us to conclude that O.E. became M.E. because O.E. was no longer suited to its speakers needs. Although a local interpretation of therapy, as noted above, stalls or blocks this criticism. Lass (1980a 84ff) states that there are many different causes for a language not to be spoken any more but no-one has ever provided conclusive evidence that maladaptation to the speaker's needs is one of them. The three following reasons for language death are usually advanced:

- (i) Externally conditioned replacement, e.g. conquest, political pressure or movement in prestige norms in bilingual communities.
- (ii) Through death of all its speakers.
- (iii) Through processes of internal change, by which it becomes 'another' language, e.g. O.E.-M.E.

Both (i) and (ii) can hardly be seen as death through dysfunction. There is nothing, nor can there be, intrinsic in, e.g. Welsh, which

has led to its decline and replacement by English. The third case of language death is also problematical in that it suggests that the entire history of language is adaptive. This implies that no language is ever either fit or adapted but continually attempting to do so and consequently in a state of flux. No-one has yet argued for this position, i.e. the only thing that provokes change is the attempt to fulfill goals. Further, 'unfit' languages are only ever identified post-hoc, i.e. they changed because they were unfit and because they were unfit they changed. This, unless independently justified, is circular.

Further, functional explanations endow all speakers of a language with pre-cognition. Speakers must be able, within functional explanations, to make comparisons between the present state of their language and any possible, but non-occurring state, and choose one or the other. Consequently, they can have global intuitions concerning the structure of their language. (A third world interpretation of language would avoid this problem but, as pointed out above, this raises its own problems.)

This is problematical for two reasons. Firstly, there is no evidence that speakers can compute, in advance, the morphosyntactic consequences of a change. Secondly, the mechanisms by which change proceeds (see Lass (1978)) are so gradual that the speaker cannot possibly know what is going on. Further, in phonology, under a Neogrammarian approach, the change would apply even in those cases where dysfunction would arise. The speakers would then, at a later date, reverse the change in those cases where dysfunction arose. However, there must have been a period when the speakers coped with the dys-

function. Again this would be avoided by a third world view of language.

Similar problems arise for functional explanations in syntax. The infinite corpus of sentences of a language are generated by a finite system, e.g. a transformational generative grammar. Changes in this finite system may have vast repercussions in the infinite corpus. The speaker would have to be able to compute the syntactic consequences of any change for his entire corpus and then decide whether or not the resulting state was less opaque than the present one. This problem could be avoided by claiming, as Lightfoot does, that therapy is involved with essentially local problems and ignores opacity created elsewhere in the system. The speaker could then compute the syntactic consequences of the change on a limited random sample of utterances and project this onto the non-finite corpus. However, this is also problematical. The random sample may not reveal the absolute consequences of any change, i.e. a small increase or decrease in opacity may have enormous consequences when projected onto the corpus. It would appear that the notion of pre-cognition must be abandoned.

#### 4.6. Necessary changes which fail

Lightfoot (1979;124) claims that the T.P. operates only when necessary and not at random. Therefore, changes occur only when necessary and not at random. This leads to a circular argument, which may be illustrated as follows:

- (i) Therapeutic re-analyses occur only when necessary and not at random.
- (ii) Therefore, because they do not occur randomly they

must be provoked by the T.P. (as such they are necessary) and only by the T.P.

(iii) The consequence of this is that all re-analyses are provoked by the T.P. and, given (i), all are necessary.

(iv) If all re-analyses are necessary all are provoked by the T.P.

It would appear that all re-analyses are necessary and therefore, are all provoked by the T.P., which operates only when necessary.

This circular argument is obviously bad for the T.P., but even worse are cases where the necessary conditions for the re-analysis to occur are met but it fails to occur. Romaine (1981b:285) refers to these cases of change not occurring where the T.P. predicts it should; or random change occurring where it is not and/or cannot be predicted by the T.P., as differential failure.

The evidence she advances of a change not occurring when predicted by the T.P. comes from a discussion of the English relative system, based on Romaine (1981a).

The change can be divided into two stages:

- 1(i) The demonstrative pronouns lost their inflections through a change which affected the case system of the entire language.
- (ii) The demonstrative pronouns underwent analogical levelling. The sibilant forms se, seo acquired fricatives in common with the rest of the paradigm; se → pe and seo → peo c.f. paet.

- (iii) The invariant þe was increasingly used as a marker of subordination. This development may have been due, in part, to the development of SVO word order in subordinate clauses which would no longer be distinguished by OV word order.
- (iv) The nominative singular demonstrative þe was now homophonous with the complementizer þe and at the same time was being extended to serve a new function as a definite article.
- 2 A catastrophic re-analysis occurs in the 15th-16th century with the Wh-relative pronouns being used as new relatives, appearing at first where the co-referential NP was in the oblique case or the object of a preposition.

Lightfoot claims that in many relative clauses it would be unclear whether þe was a nominative demonstrative article or a complementizer. This would cause difficulties in parsing and result in a lack of transparency. The changes in the relative system represent the surface manifestation of a single change in the abstract grammar, i.e. a reformulation of the Wh-movement rule. Between O.E. and M.E. there was a change in the possible (surface) constituent membership of the COMP node. In N.E. for, that and Wh-forms are mutually exclusive in the COMP node in any given clause. Until the end of E.N.E. COMP could contain two items: only a conjunction or a demonstrative pronoun in O.E.; a conjunction or a Wh-form in M.E.; and a conjunction or Wh-form in E.N.E.

