



# THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Deriving Null Functional Heads:  
A Study on Variation of Functional Structure

Aristeidis Palamaras



Doctor of Philosophy  
University of Edinburgh

2017



## Declaration

---

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Aristeidis Palamaras



## Abstract

---

The aim of this dissertation is to develop a framework of functional structure that ensures that the representation of any phrase is only as large as necessary to capture the syntactic relations relevant to it. I argue that the success of this project requires the elimination of null functional heads from the lexicon. Rather, I propose that null functional heads and their projections are dynamically created during the derivation as an extension of the projection of lexical items (i.e. lexical heads and overt functional heads). To this end, I make two proposals. Firstly, I argue that the featural specifications of lexical items are more extensive and have a more complex structure than previously thought. Secondly, I refine Giorgi and Pianesi's (1997) Feature Scattering operation so that it applies to entire segments of a featural specification, instead of individual features.

One beneficial implication of this formulation is that it reduces head movement to the incidental scattering of the phonological features of a head due to independent syntactic factors. Hence, I present an analysis of a number of cases of head movement in support of the proposed framework of functional structure. Amongst other things, I address V-to-v movement (as in the case of English main verbs), V-to-T movement (as in the case of Romance verbs and English auxiliaries) and V-to-C movement (as in the case of Germanic V2 and English residual V2). Additionally, I extend the analysis to cases of head movement to an initial position, including the movement of the verb in verb-initial clauses and the movement of the noun in noun-initial nominal phrases in Semitic and Celtic languages, which have previously received little attention in the strand of research that adopts Feature Scattering or other similar re-projective mechanisms.

Beyond head movement, I develop a uniform analysis of various subject/non-subject asymmetries, including Subject Auxiliary Inversion and *do*-support in English wh-questions and the *that*-trace effect in English embedded clauses involving

wh-extraction, with the intention to bring the relevant phenomena to bear on the overarching hypothesis that functional structure is variable.

## Lay Summary

---

The A fundamental observation in the field of syntax is that the variety and complexity of syntactic phenomena can not be accounted for solely with reference to the one-dimensional, linear order of words in a sentence. Thus, it is commonly assumed that sentences have an underlying two-dimensional, hierarchical structure. Since Chomsky (1986), theoretical developments within the tradition of generative grammar have led to the view that this underlying hierarchical structure has two components: one part of the structure accommodates words with lexical content (e.g. verbs, nouns, or adjectives) and another part accommodates words with functional content (e.g. auxiliaries like *have* and *be*, or articles like *the*), as well as other functional items (e.g. grammatical markers like the *-ed* verb ending that signifies past tense in English, or ‘null’ functional items that can be detected through their grammatical contribution to the sentence even though they are silent).

One question that has remained open since the 1980s is whether the functional component of this underlying structure is uniform across different languages and sentence types, or whether it is subject to cross- and intra-linguistic variation. The aim of this dissertation is to contribute to this theoretical debate by exploring the merits and demerits of the view that functional structure is variable. To do so, I develop a framework of functional structure that ensures that the underlying structure of any sentence is only as large as necessary to capture the syntactic phenomena evident in it. I argue that the success of this project relies on the rejection of the common assumption that null functional items, just like lexical and functional words, are part of our mental lexicon. Rather, I propose that the structural positions corresponding to those null items are an extension of the structural positions occupied by lexical and functional words. To flesh out this proposal, I make two assumptions. Firstly, I argue that the featural specification (i.e. the collection of features that determine the syntactic, semantic and phonological



properties) of lexical and functional words is more extensive and more structured than previously thought. Secondly, I refine Giorgi and Pianesi's (1997) Feature Scattering (i.e. a mechanism that allows the properties of lexical and functional words to spread across multiple structural positions) so that it applies to entire segments of a featural specification, instead of individual features. Under these assumptions, Feature Scattering becomes a recursive operation that can build a large number of structural positions (including those corresponding to null items) from a small number of lexical or functional words, as and when needed.

One beneficial implication of this approach is that it offers a new way of understanding head movement (i.e. a theoretical mechanism which accounts for the observation that the structural position in which a word is pronounced is not always the same as the position in which its syntactic or semantic properties are realised). Specifically, I argue that head movement boils down to the incidental scattering of the phonological features (i.e. the information that determines how a word is to be pronounced) of lexical and functional words due to independent syntactic factors. Hence, I present an analysis of a number of cases of head movement and the word orders they produce in support of the proposed framework of functional structure. Amongst other things, I address the sentential word order in English (where the position of the verb is best described in relative terms, as it appears after the subject and sentential adverbs), in Romance languages (where the verb appears after the subject but before sentential adverbs) and in Germanic languages with the exception of English (where the position of the verb is best described in absolute terms, as it appears second in the sentence, regardless of which element appears before it). Additionally, I extend the analysis to some word orders that have previously received little attention in the strand of research that adopts Feature Scattering or other similar mechanisms. These include the sentential word order in Semitic and Celtic languages (where the verb appears first in the sentence), as well as the word order within the nominal domain in the same languages (where the noun appears first within a nominal phrase that is comprised by articles, adjectives, numerals etc).

Beyond head movement, I address a group of syntactic phenomena collectively known as subject/non-subject asymmetries (i.e. situations where a given syntactic mechanism produces different outcomes depending on whether it involves a subject versus an object, adverb, or other modifier), with the intention to bring them to bear on the overarching hypothesis that functional structure is variable. To this end, I develop an analysis of the use of the auxiliary *do* in English interrogative sentences (where a form of *do* must be used in a simple present or past tense interrogative, but only if it enquires about the identity of a non-subject; cf. the interrogative ‘what does John play?’ enquiring about the object of the sentence with the interrogative ‘who plays chess?’ enquiring about the subject), the inversion of the subject and the finite auxiliary in the same environment (where the order of the subject and the auxiliary in an interrogative is the inverse of the order in a comparable declarative sentence, but only when it enquires about the identity of a non-subject; cf. the declarative ‘John is playing chess’ with the interrogatives ‘what is John playing?’ and ‘who is playing chess?’) and the use of the subordinator *that* in the formation of English subordinate sentences (where *that* is optionally used to introduce a subordinate sentence, unless the main sentence is an interrogative enquiring about the identity of the subject of the subordinate sentence, in which case the otherwise optional *that* must be omitted; cf. the grammatical ‘who did you say plays chess?’ with the ungrammatical ‘who did you say that plays chess?’).



## Acknowledgements

---

It's not the breadth or depth of knowledge. It's not even the hard work. It is the personal investment that makes a PhD a formative experience. And so these thanks are deeply personal.

First and foremost, my gratitude goes to my supervisor, Caroline Heycock. I have benefited equally from her incisiveness as a syntactician and her perceptiveness as an educator. I admire her ability to discern and articulate the potential of bright ideas, and her wisdom to do the same for the ones that are not quite there yet.

I would also like to thank my second supervisor, Peter Ackema, for his insightful comments on my work. Many a thing to explore, both theoretical and empirical, would have been lost to me without his input.

I am also thankful for the opportunities afforded to me by the department, school and university and all of their members. It has been a privilege to engage with academics on the top of their field, a pleasure to discuss with my fellow postgraduate students, an absolute joy to tutor the younger generation and a respite to have a coffee amongst stacks of books in the library.

This work was supported by an AHRC scholarship.



# Table of Contents

---

<b>Declaration</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Lay Summary</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>xi</b>
<b>Table of Contents</b>	<b>xiii</b>
<b>Chapter 1: Outline of the Theory</b>	<b>1</b>
1.1 Introduction	1
1.2 On the Study of Functional Structure: The Theoretical Background	6
1.2.1 The Origins of the Concept of Functional Structure	6
1.2.2 Decomposing the (Non-)Universality Question	8
1.2.3 The (Non-) Universality of Functional Structure and the (Non-)Atomicity of Functional Heads	15
1.3 A 'Variationist' Theory of Functional Structure: Giorgi and Pianesi (1997)	22
1.3.1 Syncretic Heads, Feature Scattering and the Universal Ordering Constraint	22
1.3.2 Deriving (Covert) Functional Structure	27
1.4 The Proposal: Feature Scattering and 'Extended' Featural Specifications	32
1.4.1 The 'Extended' Featural Specifications of Lexical Heads	32
1.4.2 The 'Extended' Featural Specification of Non-Derived Functional Heads	47
1.5 Conclusion	62
<b>Chapter 2: Head Movement. A Feature Scattering Analysis</b>	<b>65</b>
2.1 Introduction	65
2.2 Head Movement in Syntactic Theory	66
2.2.1 Head Movement versus Phrasal Movement	66
2.2.2 The Re-Projective Analysis of Head Movement	70
2.2.3 Deriving Head Movement from Feature Scattering	79

2.3 A Feature Scattering Analysis of Verb Movement	89
2.3.1 V-to-C, V-to-T and V-to-v	89
2.3.2 Movement of Finite versus Non-Finite Verbs	103
2.3.3 Movement in Matrix versus Embedded Clauses	115
2.4 Conclusion	130
<b>Chapter 3: Subject/Non-Subject Asymmetries</b>	<b>133</b>
3.1 Introduction	133
3.2 Subject/Non-Subject Asymmetries in English Wh-Questions	134
3.3 Subject/Non-Subject Asymmetries in the Distribution of Overt Complementizers	154
3.3.1 The <i>That</i> -Trace Effect	154
3.3.2 Amelioration and Anti- <i>That</i> -Trace Effects	162
3.4 Conclusion	169
<b>Chapter 4: Head Movement to First Position</b>	<b>173</b>
4.1 Introduction	173
4.2 Head Movement to First Position and Re-Projection	174
4.3 A Feature Scattering Analysis of Head Movement to First Position	188
4.3.1 Two Triggers for Feature Scattering	188
4.3.2 A Feature Scattering Analysis of Semitic and Celtic Nominals	191
4.3.3 A Feature Scattering Analysis of Verb-Initial Languages	218
4.4 Conclusion	226
<b>Chapter 5: Conclusion</b>	<b>229</b>
<b>Appendix</b>	<b>233</b>
<b>References</b>	<b>241</b>

### **1.1 Introduction**

This dissertation explores the proposition that functional structure displays cross- and intra-linguistic variation (Iatridou 1990, Ackema et al 1993, Thráinsson 1996, Bobaljik and Thráinsson 1998, Giorgi and Pianesi 1997, Nash and Rouveret 1997, Koenenman 2000, Bury 2003, amongst many others). While cross-linguistic variation in this area may be reasonably attributed to the hypothesis that the inventory of functional heads is language-specific, intra-linguistic variation suggests that functional structure is also constrained by narrow syntactic factors. This observation has led to the development of theories that attribute certain covert parts of the functional structure to the re-projection of functional heads (cf. Ackema et al 1993, Giorgi and Pianesi 1997, Nash and Rouveret 1997, Koenenman 2000, Bury 2003, for different versions of this concept). This dissertation proposes a more ambitious re-projective framework, according to which covert functional structure in its entirety is derived via re-projection.

This approach is supported by three observations. Firstly, the hypothesis that re-projection accounts for some, but not all, null functional heads implies a division between derived null functional heads that are created during the syntactic derivation and non-derived null functional heads that are drawn from the lexicon. Even if this division is sufficiently motivated by theory internal considerations, it is difficult to justify it from the perspective of language acquisition, considering that the learner is expected to discern between two competing analyses of null



functional heads that are practically indistinguishable. Secondly, the division between derived and non-derived null functional heads is questionable even from the narrower perspective of syntactic theory. Specifically, this division complicates the analysis of head movement, which must now allow for target heads with distinct properties. On the basis of these two observations, I conclude that it is profitable to subsume all null functional heads under a re-projective mechanism. Thirdly, Giorgi and Pianesi (1997) and Bury (2003) argue, on independent grounds, that re-projection is intrinsically linked with the creation of (non-vacuous) specifier positions. This hypothesis, in conjunction with the ancillary assumption that head movement and re-projection are closely related operations, provides an elegant analysis of the movement of a verb to C in Germanic V2 languages, to T in Romance SVO languages, or even to *v* in English. However, under this hypothesis, re-projection cannot account for the movement of a verb to an absolute clause-initial position, as in the case of Semitic or Celtic VSO languages. In theory, this limitation can be circumvented by the aforementioned division between derived and non-derived null functional heads (i.e. by assuming that the projection of non-derived heads is not constrained in the same way as the re-projection of derived heads). Nonetheless, I will argue that this solution is insufficient in at least one case, namely the movement of the noun in noun-initial nominal phrases in Semitic and Celtic languages. Thus, I will conclude that the link between re-projection and the creation of a specifier has to be weakened. By extension, this conclusion eliminates one of the few remaining obstacles for a uniform re-projective analysis of covert functional structure.

In order to eliminate non-derived null functional heads from the analysis of functional structure, I propose a new conception of the featural specifications of lexical items, as well as a new formulation of the mechanism of re-projection. With respect to the former, I propose that the specification of a lexical item consists of a sequence of functional features (e.g. *V*, *v*, *T*, *C*, etc), each one of which bears a set of un/interpretable sub-features (e.g.  $(u)\Phi$ ,  $(u)\text{Case}$ , *EPP*, etc). Furthermore, I propose that functional features and their sub-features have different functions. On one

hand, I assume that uninterpretable sub-features have to be eliminated by establishing a syntactic relation with an interpretable counterpart in an appropriate configuration (cf. Chomsky 1995). Thus, the un/interpretability of sub-features is the driving force behind syntactic operations. On the other hand, I assume that the purpose of the sequence of functional features is to determine the order of those syntactic operations (cf. Giorgi and Pianesi's 1997 Universal Ordering Constraint). With respect to the latter, I adopt a revised version of Giorgi and Pianesi's (1997) Feature Scattering. For the most part, I follow Giorgi and Pianesi in assuming that Feature Scattering is a last resort operation that takes place when the derivation fails to create an appropriate configuration for the elimination of some uninterpretable feature. In that event, Feature Scattering removes the offending feature from the specification of the relevant lexical item and re-merges it in the structure as a separate functional head. Thus, the uninterpretable feature has a second opportunity to be eliminated. However, unlike Giorgi and Pianesi (1997), I propose that Feature Scattering does not affect individual features, but entire segments of the featural specification of lexical items. With this refinement, Feature Scattering becomes a recursive operation that can derive multiple functional heads from a single lexical item. Thus, the interaction of these two simple proposals entails that even a small number of lexical items can capture complex, non-local syntactic relations. At the same time, the last resort nature of Feature Scattering entails that the functional structure of any given clause will only be as expansive as necessary to capture the syntactic relations that are relevant to it.

As I have mentioned above, the hypothesis that all null functional heads are derived via re-projection is motivated, in part, by observations relating to head movement. These same observations inform my analysis of head movement itself. Firstly, I challenge the view that re-projection is a subcase of a more general head movement operation that can target derived or non-derived functional heads (cf. Ackema et al 1993, Bury 2003). Rather, I propose that re-projection is a fundamental structure-building operation that, under the appropriate circumstances, derives head movement. Specifically, I suggest that the phonological

features of a lexical item are included in its featural specification alongside its syntactic features. Of course, this is not to say that phonological features participate in syntactic operations. However, when an independently triggered Feature Scattering operation targets the segment of the featural specification of a lexical item that happens to contain its phonological features, then those features will also be displaced from their original position. Thus, head movement boils down to the incidental scattering of the phonological features of a lexical item. Secondly, I challenge the view that re-projection, and by extension head movement, takes place solely for the purpose of creating (non-vacuous) specifier positions (cf. Giorgi and Pianesi 1997, Bury 2003). Instead, I propose that Feature Scattering may be triggered by two types of uninterpretable features that have to be eliminated under different configurations: (i) type-1 uninterpretable features have to c-command a matching interpretable feature, and (ii) type-2 uninterpretable features have to be c-commanded by a matching interpretable feature. Throughout this dissertation, I will demonstrate that this hypothesis captures a wide array of structures involving head movement. For instance, I will show that type-2 features are relevant to the analysis of verb movement to C in matrix (and some embedded) V2 clauses in Germanic languages, verb movement to T in Romance languages, as well as the movement of auxiliary verbs to T and main verbs to v in English. Moreover, I will show that type-1 features can account for verb movement in verb-initial clauses and noun movement in noun-initial nominal phrases in Semitic and Celtic languages.

Finally, I will discuss a variety of subject/non-subject asymmetries observed in English (i.e. Subject Auxiliary Inversion and *do*-support in *wh*-questions, the *that*-trace effect in embedded clauses involving *wh*-extraction and the anti-*that*-trace effect observed in non-*wh* relative clauses). I will demonstrate that the framework of functional structure developed in this dissertation provides a natural account of the unique ability of *wh*-words and relativized nominals that appear in subject position to affect Subject Auxiliary Inversion, as well as the distribution of the dummy auxiliary *do* and the overt complementizer *that*. To the extent that this

analysis is successful, it provides support for the overarching hypothesis that functional structure displays cross- and intra-linguistic variation.

The dissertation is organised as follows. In chapter 1, I outline the proposed framework of functional structure. Section 1.2 provides a preliminary discussion of the various concepts that enter into the study of functional structure. Section 1.3 focuses more narrowly on the re-projective literature on the matter. Special attention is given to Giorgi and Pianesi's (1997) contribution, which forms the basis of some of the proposals defended in this dissertation. In the same section, I outline the limitations of existing re-projective analyses and I conclude that the elimination of null non-derived functional heads is a step in the right direction. In section 1.4, I propose a new formulation of the concept of featural specifications and I refine Giorgi and Pianesi's (1997) Feature Scattering mechanism.

In chapter 2, I establish the proposed analysis of head movement. Section 2.2 provides an overview of the properties of head movement as opposed to phrasal movement, and demonstrates that re-projective analyses in general are well equipped to capture the intricacies of head movement. In the same section, I develop my theoretical account of head movement in terms of Feature Scattering. In section 2.3.1, I employ this proposal to the analysis of the V-to-C, V-to-T and V-to-v movement of main verbs in Germanic languages, Romance languages and English respectively. In section 2.3.2, I extend the analysis to the movement of auxiliary verbs in the same languages. Finally, in section 2.3.3, I explore the variation of embedded V2 patterns in various Germanic languages.

In chapter 3, I make an excursion to subject/non-subject asymmetries. Section 3.2 provides an analysis of the contrasting patterns of wh-movement, Subject Auxiliary Inversion and *do*-support in subject versus non-subject wh-questions in English. Section 3.3 extends the analysis to the *that*-trace effect and explores the complications that arise from the adverb amelioration effect and the anti-*that*-trace effect in English relative clauses.

Chapter 4 returns to the topic of head movement and focuses on verb movement to the initial position of the clause and noun movement to the initial

position of the nominal phrase. In section 4.2, I demonstrate that verb-initial clauses and, especially, noun-initial nominal phrases in Semitic and Celtic languages pose a challenge for re-projective theories of functional structure and head movement. In section 4.3.1, I propose a solution to this problem that is based on the hypothesis that Feature Scattering may be triggered by two types of features, which create different syntactic configurations. In section 4.3.2, I employ this hypothesis in the analysis of noun-initial nominal phrases in Semitic and Celtic languages. In section 4.3.3, I extend this analysis to verb-initial clauses in the same languages.

## ***1.2 On the Study of Functional Structure: The Theoretical Background***

### ***1.2.1 The Origins of the Concept of Functional Structure***

A commonplace assumption in contemporary theories of syntax within the generative tradition is that the structure of the clause consists of multiple layers of functional information (in the form of functional heads and their projections) that sit on top of a lexical head. This conception of the clause emerged soon after the advent of X-bar theory, nearly fifty years ago (Chomsky 1970, Jackendoff 1977). Needless to say that the X-bar understanding of clausal structure was, at least partly, based on pre-existing observations regarding the hierarchical nature of the clause. For example, Chomsky (1965) posited a constituent larger than the verb phrase (namely a predicate phrase that contains auxiliary verbs as well as the VP) but smaller than the root node of the clause S, and Bresnan (1970) argued for a constituent larger than S (namely an  $\bar{S}$  that contains the complementizer and S). What X-bar theory contributed to these early studies of the hierarchical organisation of the clause was the hypothesis that phrases must be headed. Initially, this treatment was reserved for the phrases of lexical heads, but eventually it was extended to functional heads. Accordingly, Chomsky (1986) reanalysed S as an IP (inflectional phrase), a projection of an inflectional head, and  $\bar{S}$  as a CP (complementizer phrase), a projection of a complementizer. Evidently, the resulting structure was already more articulated than any of its predecessors. However, the

significance of this analysis goes beyond the specifics. Chomsky (1986) effectively introduced the concept of the functional head as a building block of syntactic structures; an idea that proved to be instrumental for the systematic study of the composition and nature of clausal structure.

A sizeable part of the literature on the functional structure of the clause has historically dedicated its efforts to the identification of the components of that structure. One of the notable early works in that direction was Pollock's (1989) Split-Infl hypothesis (originally developed by Pollock and later refined in Belletti 1990 and Chomsky 1991, 1993; see also Ouhalla 1988, 1991, 1994 for a different take), which argued for a more articulated structure of the IP. Although the details of Pollock's (1989) analysis are by no means uncontroversial<sup>1</sup>, the Split-Infl hypothesis became a driving force for the relevant literature at the turn of the 1990s. Similar developments occurred in the analysis of the VP (i.e. the bipartite vP/VP of Chomsky 1995 and others, which grew out of Larson's 1988 treatment of the structure of ditransitive verbs as an instance of VP-recursion) and the CP (i.e. Rizzi 1997 and subsequent work). Another significant contribution was Cinque's (1999) treatment of adverbials as specifiers of dedicated functional heads. Additionally, Cinque (1999) marks the beginning of a project for the systematic study of the composition of functional structure cross-linguistically (i.e. the cartographic project; cf. the chapters in Cinque 2002, 2006 and Rizzi 2004, amongst many others). Finally, this strand of research has been extended beyond the confines of the functional structure of the clause, most notably in the domain of nominal structure (cf. Szabolcsi 1983, 1994, Abney 1987, Ritter 1991, Cinque 1994, amongst many others).

However, there is more to know about the properties of functional structure beyond its components. In this dissertation, I will focus on the question of whether functional structure is universal or language- (and structure-) specific. Of course, the interest in the universal properties of the syntactic system has been a main fixture

---

<sup>1</sup> Pollock (1989) posited separate functional heads hosting tense and agreement features, as well as a separate functional head expressing negation. While the Agr head is not widely accepted today, the Neg head is.

of syntactic research in the generative tradition in general and the principles and parameters era in particular. Specifically in relation to the (composition of) functional structure of the clause, the first explicit discussion of its (non-)universality can be found in Iatridou's (1990) critique of the Split-Infl hypothesis<sup>2</sup>. Since then, the increasingly detailed representations of functional structure that were emerging after Pollock (1989) were perceived by some as an area of parameterization (Ouhalla 1988, 1991, 1994, Ackema et al 1993, Thráinsson 1993, 1996, Bobaljik 1995, Bobaljik and Thráinsson 1998, to name a few), while others viewed them as a facet of the universal organization of the language system (Cinque 1999, and most of the work in the cartographic project). As we will see in the next section, functional structure is an especially multi-faceted component of syntactic theory. Thus, it should not be surprising that both, apparently polar opposite approaches have been defended in the literature.

I will discuss the issue of the (non-)universality of functional structure in more detail in the following sections. Specifically, in section 1.2.2 I will elaborate on the various components of functional structure that may be described as universal or variable, and in section 1.2.3 I will evaluate some of the main arguments in favor of the two competing views on the matter.

### ***1.2.2 Decomposing the (Non-)Universality Question***

In order to have a meaningful discussion about the universal or variable nature of functional structure, it is important to recognize that there is more than one dimension of possible variation. Thráinsson (1996), bringing together earlier work from various sources (amongst others, Iatridou 1990, Thráinsson 1993, Bobaljik 1995, Ouhalla 1988, 1991, 1994), was the first to outline explicitly the complexity of the question. Drawing from the discussion in Thráinsson (1996), as well as later work by Cinque (1999) and Starke (2001, 2004), the first distinction to be made is the following: whether languages (or structures) differ with respect to

---

<sup>2</sup> Ouhalla (1988) had previously discussed the possible parameterization of the hierarchical ordering of functional projections, a possibility that has since fallen out of favor. See section 1.2.2 for a more detailed discussion.

the functional projections they instantiate is in principle independent of whether they differ in the hierarchical ordering of those functional projections. Accordingly, the discussion in this section will have to follow these two paths separately.

Let us briefly address the hierarchy of functional projections first. Since the mid-1990s, it has been widely accepted that the hierarchy of functional projections is universal. The usual point of reference for this conclusion is Cinque's (1999) systematic, cross-linguistic study of this empirical domain (see also Cinque 2002, 2006 and Rizzi 2004; for counterarguments, see Ouhalla 1988, 1991, 1994, Nilsen 2003, Ramchand and Szevonius 2014). What remains an open question is the place of this descriptive generalization within syntactic theory. Specifically, whether the functional hierarchy can be considered a primitive of Universal Grammar (as argued by Cinque 1999), or whether it can be derived from more basic, independent properties of UG (see Nilsen 2003, Ramchand and Szevonius 2014). Not to lessen the significance of this issue, in the following discussion I will follow Adger (2013) in assuming that, at least in relation to the narrow syntactic component of a theory of functional structure, the functional hierarchy can be profitably considered axiomatic.

The second, and much more controversial, question to consider is whether the set of functional projections instantiated across languages and structures is universal or variable. It is important to clarify from the outset that this is a largely theoretical issue. Of course, it is certainly the case that the existence of a given functional projection can be empirically demonstrated in specific structures and specific languages. For instance, (1) is a small (non-exhaustive) sample of English examples that demonstrate the existence of the functional head C. In the embedded clause in (1a), C is trivially identified by the presence of the overt complementizer. In the yes/no question in (1b), C is inferred by the surface position of the auxiliary at the left of the subject, which indicates the existence of a functional head above the TP. In the embedded wh-question in (1c), C is inferred by the surface position of the wh-object at the left of the subject, which indicates the existence of a specifier, and consequently a functional head, above the TP. Finally,



in the wh-question in (1d), both the auxiliary and the wh-object surface at the left of the subject, which demonstrate the existence of C and its specifier.

- (1) a. I think that John was playing chess
- b. Was John playing chess?
- c. I wonder what John was playing
- d. What was John playing?