There are two strategies operating here. One is the adoption of the

Wh-strategy of relative clause formation. The second is the introduction of constraints on deletion of relative pronouns. Lightfoot claims that the differential introduction of the Wh-relative pronouns into various syntactic positions is consistent with the T.P. He especially claims that the lateness of who can be explained by the fact that its environment, a subject NP, was the least ambiguous. A relative clause introduced by that and a deleted subject NP presents no parsing difficulties and less opacity.

These changes in the English relative system are problematical for the T.P. because there are dialects of English, e.g. modern Scots, which have never really integrated the Wh-strategy of relativization. Some varieties of modern spoken Scots allow the relativization of NP's in all syntactic positions by the complementizer that, e.g.

4.2. He awarded the prize to the girl that fainted.

He left the book that we needed at home.

He's the man that lives upstairs.

The camera that's on the table is mine.

The change in the relative system cannot be complete if certain dialects are either in the process of implementing it or have yet to implement it. Romaine argues, and I agree, that if opacity were a sufficient and necessary condition for change it ought to happen simultaneously in all dialects of English, rather than variably and gradually. It would appear that a decline in the inflectional marking of the determiner system is not a sufficient condition for change to occur. One point Lightfoot fails to discuss concerns the difference between the written and spoken language, and that many parsing difficulties and opacity, which are responsible for re-

analysis, are really only problems for written language.

However, Lightfoot uses another blocking device to defuse this failure by saying (1979:374) that the fact that change does not occur wherever the causal factor, X, is present or, conversely that the change may occur even in those cases where the causal factor, X, is not present, does not indicate that, X, is not the causal factor.

This strikes me as an extremely odd interpretation of causation, which certainly would not be applicable in any other field of research. Causation usually takes the form of the Hempel-Oppenheim (1948) or Deductive-Nomological schema. This schema is based on deductive inference and a set of laws. It is considered 'ideal' in the sense that a well-formed explanation has the form of a deduction, and is, in principle, equivalent to a prediction. A D-N schema consists of an explanandum - the observation which is to be explained - and an explanans - a conjunction of statements specifying certain antecedent conditions and a set of general laws, e.g.

4.3.	C1, C2,.....	
	L1, L2,.....	Explanans
	E	Explanandum

Given the conditions and the laws, the observed state of affairs must follow logically. This schema represents explanation as a special case of deductive inference. The explanandum follows from the explanans by modus ponens,

4.4.	$p \supset q$	$(C \ \& \ L) \supset E$
	$p$	$(C \ \& \ L)$
	$\therefore q$	$\therefore E$

In terms of change  $q$  occurs because  $p$  caused it, i.e. whenever we find  $q$   $p$  must have occurred. We must also stringently assume that  $p$  is a nomically sufficient, necessary and antecedent condition to  $q$ . More importantly, for this explanation type, it is the case that if it adheres to all appropriate conditions and has empirical content, it cannot be denied, i.e. it is necessarily the case that  $q$ . For many philosophers of science any discipline that cannot establish true D-N schemas is neither scientific nor explanatory.

Lightfoot claims that sometimes the T.P. operates and sometimes it does not, but it still predicts and explains change. He claims (1979:345) that exact predictions will only be possible when one provides an exact characterization of the tolerable level of opacity. This he (1979:344) simply fails to do, leaving it up to someone else.

#### 4.7. Explanation and Prediction

According to Lightfoot (1979:122-123), the T.P. both predicts and explains re-analyses. However, as we have seen above, the T.P. is incapable of predicting change, e.g. the failure of Scots to implement the Wh-relativization strategy.

For some linguists and philosophers of science (see e.g. Lass (1980a, b)) failure to predict an event is equated with failure to explain. Given that the T.P. fails to predict change, it also fails to explain it and can be said to do no more than describe what happened. The repercussions of this for the T.P. is that it cannot be a properly causal explanation and is, at best, a probabilistic explanation. This has obviously occurred to Lightfoot, for he concludes (1979:408),

"Testable predictions about future changes must be

treated sceptically, given the extensive role of chance. The best one can do is to predict where a re-analysis will be required by the theory of grammar and to specify the available options for the form that the re-analysis might take. Which option will turn out to be adopted must in many instances remain a matter of chance. However, this view of change has two merits: it contributes to our understanding of the internal history of a given language by explaining (and often discovering) the simultaneity of superficially unrelated changes....."

(Emphasis mine. W.S.A.)

It would appear, from this, that Lightfoot accepts that the strongest form of explanation and prediction, i.e. the D-N schema, is outwith the scope of the T.P. The T.P. predicts only when re-analysis is required but not what form that re-analysis will take. The only conditions imposed on the form of change are those required by the theory of grammar. However, the T.P. fails in even this weaker form. As we saw above in the discussion of the Wh-relativization strategy, the T.P. fails to predict both when re-analysis is required and what form it will take.

Allotting a role to chance raises further problems for the T.P. In true causal explanations chance plays no role. Further, we cannot possibly explain anything if chance is given a free hand. There is no way of separating out the influence of chance from those factors which may be predictable but which we do not have the knowledge to locate or explain.

In the case of the T.P. prediction and explanation can only be seen as post hoc. This is something Lightfoot concedes (1979:123), "As it stands so far, this 'explanation' is after-the-fact....." There is nothing to suggest that it can ever be anything other than post

hoc.