However, the presence of a functional projection in some structures and/or languages does not suffice to make a general statement about its distribution across all structures and languages – all the more so given the common occurrence of structures and/or languages that do not provide any evidence for or against the existence of the same functional projection. To illustrate, consider that for every example in (1) one can find some comparable structure where the relevant diagnostic is missing. (2a) is an embedded clause where the overt complementizer is omitted (cf. (1a)). (2b) is a simple declarative clause where the auxiliary appears in its canonical position at the right of the subject. Therefore, the position of the auxiliary does not provide evidence for C (cf. (1b)). (2c) is an embedded question where the wh-word is the subject. Consequently, there is no comparison to be made between the surface position of the wh-word and the subject, and no conclusion to be drawn concerning C (cf. (1c)). Finally, (2d) is a wh-question where the wh-word is the subject and the auxiliary surfaces at the right of it. Therefore, similarly to the previous two examples, neither the position of the wh-subject nor the position of the auxiliary provides any information about C (cf. (1d)).

- (2) a. I think John was playing chess
- b. John was playing chess
- c. I wonder who was playing chess
- d. Who was playing chess?

Due to these limitations, the relevant literature commonly employs a top-down theoretical approach in the discussion of the universal or variable instantiation of functional projections. This is the case both for the proponents of the ‘universalist’ approach (Cinque 1999, 2006, Rizzi 2004, amongst many others)

and the proponents of the 'variationist' approach (Iatridou 1991, Thráinsson 1996, and many others). I will explore the arguments of both sides in the following section. For the time being, I return to the main thread of this section, which is to break down the question of the (non-)universality of functional structure to its constituent parts.

Still on the topic of the instantiation of functional projections, another distinction that needs to be made (perhaps a little obvious, but not any less significant) is between possible cross- and intra-linguistic variation. The earlier work on the area tended to focus on one or the other domain. For example, Thráinsson's (1993) comparison of finite versus infinitive embedded clauses in Icelandic and Travis' (1984) or Zwart's (1993) treatment of Germanic V2 clauses with a subject in first position as bare TPs focus on intra-linguistic variation. On the other hand, Ackema et al's (1993) arguments against the presence of a TP in Dutch and Bobaljik's (1995) discussion of the availability of a 'Split-Infl' in Germanic languages focus on cross-linguistic variation. However, by the time of Thráinsson (1996), the study of (arguable) variation in the composition of functional structure has been systematized enough, so that later works acknowledge the possibility of both cross- and intra-linguistic variation (to name a few: Giorgi and Pianesi 1997, Koenenman 2000, Bury 2003; even Cinque 1999 addresses these two domains separately, even though he eventually argues against any kind of variation in functional structure).

Making a distinction between the two empirical domains of (potential) variation in the instantiation of functional projections (i.e. cross- and intra-linguistic variation) is also relevant to the question of how this variation may be captured in syntactic theory. To my mind, there are two broad mechanisms that could introduce variation in the composition of functional structure: Firstly, functional structure may be determined by the set of functional heads that are available to a given language or structure and, secondly, functional structure may be determined by independent constraints on which functional heads (selected from the set of available heads) may or may not be projected in a given language or structure

(note, of course, that these two mechanisms are not mutually exclusive; it is entirely possible that both play a role in the analysis of functional structure). This distinction is found, at least tacitly, in various theoretical discussions of functional structure, including Thráinsson (1996), Cinque (1999) and, more recently, Starke (2001, 2004). However, I believe it would be profitable to make this distinction explicit, and explore some of its implications.

To clarify on the subtle distinction between the ‘available’ and actually ‘projected’ functional heads, consider the following. Let us entertain for a second the hypothesis that functional structure is indeed variable (in the sense that functional heads may be radically absent in certain syntactic environments). Further, let us focus on cross-linguistic variation. The simplest hypothesis would be to attribute this variation to the lexicon. This suggestion could be formalized along the following lines: suppose that there is a universal inventory of functional heads provided by UG, but each language selects an appropriate subset of those heads that forms a kind of language-specific functional ‘lexicon’ (cf. Chomsky 1995, Thráinsson 1996, Giorgi and Pianesi 1997, and many others). Then, certain functional heads can be radically absent from certain languages. It is, however, possible that cross-linguistic variation in functional structure may be captured in an entirely different, and more indirect, way. Suppose, as alluded above, that the projection of a functional head is sensitive to independent syntactic factors. Then, it is possible that two languages with identical inventories of functional heads nonetheless consistently display different functional structures as a result of some independent difference in their syntax. Thus, in principle, there are two distinct ways of deriving variation in a theory of functional structure (i.e. by means of a language-specific functional lexicon, or by means of syntactic constraints that operate on a universal inventory of functional heads with potential language-specific implications), as well as the possibility that these two factors work in tandem (i.e. the set of functional projection that are attested in a given language can be viewed as the intersection of the inventory of functional heads of the

language in question and the set of licit functional projections according to relevant syntactic constraints).

Before making any judgment on these formal options for the analysis of variation in functional structure, let us consider how they fare in a different domain: i.e. intra-linguistic variation. Obviously, the different clauses of a given language have access to the same inventory of functional heads, thus it is hard to see what, if any, is the role of this inventory in the analysis of intra-linguistic variation in functional structure. There is, of course, a looser sense in which this functional lexicon may indeed be relevant. Presumably, a functional head has a number of formal properties, like any other lexical item. It is, then, conceivable that the formal properties of a functional head may be relevant to its distribution in different syntactic environments. Note, however, that the previous description makes a mention of 'syntactic environments'. To my mind, it is impossible to attribute intra-linguistic variation solely to the inventory of functional heads. It can certainly be the case that the functional lexicon provides the input to syntactic derivations. However, the well-formedness of those derivations will eventually be decided by independent syntactic considerations<sup>3</sup>. Thus, I conclude that the analysis of intra-linguistic variation in functional structure must necessarily include a narrow syntactic component, while the role of the inventory of functional heads will depend on a number of other assumptions (for example, assumptions on what kinds of grammatical information can be encoded on a functional head), but it will, in any case, be secondary.

Let us now consider what the previous observations suggest for the analysis of functional structure. Thus far, I have made two points. Firstly, I have suggested that variation in the composition of functional structure may be introduced by (i) the inventory of functional heads, and (ii) the syntactic factors that determine

---

<sup>3</sup> Consider, for example, the usual assumption that the various clauses of a given language correspond to different numerations, which may include or exclude specific functional heads. This formulation of variation in syntactic derivations may be descriptively adequate, but it lacks explanatory potential. There still is a need to independently justify why certain numerations can converge while others crash, and the explanation of that can only be found in syntactic processes.

whether those heads may project or not. Secondly, I have demonstrated that the syntax of functional heads (i.e. (ii)) appears to be more significant than the inventory of functional heads (i.e. (i)) in a certain way. That is, the relevance of (i) seems to be limited to the analysis of cross-linguistic variation, while (ii) may contribute to the analysis of both cross- and intra-linguistic variation (and, indeed, it seems to be indispensable for the analysis of the latter). Thus, I conclude that a variationist theory of the composition of functional structure must incorporate some mechanism that gives rise to variation at the level of the syntactic derivation, and may additionally include a mechanism that creates variation at the inventory of functional heads. Finally, note that the distinction between (i) and (ii) as possible sources of variation is also relevant to the competing approach that maintains the universality of functional structure. Specifically, universalist theory must assume that both the inventory and the projection of functional heads are invariable.

Thus far, I have been tacitly assuming that functional structure is built out of functional heads that are atomic, in the sense that each one embodies a single categorial feature. This is a fairly common assumption in the generative literature at large, and one that is explicitly adopted in universalist theories of functional structure (e.g. Cinque 1999 and much of the cartographic literature). However, the variationist literature has explored an alternative approach with respect to the distribution of categorial features across functional and, possibly, lexical heads. Namely, Bobaljik (1995), Giorgi and Pianesi (1997), Nash and Rouveret (1997), Bobaljik and Thráinsson (1998), and others assume that the functional structure may include non-atomic heads that embody multiple categorial features. This hypothesis suggests that questions relating to the (non-)universality of functional structure apply to two levels. For example, we can question the (non-)universality of the inventory of categorial features separately from the (non-)universality of the inventory of functional heads (in the sense that the distribution of categorial features may itself be variable, resulting in language-specific non-atomic heads). Similarly, we can explore the contribution of the syntactic properties of categorial features to the (non-)universality of functional structure separately from the

contribution of the syntactic properties of functional heads. Therefore, the adoption of non-atomic heads opens new avenues of investigation into the properties of functional structure. I will explore some of the relevant possibilities in the following sections.

In this section, I have established what are the necessary components for a comprehensive theory of functional structure (whether universal or variable). With that in mind, I will proceed to a comparison of the universalist and the variationist approach to functional structure in the following section.

### ***1.2.3 The (Non-)Universality of Functional Structure and the (Non-)Atomicity of Functional Heads***

As mentioned in the previous section, the literature on the (non-)universal instantiation of functional projections usually favours a top-down theoretical approach on the matter. In this section, I attempt to evaluate the two competing views of functional structure following a similar trajectory. The discussion starts from the differences in the overarching theoretical design of the universalist versus the variationist perspectives on functional structure. As we will see, there is a theoretical trade-off between an information-rich UG (in the universalist view) and relative complexity of the language acquisition and syntactic mechanisms that determine the composition of functional structure (in the variationist view). Following this brief discussion about language design in general, I will focus more narrowly on syntax and I will consider another theoretical trade-off; one between the complexity of syntactic operations that derive functional structure and the complexity of the resulting syntactic representations. This will allow me to highlight two areas of syntactic research that lend support to the variationist approach on functional structure: head movement and subject/non-subject asymmetries. Finally, I will consider how the (non-)universality of functional structure intersects with the (non-)atomicity of functional heads, which will provide some deeper insights into the analysis of variation in functional structure.

To begin this discussion, consider the following. One of the overarching arguments in favour of the universalist view of functional structure is that there is a certain theoretical simplicity to a theory that attributes the properties of functional structure to UG. Remember that in the previous section I have broken down the question of the composition of functional structure to two components: the inventory of available functional heads and the syntactic constraints that determine the mapping of the functional lexicon to syntactic structures. Of course, if we are to assume that functional structure is determined by UG, then both the mechanisms that enable the acquisition of the functional lexicon and the mechanisms that map the functional lexicon to a syntactic structure are trivialised. The argument has been made by Cinque (2006:6 in relation to the acquisition of functional heads) and Rizzi (2004:7 in relation to the syntactic constraints that determine the projection of functional heads)<sup>4</sup>.

I believe, however, that the simplicity of the universalist approach to functional structure is illusory. Rather, what we are dealing with is a trade-off in which component of the language faculty is assumed to encode the complexity of functional structure. Note that the simplification of the relevant acquisition processes and syntactic mechanisms comes at the cost of an information-rich UG. Thus, the universalist approach is essentially relegating the explanation of the properties of functional structure to a different discipline (or disciplines); the burden of explanation for the individual's knowledge of the composition of functional structure falls to neurobiology, and the burden of explanation for the diachronic development of that knowledge falls to evolutionary biology. Of course, I am not arguing that this state of affairs is de-facto undesirable. It is, however, important to acknowledge that although an information-rich UG may not be costly for the syntactic component of the theory of functional structure, it is, nonetheless, costly from the perspective of a holistic theory of the language faculty. Thus, what

---

<sup>4</sup> Furthermore, Cinque argues that adopting the hypothesis that functional structure is universal also allows a simplification of the mechanisms that map syntactic structures to LF structures. I will not explore the semantic implications of the (non-)universality of functional structure, but I refer the reader to Cinque (1999: 128ff) for a discussion.

we are dealing with is a theoretical trade-off between a minimal syntactic theory of functional structure (supplemented by an information-rich UG) and a minimal UG (which derives rich functional structures by virtue of the complex interactions of a language-specific functional lexicon and independent, syntactic, structure-building operations).

In the previous paragraph I challenged the view that a universalist, UG-centred analysis of the composition of functional structure is more economical. Additionally, one could argue that evaluating the perceived complexity of the language acquisition and syntactic mechanisms that are presupposed in the variationist perspective of functional structure is not as straightforward as it might seem at first sight. Before reaching any conclusion, it is necessary to consider the possibility that the relevant mechanisms are independently needed. Or, from a different perspective, we have to consider whether the relevant hypotheses provide any interesting insight on empirical domains beyond the question of the composition of functional structure. Consider, for example, the issue from the perspective of language acquisition. It is certainly possible that considerations relating to the efficient or economical design of the acquisition component of a theory of functional structure can be brought to bear to the question of the (non-)universality of functional structure. However, that presupposes that we have established whether the acquisition of a functional lexicon requires acquisition mechanisms that are unique to functional heads or not. Thus, pending the results of this investigation, it is not clear that language acquisition provides any conclusive argument in favour of the universalist (or, for that matter, the variationist) approach to the composition of functional structure.

The same question can be raised in relation to the syntactic component of the analysis of functional structure: i.e. to what extent are the syntactic mechanisms of a variationist theory of functional structure (or, equally, of a universalist theory) independently motivated? In order to explore this issue, I will begin from the following observation. From a narrow syntactic perspective, the two approaches to functional structure present us with another theoretical trade-off. It is clear by now



that a universalist theory employs a minimal syntactic mechanism for the derivation of functional structure, but this comes at the cost of an expansive syntactic representation. Conversely, on the variationist view the derivation of functional structure is a much more complex state of affairs, but the resulting structure should be minimal (in the sense that functional heads and their projections can be omitted under certain circumstances). I believe that both (i) the relative complexity of syntactic derivations and (ii) the relative simplicity of syntactic representations of a variationist approach have explanatory power in different theoretical or empirical areas. In this dissertation, I will refer to head movement to defend part (i) of this claim, and subject/non-subject asymmetries to defend part (ii). In the following paragraphs, I will present a brief outline of my reasoning, in anticipation of the full analysis of these matters in chapters 2 (head movement) and 3 (subject/non-subject asymmetries).

As I have mentioned in the previous section, there has to be a syntactic dimension in the analysis of variation in functional structure. One of the better developed implementations of this observation is the concept of ‘re-projection’ (Ackema et al 1993, Giorgi and Pianesi 1997, Nash and Rouveret 1997, 2002, Koenenman 2000, Bury 2003, and others). The general idea in re-projective theories is that not every functional projection corresponds to a dedicated functional head. Rather, some functional projections arise from a properly defined process of re-projection of lexical or functional heads. Under the further assumption that this process only takes place under certain circumstances, it follows that the resulting functional structure will be variable.

Apart from the composition of functional structure, the other major focus of the re-projective literature has been head movement. According to Ackema et al (1993), Koenenman (2000), Bury (2003) and others, re-projection is closely linked with a non-standard version of head movement, which differs from more widely accepted alternatives (e.g. Chomsky 1995 and subsequent work) in two respects. Firstly, these works assume that the operation takes place to satisfy some property of the moved head itself, rather than some property of a target head as in

Chomsky's (1995) Attract version of head movement. Secondly, they assume that the moved head can project, pace Chomsky's (1995) 'target projects' constraint which states that the target head must project. Under these assumptions, the target head can be eliminated and head movement can be construed as an operation whereby a head moves from its base position, re-merges directly with its own maximal projection, and projects for a second time. In chapter 2, I will discuss this formulation of head movement in more detail and I will demonstrate that it provides a principled account of the fundamental properties of the phenomenon (see Matushansky 2006, Roberts 2002, 2022, or section 2.2.1 for an overview of the relevant properties). Thus, re-projective theories provide important insights not only into the composition of functional structure, but also into the theory of head movement. This fact, I will argue, speaks in favour of a variationist approach to functional structure (at least, one that is based on the concept of re-projection), in spite of the relative complexity of its syntactic derivations.

Let us also briefly consider how subject/non-subject asymmetries (i.e. asymmetries in Subject Auxiliary Inversion and *do*-support in English *wh*-questions, the *that*-trace effect in English embedded clauses involving *wh*-extraction, and the anti-*that*-trace effect in English non-*wh* relative clauses) relate to the study of the composition of functional structure. One of the earlier accounts of subject/non-subject asymmetries in English *wh*-questions was based on the Vacuous Movement Hypothesis (George 1980, Chomsky 1986), which places a ban on movement that does not affect the linear order of constituents. In chapter 3, I will demonstrate that the effects of the Vacuous Movement Hypothesis with respect to phrasal movement (i.e. *wh*-movement in English *wh*-questions) and head movement (i.e. Subject Auxiliary Inversion in the same structures) follow naturally from the reduced syntactic representations that are made possible by a variationist approach to functional structure. Accordingly, I will argue that such an approach provides a promising account of subject/non-subject asymmetries in English *wh*-questions. I will further consider how to extend that account to the more complex cases of the *that*-trace and the anti-*that*-trace effect. Finally, I will conclude that, to the extend

that this analysis proves to be successful, it lends support to the hypothesis that functional structure is variable.

The aim of this section thus far has been to outline the merits of the variationist approach to the composition of functional structure. In the remainder of the section, I will focus more narrowly on the merits of the subset of the variationist literature that adopts the concept of re-projection. Specifically, I will consider how the interplay between the concept of re-projection and the concept of non-atomic heads may inform the analysis of variation in functional structure. Giorgi and Pianesi (1997), Nash and Rouveret (1997, 2002) and Koenenman (2000) are particularly interesting in this respect. As I have mentioned previously, these works assume that categorial features are not necessarily distributed in a one-to-one fashion across lexical and functional heads. Rather, they assume that it is possible for a single head to have multiple categorial features. Furthermore, these works suggest that the process of re-projection allows the various categorial features of a non-atomic head to project separately. This suggestion has an interesting consequence. It entails that re-projection is not merely a case of recursion of a functional projection. Rather, it is a case of a single lexical or functional head giving rise to multiple functional projections with distinct properties. Finally, note that the relevant literature assumes that the process of re-projection is subject to cross- and intra-linguistic variation. Thus, a given non-atomic head may project a different number of functional projections under different circumstances.

Now, consider the situation from the perspective of functional structure. The approach outlined above suggests that at least some parts of the functional structure are a manifestation of the internal structure of a single head, rather than the result of the concatenation of various discrete heads. Furthermore, it suggests that functional structure may vary depending on how little or how much of the internal structure of a head is relevant to the syntax of a given structure or language.

Suppose, then, that we combine the above conclusions with a lexicalist approach to morphosyntax<sup>5</sup> (Chomsky 1970). That is, the hypothesis that syntax operates on morphologically complex lexical items that are the output of a separate, pre-syntactic, morphological component. Under this view, the statements made in the previous paragraph should be slightly rephrased as follows. Rather than associating functional structure with the ‘internal structure of a head’, it would be more accurate to associate it with the morphological structure of a lexical item<sup>6</sup>. Of course, the link between morphological and functional structure is supported by Baker’s (1985, 1988) Mirror Generalisation (i.e. the observation that the hierarchical organisation of functional heads tends to be the mirror image of the hierarchical organisation of inflectional morphemes). Finally, the appeal of this approach is that it does not treat the simplicity of syntactic representations advocated by the variationist literature as a first principle, but instead it grounds it in the organization of the linguistic system. Specifically, the claim is that lexical items provide to syntax the information needed to derive the functional structure. But this information is not exclusively (and perhaps not primarily) syntactic. Rather, the same information is also relevant to morphological structure. Therefore, it is to be expected that not all of the functional information provided by a lexical item will be useful to every structure of every language. According to this view, this is the source of variation in the composition of functional structure. Furthermore, in the analysis I will develop in the following sections, I will opt for a derivational implementation of this approach (something that is unusual amongst lexicalist theories; this is, for instance, a major differentiating factor between my analysis and Brody’s 1997, and subsequent work). The reason for doing so is to reinforce the claim that the simplicity of syntactic representations is not an extraneous constraint, but a result of loss of information in the duration of the morpho-syntactic derivation. More generally, by adopting a strongly derivational, as well as a strongly lexicalist, version

---

<sup>5</sup> This is the position adopted by Giorgi and Pianesi (1997), but see Nash and Rouveret (1997, 2002) for a structuralist alternative.

<sup>6</sup> This view has also been explored at length by Brody (1997) and subsequent work. One difference between Brody (1997) and the literature cited in the text is that Brody does not focus on the possibility of variation in functional structure.

of this approach, I hope to lay bare both the merits and the demerits of a variationist model of functional structure.

To sum up, in this section I have argued in favour of the view that functional structure is subject to cross- and intra-linguistic variation. More specifically, I suggested that a re-projective and lexicalist approach on the matter presents a number of interesting theoretical and empirical implications. I will explore this combination of hypothesis in more detail in the following sections.

### ***1.3 A 'Variationist' Theory of Functional Structure: Giorgi and Pianesi (1997)***

#### ***1.3.1 Syncretic Heads, the Feature Scattering Principle and the Universal Ordering Constraint***

In the previous section I highlighted some of the key points of the debate between the universalist and variationist perspective of functional structure. Specifically, I considered the argument that a universalist approach on the matter makes for a more economical theory of functional structure and I have determined that the argument is inconclusive. Additionally, I have argued that a variationist approach on functional structure has the potential of covering a wider theoretical and empirical domain (i.e. apart from the question of the composition of functional structure, it also offers insights on the role of head movement as a structure-building operation), at little additional cost. As I have argued previously, this potential is more prominent in a subset of variationist approaches that incorporate two independent, but complementing, assumptions: the hypothesis that the composition of functional structure is cross- and intra-linguistically variable and the hypothesis that functional structure is dynamically created during the derivation as an extension of lexical items. Accordingly, I will henceforth focus my attention on this subset of theories of functional structure.

In the following sections, I will focus in particular on Giorgi and Pianesi (1997), since they provide one of the more detailed formalisations of both of the aforementioned hypotheses. The discussion in section 1.3.1 starts with an introduction of the core concepts of their analysis and a demonstration of how the

properties of functional structure are derived from them. In section 1.3.2, I will continue with a discussion of what I perceive to be a major limitation of Giorgi and Pianesi's (1997) theory of functional structure (and, in fact, a limitation of re-projective theories of functional structure in general): the fact that re-projection is usually assumed to be relevant to a subset of functional heads which, I will argue, do not constitute a natural class. Moreover, this discussion will help set the stage for the analysis of functional structure that I will develop in section 1.4, which is partly inspired by certain central concepts of Giorgi and Pianesi's (1997) analysis.

The core of Giorgi and Pianesi's (1997) analysis of functional structure consists of three concepts: (i) syncretic categories, (ii) the Feature Scattering Principle, and (iii) the Universal Ordering Constraint. Let us consider these concepts in turn. Firstly, Giorgi and Pianesi suggest that lexical items (i.e. lexical and functional heads that are drawn from the lexicon) may have more than one categorial feature<sup>7</sup> as part of their featural specification. Giorgi and Pianesi (1997) call these lexical items 'syncretic categories', and they assume that they are not universal (cf. the quote in (3) below): Features A and B may be associated with a single lexical item in one language, but correspond to two separate lexical items in another. Obviously, the syncretic categories that are available in the lexicon of a given language play a role in determining the composition of the functional structure of said language. All else being equal, a language where features A and B are part of a syncretic category will display a less articulated functional structure compared to a language where the same features are separate heads. Following similar reasoning, syncretic categories may also be relevant to intra-linguistic variation in functional structure. Suppose that in a given language features A and B appear in the featural specification of multiple lexical items, including some (one or more) head that expresses those features syncretically and some (at least two, or possibly more) heads that express

---

<sup>7</sup> Note that I will, henceforth, refrain from using the traditional term 'categorial feature' (for features like V, T, C, etc), and I will instead refer to them as 'functional features'. This is because, according to this conception of a lexical item, the category of a head cannot be gleaned straightforwardly from the relevant features (given that there can be more than one categorial feature per head). Rather, in this system, functional features operate in more complex ways that will become clearer in the following discussion.

features A and B separately. Suppose, further, that due to independent factors (e.g. lexical idiosyncrasies of the language in question) the syncretic category has a different distribution across various clause types in comparison to the non-syncretic heads. Then we would expect that clauses involving the syncretic category would display a different functional structure from clauses that employ the non-syncretic heads.

### (3) Syncretic categories

“... [A] child learns a language by associating morphemes to features. The association could be one to one, giving agglutinative or isolating languages, or it could happen that the same morpheme is associated to more features, in this way obtaining the so-called inflected languages, such as Italian. We will call the categories obtained by means of such a multiple association *syncretic* and *hybrid*<sup>8</sup> categories.”

Giorgi and Pianesi (1997: 14)

Remember that in section 1.2.2 I drew a distinction between the inventory of functional heads that are available to a given language (or structure) and the functional heads that are actually projected in that language (or structure) as a result of independent syntactic factors. Giorgi and Pianesi's (1997) syncretic categories provide a way of understanding and formalising the contribution of the inventory of functional heads to the composition of functional structure. But their analysis also includes a hypothesis relating to the syntactic processes that manipulate the inventory of functional heads in order to derive syntactic structures. Specifically, Giorgi and Pianesi (1997) suggest that functional features are, in some sense, (semi-)independent of the head that introduces them in a syntactic structure. More concretely, they propose that a syncretic category will project at least one head (bearing all the functional features of the lexical item), but it may optionally project as many heads as the functional features of the lexical item

---

<sup>8</sup> Note that Giorgi and Pianesi (1997) suggest that there are two kinds of lexical items that bear multiple functional features (i.e. 'syncretic' and 'hybrid' categories) with subtle differences. I will not address this distinction here as it is not particularly relevant this discussion.

(where each head bears an individual functional feature). Obviously, according to this formulation, any intermediate situation is also possible. This idea is encoded in the Feature Scattering Principle, seen in (4) below. Note, of course, that whether Feature Scattering takes place or not is determined by independent syntactic factors and, therefore, its output may vary across languages or structures. Thus, two languages that happen to have the same syncretic category in their lexicon (or two structures that are derived out of numerations that happen to include the same syncretic category) may nonetheless display different functional structures if, for independent reasons, the category is scattered in one, but not the other, language (or structure).

(4) Feature Scattering Principle

“Each feature can head a projection”

Giorgi and Pianesi (1997: 15)

Thus far we have seen that Giorgi and Pianesi (1997) approach the question of the composition of functional structure from two directions: the concept of the syncretic head relates to the organisation of the inventory of functional heads of a given language, while the Feature Scattering Principle relates to the behaviour of those functional heads during syntax. These two mechanisms interact in complex ways that can capture both cross- and intra-linguistic variation. Specifically, this system predicts that variation in the composition of functional structure should display the following (rather restrictive) pattern: A language or numeration containing a category that syncretically expresses features A and B, will give rise to structures either with one or two functional projections of the relevant features. On the other hand, a language or numeration where features A and B are expressed by separate categories will necessarily give rise to structures with two functional projections of the two relevant features.