#### 4.8. Markedness

The T.P. assumes that some form of markedness theory be adopted (see Chomsky and Halle (1968, Ch.9) to allow judgement to be made on when grammars become more marked and less highly valued or vice-versa. Further, it is not simply required to make macro-pronouncements, i.e. markedness values of the grammar, but also to make micro-pronouncements, i.e. how much any particular change contributes to the overall markedness of the grammar. This assumes, correctly I believe, that not all changes contribute equally to the opacity or transparency of the grammar.

The theory of markedness has already been criticized in, e.g. Lass and Anderson (1975) and Lass (1980a). Lightfoot replies to the first of these criticisms by claiming that if markedness is to make any empirically testable claims then those claims will presumably involve perceptual strategies and diachronic change. Despite his autonomous stance, he fails to realize that the first of these is unavailable to him as evidence. However, it must be the case that if opacity reduction is the goal of the T.P. then these perceptual strategies involving, for example parsing, must be of direct interest to his claim, but this would result in the loss of the autonomous stance he takes (at least in principle if not in practice, as we have seen in 4.3 and 4.5). He further claims that one cannot simply look at one segment or rule and then claim that it is not highly valued and, therefore, likely to drop out. This criticism can be interpreted as failure to take a holistic approach. However, Lightfoot himself is guilty of failing to take a holistic approach when he states that

grammars practice therapy and not prophylaxis. He also states that:

".....grammars did not originate in the Garden of Eden in a highly marked condition, being simplified inexorably ever since and gradually becoming less marked and more highly valued."

However, compare this with:

"For example, consider the changes taking place between O.E. and N.E. It is not possible, I would claim, for N.E.: to undergo all those changes 'backwards' and arrive back at O.E. Many of the O.E.-to-N.E. changes were a function of the Transparency Principle. Therefore, to undergo these changes backwards would clearly entail massive increases in opacity."

Lightfoot (1979:152)

It appears from the latter quotation that increases in opacity/markedness are not permitted and the course of language change is towards increasing value and decreasing opacity/markedness.

A form of markedness theory is required by the T.P. to enable it to distinguish between those changes which increase opacity and those which reduce or eliminate it, but at what point does the build up of opacity cease and the re-analysis occur? If the locus of change is an abstract grammar then observation is impossible. Lightfoot's only means of distinguishing is circular.

A singular change in an abstract grammar manifests itself in the simultaneity of a set of surface changes, which eliminate the opacity. However, the set of surface changes is taken as evidence for the single change in the abstract grammar. This is framework-dependent unless it can be externally verified through some concept such as markedness. Independent verification might ensure that the simultaneity of the

changes is not a result of the method of analysis. It might also ensure that the changes are actually present and not simply that, in retrospect, we perceive a link, which is not actually present. However, given that markedness is vacuous, this independent verification is not possible.

#### 4.9. Transparency and Falsificationism

The T.P. is (Lightfoot (1979:125)), in principle, falsifiable, but, hitherto, not actually falsified. Lightfoot claims to be working within a broadly Popperian-Lakatosian framework and, therefore, we would assume that falsificationism plays a major role in his work. However, as I have pointed out above, he sets up numerous blocking devices, the effect of which is to render falsification impossible. An initial minor point would be to note his use of the term 'broadly', which clearly allows a great deal of latitude in interpreting the Popperian-Lakatosian framework.

Lightfoot maintains that a highly restricted grammar of the E.S.T. type provides the best framework for characterizing and making predictions about historical change. Such characterization will vary depending on the grammatical model one adopts, e.g. the T.P., as presently conceived, can only reduce opacity within a derivational model - a trivial but necessarily true point. Lightfoot's views of change and the T.P. must be compatible with the machinery available within E.S.T. To accept his views incurs acceptance of a derivational theory of grammar. Further, to falsify Lightfoot we must show that a particular change cannot be adequately captured by his model of grammar (see Bennett (1979)). If it is not possible to

attack the T.P. from outwith the E.S.T. model, then the possibility of falsification is drastically reduced. We could always attempt to show that an alternative set of assumptions lead to better analyses but, given our lack of criteria for comparing grammatical models globally, this would not be a fruitful line of enquiry.

Lightfoot actually rejects approaches outwith his framework, e.g.

"Therefore, when one bears in mind the abductive nature of the acquisitional process, the concept of a diachronic universal (i.e. unrelated to the theory of grammar) becomes most implausible."

This is reinforced when he claims (1979:141, 343) that there are no formally definable limits to the ways in which two consecutive grammars of a given language may differ, beyond those limits imposed by a theory of grammar. Therefore, there are no grounds to expect a theory to delimit the notion of possible syntactic change on formal grounds.

What appears to be being claimed here, is that the T.P., as part of a theory of grammar, will tell us when change is required but not what form it will take. The form of the change is left to the theory of grammar. This is highly unsatisfactory because it blocks the possibility of claiming that changes consistent with the T.P. may be given viable accounts within alternative grammatical models.

Lightfoot's characterization of the T.P. raises problems when it comes to function as an evaluation matrix, permitting a choice between two competing grammars. If there are no formal limits to possible change then given two competing analyses, unless they are both within the same framework, all the T.P. will do is to dictate the selection

of the one which is compatible with it as the correct one. Lightfoot makes no comment on what would happen if the T.P. were to be faced with two competing analyses neither of which is compatible with it.