Finally, Giorgi and Pianesi (1997) propose that there is a Universal Ordering Constraint that determines the order of (checking) operations involving the functional features of syncretic categories (see (5) below). Assuming that Feature



Scattering takes place to enable the checking of functional features that would otherwise be unable to enter the appropriate syntactic operations, entails that the order in which functional features are scattered is determined by the Universal Ordering Constraint. Furthermore, this means that this constraint (indirectly) determines the hierarchical order of functional projections in the structure of the clause. Thus, the Universal Ordering Constraint has the same effect as Cinque's (1999) Universal Hierarchy of Functional Projections or Starke's (2001, 2004) Functional Sequence. However, unlike Cinque's or Starke's formulation, the Universal Ordering Constraint (as well as the Feature Scattering Principle) does not apply directly to functional projections themselves. Rather, Giorgi and Pianesi's (1997) formulation of the functional hierarchy constrains the functional features of syncretic categories and the operations that apply to them<sup>9</sup>. Thus, Giorgi and Pianesi's analysis has a strong lexicalist flavour in that it assumes that it is the properties of lexical items that are primitive, while the properties of functional structure (especially its composition) are essentially derivative. I will return to this aspect of Giorgi and Pianesi's (1997) theory of functional structure shortly, but first, let us conclude the discussion of the hierarchy of functional projections with the following observations. The way that Giorgi and Pianesi (1997) set up the interaction between syncretic categories, the Feature Scattering Principle and the Universal Ordering Constraint entails that the hierarchical ordering of functional projections is independent of the composition of functional structure. Thus, they are able to maintain the empirically sound generalisation that the hierarchy of functional projections is universal, while exploring the descriptive and explanatory advantages of a variationist approach on the composition of functional structure.

(5) Universal Ordering Constraint

"Features are ordered so that given  $F1 > F2$ , the checking of  $F1$  precedes the checking of  $F2$ ."

Giorgi and Pianesi (1997:14)

---

<sup>9</sup> The hypothesis that the functional hierarchy is a constraint on the timing of syntactic operations is also explored, in rather different terms, in Adger (2013) and Williams (2013).

Thus far I have focused on what Giorgi and Pianesi's (1997) proposals in (3)-(5) entail for the composition and the hierarchical organisation of functional structure. There is, however, another dimension to this discussion: we can also focus on what these hypotheses entail for the organisation of the internal structure of lexical items, as well as the relevance of that internal structure to narrow syntactic mechanisms. Firstly, positing the existence of syncretic categories suggests that we treat lexical items as complex units that display numerous discrete syntactic properties. There are, of course, significant implications from the specifics of Giorgi and Pianesi's (1997) formulation, but this general idea is by no means novel. This hypothesis is strongly reminiscent of the minimalist conception of lexical items as bundles of features (Chomsky 1995). However, combining syncretic categories with the Universal Ordering Constraint implies that lexical items are more than complex units. They are, in fact, complex units with rich internal (hierarchical) structure. Furthermore, combining syncretic categories (and the Universal Ordering Constraint) with Feature Scattering suggests that (at least some parts of) the functional structure is the syntactic expression of the underlying internal structure of lexical items.

In essence, the theory of functional structure that I will propose in the following sections is an exploration of these two fundamental concepts (that lexical items have an internal structure and that functional structure is derived from it). I will begin this exploration in section 1.3.2 by raising the question of how far we could or should go in the effort of deriving (covert) functional structure from the properties of lexical items.

### ***1.3.2 Deriving (Covert) Functional Structure***

As we have seen in the previous section, Giorgi and Pianesi's (1997) theory of functional structure (as well as other re-projective theories) provides a mechanism that can reduce certain structural positions to being the output of a structure-building process (i.e. Feature Scattering) operating on lexical items with variable properties (syncretism of functional features). Thus, for every structural position

that can be dynamically constructed during the derivation we need not postulate a prefabricated functional head that is drawn from the lexicon and merged (as is) in the structure. In this section I raise the question of how far down the path of eliminating 'non-derived' functional heads we can profitably go. We can, in fact, break down this question to two parts. As it is clear by now, part of the motivation behind this approach to functional structure is the effort to eliminate structural positions that are phonologically null and syntactically inert. Thus, the first obvious question is how much of the covert parts of functional structure we can attribute to Feature Scattering (or some other similar re-projective formalism). Another, less straightforward, question is whether the same approach can be extended to overt functional positions. I will start the discussion with a brief comment on the later question and I will then go on to address the former in more detail.

Remember that from section 1.2.3 onwards I have favoured a lexicalist characterisation of the idea that functional structure is dynamically created during the derivation. To clarify, I am not simply assuming that (parts of) functional structure is derivative, but, more specifically, that it is derived from (the properties of) lexical items. Given this lexicalist perspective, it would be rather counter-intuitive to entertain the possibility that overt functional heads are derived from the lexical head of the clause (i.e. the lexical verb). It could be reasonable to assume that elements such as dummy auxiliaries are inserted post-syntactically in some structural position that is scattered from the lexical verb during the derivation, but extending this hypothesis to all functional heads (e.g. auxiliaries with a discrete semantic contribution, complementizers, etc) would signify a major departure from the basic premises of the analysis developed here<sup>10</sup>.

Let us now return to null functional heads and their standing within re-projective theories of functional structure. By and large, re-projective analyses (Ackema et al 1993, Giorgi and Pianesi 1997, Nash and Rouveret 1997, 2002,

---

<sup>10</sup> Of course, if we were to abandon the lexicalist component of the analysis, we could explore a syntactic system whereby the structural positions of both null and overt functional heads are dynamically created during the derivation. I will not pursue this approach here, but see Adger (2013) and Williams (2013) for relevant proposals.

Koenenman 2000, Bury 2003) assume that only a small number of functional heads are amenable to an analysis of the kind outlined above. The identity of the relevant functional heads varies from one work to the other, but, generally, the traditional idea that null non-derived functional heads are actual entries in our mental lexicon that, furthermore, constitute one of the building blocks of syntactic structure is not challenged. In fact, Koenenman (2000: 48) explicitly considers the possibility of entirely eliminating null non-derived functional heads, only to conclude that there is no reason to doubt that this type of functional head is acquirable, albeit in a very restricted way. In brief, he argues that null functional heads can be acquired when they are part of a paradigm that also includes overt functional heads with contrasting (semantic or syntactic) properties. Thus, Koenenman (2000) (as well as Bury 2003) assumes that null functional heads constitute a limited, but not entirely reducible, class of elements. More generally, Koenenman assumes that (covert) functional structure consists of two discrete types of functional heads (derived and non-derived) that are introduced in the structure by different syntactic operations (Feature Scattering and Merge respectively). This stance is adopted, more or less explicitly, in most of the re-projective literature. I will, however, argue below that this state of affairs is unsatisfactory for two reasons, having to do both with syntax and acquisition.

As I have just mentioned, re-projective theories of functional structure usually assume that there are two kinds of null functional heads: derived and non-derived. This distinction is not problematic or costly on its own, since we have previously concluded that within a (lexicalist) re-projective theory non-derived functional heads are independently needed to account for the existence of overt functional heads. Similarly, positing two distinct structure-building operations (one responsible for the creation of derived functional heads and one responsible for the insertion of their non-derived counterparts) is unavoidable within the kind of theory explored here. Furthermore, in the relevant literature, re-projection is typically formulated as an extension of a more fundamental structure-building operation (either Merge or Move), thus eliminating any perceived cost of adding more operations in the

toolbox of fundamental syntactic operations. There is, however, one more thing to consider: head movement. Assuming that there are two kinds of null functional heads begs the question of whether our analysis can ensure that head movement to a derived functional head is formally equivalent to the movement to a non-derived head, and if not, to provide independent justification for any formal difference that may arise. This situation is particularly problematic for theories that capture re-projection via head movement (Ackema et al 1993, Koenenman 2000, Bury 2003, but not Giorgi and Pianesi 1997<sup>11</sup>). For concreteness, consider Bury (2003). As I have mentioned previously, Bury (2003) proposes that if we were to remove the ‘target projects’ constrain, we would be able to attribute re-projection to head movement, followed by the projection of some functional feature of the moved head. However, this hypothesis creates a situation where re-projective head movement is a fundamentally different operation from head movement to a non-derived functional head. In the latter case we can presumably maintain the standard minimalist view of head movement (or movement in general) according to which the operation is triggered by some property of the target head. But this cannot apply to the former case where the target head is assumed to be dynamically created by the head movement operation itself. Interestingly, this situation not only forces us to posit two distinct head movement operations without any independent justification, but it also undermines any effort to solve the theoretical problems surrounding head movement (cf. Matushansky 2006, Roberts 2011), which is part of the motivation for exploring alternative formulations of this syntactic operation. I will discuss the properties of head movement more thoroughly in section 2.2 of chapter 2, but for the time being I conclude that the analysis of head movement poses a serious challenge to the usual stance, within the relevant literature, that

---

<sup>11</sup> Giorgi and Pianesi (1997) stand out from other re-projective theories in that they assume that head movement is (partly) dependent on the shape of functional structure, rather than the other way around. My analysis of head movement in chapter 2 follows the same direction as Giorgi and Pianesi, but goes a step farther: in brief, I will make the stronger claim that head movement is derived from (not simply constrained by) the structure-building operation Feature Scattering.

covert functional structure is the product of two distinct structure-building mechanisms: re-projection and Merge.

Before concluding this discussion, let us have another look at the acquisition of functional structure. Koenenman's (2000) argument regarding the acquirability of null non-derived functional heads is undoubtedly reasonable, but I would like to argue that there is a bigger picture to consider. Let us entertain for a moment the hypothesis that an adult language has access to both a re-projective operation that dynamically creates functional heads during the derivation and a number of null non-derived functional heads that can be introduced to a structure via Merge. Let us further assume, as I have implied in the previous paragraph, that there are few, if any, formal syntactic differences between derived and non-derived null functional heads. Finally, following Koenenman (2000), let us assume that non-derived functional heads differ from their derived counterparts only in that they fill in the gaps of some paradigm of overt functional heads; in one sense, null non-derived functional heads are structural positions that have been lexicalised due to the fact that there is indirect evidence for their existence. Presumably, this situation entails that when a learner is faced with language data that suggest the existence of covert functional structure, s/he will have two possible analyses for that covert structure to choose from. My question, then, is the following: if the learner has the option to analyse a covert functional head via re-projection, wouldn't that undermine their ability to instead acquire a null non-derived functional head, if that happens to be the correct analysis in the target language? Of course, the question can equally well be formulated the other way around: if the learner has the option to posit a null non-derived functional head, wouldn't that undermine the alternative, re-projective analysis? To put it in other words, in both cases what the learner has to do is to associate syntactic properties relevant to the functional structure with a lexical entry. In one case, the relevant functional features are syncretically associated with a single lexical item (from which they can be scattered during the derivation to create null derived functional heads), while in the other case the same features are separately associated with several lexical items (including some null non-derived

functional heads). Assuming that there is no formal distinction between these two options, how can a learner choose one over the other? I believe that the only way out of this conundrum is a uniform analysis of covert functional structure. Thus, I conclude that if we are to explore the merits of a re-projective theory of functional structure, we have to take the necessary steps to entirely eliminate null non-derived functional heads and, instead, pursue the hypothesis that covert functional structure is exclusively the product of a re-projective mechanism.

In the following section I will develop a re-projective theory of functional structure, starting specifically from the conclusion that I have reached in the last few paragraphs: that a successful re-projective theory must be able to replace, not merely amend, the traditional conception of covert functional structure.

#### ***1.4 The Proposal: Feature Scattering (Giorgi and Pianesi 1997) and ‘Extended’ Featural Specifications***

##### ***1.4.1 The ‘Extended’ Featural Specifications of Lexical Heads***

In the remainder of this chapter, I will bring together the various conclusions from the previous discussions in order to develop a novel theory of functional structure. Building on the re-projective literature, the overarching theme of this proposal is that functional structure is a dynamic syntactic object (one that is built according to the specific requirements or properties of a given derivation, with a potentially variable outcome across languages and structures) that is derivative of the internal structure of lexical items. More specifically, this analysis extends, and in some cases redefines, previous re-projective theories (drawing especially from Giorgi and Pianesi 1997) with the aim to achieve the goal outlined in section 1.3.2: namely to eliminate null non-derived functional heads and derive the entirety of covert functional structure via Feature Scattering. In sections 1.4.1 and 1.4.2, I set out my view of the featural specifications of lexical items (i.e. the lexical and overt non-derived functional heads that are part of the mental lexicon), which departs from the traditional conception of lexical items as well as Giorgi and Pianesi’s (1997) syncretic categories. Additionally, in the same discussion, I refine Giorgi and

Pianesi's (1997) concept of Feature Scattering, as well as the Universal Ordering Constraint.

I will preface the discussion in this section with the following observation: considering that the stated goal of this analysis is the radical elimination of null non-derived functional heads and that, furthermore, those null functional heads must in principle be derivable via Feature Scattering, it is obvious that we need to hypothesize an equally dramatic increase in the size and detail of the featural specifications of lexical items. To clarify, Giorgi and Pianesi (1997) (as well as other re-projective theories like Nash and Rouveret 1997, 2002, Koenenman 2000, Bury 2003) already assume that the entire array of functional features corresponding to the atomic functional heads of a cartographic approach are present in the numeration, but they are part of the featural specification of a handful of lexical and (overt or covert) functional heads. I propose that it is necessary to push this reasoning farther, so that the entire array of functional features is mapped onto the subset of overt heads (i.e. lexical and overt non-derived functional heads). Furthermore, the fact that the number and category of overt non-derived functional heads varies from one structure to the other (and from one language to the other) implies that the mapping of functional features to lexical items cannot be static. For ease of exposition, I will start from a case that I consider to be the most straightforward in terms of the logic of the argument, as well as the most distant from traditional analyses: a clause where the lexical verb (or, equally, the lexical head of an extended projection of a different category; i.e. a noun in a nominal phrase, an adjective in an adjectival phrase, etc) is the only available lexical item and, therefore, the entire array of functional features must be present in its featural specification. I will return to the complications that arise when a clause is built out of numerous (and varying) lexical items in section 1.4.2.

I will start this discussion with a brief excursion into usual minimalist assumptions about the properties of the lexical and functional heads that make up the structure of extended projections, in the sense of Grimshaw (1991). First of all, it is generally accepted that an extended projection consists of a single lexical head



$X_0$  (for example, in the case of the clause, the lexical head is the lexical verb  $V$ ) and a sequence of functional heads  $X_1, X_2, \dots, X_n$  (in the same example, the first functional head in the sequence, and lowest in the structure, is little  $v$ ; the last and structurally highest functional head  $X_n$  will be Rizzi's 1997 Force<sup>12</sup>; finally, the intermediate positions in the sequence include heads for Tense, Aspect, Mood, Negation, Topic, Focus, etc). Secondly, it is common practise to express one part of the syntactic properties of lexical and functional heads by means of their 'category' or 'categorical feature' (e.g.  $V$  or  $v$  in the previous example). As the term implies, the category of a head places it into a group of lexical items with common properties, mostly in terms of their distribution. Finally, another part of the syntactic properties of a head (broadly speaking, those properties that have to do with the relations established between the head in question and other heads or phrases in a structure; e.g. case or agreement) is usually expressed by means of one or more features that comprise the featural specification of the relevant lexical item. There are various feature systems proposed in the literature that flesh out how and when the aforementioned features establish appropriate syntactic relations. Although there certainly are more elaborate analyses, Chomsky's (1995) version will be sufficient for the purposes of this discussion. According to Chomsky (1995), features can be interpretable or uninterpretable. Syntactic operations are assumed to be triggered by uninterpretable features (annotated as  $uF$ ). The justification for this hypothesis is that the semantic component cannot operate on uninterpretable features, thus those features need to be eliminated before the structure is handed over from narrow syntax to the semantics interface. The elimination of an uninterpretable feature is thought to be achieved by establishing a syntactic relation with a matching interpretable feature (i.e. an  $iF$ , or simply  $F$ ).

Putting all of the above together, the information that is needed to construct an extended projection according to these assumptions can be summarised schematically with the following:

---

<sup>12</sup> This, of course, depends on the analysis, but for ease of exposition I will use Rizzi's (1997) version of the structure of the 'left periphery' (i.e. Force > Topic > Focus > Finite) as a guideline.

(6) In order to derive an extended projection, a numeration must include:

- a. A lexical head:  $X_0\{(u)F_x, \dots\}$
- b. A number of functional heads:  $X_1\{(u)F_y, \dots\}$
- ...
- $X_n\{(u)F_z, \dots\}$

(In the above,  $X_0 \dots X_n$  represent the category of the relevant heads. The subscripts 0 ... n indicate that these categories and the corresponding heads are organised according to an appropriate universal functional hierarchy.  $(u)F_x \dots (u)F_z$  are the interpretable and uninterpretable features of each head. For clarity, the subscripts x ... z do not represent any hierarchical organisation, unlike subscripts 0 ... n. They simply indicate that the featural specification of each head is distinct. Finally, the '...' in the featural specification of each head indicates that it may contain more than one un/interpretable feature.)

As a first step towards developing a Feature Scattering theory of functional structure that eliminates null non-derived functional heads, I will propose a formalism that incorporates all of the information in (6) in the featural specification of the lexical head (which, I have argued previously, is necessary in a structure that does not involve any overt non-derived functional heads). My proposal is as follows:

(7) In order to derive an extended projection via Feature Scattering, the numeration must include a lexical head with the following 'extended' featural specification (in the absence of any overt non-derived functional heads):

Lexical head:  $X_0\langle \text{fn}F_n\{(u)F_z, \dots\}, \dots, \text{fn}F_1\{(u)F_y, \dots\}, \text{fn}F_0\{(u)F_x, \dots\} \rangle$

Before looking at how functional structure can be derived from this featural specification, let us unpack the assumptions made in the formalism in (7). Notice, first of all, that the featural specification in (7) includes two types of features, namely the functional features  $\text{fn}F_0 \dots \text{fn}F_n$  and the un/interpretable features  $(u)F_x \dots (u)F_z$ . In the following paragraphs I will propose that those features differ on a number of properties, but initially the distinction is relevant to their organisation within the featural specification. Specifically, I assume that the un/interpretable

features (u)F are, essentially, sub-features of various functional features fnF. Moreover, I assume that the sub-features of a given functional feature form a set (i.e. if there are more than one sub-features associated with a single functional feature, then the former are not ordered in relation to each other). Finally, following Giorgi and Pianesi (1997), I propose that the functional features of a given head form a sequence (i.e. they are ordered). To summarise all of the above, I propose that the featural specification of a lexical item is a sequence whose members are sets, where the sets are labelled by a functional feature and have un/interpretable features as their own (i.e. the set's) members.

The description of un/interpretable features as sub-features is a unique aspect of the current analysis. But, apart from that, I will adopt a rather uncontroversial view of their properties. For consistency, I will assume that un/interpretable features operate as described by Chomsky (1995). Thus, I am assuming that syntactic relations are fundamentally relations between matching un/interpretable features that are triggered by the uninterpretable part of the pair. Note, however, that nothing hinges on 'uninterpretability' being the trigger of syntactic operations. It would be a rather trivial task to replace it with some other property of sub-features, without affecting the rest of the analysis. What is actually important for the theory developed here is that syntactic operations are triggered by (and apply to) sub-features. Thus, I will usually (although not exclusively) prefer to use the term 'sub-feature' to refer to those features that are relevant to syntactic operations. Finally, considering that syntactic relations vary across structures and languages, it is necessary to assume that the set of sub-features that is associated with a given functional feature is (cross- and intra-linguistically) variable.

Let us now have a closer look at functional features. Tentatively, functional features are similar to the traditional concept of a category in that they are relevant to the distribution of lexical items. Nonetheless, there are some significant differences in the formalisation of those concepts. Most notably, while traditional analyses ascribe a single category to any head, in this proposal it is possible for a single head to be associated with a multitude of functional features. How could that

be? Remember that one of the underlying goals of this analysis is to develop a set of theoretical tools that ensure ‘minimal syntactic representations’. Practically speaking, this means that structural (functional) positions should only be created when necessary. Consider, then, the scenario where a clause consists of a single maximal projection. This would be the case of a clause that does not include any other verbal or clausal heads (i.e. auxiliary verbs, complementizers, etc) apart from the lexical verb, and no other constituents (i.e. adverbs, negative markers, etc) apart from the arguments of the verb. What is interesting in this scenario is that the (distributional) properties of this maximal projection is an amalgamation of the properties of a verb and a clause: on one hand it is headed by a lexical head that is capable of selecting arguments, while on the other hand it behaves as a full clause that can be embedded under a matrix verb. Compare, now, this situation with a clause that consists of two maximal projections, namely the projection of the lexical verb and the projection of a null derived functional head that corresponds to the traditional C. This could be, for example, a *wh*-question, where the derivation needs to create an additional functional projection (via some version of a Feature Scattering operation) in order to accommodate the fronted *wh*-word. In this case, we would expect that the verbal and clausal properties of this extended projection are spread across the two maximal projections: the arguments of the verb are still selected by the lexical head of the lower maximal projection, but, presumably, it is the higher maximal projection (the one that is headed by the null derived functional head) that would be selected by a matrix verb. Therefore, there is a certain sense that the distributional properties of (lexical and functional) heads are malleable<sup>13</sup> within the system developed here. With that observation in mind, the intended

---

<sup>13</sup> Note that the situation is actually more complicated than what I have suggested so far. Firstly, aside from its verbal and clausal distributional properties, the verb in the above example may also display the distribution of any of the intermediate functional projections of a verbal extended projection, in the sense that it can be the complement of a (derived or non-derived) functional head belonging to the same extended projection. This situation will be the main topic of the following section (1.4.2). Secondly, it is necessary to flesh out the mechanism that determines the distributional properties of a null derived functional head (to which I will return after introducing my version of Feature Scattering in (9) below) and to compare it to the distributional properties of an overt non-derived functional head (to which I will return in section 1.4.2).

interpretation of the featural specification in (7) is that the distributional properties of a head are not determined by any one functional (or categorial) feature, but rather by a sequence of functional features (note, further, that under these assumptions the category of a head becomes a descriptive term that picks out heads that have identical sequences of functional features in their featural specifications). Of course, the ultimate goal is to combine this formalism of featural specifications (as sequences of functional features) with a Feature Scattering operation similar to that of Giorgi and Pianesi (1997), thus deriving the flexibility of the distributional properties of (lexical and functional) heads. I will flesh out the details of these mechanisms in the following discussion.

To conclude the discussion on functional features, we need to clarify another two properties that distinguish their syntactic behaviour from that of un/interpretable features. Firstly, I will assume that, unlike uninterpretable features, functional features do not trigger syntactic operations, although they do participate in them in different ways that will be discussed below. Secondly, as I have implied in the previous paragraph, the sequence of functional features provides the 'blueprint' from which Feature Scattering derives the functional structure. Thus, I will assume, tentatively for the moment, that the sequence of functional features displays the same properties as Cinque's (1999) Universal Hierarchy of Functional Projections (apart from the obvious difference that the sequence of functional features expresses the hierarchical relation between features, not heads). Specifically, I will assume that both the order and the membership of the sequence of functional features do not vary across structures or languages. Finally, notice that this suggestion is very similar to the reasoning behind Giorgi and Pianesi's (1997) Universal Ordering Constraint. However, for independent reasons that I will discuss below, I will formalise the constraint that determines the sequence of functional features of various lexical items in a slightly different way than Giorgi and Pianesi (1997). I will return to this topic briefly in the following discussion and again in section 1.4.2.

Now that we have described the properties of the featural specifications of the lexical items on which syntactic operations are applied, we can flesh out the operation of Feature Scattering in more detail. As a first general description, I assume that Feature Scattering is a syntactic operation that is triggered when some uninterpretable feature of a head is unable to enter a relation with a matching interpretable feature, because the syntactic configuration (at that particular point during the derivation) does not allow it. Feature Scattering moves the relevant uninterpretable feature to a newly created functional head, thus altering the configuration and allowing the appropriate syntactic relation between the matching (interpretable and uninterpretable) sub-features to be established. To be more precise, I will argue that uninterpretable features fall in two classes depending on the configuration that is necessary for them to enter a syntactic relation. The first type of uninterpretable feature must be c-commanded by the matching interpretable feature in order to successfully establish a syntactic relation with it. Assuming that derivations progress from the bottom up, and that uninterpretable features need to be eliminated immediately, features of this type essentially establish Specifier-Head relations. Most re-projective theories assume that the main function of Feature Scattering (or other re-projective mechanism) is to generate pairs of functional heads and specifiers so as to enable those Specifier-Head relations. Indeed, Giorgi and Pianesi (1997) and Bury (2003) argue that this is the sole purpose of Feature Scattering. However, in chapter 4, I will discuss some empirical evidence suggesting that this is insufficient. Thus, I will propose that there is a second type of uninterpretable feature that must c-command the matching interpretable feature in order to establish a syntactic relation with it. This proposal ascribes a second function to the operation of Feature Scattering: the generation of functional heads that enter in a syntactic relation with some element inside their complement.

As I have alluded previously, although functional features cannot trigger syntactic operations, they are nonetheless relevant to Feature Scattering in certain ways. Firstly, I will suggest that the sequence of functional features determines the

timing of syntactic operations, including Feature Scattering. Specifically, given the featural specification in (7), I assume that syntax will perform any and all operations involving the uninterpretable features associated with the functional feature  $fnF_0$ , then it will move on to the operations involving the uninterpretable features of the functional feature  $fnF_1$ , and so on. Thus, I will propose that the timing of syntactic operations is determined by the Sequence of Operations in (8). Obviously, (8) is a spiritual successor to Giorgi and Pianesi's (1997) Universal Ordering Constraint, in so far as it derives the hierarchy of functional projections from the order of Feature Scattering operations. However, due to the general architecture of the analysis developed here (i.e. the fact that I am explicitly assuming that lexical items have an internal hierarchical structure which, furthermore, determines the order of syntactic operations) the content of Giorgi and Pianesi's (1997) Universal Ordering Constraint has to be broken down into two parts: a constraint that determines the order of syntactic operations (i.e. the Sequence of Operations in (8)) and some constraint that determines the sequence of functional features of lexical items. I will return to the latter topic in section 1.4.2.

#### (8) Sequence of Operations

The order of syntactic operations triggered by the uninterpretable features of a head is such that the uninterpretable sub-features of  $fnF_x$  are eliminated before the uninterpretable sub-features of  $fnF_y$  (where  $fnF_x$  and  $fnF_y$  are members of the sequence of functional features of the head in question), iff  $x < y$ .

Secondly, I will propose that a Feature Scattering operation that was triggered by a given uninterpretable feature does not displace the target sub-feature on its own. Rather, Feature Scattering displaces a segment of the featural specification of the relevant head, which contains the functional feature hosting the targeted sub-feature (including the entire set of sub-features associated with this functional feature) and all of the superordinate functional features (including their sets of sub-features) in that same featural specification. Of course, I assume that the displaced segment of the featural specification of the original head will constitute the featural specification of the newly created functional head. Thus, the proposal is

that Feature Scattering operates as described in (9). Finally, note that the hypothesis that Feature Scattering may affect entire segments of a featural specification in a single step is a unique aspect of this analysis that arises naturally from the rich internal structure of lexical items proposed in (7).

(9) Feature Scattering

Assume that at a given (intermediate) step of a derivation, an uninterpretable feature  $uG$  of a head  $X_0$  - where  $X_0$  has the featural specification in (i) - cannot establish the necessary relation with a matching interpretable feature.

$$i. X_0 \langle fnF_n\{(u)F_z, \dots\}, \dots, fnF_m\{uG, \dots\}, \dots, fnF_1\{(u)F_y, \dots\}, fnF_0\{(u)F_x, \dots\} \rangle$$

Then Feature Scattering will strip a segment of the featural specification of  $X_0$ , leaving  $X_0$  with the altered featural specification in (ii) and creating a new functional head  $X_1$  with the featural specification in (iii) that merges with  $X_0P$ .

$$ii. X_0 \langle \dots, fnF_1\{(u)F_y, \dots\}, fnF_0\{(u)F_x, \dots\} \rangle$$

$$iii. X_1 \langle fnF_n\{(u)F_z, \dots\}, \dots, fnF_m\{uG, \dots\} \rangle$$

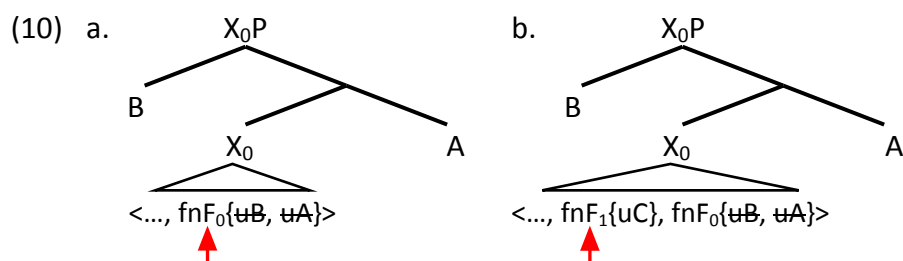
Before concluding this section, let us go over a few toy examples to demonstrate the fundamentals of this proposal. Assume that we have a lexical item  $X_0$  with a featural specification consisting of three functional features  $fnF_0$ ,  $fnF_1$ ,  $fnF_2$ . In principle, this featural specification can derive a structure with one, two or three maximal projections depending on various factors. The latter case is the most straightforward. I demonstrate the steps of the relevant derivation in (10) below. As I have suggested, the derivation will start with the sub-features of  $fnF_0$ . Let us assume that  $fnF_0$  carries two uninterpretable sub-features ( $uA$  and  $uB$ ) that trigger the merger of two constituents with matching interpretable features ( $A$  and  $B$  respectively) at the position of the complement and the specifier of  $X_0$ <sup>14</sup>. This will

---

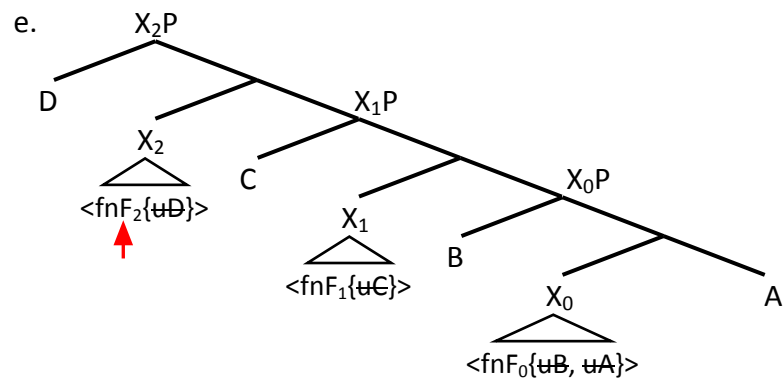
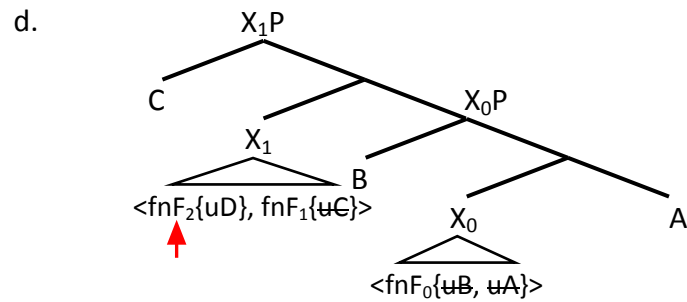
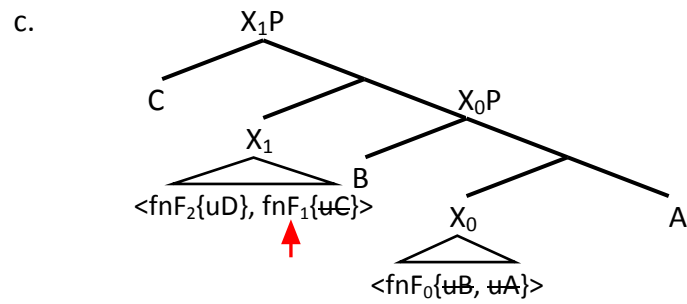
<sup>14</sup> At this point I am glossing over the derivation of the internal structure of  $X_0P$  in order to focus on the effects of Feature Scattering. However, the derivation of  $X_0P$  raises one important question: given that the sub-features  $uA$  and  $uB$  of  $fnF_0$  are assumed to be an unordered set, what determines the order of merger of  $A$  and  $B$  such that  $A$  appears as a complement and  $B$  as a specifier of  $X_0$ ? I will address this issue in section 2.3.1 in connection with the analysis of ditransitive verbs, which provide a concrete example of the relevant structure.



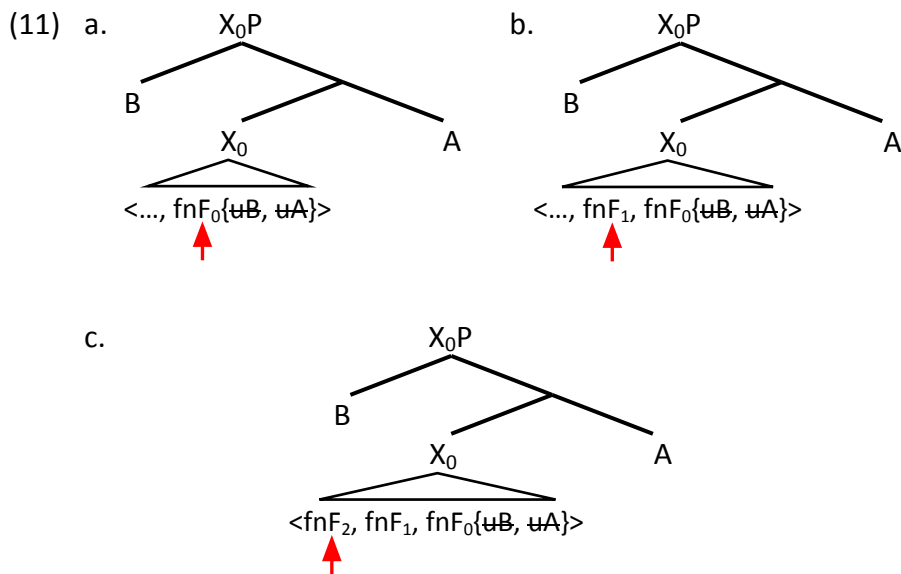
result in an (intermediate) structure as in (10a). Then the derivation will move on to the sub-features of  $fnF_1$ . Let us assume that  $fnF_1$  carries an uninterpretable sub-feature  $uC$  that cannot be matched against the interpretable features of any of the existing constituents of the structure in that point of the derivation (i.e. (10b)). Note, further, that in (10b) both the complement and the specifier of  $X_0$  are filled, so there are no available structural positions<sup>15</sup> to merge an appropriate constituent with a matching interpretable feature (i.e. C). Thus, Feature Scattering removes the segment of the featural specification of  $X_0$  that contains the target uninterpretable sub-feature (i.e.  $\langle fnF_2\{uD\}, fnF_1\{uC\} \rangle$ ) and it creates a new functional head  $X_1$  out of the scattered features. Then, the uninterpretable feature  $uC$ , which is now part of the featural specification of the head  $X_1$  rather than  $X_0$ , can trigger the merger of a constituent C at the specifier of  $X_1$ . The result of these operations will be the (intermediate) structure in (10c). Finally, the derivation will proceed to the sub-features of  $fnF_2$ . If we assume that  $fnF_2$  carries an uninterpretable sub-feature  $uD$  that, similarly to  $uC$  earlier, does not match any of the interpretable features present in the structure in (10d), then the same chain of events will repeat itself. Thus, we arrive at the final, fully articulated functional structure in (10e).



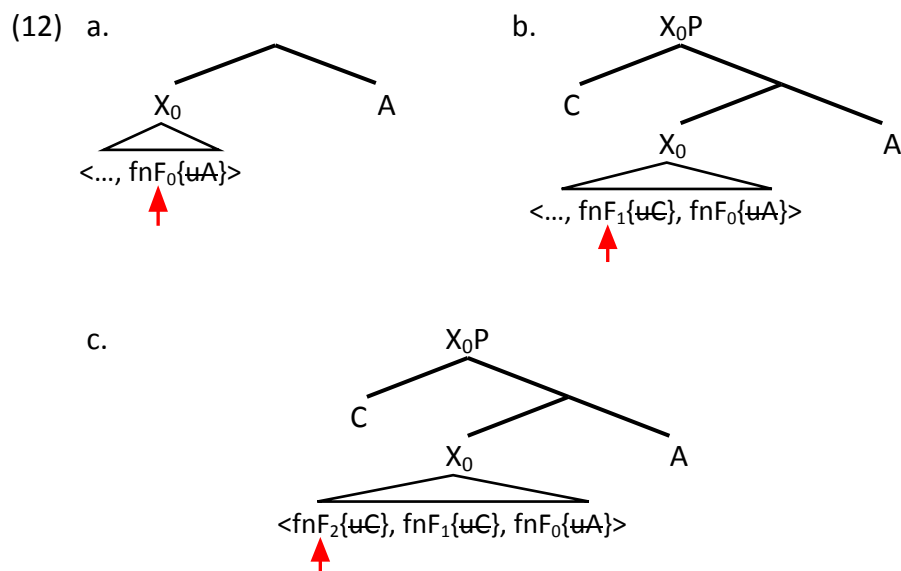
<sup>15</sup> Note that I am following Kayne (1994) and Cinque (1999) in assuming a single specifier per head, pace Chomsky (1995).



The same lexical item can give rise to a functional structure consisting of a single maximal projection, as long as its variable properties (i.e. its sub-features) allow it. This situation can arise in more than one ways. The simplest case would be if the functional features  $fnF_1$  and  $fnF_2$  did not have any uninterpretable sub-features, which entirely eliminates the possibility that they would undergo Feature Scattering. The derivation in (11) demonstrates this possibility (the step of the derivation shown in (11a) is the same as (10a) above; then the derivation proceeds to the functional features  $fnF_1$  and  $fnF_2$ , in (11b) and (11c) respectively, but there are no operations to be performed in either of those steps).



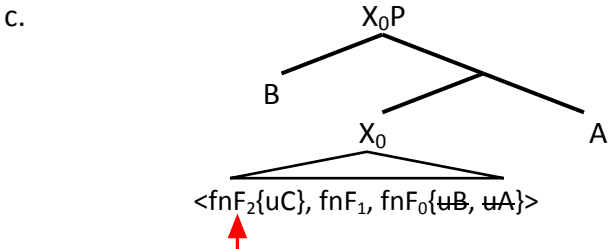
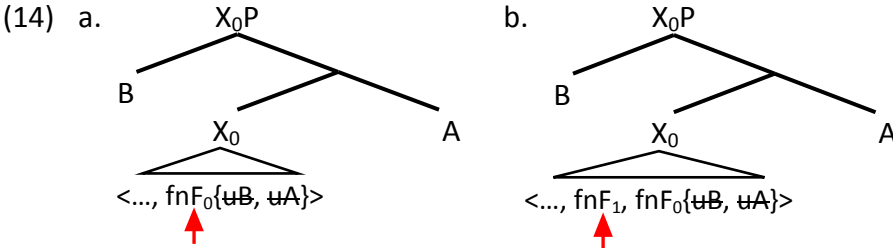
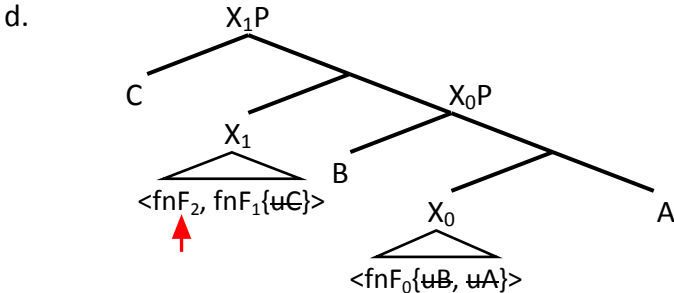
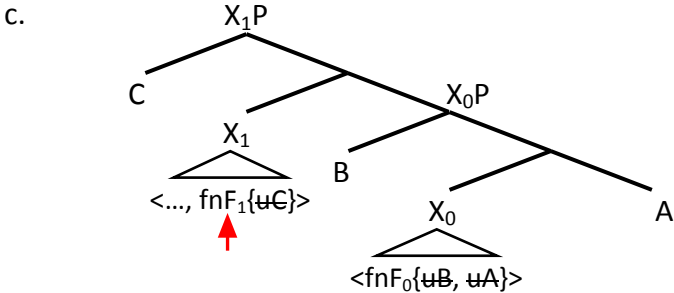
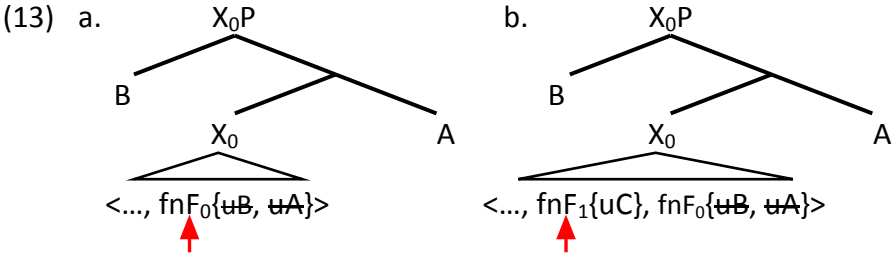
But the same result can arise even if  $fnF_1$  and  $fnF_2$  have uninterpretable sub-features, as long as those features can establish the necessary syntactic relations with their interpretable counterparts from the position of the head  $X_0$  (i.e. without needing to be scattered to a different structural position). Assume, for example, that the functional feature  $fnF_0$  has only one uninterpretable sub-feature (e.g.  $uA$ ). Accordingly, the first step of the derivation will produce the structure in (12a), where the specifier of  $X_0$  is still available. Then the derivation will proceed to the sub-features of  $fnF_1$ , for example an uninterpretable  $uC$ . Since a constituent with a matching interpretable feature (i.e.  $C$ ) can be directly merged as the specifier of  $X_0$ , the syntactic relation between  $uC$  and  $C$  can be established without resorting to Feature Scattering, as shown in (12b). In order to demonstrate yet another circumstance where Feature Scattering is pre-empted, let us assume that the remaining functional feature  $fnF_2$  has an uninterpretable sub-feature that matches one of the interpretable features already present in the structure. For example, assume that  $fnF_2$  has another instance of the uninterpretable feature  $uC$ . Then, (the second instance of)  $uC$  can enter a syntactic relation with its interpretable counterpart that is already available at the specifier of  $X_0$ , without expanding the structure in any way (cf. (12c)). Thus, (12b) and (12c) show two additional situations where Feature Scattering is superfluous and, therefore, excluded.

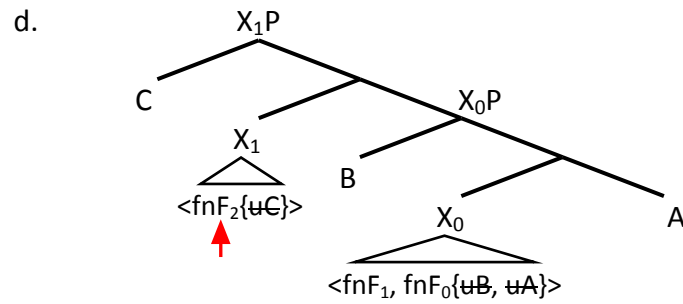


Finally, the same lexical item, with its three functional heads, may result in a functional structure with two maximal projections if one step of the derivation involves Feature Scattering (as in (10)), while the other does not (as in (11)-(12)). Note, however, that the resulting structure can take different forms depending on the timing of the Feature Scattering operation. I demonstrate the two possibilities in the derivations in (13) and (14) below. In (13), Feature Scattering is targeting the functional feature  $fnF_1$  (in steps (13b-c) of the derivation), which is assumed to have an uninterpretable sub-feature, while  $fnF_2$  does not have any sub-features to be scattered (step (13d)). Conversely, in (14), I assume that it is  $fnF_2$  that has an uninterpretable sub-feature triggering Feature Scattering (in steps (14c-d)), while  $fnF_1$  is unaffected (step (14b)). Observe that, as a result of this difference in the timing of operations, the heads  $X_0$  and  $X_1$ <sup>16</sup> end up with different featural specifications in the two derivations. At the end of the derivation in (13) the lexical head  $X_0$  is specified as  $\langle fnF_0\{\dots\} \rangle$  and the derived functional head  $X_1$  is specified as  $\langle fnF_2\{\dots\}, fnF_1\{\dots\} \rangle$ , while in (14) the lexical head is  $X_0\langle fnF_1\{\dots\}, fnF_0\{\dots\} \rangle$  and the derived functional head is  $X_1\langle fnF_2\{\dots\} \rangle$ . This subtle difference, in two otherwise very

<sup>16</sup> Note that the subscript on X simply indicates how many derived heads ( $X_1$ ,  $X_2$ , etc) have been scattered from a certain lexical item ( $X_0$ ) and it is independent of the subscript of functional features  $fnF$ .

similar structures, is an excellent demonstration of the flexible nature of heads in the theory developed here.





To summarize this section, the proposals in (7) to (9) outline a theory of functional structure according to which the syntactic relations and distribution of lexical items are not facilitated by prefabricated covert functional positions. Rather, the local syntactic relations involving a given (lexical or functional) head are determined by the un/interpretable features in its featural specification, the non-local relations involving the same head result from the scattering of un/interpretable features, and, finally, the distribution of said head is determined by the sequence of functional features in its featural specification.

#### **1.4.2 The 'Extended' Featural Specification of Non-Derived Functional Heads**

In the previous section I made two fundamental proposals. The first one was that lexical items are sequences of functional features that have a set of un/interpretable features as their sub-features. The second one was a revision of Giorgi and Pianesi's (1997) Feature Scattering that is compatible with the new conception of the featural specification of lexical items. I developed these proposals on the basis of a simple clause that contains a single non-derived head, the lexical verb, and a number of null derived functional heads (or, equally, the extended projection of a different category that happens to contain a single lexical item). In this section, I will extend the analysis to account for structures that contain one or more overt non-derived functional heads (e.g. auxiliary verbs, complementizers, etc) aside from the lexical head of the extended projection. There are two issues that need to be addressed to achieve this goal. Firstly, we need to posit a featural specification for overt non-derived functional heads that is consistent with the feature system developed in the previous section. In the following discussion, I will

argue that overt non-derived functional heads are a sequence of functional features (with un/interpretable sub-features), just like lexical heads. This hypothesis will lead us to the second issue: ensuring that there is no overlap in the sequence of functional features of an overt non-derived functional head and a lexical head that are part of the same derivation. In order to achieve this, I will propose some further amendments to the hypotheses relating to the nature and properties of lexical items that I have developed so far.

As I have argued in section 1.3.2, the hypothesis that overt functional heads are derived in the same way as their null counterparts (i.e. via Feature Scattering) is incompatible with the fundamental assumptions of this analysis. Rather, I treat overt functional heads as lexical items (i.e. non-derived heads that are stored in the mental lexicon). Thus, it is reasonable to expect that overt non-derived functional heads have their own set of syntactic properties, presumably expressed by means of a featural specification. The question then is what this featural specification looks like. There are two things to take under consideration. Firstly, an overt non-derived functional head (of a given category) has its own distributional properties that distinguish it from other (non-derived) functional or lexical heads. Secondly, it is generally assumed that derivations proceed by extending the root of the structure in a bottom-up fashion (cf. Chomsky 1995). Within the feature system adopted here, this view translates to the assumption that syntactic operations are always triggered by the (features of the) head closest to the root of the structure. Thus, we would expect that the distribution of a non-derived functional head is determined by its own featural specification, not by the lexical head (or any other functional head) below it. Consider, for example, the distribution of an overt non-derived functional head in relation to a null derived functional head above it<sup>17</sup>, which in this system is established via Feature Scattering. To make the example more concrete,

---

<sup>17</sup> Note that this is a rather narrow example and there are another two cases to consider: firstly, the distribution of a non-derived functional head in relation to other non-derived heads above it and, secondly, the distribution of a non-derived functional head in relation to the (non-derived or derived) heads below it. The later case is addressed in the next paragraph, while the former will be addressed in the following discussion (cf. the discussion around the Inventory of Functional Features introduced in (15) and amended in (31)).

assume a clause consisting of a lexical verb, an overt non-derived functional head corresponding to T (i.e. some auxiliary or modal verb), and a null derived functional head corresponding to C. Obviously, the derived head C must be scattered from the non-derived head T, not the lexical verb that is lower down in the structure. The details of this derivation will be fleshed out later on, but for the time being this initial conclusion suffices to suggest that the featural specification of an overt non-derived functional head must contain its own sequence of functional features (needless to say that these functional features must also have their own sub-features, because otherwise the head would not be able to participate in any syntactic operation).

We can, in fact, be more precise about the featural specification of a non-derived functional head. Note that the sequence of functional features of a non-derived functional head must be fundamentally similar to the sequence of the lexical head of the same extended projection. This happens because the sequence of functional features of the two heads under consideration must be able to capture their distribution in relation to the same set of (derived or non-derived) heads that make their extended projection. Thus, we can model the featural specification of a non-derived functional head after the specification of the lexical head of the extended projection. However, that does not mean that the two featural specifications are identical. If that were the case, the distribution of the relevant lexical items would be indistinguishable. Since the featural specification of a non-derived functional head and that of a lexical head (of the same extended projection) cannot differ in the order of functional features, I will suggest that what differentiates them is their size. Specifically, taking once again under consideration the bottom-up nature of syntactic derivations, I assume that a non-derived functional head only needs to have access to the information relating to the hierarchical ordering of other functional heads above it (i.e. the ones that may be scattered from it). Or, to put it another way, the non-derived functional head in question does not need to have any information about the functional structure below it, because that part of the structure will be derived from some lower lexical



item (i.e. the lexical head or, possibly, some intermediate non-derived functional head). Thus, as a first approximation, I propose the following: assuming that the functional structure of a given extended projection is derived from the sequence of functional features  $\langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_m\{\dots\}, \dots, \text{fnF}_0\{\dots\} \rangle$ , the featural specification of a non-derived functional head  $Y_0$  is a sequence of functional features that includes  $\text{fnF}_m$  and all of its superordinate functional features (as shown in (15b)), where  $\text{fnF}_m$  is what traditionally would be considered the categorial feature of  $Y_0$ . For comparison, (15a) is the featural specification of a lexical head  $X_0$  which, according to the proposal in (7), contains the entire sequence of functional features of the relevant extended projection.

- (15) a. Lexical head:  $X_0 \langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_m\{\dots\}, \dots, \text{fnF}_0\{\dots\} \rangle$   
 b. Non-derived functional head:  $Y_0 \langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_m\{\dots\} \rangle$

Note, however, that there is one residual problem with the featural specifications in (15): there is some overlap in the sequence of functional features of the lexical head  $X_0$  and the functional head  $Y_0$ , which entails that some of those functional features are reduplicated. In fact, this is not simply a matter of redundancy in the proposed featural specifications. The problem runs deeper. Remember that in the previous section I argued that a lexical head must have a featural specification like the one in (15a) in a derivation that does not include any non-derived functional heads. But, in a derivation that also includes a non-derived functional head, the featural specification in (15a) becomes problematic because a derived functional head corresponding to, for example,  $\langle \text{fnF}_n\{\dots\} \rangle$  can, in principle, be scattered twice: once from the lexical head  $X_0$  (if it is specified as in (15a)) and once from the non-derived functional head  $Y_0$ . Furthermore, a derived functional head that is scattered from the lexical head  $X_0$  and is specified as  $X_1 \langle \text{fnF}_n\{\dots\} \rangle$  will necessarily appear below the non-derived functional head  $Y_0$ , which entails that the hierarchical ordering of  $Y_0$  and  $X_1$  will be incorrect. Thus, I conclude that we need some mechanism that dynamically adjusts the featural specification of lexical items according to the other lexical items that participate in a given derivation.

The first step towards solving the aforementioned problem is the following. I propose that alongside the lexicon proper, there exists a functional ‘lexicon’ that I will call the Inventory of Functional Features (IFF). The IFF contains functional features that are listed together with their sub-features. As I have suggested in the previous section, the sub-features of any given functional feature may vary across languages, or even structures. Accordingly, I will assume that the IFF is a language-specific inventory. For example, the IFF of a language that displays wh-fronting in interrogative clauses will include a functional feature C with a uWh sub-feature, whereas the IFF of a language that does not display wh-fronting will include a different version of C that does not bear a uWh. Furthermore, the IFF of a given language may contain multiple versions of a single functional feature that are differentiated by their sub-features. For example, the IFF of a wh-fronting language will include the aforementioned interrogative version of C that bears uWh, but it will also include a declarative version of C that lacks uWh. Finally, I will propose that functional features inside the IFF are organised in a number of separate sequences that correspond to the extended projections of various categories. I will schematically represent the IFF as in (16), where (16a) is the sequence of functional features corresponding to the extended projection of a lexical item of category F, (16b) is the sequence of an extended projection of category G, etc.