Although claiming, at several points (1979:74, 125, 140) to be working within a broadly Popperian-Lakatosian framework, he claims that if Popper's criteria for falsification are too rigorous for physics then it is unreasonable to apply them to less mature research programmes like linguistics. This prevents falsification on a global scale. It also seems, to me, rather paradoxical to claim to be operating within a particular framework which you then reject as being too stringent.

Another blocking device is the claim that (1979:76) the crucial factor (in science) is depth of explanation and not data coverage. This I would firmly reject, for unless we are fully accountable to the data then we run the risk of making totally vacuous claims. Surely we can only arrive at an explanation, on which we might make predictions, by covering all relevant data. Failure to do so will almost certainly lead to us making false claims and predictions.

This leaves us with only two possible methods of falsification, both of which require the acceptance of models of syntactic change inconsistent with E.S.T. and the T.P. These possibilities, which we have discussed above (4.4 and 4.6) are:

- (i) the gradualness of syntactic change
- (ii) differential failure, i.e. the conditions for change to occur are met, the T.P. predicts change but none occurs, or random change, not predicted

by the T.P. occurs.

The concept of gradual change is in conflict with Lightfoot's claim that change is sudden, wholesale and cataclysmic. We have already noted that in the case of the development of the English passive, he dismisses the gradualness issue as uninteresting and notes that the gradual spread of some of the English passives does not follow from E.S.T., i.e. it is inconsistent with it, but is probably an artifact of restricted data. This cannot be the case when we consider that gradualness or diffusion in syntactic change is not without support, e.g. the discussion of the development of the modals in chapter two, or the papers in Li (1977).

We have already seen that differential failure occurs in Scots in the case of the modals and the implementation of the Wh-relativization strategy. Lightfoot's response was to claim that if a change fails to occur even when causal factor, X, is present or occurs when causal factor, X, is not present this does not necessarily imply the X is not the causal factor. This blocking factor puts him in conflict with his aims and the framework within which he claims to be operating. He says:

".....hypotheses will be evaluated on the basis of the predictions they make, the extent to which those predictions are both correct and non-trivial."

Removal of the blocking device, discussed in 4.6 and outlined above, would render the T.P. falsified with respect to the Wh-relativization strategy in Scots, when tested against this statement.

Finally, I would like to claim that Lightfoot does not operate within even a broadly Popperian-Lakatosian framework. Popper believed that

one ought to be bold in conjectures but austere in refutation, and specify precisely under what conditions one is willing to give up ones position. Commitment, for Popper, is an outright crime.

Lightfoot fails to specify under what conditions he is willing to give up his position but implicitly recognizes what they are and blocks them whenever they occur. This too, however, is blocked from criticism, for while claiming to be operating within a broadly Popperian-Lakatosian framework he rejects that framework as being too stringent.

I think that the above discussion seriously undermines Lightfoot's claims concerning the T.P. and its ability to both explain and predict change. Further, there are many inconsistencies in his position. Most striking of these is his claim to be operating within the framework laid down by Popper and Lakatos, which he explicitly rejects, and his failure to maintain his autonomous stance. These, along with the other criticisms discussed above, render the T.P. vacuous.

BIBLIOGRAPHYAbbreviations

- AL Acta Linguistica Hafniensia
- CLS Papers from the Chicago Linguistic Society
- FL Foundations of Language
- FLH Folia Linguistica Historica
- IS Information Sciences
- JL Journal of Linguistics
- LA Linguistic Analysis
- Lg Language
- LIn Linguistic Inquiry
- LP Linguistics and Philosophy
- MT Machine Translation
- TPS Transactions of the Philological Society

- Abraham, W. (1978) Valence, Semantic Case and Grammatical Relations. Amsterdam: John Benjamins.
- Aitchison, J. (1980) 'Review of Lightfoot (1979)' Linguistics 18: 137-146.
- Aitken, A. J., A. McIntosh and H. Palsson (1971) Edinburgh Studies in English and Scots. London: Longman.
- Akmajian, A., S. Steele and T. Wasow (1979) 'The category AUX in universal grammar' LIn 10: 1-64.
- Andersen, H. (1973) 'Abductive and deductive change' Lg 49.4: 765-793.
- (1974) 'Towards a typology of change: bifurcating changes and binary relations' In Anderson and Jones (1974: 17-60).
- Anderson, J. M. (1970) The Grammar of Case. Cambridge: Cambridge University Press.
- (1971a) 'Some proposals concerning the modal verb in English' In Aitken et al (1971).
- (1971b) 'Dependency and grammatical functions' FL 7: 30-37.
- (1976) On Serialization in English Syntax. Ludwigsburg: Ludwigsburg Studies in Language and Linguistics.
- (1977) On Case Grammar: Prolegomena to a Theory of Grammatical Relations. London: Croom Helm.
- (1978) 'On the derivative status of grammatical relations' In Abraham (1978: 661-694)
- (1979a) 'On being without a subject' Indiana University Linguistics Club.
- (1979b) 'Subjecthood' Hungarian Studies in English 12: 121-138.
- (1979c) 'Serialization, dependency and the syntax of possessives in Moru' Studia Linguistica 33: 11-25.
- (1980) 'Anti-unaccusative, or: relational grammar is case grammar' Revue Roumaine de Linguistique.
- (1982) Language Form and Linguistic Variation. Amsterdam: John Benjamins.
- (1984) 'Objecthood' In Plank (1984: 29-54).
- (in press) 'Case grammar and the lexicon: concerning the irrelevance of grammatical relations to the expression of lexical regularities' New University of Ulster.

- Anderson, J. M. and C. Jones (1974) Historical Linguistics. Proceedings of the First International Conference on Historical Linguistics. Amsterdam: North Holland.
- Anderson, S. (1977) 'Comments on the paper by Wasow' In Culicover, Wasow and Akmajian (1977: 361-377).
- Aronoff, M. (1976) Word Formation in Generative Grammar. Cambridge, Mass: M.I.T. Press.
- (1980) 'Contextuals' In Hoekstra, Moorgat and van der Hulst (1980: 49-72).
- Awbery, G. (1976) The Syntax of Welsh: A Transformational Study of the Passive. Cambridge: Cambridge University Press.
- Bach, E. and R. T. Harms (1968) Universals in Linguistic Theory. New York: Holt, Rinehart and Winston.
- Bailey, Charles-James N. and Roger W. Shuy (1973) New Ways of Analysing Variation in English. Washington, D.C.: Georgetown University.
- Baker, C. L. (1978) Introduction to Generative Transformational Grammar. Englewood Cliffs: Prentice Hall.
- Bartsch, R. and T. Vennemann (1972) Semantic Structures. Frankfurt: Athenäum.
- Bennett, P. A. (1979) 'Observations on the transparency principle' Linguistics 17: 843-861.
- (1981) 'Is syntactic change gradual?' Glossa 15: 115-135.
- Bickerton, D. (1971) 'Inherent variability and variable rules' FL 7: 457-92.
- (1977) 'Change and variation in Hawaiian English vol.II: Creole syntax' University of Hawaii Social Sciences and Linguistics Institute.
- Bierwisch, M and K-H. Heidolph (1970) Progress in Linguistics. The Hague: Mouton.
- Bosworth, J. and T.N. Toller (1882-98) An Anglo Saxon Dictionary. Oxford: O.U.P.
- Botha, R. P. (1968) The Function of the Lexicon in Transformational Grammar. The Hague: Mouton.
- (1971) Methodological Aspects of Transformational Generative Phonology. The Hague: Mouton.
- Boyd, J. and J. P. Thorne (1969) 'The deep grammar of modal verbs' JL 5: 57-74.
- Brame, M. (1976) Conjectures and Refutations in Syntax and Semantics. New York: Elsevier North Holland.

- Brame, M. (1978) 'The base hypothesis and the spelling prohibition'  
LA 4: 1-30.
- Bresnan, J. (1970) 'On complementizers: toward a syntactic theory of  
complement types' FL 6: 297-321.
- (1972) Theory of Complementation in English Syntax.  
Doct. Diss. M.I.T.
- (1976) 'Evidence for an unbounded theory of transforma-  
tions' LA 2: 353-93.
- (1978) 'A realistic transformational grammar' In Halle,  
Bresnan and Miller (1978: 1-59).
- (1979) 'A theory of grammatical representation'  
Duplicated lecture notes, Department of Linguistics and  
Philosophy. M.I.T. Camb, Mass.
- (1982) 'Control and complementation' LIn 13: 343-434.
- Brown, E. K. and M. Millar (1980) 'Auxiliary verbs in Edinburgh  
speech' TPS: 81-133.
- Brown, E. K. and J. E. Miller (1980) Syntax: A Linguistic Intro-  
duction to Sentence Structure. London: Hutchinson.
- (1982) Syntax: Generative Grammar. London: Hutchinson.
- Brunner, K. (1963) An Outline of Middle English Grammar (translated  
by G. K. W. Johnston) Oxford: Basil Blackwell.
- Butters, R. R. (1973) 'Acceptability judgements for double modals in  
southern dialects' In Bailey and Shuy (1973: 276-286).
- Callaway, M. (1913) The Infinitive in Anglo-Saxon. Washington, D.C.  
Carnegie Institution.
- Cann, R. (1984) 'Bars and heads: reformulating X-bar syntax'  
Unpublished paper.
- Carlson, G. and T. Roeper (1980) 'Morphology and subcategorization:  
case and the unmarked complex verb' In Hoekstra, Moorgat and  
van der Hulst (1980: 123-164).
- Chen, M. (1972) 'The time dimension: contribution towards a theory  
of sound change' FL 8: 457-98.
- Chen, M. and W. S-Y. Wang (1975) 'Sound change: actuation and  
implementation' Lg 51: 255-281.
- Chomsky, N. (1957) Syntactic Structures. The Hague: Mouton.
- (1965) Aspects of the Theory of Syntax. Camb. Mass.:  
M.I.T. Press.
- (1970) 'Remarks on nominalization' In Jacobs and  
Rosenbaum (1970:184-221).