- (16) Inventory of Functional Features
- a. <fnF<sub>n</sub>{...}, ... , fnF<sub>m</sub>{...}, ... , fnF<sub>0</sub>{...}>
  - b. <fnG<sub>k</sub>{...}, ... , fnG<sub>j</sub>{...}, ... , fnG<sub>0</sub>{...}>
  - c. ...

As the above description suggests, the IFF essentially provides templates for the featural specifications of various lexical items. With that in mind, I will propose that lexical items are stored in the lexicon with an underspecified featural specification, since missing information can be recovered from the IFF. For the purposes of this discussion, I will simply note that the featural specification of a lexical item has to include at a minimum one functional feature corresponding to its traditional category (but see section 2.3.2 for an important amendment). So, the

heads in (15) will be stored in the lexicon with the featural specifications shown in (17), where the lexical head  $X_0$  is specified with  $fnF_0$  and the non-derived functional head  $Y_0$  is specified with  $fnF_m$ .

- (17) Lexicon
- a. Lexical head:  $X_0 <fnF_0\{\dots\}>$
  - b. Non-derived functional head:  $Y_0 <fnF_m\{\dots\}>$

I will then propose that the numeration of a given derivation includes a number of lexical items and one or more sequences of functional features taken from IFF as necessary. For example, (18) represents a (partial) numeration from which we can derive a structure consisting of two lexical items that belong to the same extended projection of category F. Finally, I will propose a pre-syntactic operation (i.e. one that takes place in the numeration) called Feature Bundling that maps each sequence of functional features selected in the numeration onto the lexical items with a matching functional feature that are available in the same numeration. To be more precise, I assume that this process goes through each sequence of functional features (e.g. (18a)), starting from the functional feature with the lower subscript (in this example  $fnF_0$ ) and going upwards. Therefore, the first step of this process will map  $fnF_0$  onto the lexical head of the extended projection (i.e.  $X_0$ ), which, presumably, cannot be absent from the numeration. Then, Feature Bundling will continue mapping the functional features of the same sequence onto  $X_0$  (in the appropriate order) until it reaches a functional feature  $fnF_m$ , such that there is a non-derived functional head with a matching feature (in this example  $Y_0$ ) available in the numeration. Obviously,  $fnF_m$  will be mapped onto  $Y_0$ . After that point, Feature Bundling will keep mapping functional features onto  $Y_0$  (always in the appropriate order), either until the sequence of functional features is exhausted, or the operation comes across another functional feature that can be mapped onto a lexical item with a matching feature. In our example, the resulting featural specifications, which will eventually be used in the derivation, will have the

form in (19). Finally, (20) provides a formal description of the operation of Feature Bundling<sup>18</sup>.

(18) Numeration

- a.  $\langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_m\{\dots\}, \dots, \text{fnF}_0\{\dots\} \rangle$
- b.  $X_0 \langle \text{fnF}_0\{\dots\} \rangle$
- c.  $Y_0 \langle \text{fnF}_m\{\dots\} \rangle$

(19) Numeration (after Feature Bundling)

- a.  $X_0 \langle \text{fnF}_{m-1}\{\dots\}, \dots, \text{fnF}_0\{\dots\} \rangle$
- b.  $Y_0 \langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_m\{\dots\} \rangle$

(20) Feature Bundling

In a numeration containing a sequence of functional features  $\langle \text{fnF}_n\{\dots\}, \dots, \text{fnF}_0\{\dots\} \rangle$ , a functional feature  $\text{fnF}_x\{\dots\}$  (where  $n \geq x \geq 0$ ) is mapped onto a lexical item bearing a functional feature  $\text{fnF}_y$ , iff:

- i.  $x \geq y$ , and
- ii. there is no lexical item bearing a functional feature  $\text{fnF}_z$  where  $x \geq z > y$ .

Now that we have seen exactly how the interaction of the Inventory of Functional Features and Feature Bundling determines the featural specification of various heads, I would like to take a step back and provide some additional comments and clarifications on the concept of the IFF. Firstly, remember that I have previously acknowledged that my formulation of the Sequence of Operations (cf. (8) in section 1.4.1) captures only one aspect of Giorgi and Pianesi's (1997) Universal Ordering Constraint; i.e. the timing of syntactic operations. What remained an open question in that earlier discussion is what determines the internal structure of a head (i.e. its sequence of functional features). As we have just seen, the IFF addresses exactly that omission. Secondly, given the current status of the literature on the hierarchy of the functional structure, the null hypothesis would seem to be

---

<sup>18</sup> Giorgi and Pianesi (1997:231) propose that there is a second version of Feature Scattering that takes place in the numeration, rather than narrow syntax. They call this operation Feature Scatter B. Feature Bundling is the opposite of Feature Scatter B, in the sense that the later decomposes syncretic heads (to use Giorgi and Pianesi's terminology) into atomic heads in the numeration, while the former combines atomic elements (i.e. functional features) into a complex head.

that the ordering of functional features within the IFF is universal (cf. section 1.2.2). Of course, as I have mentioned in section 1.2.2, I will leave open the question of whether this order is a primitive of UG or derived from more fundamental properties of the language faculty, since it extends far beyond the scope of this dissertation. Similarly, I will not address the possibility that there is room for some limited variation in the ordering of functional features within IFF (as suggested by Ouhalla 1988, 1991, 1994, Nilsen 2003, Ramchand and Svenonius 2014), although it is, in principle, compatible with the analysis developed here.

With those clarifications out of the way, we can now go over a few example derivations to see how this analysis works in practice. Let us assume that the IFF includes (amongst other sequences) a sequence of three functional features as in (21), which corresponds to an extended projection similar to the one in examples (10)-(14) above.

- (21) Inventory of Functional Features  
 <fnF<sub>2</sub>{...}, fnF<sub>1</sub>{...}, fnF<sub>0</sub>{...}>

The derivations that may arise from this starting point depend on the number of lexical items that are included in the numeration. For example, the numeration in (22) contains the sequence of functional features from (21) and a single lexical item, namely the lexical head X<sub>0</sub> of the extended projection (which, presumably, will initially be specified with the functional feature fnF<sub>0</sub>, as in (22b)). In this case, Feature Bundling will map the entire sequence of functional features onto X<sub>0</sub>, producing the featural specification in (23). Of course, this is the same basic featural specification that gave rise to all of the derivations in (10)-(14) in the previous section, depending on the sub-features that were associated with each functional feature.

- (22) Numeration 1  
 a. <fnF<sub>2</sub>{...}, fnF<sub>1</sub>{...}, fnF<sub>0</sub>{...}>  
 b. X<sub>0</sub><fnF<sub>0</sub>{...}>
- (23) Numeration 1 (after Feature Bundling)  
 a. X<sub>0</sub><fnF<sub>2</sub>{...}, fnF<sub>1</sub>{...}, fnF<sub>0</sub>{...}>

Consider now the numeration in (24), which includes three lexical items: the lexical head  $X_0$  and the overt non-derived functional heads  $Y_0$  and  $Z_0$ . Here, Feature Bundling will map each one of the functional features of the sequence in (24a) to a separate head, producing the numeration in (25).

- (24) Numeration 2
- a.  $\langle \text{fnF}_2\{\dots\}, \text{fnF}_1\{\dots\}, \text{fnF}_0\{\dots\} \rangle$
  - b.  $X_0\langle \text{fnF}_0\{\dots\} \rangle$
  - c.  $Y_0\langle \text{fnF}_1\{\dots\} \rangle$
  - d.  $Z_0\langle \text{fnF}_2\{\dots\} \rangle$

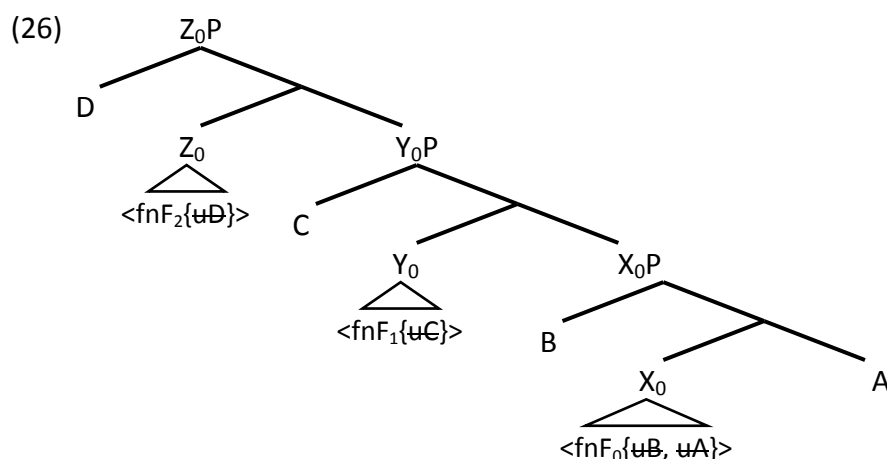
- (25) Numeration 2 (after Feature Bundling)
- a.  $X_0\langle \text{fnF}_0\{\dots\} \rangle$
  - b.  $Y_0\langle \text{fnF}_1\{\dots\} \rangle$
  - c.  $Z_0\langle \text{fnF}_2\{\dots\} \rangle$

The derivation that will arise from this numeration will be much more similar to the derivations of a standard minimalist analysis, than to the derivations demonstrated in the previous section. As we can see in (26), the uninterpretable features  $uA$  and  $uB$  of  $X_0$  are eliminated against the matching interpretable features of the constituents  $A$  and  $B$  that are merged in the complement and specifier position of  $X_0$ , thus projecting an  $X_0P$ . Then,  $X_0P$  will be merged with  $Y_0$ , which takes a specifier of its own and projects a  $Y_0P$ . Finally, the same steps are repeated with  $Z_0$ , which is merged with  $Y_0P$  and a specifier and projects  $Z_0P$ . What is interesting about (26) is that the overt non-derived functional heads  $Y_0$  and  $Z_0$  are essentially of the same category as the null derived functional heads  $X_1$  and  $X_2$  (respectively) in (10), by virtue of the fact that the latter have, by the end of the derivation, the same functional features as the former. Thus, in spite of their different origin, in terms of their distribution null derived functional heads are fundamentally the counterparts of overt non-derived functional heads<sup>19</sup>. Finally, note that the merger

---

<sup>19</sup> Note that in the schematic representation I annotate null derived functional heads with the same letter (and an additional numeric subscript) as the head they are scattered from (e.g.  $X_1$  is scattered from  $X_0$ ,  $X_2$  is scattered from  $X_1$ , etc). On the other hand, overt non-derived functional heads are annotated by their own letter (e.g.  $Y_0$  and  $Z_0$ ). Note, further, that the maximal projection of a null derived functional head is connected with a continuous line with the maximal projection of the head it is scattered from, while the

of an overt non-derived functional head does not depend on its uninterpretable features in the way that the scattering of a null derived functional head depends on the uninterpretable features of its source head. Thus, in (26), any of the uninterpretable features  $uA$ - $uD$  may be missing (and, consequently, any of the structural positions A-D may remain unfilled), without altering the status of  $Y_0$  and  $Z_0$  as independent heads.



The last possibility to consider is a numeration that includes two lexical items: the lexical head and one overt non-derived functional head. In fact, there are two subcases here. The numeration in (27) includes an overt non-derived functional head  $Y_0$  specified with  $fnF_1$ , while the numeration in (29) includes a head  $Z_0$  specified with  $fnF_2$ . Consequently, the numerations that will arise after Feature Bundling will also be different. As shown in (28) and (30) (which result from (27) and (29), respectively), both the featural specification of the lexical head  $X_0$  and the featural specifications of the non-derived functional heads  $Y_0$  and  $Z_0$  are different between the two numerations.

- (27) Numeration 3
- a.  $\langle fnF_2\{\dots\}, fnF_1\{\dots\}, fnF_0\{\dots\} \rangle$
  - b.  $X_0 \langle fnF_0\{\dots\} \rangle$
  - c.  $Y_0 \langle fnF_1\{\dots\} \rangle$

---

maximal projection of an overt non-derived functional head is not connected in this way (compare the trees in (8) and (24)). These are simply annotation devices to easily identify null derived functional heads, as well as the head from which they are scattered.

(28) Numeration 3 (after Feature Bundling)

- a.  $X_0 <fnF_0\{\dots\}>$
- b.  $Y_0 <fnF_2\{\dots\}, fnF_1\{\dots\}>$

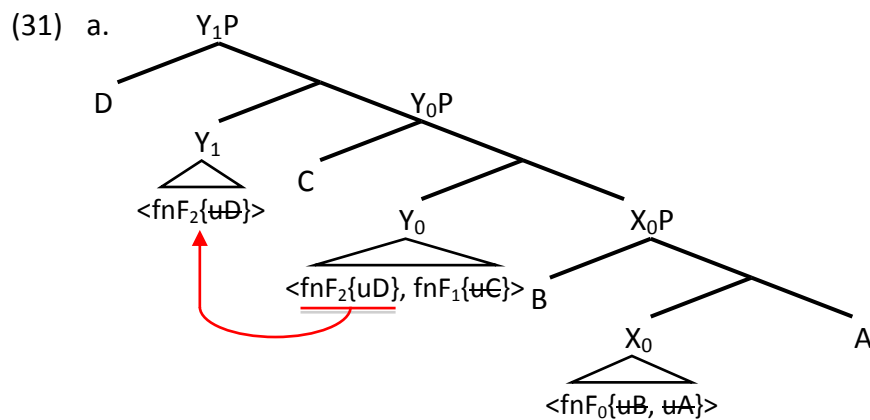
(29) Numeration 4

- a.  $<fnF_2\{\dots\}, fnF_1\{\dots\}, fnF_0\{\dots\}>$
- b.  $X_0 <fnF_0\{\dots\}>$
- c.  $Z_0 <fnF_2\{\dots\}>$

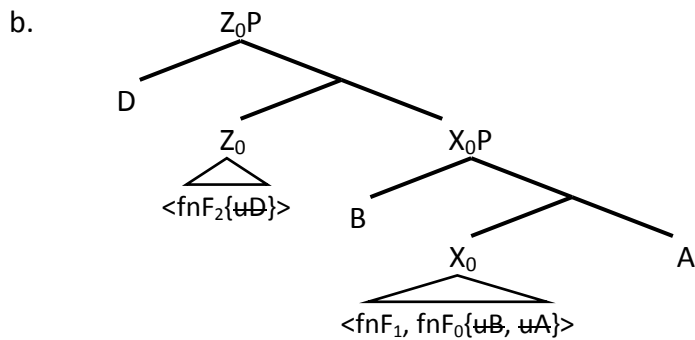
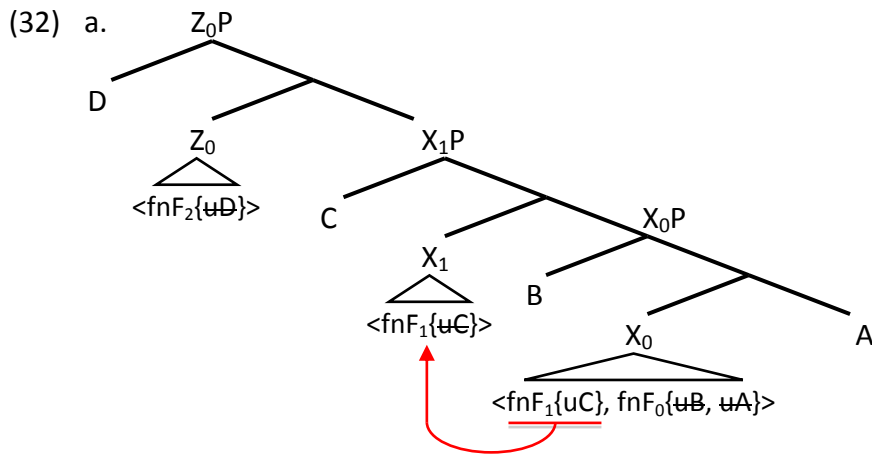
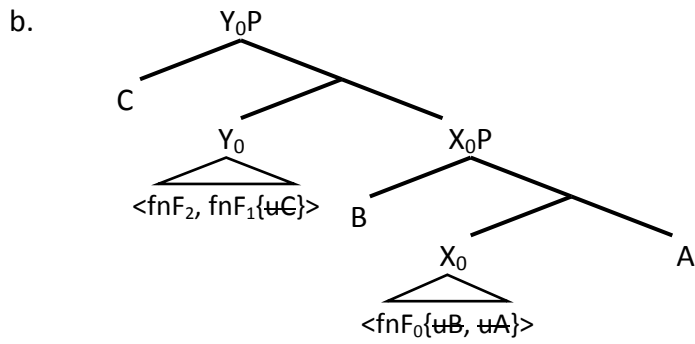
(30) Numeration 4 (after Feature Bundling)

- a.  $X_0 <fnF_1\{\dots\}, fnF_0\{\dots\}>$
- b.  $Z_0 <fnF_2\{\dots\}>$

Furthermore, the two numerations in (28) and (30) may give rise to two derivations each, depending on whether the option of Feature Scattering will be taken or not (in the first case, it is  $Y_0$  that is a potential candidate for Feature Scattering, while in the second case it is  $X_0$ ). Thus, the structures in (31a) (where  $Y_1$  has been scattered from  $Y_0$ ) and (31b) (which does not involve Feature Scattering) are derived from (28), while the structures in (32a) (which involves the scattering of  $Z_1$  from  $Z_0$ ) and (32b) (with no Feature Scattering) are derived from (30).







I will conclude this section with a brief discussion of a residual issue of the analysis developed above. In the previous derivations I took the merger of overt non-derived functional heads for granted. It is, however, generally acknowledged that providing an appropriate formalism for that operation is not a straightforward task, given that any such mechanism must be capable of capturing the correct hierarchical order of functional heads. One approach, which builds on Grimshaw's (1991, 2000) observation that the complement-taking properties of functional heads differ from those of lexical heads, is to posit a formal distinction between

Head-Complement relations within an extended projection and Head-Complement relations across the boundaries of an extended projection. In the later case, merger is assumed to be triggered by the c-selectional properties of a lexical head (expressed in terms of uninterpretable feature or some similar technology), while in the former case merger is triggered by a universal constraint on the well-formedness of functional structure (akin to Cinque’s 1999 Universal Hierarchy of Functional Projections or Starke’s 2001, 2004 Functional Sequence). This approach is rather unappealing under the system developed here for two reasons. Firstly, the analysis defended here already differentiates between two distinct structure-building mechanisms (i.e. the introduction of a null derived functional heads in the structure is attributed to Feature Scattering and the merger of a head with a specifier or complement not belonging to its extended projection is attributed to standard c-selection). Thus, positing a third structure-building mechanism within this analysis seems like an unwelcomed addition. Additionally, this analysis already includes a constraint on the hierarchical organisation of functional features (i.e. the Inventory of Functional Features), which has been argued to determine the hierarchical order of one type of functional heads (namely null derived functional heads). Again, positing an additional constraint on the hierarchical order of overt non-derived functional heads (which is, of course, the same as the order of their null counterparts) would be a rather costly reduplication of the IFF.

Thus, it seems that the least costly assumption is to attribute the merger of overt non-derived functional heads to c-selection. Accordingly, I will make the tentative proposal that every functional feature  $fnF_m$  has one non-optional uninterpretable sub-feature  $ufnF_{m-1}$ , which is eliminated when it occupies a structural position that c-commands a matching interpretable feature  $fnF_{m-1}$  (i.e. the immediately subordinate functional feature in the sequence). Thus, the IFF in (16) should be amended as in (33) below.

- (33) Inventory of Functional Features - amended:  
 $\langle fnF_n\{ufnF_{n-1}, \dots\}, \dots, fnF_m\{ufnF_{m-1}, \dots\}, \dots, fnF_0\{\dots\} \rangle$

The uninterpretable feature  $u\text{fn}F_{m-1}$  may behave in one of two ways, depending on its placement in relation to the matching interpretable feature  $\text{fn}F_{m-1}$ . The first case is when the functional features  $\text{fn}F_m$  (together with its sub-feature  $u\text{fn}F_{m-1}$ ) and  $\text{fn}F_{m-1}$  are mapped onto different lexical items in the numeration. The derivation in (26) above provides us with a couple of examples of this situation. In this derivation, the non-derived functional head  $Y_0$  will have the featural specification  $\langle \text{fn}F_1\{u\text{fn}F_0, \dots\} \rangle$ , where  $u\text{fn}F_0$  matches the functional feature of the lexical head  $X_0\langle \text{fn}F_0\{\dots\} \rangle$ . Similarly, the non-derived functional head  $Z_0$  will be specified as  $\langle \text{fn}F_2\{u\text{fn}F_1, \dots\} \rangle$ , where  $u\text{fn}F_1$  matches the functional feature of  $Y_0$ . Thus, in (26), we can attribute the merger of  $Y_0$  with  $X_0P$ , as well as the merger of  $Z_0$  with  $Y_0P$ , to their respective uninterpretable features. Interestingly, there is another example of this situation that is worth consideration. In (32b), the non-derived functional head  $Z_0$  again has the featural specification  $\langle \text{fn}F_2\{u\text{fn}F_1, \dots\} \rangle$ . In this derivation, however,  $Z_0$  is merged with  $X_0P$ , rather than  $Y_0P$  (since the numeration does not include  $Y_0$ ). This is made possible by the fact that in this numeration the functional feature  $\text{fn}F_1$  is mapped onto  $X_0$  (alongside the  $\text{fn}F_0$ ). This is a welcome result, since we manage to capture the fact that there is a certain degree of optionality in what constitutes a licit complement of a functional head without having to stipulate any variation in its uninterpretable features (something that is quite difficult to achieve in a standard minimalist analysis). Rather, it is the featural specification of the complement that varies, as a natural consequence of the analysis developed earlier in this section.

The second case that we need to consider is a numeration where the functional features  $\text{fn}F_m$  (with its sub-feature  $u\text{fn}F_{m-1}$ ) and  $\text{fn}F_{m-1}$  are mapped onto the same lexical item. We have seen various examples of this, but let us consider the lexical head  $X_0$  in (32b), which has two functional features and, therefore, will have the featural specification  $\langle \text{fn}F_1\{u\text{fn}F_0, \dots\}, \text{fn}F_0\{\dots\} \rangle$ . The puzzle we face here is this: how is the uninterpretable feature  $u\text{fn}F_0$  eliminated, considering that the matching interpretable feature is located in the same head? In order to uphold the proposed derivation of (32b), it has to be the case that  $u\text{fn}F_0$  (together with the

host functional feature  $fnF_1$ ) does not need to be scattered to a position c-commanding  $fnF_0$ . To see why this is important, consider that if we extrapolate from this example to the general case we arrive at a situation where the scattering of functional features is no longer optional, which is a conclusion that runs contrary to the goals of this analysis. Thus, it seems that we have to allow for the possibility that an uninterpretable feature can be trivially eliminated against a matching interpretable feature that happens to appear on the same head. In principle, we can achieve this in two ways: (i) we can stipulate that the domain of the operation triggered by the uninterpretable feature  $unfnF_{m-1}$  includes the c-command domain of the head hosting  $unfnF_{m-1}$ , as well as the head itself, or (ii) we can construe c-command as a reflexive relation, so that the head hosting  $unfnF_{m-1}$  c-commands itself by definition<sup>20</sup>. Since nothing significant hinges on this point, I will tentatively accept the later, less stipulative, approach (although I will acknowledge that this approach might also have more subtle and far-reaching implications).

To recap, in this section I have made a number of proposals concerning the featural specification of lexical items before they enter the syntactic derivation. The goal of these proposals is twofold: Firstly, we need to ensure that any null functional head can, in principle, be derived from the featural specification of the lexical items (i.e. lexical heads and overt non-derived functional heads) that are available in a given numeration and, secondly, we need to avoid any redundancy in the featural specification of said lexical items which could potentially lead to the projection of redundant functional positions. To achieve this, I have suggested that lexical items are, in a sense, underspecified in the lexicon. Their full featural specification is determined in (and depends on) the numeration. The process by which lexical items acquire their full featural specification involves two hypotheses. Firstly, I have proposed that, apart from lexical items, the numeration also includes

---

<sup>20</sup> In turn, there are at least two straightforward ways to define c-command as a reflexive relation. Starting from a usual definition of c-command according to which a node  $\alpha$  c-commands a node  $\beta$  iff (i)  $\alpha$  and  $\beta$  do not dominate each other and (ii) every node dominating  $\alpha$  also dominates  $\beta$  (cf. Chomsky 1995), we can either adopt a non-reflexive definition of dominance so that  $\alpha$  does not dominate itself, therefore satisfying (i), or, alternatively, we can entirely remove the stipulation in (i) (cf. Reinhart 1981).

functional features drawn from the Inventory of Functional Features in (16). Secondly, I have proposed that the pre-syntactic operation Feature Bundling in (20) maps the functional features of the IFF onto the lexical items that are available in a given numeration. The mapping is such that each lexical item has the necessary functional features to derive (provided that the appropriate conditions for Feature Scattering are met) the covert functional structure between itself and the next higher lexical item. Thus, Feature Bundling ‘collapses’ functional features on the non-derived (lexical or functional) heads that are available in a given numeration, ensuring that there are no superfluous (null) functional heads at the beginning of a derivation. As the derivation proceeds, Feature Scattering may ‘expand’ those functional features into derived functional heads, but only when the additional structural positions are independently needed.

### **1.5 Conclusion**

In this chapter, I proposed a novel analysis of functional structure that follows the tradition of re-projective theories on the matter. The discussion started with an overview of some of the early literature on functional structure (section 1.2.1), which served as an introduction to the wide array of theoretical issues that are relevant to this field of inquiry. In section 1.2.2, I introduced the question of the (non-)universality of functional structure and I clarified that it pertains to (i) the hierarchical ordering of functional projections, and (ii) the instantiation of those functional projections. The relevant literature generally accepts that the hierarchical ordering of functional projections is indeed universal. On the contrary, the (non-)universality of the instantiation of functional projections is a hotly debated topic. In section 1.2.3, I demonstrated that the issue of the (non-)universality of functional structure intersects in important ways with another theoretical issue: namely our understanding of what is the smallest component of a syntactic structure (i.e. whether heads are atomic elements that express a singular feature or complex elements that express a multitude of features).

From section 1.3 onwards, the discussion focused on a subset of those theoretical questions and assumptions. Specifically, I have pursued the path of a variationist (i.e. the view that the instantiation of functional heads is variable) and lexicalist (i.e. the view that heads are complex, internally structured objects) approach to functional structure. In section 1.3.1, I provided an overview of Giorgi and Pianesi's (1997) analysis of functional structure as an example of the re-projective approach on the matter. In section 1.3.2, I challenged the commonly held view that the mechanism of re-projection is only relevant for a subset of covert functional heads. I have argued that this approach is costly from a theoretical standpoint (especially in relation to the analysis of head movement), as well as untenable from the point of view of language acquisition (because it creates a situation where there are two competing, but practically indistinguishable, analyses of covert functional structure). On these grounds, I have concluded that it is profitable to extend the formulation of re-projection so as to capture covert functional structure in its entirety.