- Chomsky, N. (1972) Studies on Semantics in Generative Grammar.  
The Hague: Mouton.
- (1980) 'Markedness and core grammar' Unpublished paper.
- (1981) Lectures on Government and Binding. Dordrecht:  
Foris.
- Chomsky, N. and M. Halle (1968) Sound Pattern of English. New York:  
Harper Row.
- Christie, W. (1976) Current Progress in Historical Linguistics.  
Amsterdam: North Holland.
- Cole, P. and J. Sadock (1977) Grammatical Relations. (=Syntax and  
Semantics, vol 8). New York: Academic Press.
- Culicover, P. (1976) Syntax. New York: Academic Press.
- Culicover, P., T. Wasow and A. Akmajian (1977) Formal Syntax. New  
York: Academic Press.
- Curtius, G. (1880) The Greek Verb. London: Murray.
- Dingwall, W. O. (1971) A Survey of Linguistic Science. Maryland:  
University of Maryland Press.
- Dowty, D. (1978) 'Governed transformations as lexical rules in  
Montague Grammar' LIn 9: 393-427.
- Dressler, W. U., O. E. Pfeiffer and J. R. Rennison (1981) Phonologia  
1980.
- Durkheim, E. (1895) The Rules of Sociological Method. English ed  
London (1966): Collier-Macmillan.
- Emonds, J. (1976) A Transformational Approach to English Syntax.  
New York: Academic Press.
- Fillmore, C. J. (1968) 'The case for case' In Bach and Harms (1968:  
1-88).
- Fisiak, J. (1978) Recent Developments in Historical Phonology. The  
Hague: Mouton.
- Flasdeick, H. M. (1937) 'Das Verbum wollen in Altergermanischen (unter  
besonder Berücksichtigung des Altenglischen). Anglia 61:1-42.
- Freidin, R. (1974) 'Transformations and interpretive semantics' In  
Shuy and Bailey (1974).
- (1975) 'The analysis of passives' Lg 51: 384-405.
- Fries, C. C. and K. L. Pike (1949) 'Coexistent phonemic systems'  
Lg 25: 29-50.

- Gazdar, G. (1979) 'Constituent structures' Unpublished m.s. University of Sussex.
- Gazdar, G., G. K. Pullum and I. Saq (1981) 'Auxiliaries and related phenomenon in a restrictive theory of grammar' Indiana University Linguistics Club.
- Ginsburg, S. and B. Hall-Partee (1969) 'A mathematical model of transformational grammar' Information and Control 15: 297-334.
- Grant, W. and J. M. Dixon (1921) Manual of Modern Scots. Cambridge: Cambridge University Press.
- Greenberg, J. H. (1963) 'Some universals of grammar with particular reference to the order of meaningful elements' In Greenberg (1963: 58-90).
- (1963) Universals of Language. Camb., Mass.: M.I.T. Press.
- (1978) 'Diachrony, synchrony and language universals' In Greenberg (1978).
- (1978) Universals of Human Language vol I. Stanford: Stanford University Press.
- Grinder, J. T. and S. H. Elgin (1973) Guide to Transformational Grammar. New York: Holt Rinehart and Winston.
- Gruber, J. (1965) Studies in Lexical Relations. Doct. Diss. M.I.T.
- Halle, M. (1959) Sound Pattern of Russian. The Hague: Mouton.
- (1973) 'Prolegomena to a theory of word formation' LIn 4: 13-16.
- Halle, M., J. Bresnan and G. Miller (1978) Linguistic Theory and Psychological Reality. Camb., Mass.: M.I.T. Press.
- Harris, Z. S. (1957) 'Co-occurrence and transformation in linguistic structure' Lg 33: 283-340.
- Hays, D. (1964) 'Dependency grammar, a formalism and some observations' Lg 40: 511-525.
- Hemple, C. and P. Oppenheim (1948) 'Studies in the logic of explanation' Philosophy of Science 15: 133-175.
- Hinds, J. V. (1973) 'On the status of the VP node in Japanese' Language Research 9: 44-57.
- Hockett, C. F. (1954) 'Two models of grammatical description' Word 10: 210-234.
- (1958) A Course in Modern Linguistics. New York: Macmillan.

- Hoekstra, T., H. van der Hulst., and M. Moorgat (1980) 'Introduction'  
In Hoekstra, van der Hulst and Moorgat (1980: 1-48).
- (1980) Lexical Grammar. Dordrecht Foris.
- Huddleston, R. D. (1974) 'Further remarks on the analysis of  
Auxiliaries as main verbs' FL 11: 215-229.
- (1976) An Introduction to English Transformational Syntax.  
London: Longman.
- (1979) 'Would have become': empty or modal WILL?' JL 15:  
335-340.
- Jackendoff, R. S. (1972) Semantic Interpretation in Generative  
Grammar. Camb., Mass.: M.I.T. Press.
- (1975) 'Morphological and semantic regularities in the  
lexicon' Lg 51: 639-671.
- (1977) X-Syntax: A Study of Phrase Structure. Camb.,  
Mass.: M.I.T. Press.
- Jacobs, R. A. and P. S. Rosenbaum (1970) Readings in English  
Transformational Grammar. Waltham, Mass.: Ginn.
- Jespersen, O. (1949) A Modern English Grammar on Historical  
Principles. vols I-VII. London: Allen and Unwin.
- Johnson, D. E. (1976) 'Toward a theory of relationally based grammar'  
Indiana University Linguistics Club.
- (1977a) 'On relational constraints in grammars' In Cole  
and Sadock (1977: 151-178).
- (1977b) 'On Keenan's definition "subject of"' LIn 8:  
673-692.
- Katz, J. J. (1972) Semantic Theory. New York: Harper Row.
- Keenan, E. L. (1976) 'Toward a universal definition of "subject"' In  
Li (1976: 303-333).
- Keenan, E. L. and L. Faltz (1981) 'Logical types for natural  
language' U.C.L.A. Occasional Papers in Linguistics, No. 3.  
U.C.L.A. Department of Linguistics.
- Kimball, J. P. (1972) Syntax and Semantics I. New York: Seminar  
Press.
- (1973) Syntax and Semantics II. New York: Seminar Press.
- King, R. D. (1969) Historical Linguistics and Generative Grammar.  
Englewood Cliffs, N.J.: Prentice Hall.

- Kiparsky, P. (1971) 'Historical linguistics' In Dingwall (1971: 577-649).
- Kiparsky, P. and C. Kiparsky (1970) 'Fact' In Steinberg and Jacobovits (1971: 345-369).
- Kurath, H. and S. A. Kuhn (1956) Middle English Dictionary. Ann Arbor.
- Labov, W., M. Yaeger and R. Steiner (1972) A quantitative study of sound change in progress. Final report on N.S.F. Contract 3287, 2 vols. U.S. Regional Survey, Philadelphia.
- Lakatos, I. (1970) 'Falsification and the methodology of scientific research programmes' In Lakatos and Musgrave (1970: 91-196).
- Lakatos, I. and I. Musgrave (1970) Criticism and the Growth of Knowledge. Cambridge: Cambridge University Press.
- Lakoff, G. (1970) Irregularity in Syntax. New York: Holt, Rinehart and Winston.
- Lakoff, R. (1968) Abstract Syntax and Latin Complementation. Camb., Mass.: M.I.T. Press.
- Langendoen, T. D. (1975) 'Finite state parsing of phrase-structure languages and the status of re-adjustment rules in the grammar' LIn 6: 533-554.
- Lapointe, S. (1980) 'A lexical analysis of auxiliary verbs' In Hoekstra, Moorgat and van der Hulst (1980: 215-254).
- Lass, R. (1978) 'Mapping constraints in phonological reconstruction: on climbing down trees without falling out of them' In Fisiak (1978: 245-286).
- (1980) On Explaining Language Change. Cambridge: Cambridge University Press.
- (1981) 'Explaining sound change: the future of an illusion' In Dressler, Pfeiffer and Rennison (1981: 257-273).
- (1984) Phonology: An Introduction to Basic Concepts. Cambridge: Cambridge University Press.
- Lees, R. D. (1960) The Grammar of English Nominalizations. The Hague: Mouton.
- Lees, R. D. and E. Klima (1963) 'Rules for English pronominalizations' Lg 39: 17-29.
- Li, C. (1976) Subject and Topic. New York: Academic Press.
- (1977) Mechanisms of Syntactic Change. Austin: University of Texas Press.

- Lightfoot, D. W. (1974) 'The diachronic analysis of the English modals'  
In Anderson and Jones (1974: 219-250).
- (1976) 'The base component as a locus of syntactic change'  
In Christie (1970).
- (1979) Principles of Diachronic Syntax. Cambridge:  
Cambridge University Press.
- Lineell, P. (1979) Psychological Reality in Phonology. Cambridge:  
Cambridge University Press.
- Lipka, L. (1975) 'Rediscovery procedures in the lexicon' Lingua 36:  
197-224.
- Lyons, J. (1968) An Introduction to Theoretical Linguistics.  
Cambridge: Cambridge University Press.
- McCawley, J. D. (1968) 'Lexical insertion in a grammar without deep  
structure' CLS 4: 71-80.
- (1970) 'English as a VSO language' Lg 46: 286-299.
- (1975) 'The category status of the English modals' FL 12:  
597-6 .
- Malkiel, Y. (1967) 'Multiple vs simple causation in linguistic change'  
In To Honour Roman Jakobson: Essays on the Occasion of his  
Seventieth Birthday. The Hague: Mouton.
- Marchand, H. (1969) The Categories and Types of Present-Day English  
Word-Formation. 2nd ed. München: Beck.
- Matthews, P. H. (1965a) 'The inflectional component of a word-and-  
paradigm grammar' JL 1: 139-171.
- (1965b) 'Some concepts in word-and-paradigm morphology'  
FL 1: 268-289.
- (1966) 'A procedure for morphological encoding' MT 9:  
15-21.
- (1967) 'The main features of modern Greek verb inflection'  
FL 3: 262-284.
- (1972) Inflectional Morphology. Cambridge: Cambridge  
University Press.
- (1974) Morphology. Cambridge: Cambridge University Press.
- (1981) Syntax. Cambridge: Cambridge University Press.
- Miller, G. A. (1962) 'Some psychological studies of grammar' American  
Psychologist 17: 748-762.
- Mitchell, B. (1975) A Guide to Old English. Oxford: Blackwell.