In section 1.4.1, I proposed a new analysis of functional structure that achieves the goal of eliminating null non-derived functional heads. This analysis is based on two fundamental hypotheses. Firstly, I proposed that the featural specification of lexical items is more expansive and more structured than previously assumed. Specifically, I suggested that lexical items consist of a sequence of functional features that each bears a set of un/interpretable sub-features. Furthermore, I posited that the purpose of the sequence of functional features is to determine the order of syntactic operations (an idea that is similar to Giorgi and Pianesi's (1997) Universal Ordering Constraint), while the function of uninterpretable sub-features is to trigger syntactic operations. Secondly, I proposed a refinement of Giorgi and Pianesi's (1997) Feature Scattering, so as to accommodate for the 'extended' featural specifications I adopted earlier. According to this formulation, Feature Scattering operates, in a recursive manner, on entire segments of the featural specification of a given head, rather than individual features. Finally, in section 1.4.2, I addressed the issue of the redundancy of

functional features that arises from the adoption of 'extended' featural specifications. I proposed that lexical items are stored in the lexicon with an underspecified featural specification. Furthermore, I have formulated a pre-syntactic operation called Feature Bundling that assigns complete featural specifications to lexical items, depending on the number and category of the lexical items that are available in a given numeration.

## Chapter 2

---

# Head Movement: A Feature Scattering Analysis

### **2.1 Introduction**

This chapter explores head movement, mostly from the perspective of the re-projective literature on functional structure. I have previously argued in favour of the view that all null functional heads ought to be derived by a re-projective mechanism. If this argument is on the right track, it implies that the mechanism of re-projection has a wider application than head movement. This observation stands in contrast to theories that derive re-projection from head movement (Ackema et al 1993, Bury 2003 and others). Rather, this view aligns with works that assume that head movement is constrained by an independent re-projective mechanism (Giorgi and Pianesi 1997, Nash and Rouveret 1997, 2002). In this chapter, I develop a stronger version of this approach, according to which head movement is derived from re-projection. This analysis is based on the simple hypothesis that the featural specification of lexical items also contains their phonological features. Therefore, even though phonological features do not participate in syntactic operations, they can be displaced incidentally by Feature Scattering operations that have been triggered for independent reasons. In the following discussion, I will flesh out this proposal and I will employ it in the analysis of V-to-C movement in Germanic V2 languages, V-to-T movement in Romance languages and V-to-v movement in English.

This chapter is organised as follows. Section 2.2.1 provides an outline of the properties of head movement that distinguish it from phrasal movement. Section



2.2.2 demonstrates that re-projective theories in general provide a principled account of many of the distinctive properties of head movement. However, in section 2.2.3, I argue that earlier re-projective theories face certain challenges. In the same section, I propose an alternative re-projective account of head movement, which resolves these issues. In section 2.3, I explore how this proposal applies to various cases of verb movement. Section 2.3.1 is focused on the V-to-C, V-to-T and V-to-v movement of main verbs in Germanic V2 languages, Romance languages and English respectively. In section 2.3.2, I address the movement of auxiliary verbs to C or T in the same languages. Finally, section 2.3.3 explores V-to-C movement, or lack thereof, in embedded clauses in Germanic languages.

## ***2.2 Head Movement in Syntactic Theory***

### ***2.2.1 Head Movement versus Phrasal Movement***

In the previous chapter (section 1.2.3), I have mentioned that within the broader literature adopting a variationist view of functional structure there is a strand of research that focuses on the intersection of the properties of functional structure with the properties of head movement. I have used the cover term ‘re-projective’ theories of head movement to refer to these works (Ackema et al 1993, Koenenman 2000, Bury 2003, Surányi 2005, amongst others). Moreover, I have argued that this approach, to the extent that it is successful in providing a uniform analysis of functional structure and head movement, provides an independent argument in favour of variation in functional structure. In the following sections, I will discuss head movement and its relation to the theory of functional structure in more detail. I will start the discussion with a brief overview of the properties of head movement, as understood in the generative tradition (section 2.2.1). I will then consider how the re-projective head movement approach fares in capturing the properties of head movement and I will conclude that the results are very promising (section 2.2.2). I will however point out that current re-projective analyses face certain limitations (section 2.2.3) and I will propose a different analysis of head movement that is based on the theory of functional structure

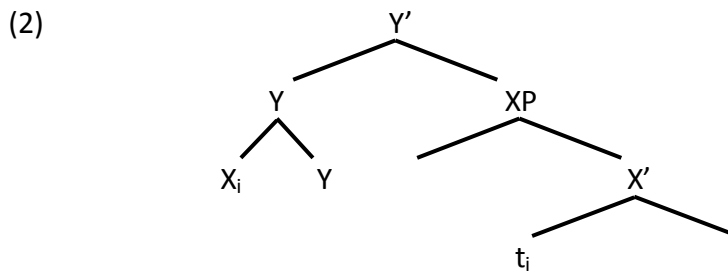
developed in section 1.4. In brief, I will show that head movement can be derived from Feature Scattering, as formulated in the previous chapter, and, therefore, there is no need to treat it as an independent syntactic operation. Finally, I will conclude this discussion with an outline of the advantages of the proposed analysis over the existing re-projective head movement theories.

At the core of the study of head movement lies the following point of friction. On one hand, there is reason to believe that there is a fundamental similarity between head movement and phrasal movement, in as much as both express the dissociation of the underlying and the surface position of various syntactic objects (either heads or phrases). This view has been particularly prominent in the government and binding literature, where head movement was commonly treated as an instance of the generalised Move  $\alpha$  operation (the extension of Move  $\alpha$  to head movement was pioneered by Koopman 1984, Travis 1984, and others). On the other hand, since the 1980s there have been mounting arguments (both empirically and theoretically motivated) demonstrating significant contrasts in the properties of head movement versus phrasal movement. The list in (1) summarizes the most widely acknowledged properties of head movement that differentiate it from phrasal movement ((1) is a compilation of Matushansky 2006 and Roberts 2001, 2011; cf. these works for more detailed discussion and references).

- (1) Properties of head movement:
  - a. Moved heads land in head positions whereas moved phrases land in specifier positions.
  - b. Movement of head X to head Y creates a constituent that excludes Y's complement.
  - c. After head movement of head X to head Y, subsequent instances of head movement can only move the X+Y complex; movement of head X while Y is stranded, or vice versa, is not allowed (i.e. excorporation is disallowed). Hence, head movement is roll-up rather than successive cyclic.
  - d. Movement of head X to head Y cannot skip intervening heads.

The property in (1a) was originally formalised by Emonds' (1970, 1976) Structure Preservation Hypothesis. Although this constraint has evolved with time (cf. Chomsky's 1995 Chain Uniformity Condition, for example), it has remained a

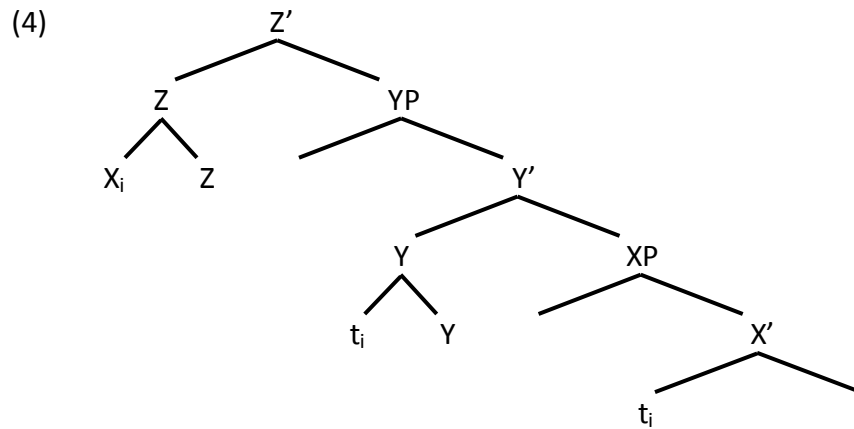
core concept in the theory of head movement. Regarding the property in (1b), in traditional theories it is captured by the head adjunction analysis of head movement originating in Baker (1988), which I have briefly outlined in section 1.2.3 of chapter 1 (i.e. the idea that a head X moves to, and adjoins with, a target head Y). The tree in (2) is a schematic representation of the head adjunction analysis, which demonstrates how this understanding of head movement captures the generalisation in (1b).



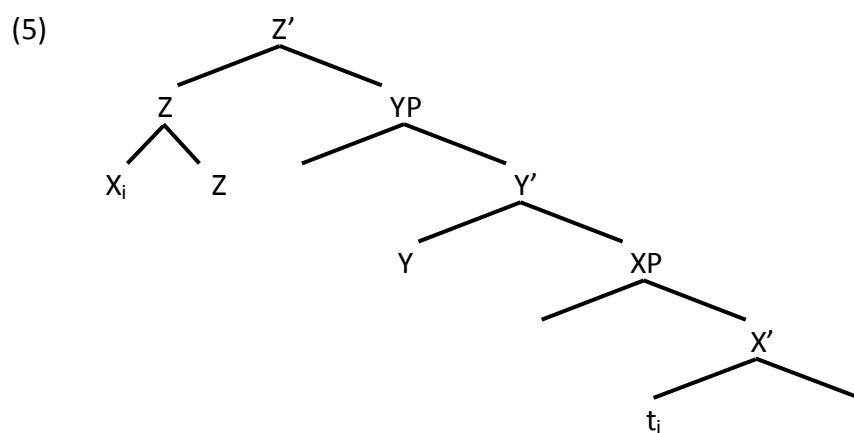
Interestingly, the head adjunction analysis of head movement leads to two additional theory internal problems, as noted by Chomsky (2001), which I list in (3). Note that these are only problematic for head movement and not phrasal movement.

- (3) Theory internal problems of the head adjunction analysis:
- a. Head movement violates the Extension Condition (i.e. adjunction of head A to head B does not extend the root of the tree).
  - b. A moved head does not c-command its trace.

Returning to (1), the generalisation in (1c) was originally put forward by Baker (1988) and attributed to a morphological constraint. Namely, Baker (1988) proposes that words cannot contain traces. Thus, on the assumption that head movement feeds affixation, the intermediate trace in (4) (which represents a hypothetical structure involving successive cyclic head movement) is illicit. Later analyses have attempted to derive (1c) from the Empty Category Principle (cf. Roberts 1991), but with the advent of minimalism and the abandonment of the ECP an updated narrow syntactic explanation of the ungrammaticality of the structure in (4) has not been forthcoming.



Finally, the generalisation in (1d) was originally formalised by Travis' (1984) Head Movement Constraint. Once again, the Head Movement Constraint has been derived from the ECP in the government and binding literature (cf. Roberts 1991). Note that in a structure like (5), where X has moved directly to Z skipping over Y, the head X does not properly govern its trace (assuming that Y is a barrier), thus violating the ECP. However, this is another instance of an analysis based on the ECP that that did not survive into the minimalist literature and has not been successfully replaced with a minimalist alternative. Finally, note that (1c) and (1d) are jointly responsible for excluding structures of the type  $[_{ZP} Z+X_i [_{VP} Y [_{XP} t_i]]]$  (i.e. long head movement): (1c) precludes the possibility of deriving such a structure via successive cyclic movement and (1d) precludes the possibility of deriving it via a single, long distance head movement operation.



Before concluding this section, I will briefly outline some further properties of head movement that I did not list previously because they are not strictly relevant to the following discussion, either due to the fact that they can be derived rather trivially in all the alternative analyses of head movement that I will consider or because they are rather controversial. Firstly, according to the government and binding literature, the ECP has the further consequence that the movement of a head X to a head Y can only take place from the complement of Y, but not from the specifier or an adjunct of Y. This theoretical conclusion originates in Baker (1988), who also provides empirical data supporting this generalisation. In a similar vein, the ECP also places a restriction on the direction of head movement so that a head X can only move upwards in the structure, but not downwards. Finally, Chomsky (2001) has argued that head movement differs from phrasal movement in that it does not affect semantic interpretation. This view, however, has been challenged convincingly in Matushansky (2006), Roberts (2010) and Lechner (2006) (cf. these works for a detailed discussion).

To summarise, in this section I have outlined the contemporary understanding of the properties of head movement. As I have pointed out, these conclusions have risen from a long tradition of comparing the properties of head movement with those of phrasal movement, which has identified historically a number of theoretical and empirical problems. In the following section, I will outline some of the more recent approaches to the puzzle of head movement. Furthermore, I will present a more in depth discussion of one particular strand of research on head movement, namely the re-projective approach to head movement, which is particularly relevant to the overarching topic of this dissertation (i.e. the properties of functional structure).

### ***2.2.2 The Re-Projective Analysis of Head Movement***

The previous section has established that reconciling the properties of head movement with the properties of phrasal movement under a uniform theory of movement is not a trivial task. In recent years there has been a large body of

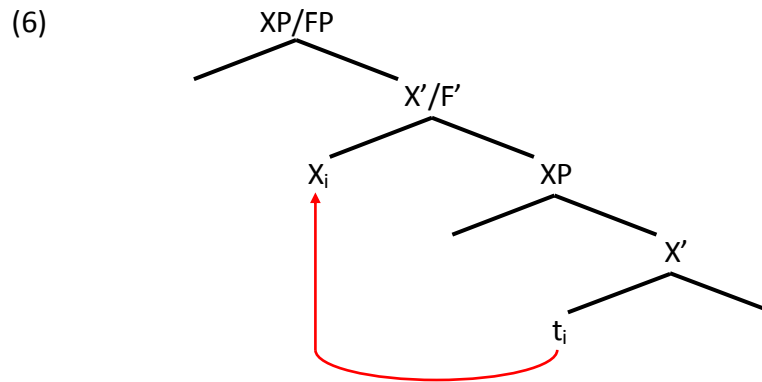
research dedicated to this issue. In the most general terms, the relevant literature has pursued one of two possible approaches. The first option is to attempt to reformulate head movement (or movement in general) in such a way that the idiosyncrasies of head movement are eliminated or, at least, explained in a principled way (Ackema et al 1993, Koenenan 2000, Bury 2003, Surányi 2005, Donati 2006, Matushansky 2006, amongst others, have pursued various revisions of head movement in that vein). The second option is to eliminate verb movement from narrow syntax, thus removing the need to compare the properties of head and phrasal movement. There are, in fact, different directions that one could follow from that starting point. For example, in the 'remnant movement' literature (Koopman and Szabolcsi 2000, Mahajan 2000, 2003, Müller 2004, Kayne and Pollock 2001, Poletto and Pollock 2004, Pollock, Poletto and Munaro 2001, Pollock 2006, Nilsen 2003, amongst many others) head movement is analysed as a composite of independent phrasal movement operations, which initially displace all the material of a given phrase apart from its head and subsequently displace the remnant of that phrase (i.e. the head itself). This analysis essentially eliminates head movement as a distinct syntactic operation. Alternatively, Chomsky (2001) suggests that head movement ought to be relegated to the PF interface. Thus, head movement is not entirely eliminated, but, nonetheless, it is not comparable to phrasal movement since the two operations take place in different components of the derivation (Boeckx and Stjepanović 2001 and Harley 2004 explore Chomsky's suggestion in more detail).

In this section, I will focus my attention on the former approach (i.e. the reformulation of head movement), since this strand of research has generated the most interesting insights into the topic of functional structure. I will focus, in particular, on the literature that has explored the reformulation of head movement in terms of re-projection (i.e. Ackema et al 199, Koenenan 2000, Bury 2003, Surányi 2005). Towards the end of this section, I will comment separately on Donati (2006) and Matushansky (2006) who defend an alternative reformulation of head movement. Regarding the remnant movement and the PF movement analyses, I

refer the reader to Roberts (2010, 2011) for a detailed overview, as well as a series of arguments as to why neither of these approaches eliminates (syntactic) head movement in toto.

I will start this discussion with a brief outline of how the re-projective literature departs from the traditional minimalist conception of head movement stemming from Chomsky (1995). According to Chomsky's (1995) Attract version of movement, head movement is triggered by some property of a target head. That is, in the case of a head X moving to a target head Y, the trigger of the operation is Y. Note that this state of affairs presupposes that, in a derivational model of syntax, the timing of the relevant operations is such that the merger of the target head Y precedes the movement of head X. This further suggests that X necessarily moves to a structural position that is already filled by Y. Thus, according to traditional analyses, head movement results in the adjunction of the moved head X to the target head Y. Finally, another feature of this analysis that will be relevant to our discussion is the hypothesis that after a head movement operation (or movement in general) has occurred, it is always the target head Y that projects (i.e. the 'target projects' constraint; cf. Chomsky 1995).

Having established this background, we can now return to the re-projective approach on head movement. The analyses of Koenenman (2000), Bury (2003) and Surányi (2005) differ significantly in their scope, goals, supporting arguments and the details of their formalisms. Nonetheless, these works paint a rather coherent picture of re-projective head movement. Their common ground could be summarised as follows: head movement is conceived as an operation that displaces a head X from its base position and directly re-merges X with its own maximal projection XP, allowing X to re-project a maximal projection (note that the resulting projection could either be a recurring XP or an FP, depending on whether one assumes that the head X has a single categorial feature that may project or a multitude of them). The tree in (6) provides a schematic representation of re-projective head movement.



From the above description it is clear that this formulation of head movement consists of two complementary hypotheses, or, perhaps more accurately, it arises from the rejection of two fundamental assumptions of the Attract version of head movement. Firstly, re-projective analyses propose that head movement is triggered by some property of the moved head itself, rather than some property of the target. One consequence of this proposal is that the target position is created during the head movement operation, not prior to it. Therefore, this approach renders head adjunction redundant, since the position targeted by the moved head  $X$  is not filled by any other head  $Y$ . Secondly, re-projective analyses depart from traditional views of head movement in that they reject the target projects constraint and, instead, they allow for the moved head itself to project.

Thus far I have merely outlined the contrasts between Chomsky's (1995) Attract formulation of head movement and the re-projective alternative of Koenenman (2000), Bury (2003) and Surányi (2005). Let us now consider the merits of re-projective head movement. The most obvious benefit of this analysis pertains to the generalisations in (1b) (i.e. the moved head and the target head form a constituent) and (1c) (i.e. the moved head and the target head cannot undergo further movement independently). Observe that in a structure like (6), resulting from re-projective head movement, the target position is occupied by a single head  $X$ , not a complex  $X+Y$  head as in a traditional analysis of head movement. Thus, the effects of (1b) and (1c) are trivially derived. Furthermore, since the re-projective analysis of head movement does not invoke head adjunction, it does not face the theoretical problems mentioned in (3a) (i.e. head movement in (6) does extend the



root, so there is no violation of the Extension Condition) and (3b) (i.e. the moved head X in (6) does c-command its trace).

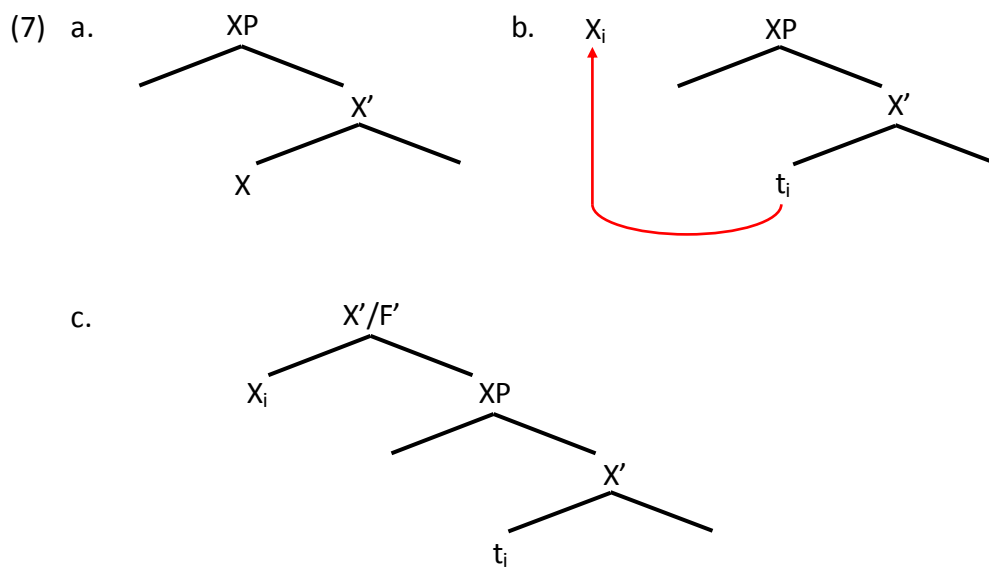
It could also be argued that a re-projective formulation of head movement can capture the Head Movement Constraint (cf. (1d)), provided that the analysis incorporates one additional assumption. Following the rationale I have outlined earlier, suppose that re-projective head movement is triggered by some property (i.e. an uninterpretable feature) of the moved head itself. Suppose, further, that we adopt some version of Pesetsky and Torrego's (2001) Earliness Principle, which states that uninterpretable features must be eliminated as soon as possible. Consider, then, the intermediate point of the derivation in (6) at which the lower XP has just been projected, but X has not yet moved. At that point, X still has an uninterpretable feature that has not been eliminated. Furthermore, according to the Earliness Principle, the elimination of this uninterpretable feature must take precedence over any other operation. Therefore, the movement of the head X will necessarily precede the merger of any head Z. Thus, under these assumptions, a structure where head X moves over head Z cannot be derived. This reasoning demonstrates that the Head Movement Constraint does not need to be stipulated within a re-projective theory of head movement, since it can be derived from more general hypotheses pertaining to the nature of syntactic operations.

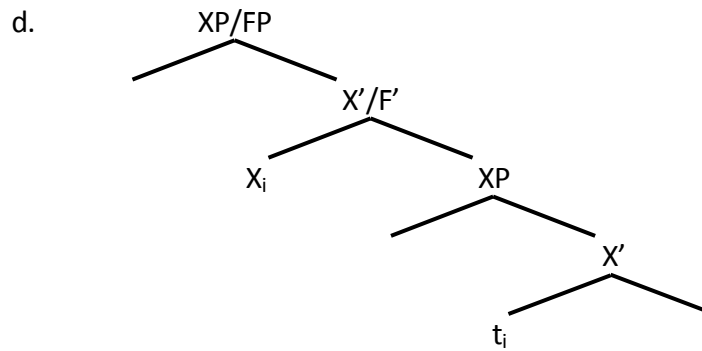
Finally, re-projective theories of head movement provide some equally interesting, but more subtle, insights into Structure Preservation (cf. (1a)). Observe that the head X in (6) does indeed surface in a head position after head movement (i.e. it surfaces in a position that we would straightforwardly identify as the head of the higher maximal projection XP/FP). But how does a re-projective analysis ensure that this will always be the case for head movement, and how does it ensure that phrasal movement behaves differently? Remember that according to this analysis the target position of head movement is not constructed independently of the movement operation. Therefore, it should be the case that the effects of Structure Preservation follow from the properties of the syntactic objects undergoing movement, rather than the properties of the target position.

In order to flesh out this suggestion, consider the following. First of all, let us clarify our background assumptions on the properties of heads and phrases, outside the context of movement. According to an X-bar theoretic understanding of syntactic structure, we define heads as syntactic objects that can (i) take complements, (ii) take specifiers, and (iii) project a maximal projection. The other side of the coin is that maximal projections are the syntactic objects that appear as complements or specifiers (or adjuncts) of heads, which cannot themselves project. Keeping that in mind, we can easily demonstrate that in a traditional analysis of movement Structure Preservation need not be stipulated with respect to phrasal movement. Rather, its effects on phrasal movement can be straightforwardly derived from independent factors. Assuming an Attract version of phrasal movement, a moved phrase will necessarily merge with the intermediate projection of the attracting head (note that if the moved phrase were to merge in any other position in the structure, it would violate the Extension Condition). Furthermore, the moved phrase (i.e. a maximal projection) is unable to project, by definition. It is then necessary that the attracting head projects. Thus, the position of a moved phrase is the characteristic position of a specifier (i.e. the daughter of a maximal projection of a different category and the sister of an intermediate projection).

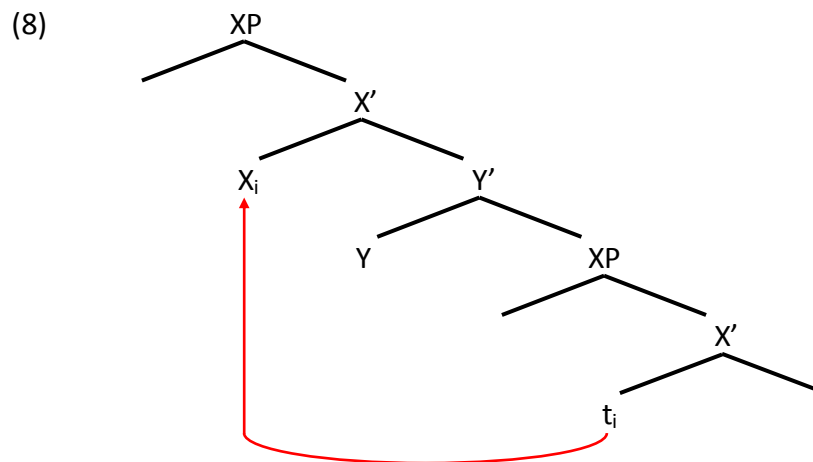
On the contrary, under a version of head movement based on Attract and head adjunction, the properties of a moved head do not align with the fundamental properties of a (non-moved) head: the moved head does not take a complement or a specifier and it does not project. Rather, it is assumed that there is a separate head in the target position which takes a complement and specifier, and projects a maximal projection. Thus, these constraints on the properties of the moved head need to be independently stipulated. The re-projective literature essentially contends that if we eliminate these restrictions on the behaviour of moved heads, then the effects of Structure Preservation in relation to head movement can be derived from the inherent properties of the head, in a similar way that Structure Preservation as it pertains to phrasal movement can be derived from the properties of a maximal projection. To demonstrate, let us consider in detail the steps of the

derivation that lead up to the structure in (6). Let us take (7a) (i.e. the point in the derivation where an XP has just been constructed) as our starting point. Suppose then that the head X has some property that triggers head movement. (7b) is a visual representation of the idea that the head movement operation does not target some predefined and preconstructed structural position. Rather, after being displaced from its base position, the head X will be 're-attached' to the structure in a manner consistent with its properties. That is, the head X will take the XP as its complement (while also projecting an intermediate projection), as in (7c). Subsequently, the head X will be able to take a specifier and project a maximal projection (cf. (7d)), always according to the fundamental properties of heads. For a different presentation of these ideas, I refer the reader to Bury (2003) and Surányi (2005). It is also of note that the relevant arguments trace back to Ackema et al (1993), albeit set in a different framework.





Before concluding this section, I will also briefly comment on Donati (2006) and Matushansky (2006), who pursue the same general reasoning outlined above but implement head movement in a rather different way. According to these works, head movement is a composite operation that consists of two independent components: a syntactic operation that moves the head  $X$  to the specifier of the target head  $Y$ , and a post-syntactic operation (M(orphological)-merger in Matushansky's 2006 terms) that merges the two adjacent heads into a single unit. Schematically, this formulation of head movement creates a structure like (8) below.



Consider how this analysis compares with the version of head movement proposed by Koenenman (2000), Bury (2003) and Surányi (2005). One difference is that Donati (2006) and Matushansky (2006) maintain the assumption that head movement is triggered by a target head  $Y$ , while Koenenman (2000), Bury (2003) and Surányi (2005) eliminate the target head. On the other hand, all of these works

abandon the traditional head adjunction analysis of head movement. Furthermore, they all reject the target projects constraint. Interestingly, the set of assumptions defended by Donati (2006) and Matushansky (2006), leads to a structure where the moved head lands in a specifier position. This is tantamount to a rejection of the letter of Structure Preservation (i.e. (1a)). Nonetheless, the structure in (8) still accounts for the observable facts expressed by Structure Preservation, in a fashion that is rather similar to the reasoning demonstrated in (7). According to the analyses of Donati and Matushansky, the fact that heads and phrases appear to occupy different structural positions is a mere epiphenomenon of the different properties of heads and maximal projections. The former can project and they can also undergo M-merger, while the latter cannot do either of these things.

Another significant difference between the M-merger analysis of Donati (2006) and Matushansky (2006) and the re-projection analysis of Koeneman (2000), Bury (2003) and Surányi (2005), relates to property (1c) of head movement (i.e. the ban on excorporation). Matushansky (2006) argues that head movement and M-merger are independent of each other. Thus, if a derivation were to involve successive steps of head movement not followed by M-merger, the resulting structure would display excorporation. Therefore, the M-merger analysis is more permissive than the re-projective approach with respect to excorporation. One related point is that there seems to be no straightforward way of deriving the properties in (1b) (i.e. the constituency of moved and target head after head movement) and (1d) (i.e. the Head Movement Constraint) from the basic assumptions of the M-merger analysis of head movement. Thus, it appears that these properties need to be stipulated, derived from some different property of narrow syntax, or demonstrated to be empirically false.

To conclude, this section has provided a broad outline of the re-projective analysis of head movement originating in Ackema et al (1993) and updated, most recently, in works such as Koeneman (2000), Bury (2003) and Surányi (2005). Furthermore, I have discussed in some detail the merits of this analysis in deriving the properties of head movement as outlined in section 1.2.1. In this respect, I have

concluded that re-projection is a rather successful theoretical tool that addresses many of the historically unresolved problems of head movement. In the following section, I will present a critique of the current re-projective analyses that is focused on their implications in relation to the properties of functional structure, rather than head movement per se. This will prompt me to present a novel analysis of head movement that maintains most of the fundamental assumptions and conclusions of the re-projective literature, but implements them in a distinct way based on the analysis of functional structure developed in chapter 1.

### ***2.2.3 Deriving Head Movement from Feature Scattering***

The previous two sections have outlined the theoretical background that will be relevant to the analysis of head movement that I will defend in this dissertation. As I will demonstrate in the following discussion, this proposal has many of the hallmarks of a re-projective analysis of head movement, with one significant exception. Re-projective analyses maintain that head movement is an independent syntactic operation, from which we can derive certain aspects of the functional structure. I will, instead, approach the matter from the opposite direction: I will explore the hypothesis that head movement is derived from a more fundamental structure-building operation (namely Feature Scattering). This section is organised as follows. I will start with a brief discussion of the limitations of re-projective analyses in capturing (covert) functional structure, as well as head movement itself. I will then demonstrate how one can derive head movement from the theory of functional structure developed in chapter 1. Finally, I will discuss the implications of this proposal in comparison to competing analyses of head movement.

As I have acknowledged in the previous section, the re-projective approach of Koenenman (2000), Bury (2003), Surányi (2005) and others, offers some very promising solutions to long-standing problems in the theory of head movement. However, I will argue that these analyses are not entirely satisfactory when it comes to the intersection between head movement and covert functional structure. To clarify, consider the following points. Null functional heads can be divided, on a

language-by-language basis, in those that can serve as a (final or intermediate) target for head movement and those that cannot. Furthermore, the aforementioned analyses assume that there is a distinction between null functional heads that are derived via re-projection and null non-derived functional heads that are drawn from the lexicon. The question then is how these properties relate with each other. Specifically, we have to consider (i) whether head movement can target derived and/or non-derived functional heads and (ii) whether derived and/or non-derived functional heads can exist independently of head movement.

Let us start with the first question. It is trivial to observe that head movement can indeed target null derived functional heads, considering the very definition of re-projective head movement. Whether head movement can target null non-derived functional heads is a more contentious issue. Bury (2003) adopts an affirmative answer to this question (note that this is not the case for Koenenman 2000 and Surányi 2005). This assumption, however, has the serious implication that the re-projective analysis of head movement is only relevant to a subset of the instances of head movement. This means, by extension, that re-projection only accounts for the properties of head movement for the same subset of head movement cases. To address this limitation, one would have to allow for a second type of head movement operation (one that more closely resembles the traditional Attract version of head movement) that is employed in the case of movement to null non-derived functional heads. However, as I have argued in section 1.3.2 of chapter 1, this is a costly solution, since it introduces redundancy to the analysis by reduplicating the head movement operation. The situation is further exacerbated by the fact that head movement displays a very narrow set of non-trivial properties that need to be accounted for, as discussed in section 2.2.1. Therefore, I conclude (pace Bury 2003) that it is preferable to limit the possible targets of head movement to null derived functional heads.

Let us also consider the other side of this coin. Namely, whether it is possible to have derived and non-derived null functional heads that are not targeted by head movement. I believe that the first part of this question (i.e. whether there are

derived null functional heads that are not involved in head movement) should be answered positively. Consider, for example, topicalization in English. Following Rizzi's (1997) analysis of the left periphery, the topicalized constituent appears on the specifier of the null functional head Top(ic). In English, Top does not have any overt counterpart. Therefore, according to Koenenman's (2000) criteria for the acquisition of null functional heads (i.e. the idea that null functional heads can be acquired when they are part of a paradigm that also includes one or more overt functional heads with distinct semantic interpretation), Top should be treated as a derived functional head. But, of course, Top in English is not targeted by any head movement operation. Therefore, I conclude that null functional heads can be derived, regardless of their involvement into head movement operations. Or, to put it more precisely, the functional heads that appear as (final or intermediate) landing sites for head movement are a subset of the null derived functional heads. If this observation is on the right track, then we can further conclude that re-projective head movement ought to be formulated as a subcase of a more general re-projective, structure-building operation. Accordingly, in the remainder of this section I propose an analysis that derives head movement from the (essentially, re-projective) mechanism of Feature Scattering that was developed in chapter 1. Note that this approach is not, in principle, incompatible with the analysis of Koenenman (2000) or Surányi (2005), but it is something that has not been sufficiently explored within re-projective theories in the past. Finally, the last question is whether null functional heads that are not targeted by head movement can be non-derived. I suggest that once we accept the conclusion that re-projection is a fundamental structure-building operation (i.e. an operation that is broader than head movement) there remains little motivation to assume that null non-derived functional heads exist at all (see also the discussion in section 1.3.2 for a further argument from the perspective of language acquisition). Thus, I argue that this view should be considered as the null hypothesis, at least within a re-projective analysis of functional structure. Note, however, that the analysis of head movement that I will propose in this section rests on the weaker hypothesis that head movement is



derived from re-projection and not on the stronger hypothesis that the entirety of covert functional structure is derived via re-projection.

I will preface this discussion with a brief reminder of the relevant points of the analysis in chapter 1. Remember that I have proposed that lexical items have a featural specification consisting of a (universal) sequence of functional features  $fnF$ , which may carry a number of un/interpretable sub-features  $(u)F$  (subject to cross- and intra-linguistic variation). The general format of these featural specifications is represented in (9) (repeated from chapter 1). Furthermore, I have proposed that un/interpretable features and functional features serve different roles in syntactic operations: uninterpretable features function as triggers for syntactic operations, while the sequence of functional features determines the order of those operations.

(9)  $X_0 \langle fnF_n \{(u)F_z, \dots\}, \dots, fnF_1 \{(u)F_y, \dots\}, fnF_0 \{(u)F_x, \dots\} \rangle$

Finally, I have adopted and refined Giorgi and Pianesi's (1997) concept of Feature Scattering. Following Giorgi and Pianesi (1997), I assume that Feature Scattering is a last resort syntactic operation that is triggered when the syntactic configuration (at some intermediate point in the derivation) does not allow for a given uninterpretable feature to establish an appropriate syntactic relation with a matching interpretable feature. In that event, Feature Scattering displaces the offending uninterpretable feature and re-merges it in the structure in the form of a null derived functional head. However, unlike Giorgi and Pianesi (1997), I have proposed that Feature Scattering does not affect individual features, but entire segments of the relevant featural specification. Specifically, I have proposed that Feature Scattering affects the functional feature hosting the offending uninterpretable sub-feature and all of its superordinate functional features in the sequence, as described in (10) (repeated from chapter 1).

(10) Feature Scattering

Assume that at a given (intermediate) step of a derivation, an uninterpretable feature  $uG$  of a head  $X_0$  - where  $X_0$  has the featural

specification in (i) - cannot establish the necessary relation with a matching interpretable feature.

i.  $X_0 < \text{fnF}_n\{\text{u}\text{F}_z, \dots\}, \dots, \text{fnF}_m\{\text{u}\text{G}, \dots\}, \dots, \text{fnF}_1\{\text{u}\text{F}_y, \dots\}, \text{fnF}_0\{\text{u}\text{F}_x, \dots\} >$

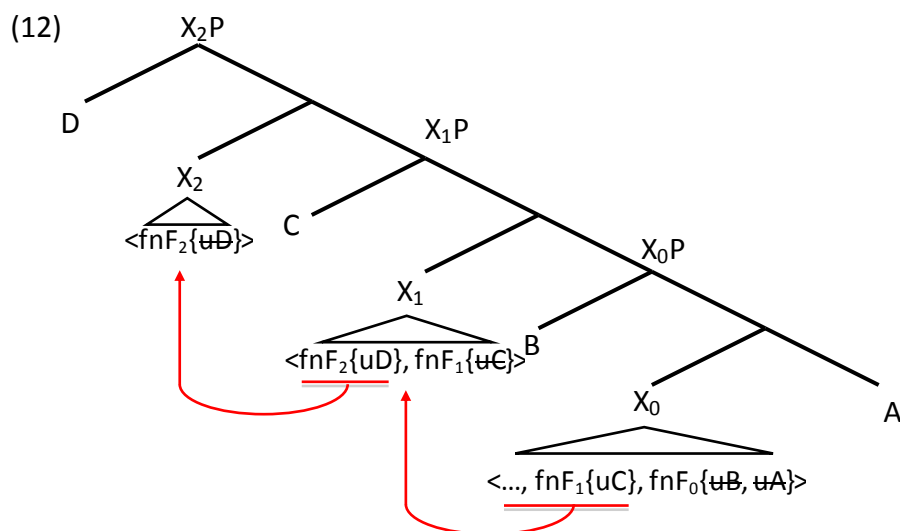
Then Feature Scattering will strip a segment of the featural specification of  $X_0$ , leaving  $X_0$  with the altered featural specification in (ii) and creating a new functional head  $X_1$  with the featural specification in (iii) that merges with  $X_0\text{P}$ .

ii.  $X_0 < \dots, \text{fnF}_1\{\text{u}\text{F}_y, \dots\}, \text{fnF}_0\{\text{u}\text{F}_x, \dots\} >$

iii.  $X_1 < \text{fnF}_n\{\text{u}\text{F}_z, \dots\}, \dots, \text{fnF}_m\{\text{u}\text{G}, \dots\} >$

The combination of the two hypotheses in (9)-(10) allows us to derive the covert parts of the functional structure from the lexical items that are available in a given numeration in an exhaustive and economical manner (i.e. all functional heads that are necessary in a given structure will be projected, but any superfluous functional head will not). The tree in (12) demonstrates how the uninterpretable sub-features of a single lexical item  $X_0$  with three functional features (i.e. (11)) may lead to a fully expanded functional structure. Briefly, the uninterpretable features of  $\text{fnF}_0$  are eliminated against the complement and specifier of  $X_0\text{P}$ ; the uninterpretable feature of  $\text{fnF}_1$  cannot be eliminated at the  $X_0\text{P}$  level and is, therefore, scattered to  $X_1$  to be eliminated against the specifier of  $X_1\text{P}$ ; similarly, the uninterpretable feature of  $\text{fnF}_2$  cannot be eliminated at the  $X_1\text{P}$  level and is scattered to  $X_2$ , where it is eliminated against the specifier of  $X_2\text{P}$ . In the following discussion I will use this derivation to demonstrate my analysis of head movement.

(11)  $X_0 < \text{fnF}_2\{\text{u}\text{D}\}, \text{fnF}_1\{\text{u}\text{C}\}, \text{fnF}_0\{\text{u}\text{B}, \text{u}\text{A}\} >$

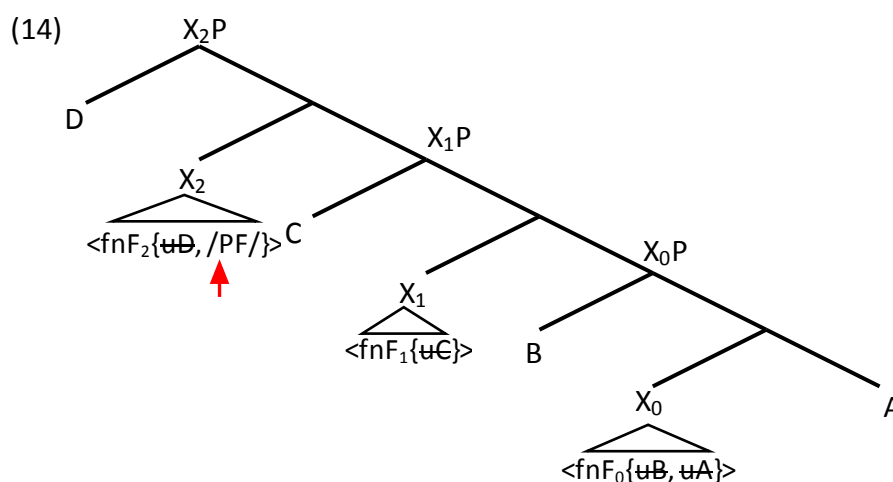


Let us now see how head movement comes into the picture. I propose that the featural specification of a lexical item contains its phonological features alongside its syntactic features. Specifically, I will assume that the phonological features of a lexical item appear as sub-features of one of its functional features. Of course, this is not to say that phonological features contribute in any way to narrow syntactic operations. However, under the assumption that Feature Scattering affects entire segments of a given featural specification, it stands to reason that phonological features may also be affected by scattering under specific circumstances. Namely, given a featural specification where the phonological features are hosted by some functional feature  $fnF_x$ , a Feature Scattering operation that targets an uninterpretable feature hosted by  $fnF_x$ , or any of its subordinate functional features, will result in the displacement of said phonological features. Consequently, the phonological features of a lexical item may be realized in the position of the lexical head or in the position of any of the functional heads scattered from it<sup>1</sup>, depending on their placement in the relevant featural specification and the Feature Scattering operations that take place in a given derivation. To demonstrate, suppose that the phonological features of the head  $X_0$  in (11) are situated on the functional feature  $fnF_2$ , as shown in (13). Of course, the presence of the phonological features does not have any effect on the syntactic

<sup>1</sup> See Brody (1997, and subsequent work) for a similar analysis of head movement, albeit set in a representational rather than a derivational model of syntax.

derivation, which will produce the exact same structure as in (12). Rather, the placement of the phonological features at the end of the derivation is only relevant in determining the surface position of the lexical item. As we can see in (14), in this case the lexical item will surface in  $X_2$ , thus deriving the head movement of  $X_0$  to  $X_2$ <sup>2</sup>.

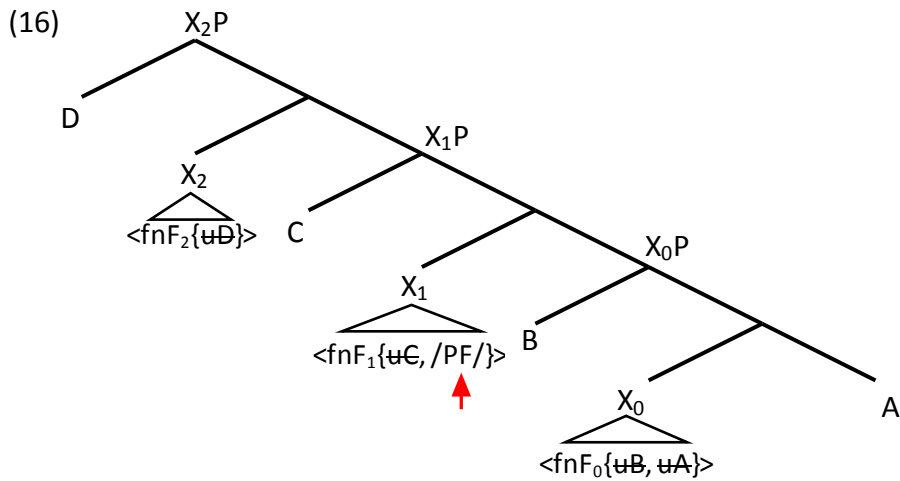
(13)  $X_0 \langle \text{fnF}_2\{\text{uD}, /PF/\}, \text{fnF}_1\{\text{uC}\}, \text{fnF}_0\{\text{uB}, \text{uA}\} \rangle$



Consider now an alternative placement of the phonological features of  $X_0$ . Suppose that the phonological features are situated in the functional feature  $\text{fnF}_1$ , as shown in (15). The derivation in (16) is identical to the previous one, but this time the lexical item will surface in the position of  $X_1$ . Therefore, this is a case of head movement of  $X_0$  to the intermediate functional head  $X_1$ .

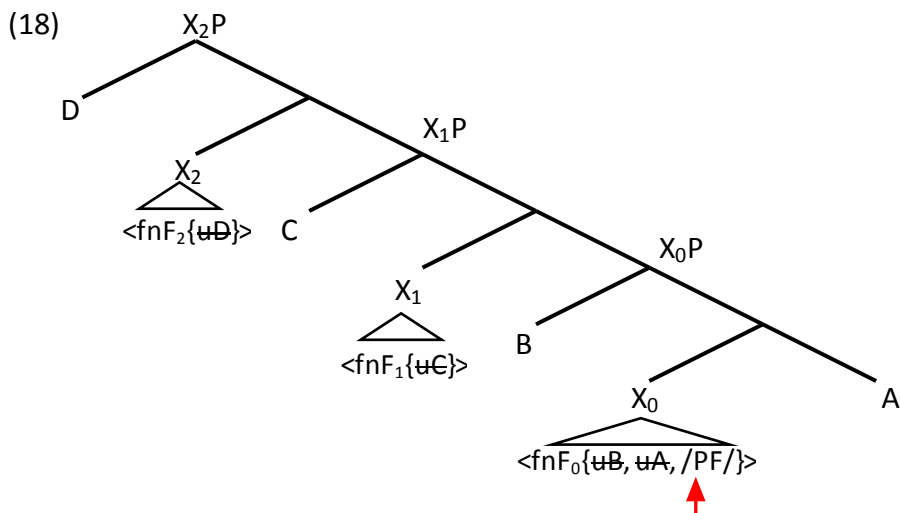
(15)  $X_0 \langle \text{fnF}_2\{\text{uD}\}, \text{fnF}_1\{\text{uC}, /PF/\}, \text{fnF}_0\{\text{uB}, \text{uA}\} \rangle$

<sup>2</sup> Note that in a traditional analysis this structure would involve two head movement operations. At first blush, the structure in (14) seems to imply that head movement happens in a single step. However, a closer look at the individual steps of this derivation (as shown in (12)) reveals that the phonological features of the lexical item are displaced twice, by two separate Feature Scattering operations. Thus, there is no significant difference in this respect.



Finally, let us consider the effect of placing the phonological features of  $X_0$  in the functional feature  $fnF_0$ , as in (17). As we can see in (18), in this case the phonological features of the lexical item are not affected by any of the Feature Scattering operations that take place during the derivation and they remain in the position of  $X_0$ . Thus, the lexical item will surface in-situ.

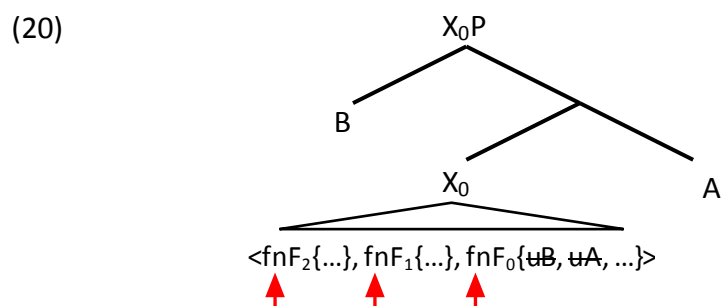
(17)  $X_0\langle fnF_2\{uD\}, fnF_1\{uC\}, fnF_0\{uB, uA, /PF/\}\rangle$



Notice, however, that according to this analysis the surface position of a lexical item does not depend solely on the placement of the phonological features, but also on the Feature Scattering operations that may or may not take place in a derivation. For example, in a structure like (20), where none of the functional features of  $X_0$  are scattered because they do not have any uninterpretable features

(cf. (19)), the placement of the phonological features is irrelevant. Whether they are placed in  $fnF_0$ ,  $fnF_1$  or  $fnF_2$ , the lexical item will surface in the position of the only available head  $X_0$ . Generally speaking, this example demonstrates that head movement will only be detectable if the relevant parts of the syntactic structure are independently projected, which is a natural consequence of the overarching variationist perspective to functional structure adopted in this dissertation. I will return to this point with some concrete examples in the following section.

(19)  $X_0\langle fnF_2\{\dots\}, fnF_1\{\dots\}, fnF_0\{uB, uA, \dots\}\rangle$



The head movement analysis developed above presents one more intriguing implication. According to a traditional theory of head movement, a moved head is inserted in some base position and subsequently displaced to a different structural position (subject to cross- and intra-linguistic variation). This suggests that placing a head in any moved position is more marked or costly than placing it in the base position. Thus, head movement needs to be independently motivated. However, according to the analysis proposed here, head movement operates on different assumptions. The fundamental claim throughout this dissertation is that the span of structure starting from a lexical head and including all of the functional projections that are scattered from it corresponds to a single, internally structured lexical item. Thus, it is reasonable to conclude that there is a de-facto multitude of structural positions on which the head can be phonologically realized. Or, to put it differently, none of the possible surface positions of a head is more costly<sup>3</sup>. One has to wonder, then, if head movement needs to be independently motivated, or if it can be

---

<sup>3</sup> The same argument has been made by Brody (1997).

treated as, essentially, a distributional accident: the phonological material of a lexical item needs to be linearized in one way or another and languages may choose any of the available structural positions to serve as the surface position for the lexical item in question - the choice then is maintained (or changed) historically via acquisition. Note, however, that this is not to say that head movement may be inconsistent in a given language or structure. Rather, the point is that there are no independent factors motivating head movement, apart from those that are relevant to the derivation of the functional structure.

Finally, a brief comment on the properties of head movement is in order. Notice that Feature Scattering is, at its core, a re-projective operation. By extension, head movement (which is described here as a consequence of Feature Scattering) should also display all the characteristics of a re-projective operation. Therefore, this analysis is able to derive all the properties of head movement (as summarized in (1) in section 2.2.1) in exactly the same manner as other re-projective head movement analyses (cf. the discussion in section 2.2.2).

To summarize, in this section I have claimed that the concept of re-projection is relevant in equal measure to the analysis of head movement and the analysis of (covert) functional structure. Furthermore, I have argued that from this perspective re-projective head movement ought to be understood as a subcase of a generalized re-projective structure-building operation. In chapter 1, I have formalised this fundamental structure-building operation in terms of Feature Scattering (as per Giorgi and Pianesi 1997). Accordingly, in this section I have outlined an analysis that derives head movement from Feature Scattering. In fact, this proposal is slightly stronger in the sense that head movement is not considered to be an independent syntactic operation but, rather, a mere epiphenomenon of the incidental scattering of the phonological features of a lexical item due to independent reasons. The remainder of this chapter (as well as part of chapter 4) will be dedicated to the analysis of head movement across different languages and structures. Specifically, the following sections will compare the patterns of verb movement associated with V2 word orders, V-to-T movement and V-to-v movement. In chapter 4, I will return

to verb movement to address verb-initial word orders, and I will also discuss some cases of head movement in the nominal domain.

### ***2.3 A Feature Scattering Analysis of Verb Movement***

#### ***2.3.1 V-to-C, V-to-T and V-to-v***

Thus far, I have demonstrated how head movement may be derived from Feature Scattering in the abstract case. In the following, I will further substantiate this proposal by addressing a number of concrete cases of verb movement. Specifically, in this section I will address three cases of verb movement: (i) the V-to-C movement of Germanic V2 languages (as well as English clauses displaying residual V2, as per Rizzi 1997), (ii) the V-to-T movement of Romance languages (as well as English auxiliary verbs), and (iii) the V-to-v movement of English lexical verbs. For ease of presentation, I will break down the analysis in three parts. In section 2.3.1, I will focus on verb movement in simple matrix clauses with a finite lexical verb (i.e. clauses that lack any auxiliary verb), in order to demonstrate the basics of the analysis. As we will see in due course, because of the absence of fixed structural positions under the proposed analysis of functional structure, the contrasting pattern of verb movement of finite versus non-finite verbs requires special attention. Thus, I will separately address complex matrix clauses where the finite verb is an auxiliary in section 2.3.2. Finally, in section 2.3.3, I will discuss the contrasting patterns of verb movement in main versus embedded clauses. I will return to one remaining major case of verb movement in chapter 4. Namely, the V-to-C movement of verb-initial languages found, for example, in Semitic and Celtic languages. Additionally, in the same chapter, I will address the head movement of nouns to an initial position within the nominal phrase, which is also a characteristic of the same Semitic and Celtic languages.