- O'Neill, J. (1973) Modes of Individualism and Collectivism. London: Heinemann.
- Palmer, F. R. (1974) The English Verb. London: Longman.
- (1977) 'Modals and actuality' JL 13: 1-23.
- (1978) 'Past tense transportation: a reply' JL 14: 77-81.
- (1979) 'Modality and the English modals' London: Longman.
- Palmer, L. R. (1954) The Latin Language. London: Faber.
- Perkins, C. (1982) 'Core meanings of the modals' JL 18: 245-273.
- Perlmutter, D. M. and P. M. Postal (1977) 'Toward a universal characterization of passivization' Proceedings of the 3rd annual meeting of Berkeley Linguistics Society. 394-417.
- Peters, P. S. and R. W. Ritchie (1971) 'On restricting the base component of transformational grammars' Information and Control 18: 483-501.
- (1973) 'On the generative power of transformational grammars' Information Sciences 6.
- Pike, K. L. (1963) 'A syntactic paradigm' Lg 39: 216-230.
- Popper, K. R. (1959) The Logic of Scientific Discovery. London: Hutchinson.
- (1963) Conjectures and Refutations. London: Routledge and Kegan Paul.
- (1973) Objective Knowledge. Oxford: Clarendon
- Pullum, G. K. (1976) 'The Duke of York gambit' JL 12: 83-102.
- (1981) 'Evidence against the AUX node in Luiseno and English' LIn 12: 435-464.
- Pullum, G. K. and D. Wilson (1977) 'Autonomous syntax and the analysis of auxiliaries' Lg 46: 259-285.
- Quirk, R. and S. Greenbaum (1973) A University Grammar of English. London: Longman.
- Quirk, R. and J. Rusiecki (1982) 'Supplementing corpus elicitation' In Anderson (1982: 379-394).
- Quirk, R. and C. L. Wrenn (1955) An Old English Grammar. London: Methuen.
- Radford, A. (1981) Transformational Syntax: A Students Guide to Chomsky's Extended Standard Theory. Cambridge: Cambridge University Press.

- Robinson, J. J. (1967) 'On some alternative sets of dependency rules equivalent to the re-write of grammar II' Yorktown Heights: New York: I.M.B.
- (1970a) 'Dependency structures and transformational rules' Lg 46: 259-285.
- (1970b) 'Case category and configuration' JL 6: 57-80.
- Ross, J. R. (1969) 'Auxiliaries as main verbs' In Todd (1969).
- (1970) 'Gapping and the order of constituents' In Bierswisch and Heidolph (1970).
- Romaine, S. (1980) 'Syntactic complexity, relativization and stylistic levels in Middle Scots' FLH 2: 56-77.
- (1981) 'The transparency principle' Lingua 55: 277-300.
- Rosenbaum, P. (1967) The Grammar of English Predicate Complement Constructions. Camb., Mass.: M.I.T. Press.
- Sankoff, G. B. and P. Brown (1976) 'The origins of syntax in discourse: the case of Tok Pisin relatives' Lg 52: 631-666.
- Saunders, G. A. (1972) Equational Grammar. The Hague: Mouton.
- Schachter, P. (1976) 'A non-transformational account of gerundive nominals in English' LIn 7: 205-241.
- Schieber, S. M. (1984) 'Direct parsing of IP/LP grammars' LP 7.2: 135-154.
- Schwartz, A. (1972) 'The VP-constituent in SVO languages' In Kimball (1972).
- Šćur, G. S. (1968) 'On the non-finite forms of the verb can in Scottish' AL 11: 211-218.
- Shuy, R. and N. C.-J. Bailey (1974) Towards Tomorrow's Linguistics. Washington D.C.: Georgetown University Press.
- Siegel, D. (1973) 'Non-sources of unpassives' In Kimball (1973).
- (1974) Topics in English morphology. Doct. Diss. M.I.T.
- (1978) 'The adjacency constraint and the theory of morphology' NELS 8: 189-197.
- Steele, S. (1978) 'The category AUX as a language universal' In Greenberg (1978)
- Steele, S. et al (1981) An Encyclopedia of AUX. Camb. Mass.: M.I.T. Press.
- Steinberg, D. and L. Jacobovits (1971) Semantics: An Interdisciplinary Reader. Cambridge: Cambridge University Press.

- Stratmann, F. H. (1891) Middle English Dictionary. Oxford: Oxford University Press.
- Stockwell, R. P. (1976) 'Reply to Lightfoot (1976)' In Christie (1976).
- Todd, W. (1969) Studies in Philosophical Linguistics, Series 1. Evanston: Great Expectations.
- Traugott, E. (1972a) The History of English Syntax. New York: Holt, Rinehart and Winston.
- (1972b) 'On the notion of "restructuring" in historical syntax' In Proceedings of the XIth International Conference of Linguistics, Bologna.
- (1977) 'From space to time to logical relation: Evidence for a dynamic theory of language' Third International Conference on Historical Linguistics. Hamburg: August 1977.
- Vennemann, Theo (1973) 'Explanation in syntax' In Kimball (1973).
- Visser, F. Th. (1963-73) An Historical Syntax of English vols. I-IIIb. Leiden, Brill.
- Wang, W. S.-Y. (1977) The Lexicon in Phonological Change. The Hague: Mouton.
- Wasow, T. (1977) 'Transformations and the lexicon' In Culicover, Wasow and Akmajian (1977: 327-360).
- (1978) 'Remarks on processing, constraints, and the lexicon' Proceedings of the Second Theoretical Issues in Natural Language Processing Conference.
- (1980) 'Major and minor rules' In Hoekstra, Moorgat and van der Hulst (1980: 285-312).
- Williams, E. (1978) 'Notes on lexical theory' m.s.
- Woodcock, A. and M. Davis (1980) Catastrophe Theory. Harmondsworth: Pelican.
- Wright, J. (1910) Wright's Grammar of the Gothic Language. Oxford: Oxford University Press.