I will preface this discussion with a brief overview of the verb movement phenomena that I will address in this section. As is well known, V2 refers to a word order pattern that is common in Germanic languages (with the exception of English) whereby finite verbs appear as the second element in the clause, following a



constituent of any category in the first position and preceding the subject in its canonical position (unless it is the subject itself that surfaces in the initial position). Since den Besten (1983), V2 is commonly seen as an instance of verb movement to a null complementizer, followed by movement of the fronted constituent to the specifier of said complementizer. The V-to-C analysis of V2 has several benefits. Firstly, it captures the order of all relevant constituents (the initial constituent, the verb and the subject). Secondly, assuming that the complementizer is the highest functional head in the clause, it captures the fact that there is exactly one constituent preceding the verb in its moved position. Finally, it captures the complementarity of V2 and an overt complementizer in embedded clauses in German (den Besten 1983). To clarify on the last point, although the verb typically appears in a clause-final position in German embedded clauses, embedded V2 is observed in clauses that are embedded under a limited class of verbs. Interestingly, the clauses that display limited embedded V2 also require complementizer omission, which is an option that is otherwise unavailable in German. That is, V2 and the overt complementizer are in complementary distribution in German. Note, however, that this is not the case in all Germanic languages, which suggests that there is more to say about the analysis of embedded V2: for instance, Icelandic and Yiddish display V2 in all embedded clauses even though they have an obligatory overt complementizer and Mainland Scandinavian languages are similar to German in that they allow embedded V2 only under certain verbs, but the overt complementizer is not necessarily omitted in that context (I will discuss these contrasts in more detail in section 2.3.3). Nonetheless, the core of den Besten's (1983) analysis of V2 (i.e. the suggestion that it involves the displacement of the verb from its base position in the VP, over the surface position of the subject in the middle field of the clause, and into a structural position situated high in the functional structure of the clause) remains generally accepted and influential<sup>4</sup>. In this section I will focus on simple examples of V2 in German matrix clauses like the ones in (21) below.

---

<sup>4</sup> See Holmberg (2015) and Koenenman (2000:9ff) for recent reviews of the empirical and theoretical advancements in our understanding of V2 from the 1980s onwards.

- (21) a. V2 with a subject in first position  
 Ich las (schon letztes Jahr) diesen Roman  
 I read (already last year) this book
- b. V2 with an object in first position  
 diesen Roman las ich (schon letztes Jahr)  
 this book read I (already last year)
- c. V2 with an adjunct in first position  
 schon letztes Jahr las ich diesen Roman  
 already last year read I this book

The other case of verb movement that has historically informed the study of head movement is the movement of V-to-T, commonly exemplified by French or Italian. The generalisation regarding the distribution of the verb in these languages is that the finite verb precedes certain elements like adverbs, negation (in the case of French), and floating quantifiers. The fact that these elements are vP-external suggests that the verb moves from its base position to some vP-external functional head. Furthermore, in these languages the verb appears after the subject, which narrows down the surface position of the verb to T. Example (22) demonstrates the distribution of the verb in French<sup>5</sup>. Rather unsurprisingly, the same diagnostic elements (adverbs, negation and floating quantifiers) can be used to demonstrate yet another pattern of verb movement: i.e. V-to-v. For instance, English lexical verbs (but not English auxiliaries, which will be discussed in section 2.3.1) appear after those elements, indicating that they remain in some vP-internal position. Example (23a), the English counterpart of (22), demonstrates this point. Furthermore, the fact that English ditransitive verbs precede both of their objects, as shown in (23b), suggests that their surface position is v.

- (22) Jean lit souvent des romans  
 John reads often novels  
 'John often reads novels'

- (23) a. John often reads novels  
 b. John gave Mary the book

---

<sup>5</sup> Note that I am focusing on examples involving adverbs for ease of exposition. In the following discussion I will indicate which parts of the analysis carry over to negation straightforwardly and which do not, as appropriate.

In view of the analysis of head movement outlined in section 2.2.3, the most natural way to capture the cross-linguistic parameterization of verb movement described in the previous paragraphs would be to attribute it to variation in the featural specification of the verbs themselves. Roughly speaking, I will propose that V-to-C movement in German is the result of a featural specification where the phonological features of the verb are hosted by some functional feature that is quite high in the relevant sequence. As a result, the phonological features of the verb will be (incrementally) displaced by each and every Feature Scattering operation that may take place. Thus, the verb will surface in the position of the highest functional head created by Feature Scattering. The exact opposite hypothesis can be used to capture V-to-v in English. I will suggest that the phonological features of English lexical verbs are hosted by some functional feature that is relatively low in the sequence, so that they are unaffected by any Feature Scattering operation. Consequently, the verb will surface in its base position rather than the position of any of the functional heads that may arise due to Feature Scattering. Finally, the French V-to-T movement can be treated as an intermediate situation between V2 and V-to-v. Thus, I will propose that the phonological features of the verb are hosted by some intermediate functional feature in the sequence, so that they will be affected by some, but not all, Feature Scattering operations (i.e. the phonological features will be displaced by any scattering operation targeting their host or its subordinate functional features, but they will not be affected by any scattering operation targeting functional features above their host). Accordingly, the surface position of the verb in this case will be the relevant intermediate derived functional head.

Let us have a closer look at this proposal. For the purposes of this discussion, I will assume a simplified version of the featural specification of lexical verbs consisting of the sequence of functional features  $\langle C\{\dots\}, T\{\dots\}, v\{\dots\}, V\{\dots\} \rangle$ . As mentioned earlier, German V2 involves the displacement of two syntactic objects: the verb surfaces in C and the fronted constituent surfaces in the specifier of C. To capture the former, I assume that the phonological features of the lexical verb

appear in its featural specification as sub-features of the functional feature C. For the latter, I will assume that C has some uninterpretable sub-feature  $uX$  that triggers phrasal movement. Furthermore, we need to assume that the functional feature T has some uninterpretable feature (call it EPP) to ensure that the subject will surface at its specifier. Finally, a transitive verb, like the one in (21b), will have one uninterpretable sub-feature  $uD$  on V and a second one on  $v$ , in order to select its object and subject. Thus, the featural specification of a German lexical verb will look like (24) below.

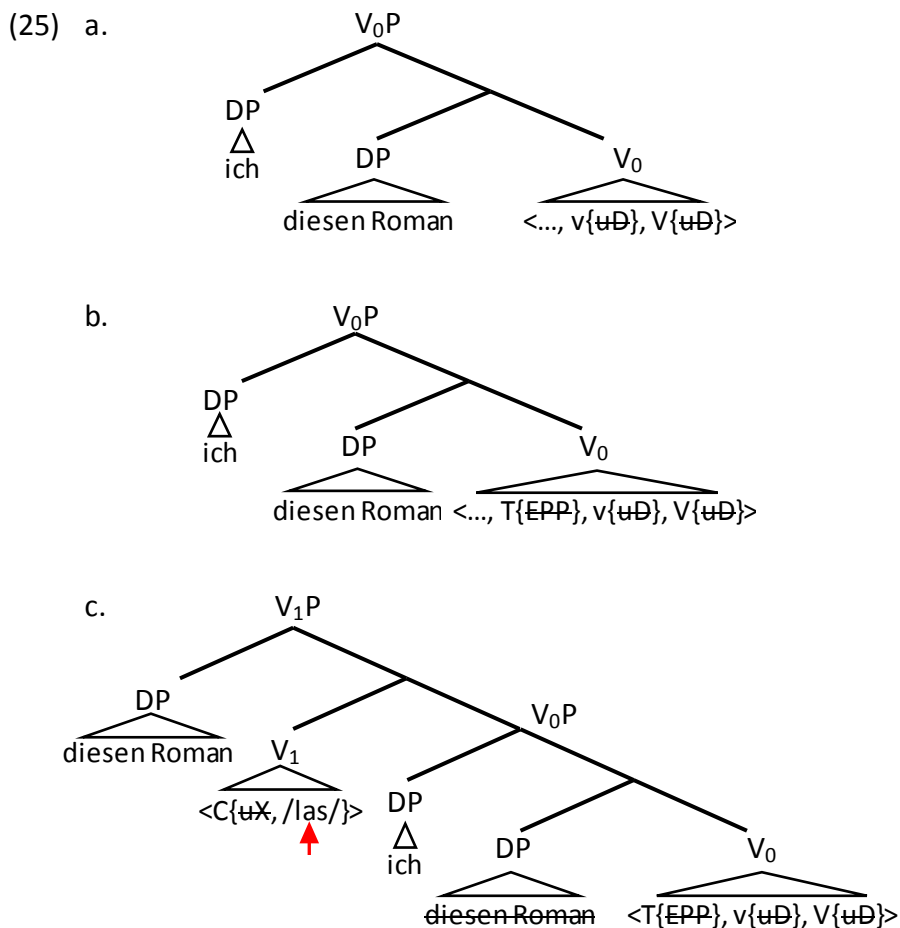
(24) Lexical verb (in German)  
 $\langle C\{uX, /las/\}, T\{EPP\}, v\{uD\}, V\{uD\}\rangle$

Given the featural specification in (24), the derivation of a V2 clause with an object in the first position will proceed as follows. As one would expect, the first few steps of the derivation will form the  $V_0P$  (as shown in (25a)), in order to eliminate the uninterpretable features of V and  $v$  against the complement and specifier of  $V_0$ . The next uninterpretable feature in the sequence is the EPP, which is hosted by T. Since the EPP is already in a Specifier-Head relation with the subject at the  $V_0P$  level, it can be eliminated without resorting to Feature Scattering (cf. (25b)). The final uninterpretable feature in the sequence is the  $uX$  appearing on C. Assuming that  $uX$  needs to establish a Specifier-Head relation with the object, it cannot be eliminated at the  $V_0P$  level<sup>6</sup>. Therefore,  $uX$  has to be scattered to  $V_1$  and

---

<sup>6</sup> One problem that arises here is why the Specifier-Head relation is needed when there already exists a Head-Complement relation between  $uX$  and the object of the verb at the  $V_0P$  level. One solution to this problem is to adopt Brody's (1997) view on Specifier-Head and Head-Complement relations. According to Brody's (1997) Mirror Principle, Head-Complement relations are an expression of the morphological structure of a lexical item, while Specifier-Head relations hold between separate lexical items (or, more accurately, between one lexical item and the projection of a second lexical item). If we were to translate this analysis to the system developed here, we would arrive at the suggestion that Feature Scattering creates Head-Complement relations, while traditional Merge creates Specifier-Head relations. This approach would lead to a more articulated structure for (25), where the object initially appears at the specifier of  $V_0$ , the subject appears at the specifier of a scattered  $V_1$  and, finally, the object moves to the specifier of a scattered  $V_2$ . In that structure, the scattering of  $V_2$  is not problematic, because at the  $V_1P$  position  $uX$  is not in a local relation with the object. Note, however, that this more articulated structure does not

attract the object to its specifier, as shown in (25c). To be more precise, under the formulation of Feature Scattering in (10), the derived functional head  $V_1$  will have a featural specification consisting of the functional feature C and its sub-features, which include  $uX$  and the phonological features of the verb. Thus, the verb will surface in  $V_1$ , which is, of course, the second position of the clause.



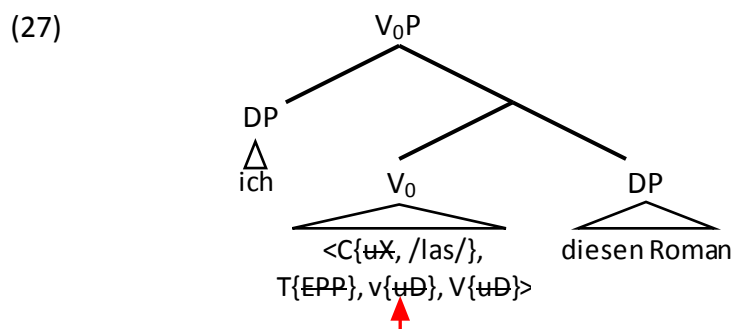
Note that in (25) I take for granted that  $V_0P$  is head-final and  $V_1P$  is head-initial. This seems like a reasonable assumption considering that  $V_0P$  is an amalgam of a traditional VP, vP and TP, while  $V_1P$  is the counterpart of a traditional CP, given their respective functional features. The situation, however, is less straightforward in a V2 clause with a subject in the first position like (21a). Let us consider the relevant derivation. Until a certain point, the derivation will be

---

alter in any way the order of relevant operations. For that reason, in the main text I will prefer the traditional and simpler syntactic representations that I have been using so far.



will be head-initial, while if the topmost functional feature is V, v, or T (like  $V_0$  in (25)), the maximal projection will be head-final.



Finally, note that this analysis suggests that there is an asymmetry between a V2 clause with a subject in first position and a V2 clause with a different fronted constituent. Specifically, the structure of the former is smaller than the structure of the latter. In that broad sense, this analysis is closer to works that treat V2 clauses with a subject in first position as a bare TP (amongst others Travis 1984, Zwart 1993), rather than works that treat them as full-fledged CPs (amongst others den Besten 1983, Weerman 1989, Vikner 1990).

The next case of head movement to consider is French V-to-T. As I have mentioned before, this movement is observable when the verb moves across some diagnostic element, like an adverb or a negative marker. At this point, I will demonstrate the workings of the proposed analysis using an example of verb movement over an adverb, as in (22) (repeated in (28) below for convenience). This analysis carries over straightforwardly to verb movement over a negative marker. Before addressing head movement itself, it is necessary to clarify how our diagnostic element (i.e. an adverbial phrase) is introduced into the structure. I will follow Cinque (1999) in assuming that adverbials are specifiers of dedicated null functional heads. However, unlike Cinque I will assume that the relevant functional heads are created during the derivation as a result of Feature Scattering. Thus, I propose that the featural specification of a lexical verb contains a number of Mod(ifier) functional features, each one with a different semantic contribution, that are dotted along the sequence of functional features. Each Mod functional feature may optionally have an uninterpretable sub-feature (e.g. a  $uA$  if the modifier is an

adverbial phrase, a uP if the modifier is a prepositional phrase, etc). When the uninterpretable feature is present, it will require the insertion of an adverb or other modifier at the specifier of the head that carries the host functional feature. Of course, this requirement may give rise to Feature Scattering in the appropriate circumstances. On the other hand, if the uninterpretable feature is absent, then the contribution of the Mod functional feature will not be detectable in any practical sense. In relation to the discussion of V-to-T movement, the relevant Mod functional feature is the one that is situated between the functional features T and v. Besides the addition of the Mod functional feature, I will make the following assumptions in relation to the featural specification of a French lexical verb. Firstly, as I have already suggested, V-to-T movement can be accounted for if the phonological features of the verb are hosted by the functional feature T. Secondly, I assume that T has an uninterpretable EPP sub-feature, as per usual. Finally, depending on the number of the verb's arguments, the functional features v and V will have the appropriate number of uninterpretable uD sub-features. Taking all of the above under consideration, I suggest that the verb in (28) will have the featural specification shown in (29) below.

(28) Jean lit souvent des romans  
 John reads often novels  
 'John often reads novels'

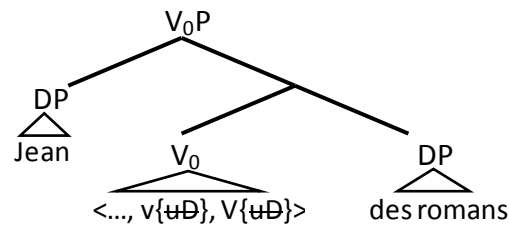
(29) Lexical verb (in French)  
 <C, T{EPP, /lit/}, Mod{uA}, v{uD}, V{uD}>

Based on this featural specification, the derivation will proceed as follows. As expected, the uninterpretable features of V and v are responsible for the formation of V<sub>0</sub>P, as shown in (30a). The next uninterpretable feature in the featural specification of the verb is the sub-feature uA of Mod. However, this sub-feature cannot be eliminated at the V<sub>0</sub>P level, since at this point of the derivation there is no empty specifier position for the adverb. Thus, Feature Scattering creates a V<sub>1</sub>, which can accommodate the adverb in its specifier so that uA may be eliminated (cf. (30b)). Note, further, that this Feature Scattering operation will also displace the

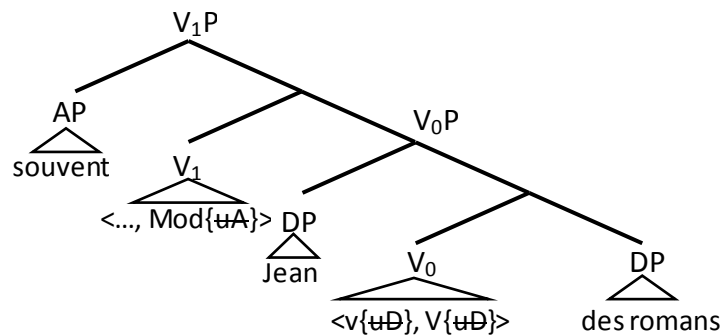


functional feature T, which disrupts the Specifier-Head relation between the EPP sub-feature and the subject. Therefore, Feature Scattering will be triggered again in order to create a V<sub>2</sub>, which will attract the subject to its specifier so that EPP may be eliminated (cf. (30c)). Finally, note that the phonological features of the verb, which appear alongside the EPP as sub-features of T, will surface at the position of V<sub>2</sub>. Thus, the derivation in (30) captures the movement of the verb across the adverb and into T<sup>7</sup>.

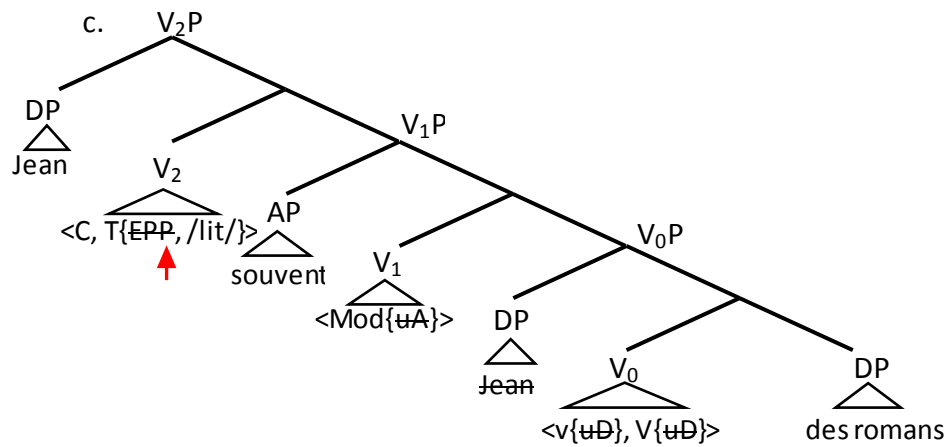
(30) a.



b.



<sup>7</sup> This derivation carries over to the analysis of French V-to-T movement over negation, if we simply replace the functional feature Mod and its uninterpretable sub-feature uA with a functional feature Pol(arity) carrying an uninterpretable sub-feature uNeg. As a result, the specifier of V<sub>1</sub> would be filled with a negative marker rather than an adverb. Note, however, that these differences do not alter the derivation in any significant way.

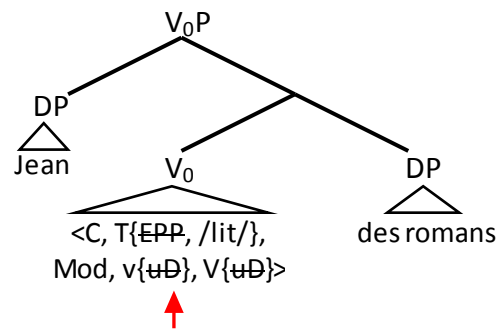


It would be interesting to briefly compare the above with the derivation of a French clause that does not include any of the elements that serve as a diagnostic for head movement. Consider, for example, the sentence in (31). In this case, the lexical verb will have a featural specification as in (32), where the functional feature Mod does not have any sub-feature. As a result, Mod will not undergo Feature Scattering during this derivation. This has the further implication that the functional feature T will still be situated at V<sub>0</sub> during the next step of the derivation, when the uninterpretable EPP feature needs to be eliminated. Consequently, T will not undergo Feature Scattering either, since the necessary Specifier-Head relation between EPP and the subject can be established at the V<sub>0</sub>P level. Therefore, the derivation will result in a single maximal projection, as shown in (33). Furthermore, the structure in (33) does not actually involve head movement. But the simplicity of this structure does not mean that it encompasses less syntactic information. On the contrary, the inherent flexibility of the proposed analysis of functional structure allows us to capture all the syntactic relations that are relevant to this example, while minimizing both syntactic structure and syntactic operations.

- (31) Jean lit des romans  
 John reads novels  
 'John reads novels'

- (32) <C, T{EPP, /lit/}, Mod, v{uD}, V{uD}>

(33)



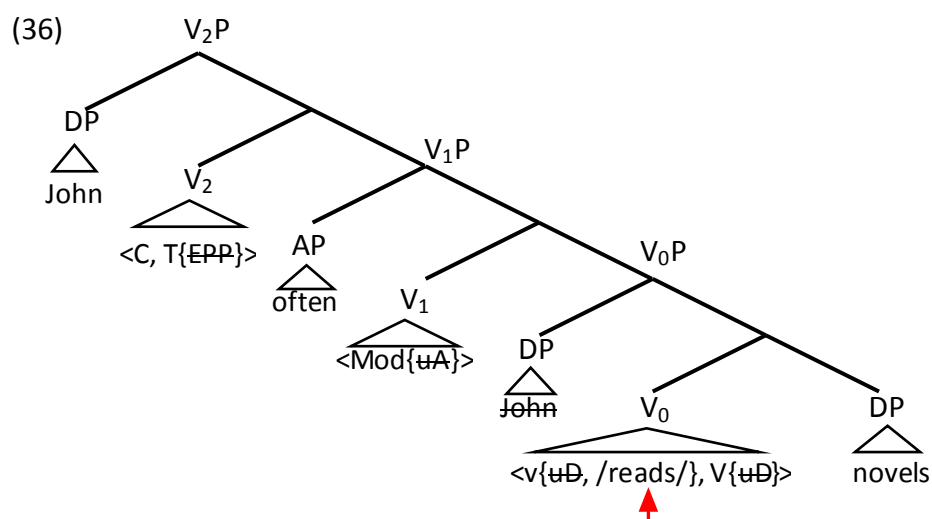
Finally, we can compare French V-to-T movement with the behaviour of English main verbs. Let us consider the example in (23a) (repeated in (34) below for convenience), which is the English counterpart of the French example in (28). The featural specification of the English main verb will be nearly identical to the one of the French verb in (29), with one exception. In order to capture the absence of V-to-T movement in the English case, I will assume that the phonological features of the verb are hosted by  $v$ . Thus, the relevant featural specification will be as in (35). Furthermore, due to the significant similarity of these featural specifications, the English derivation will proceed in exactly the same manner as the French case in (30). In (36) below I demonstrate the final product of this derivation for comparison<sup>8</sup>.

(34) John often reads novels

(35) Lexical verb (in English)  
 $\langle C, T\{EPP\}, Mod\{uA\}, v\{uD, /reads/\}, V\{uD\} \rangle$

---

<sup>8</sup> Note that this derivation will not suffice to account for an English clause involving sentential negation because it does not provide any insight on *do*-support. I will return to this issue in the following section.



Note, however, that although there is no structural difference between the English and French derivation in (36) and (30) respectively, there is a difference in the placement of the phonological features of the verb. While the French verb surfaces in the final position of the functional feature T (i.e. in  $V_2$ ), the English verb will surface in the position of  $v$  (i.e.  $V_0$ ). Thus, the English verb appears after the adverb, unlike its French counterpart.

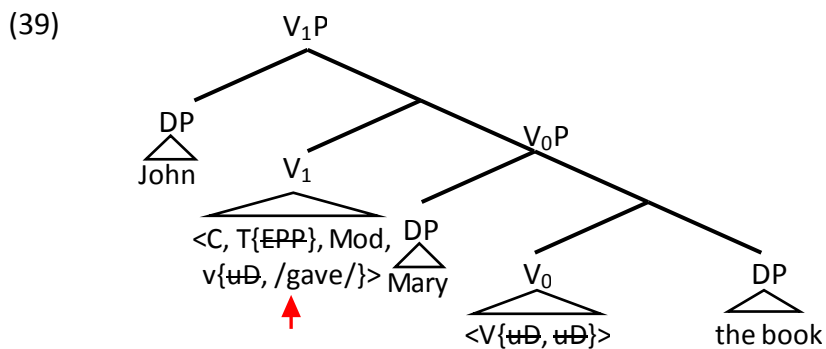
Before concluding this section, let us briefly consider the derivation of a clause involving a ditransitive verb, as in (23b) (repeated in (37) below). In this case, the functional feature  $V$  will have two uninterpretable  $uD$  features, as shown in the featural specification in (38). As a result, the complement and specifier of  $V_0$  will be filled by the two objects<sup>9</sup> of the verb, while the subject will be inserted at the specifier of a  $V_1$ , which will carry the scattered functional feature  $v$  and its sub-features. After that, the derivation will proceed as expected (in this example, the only remaining uninterpretable feature is  $T$ 's  $EPP$ , which can be eliminated at

<sup>9</sup> One residual question is what determines the order of merger of the two objects of a ditransitive verb, given that I have been assuming that the sub-features of a functional feature constitute an unordered set. There are two possible solutions to this problem. The first one is to revise the assumption that the sub-features of a functional heads are not ordered and, instead, assume that they constitute an ordered sequence. The second option is to adopt a tri-partite analysis of the traditional verb phrase, such that it consists of three functional features that may have at most one uninterpretable sub-feature each. However, in this dissertation I will not commit to either of those suggestions, considering that this problem is not unique to this analysis (i.e. the assumption that the uninterpretable features of a head are an unordered set is common throughout the minimalist literature).

the  $V_1P$  level without being scattered). The structure in (39) demonstrates the result of this derivation, which verifies that placing the phonological features of the verb in  $v$  captures the correct order of the verb and its objects. Thus, the analysis succeeds in capturing V-to-v movement while also restricting it to structures where the functional head corresponding to  $v$  is projected for independent reasons, similarly to the French V-to-T movement.

(37) John gave Mary the book

(38)  $\langle C, T\{EPP\}, Mod, v\{uD, /gave/\}, V\{uD, uD\}\rangle$



To sum up, in this section I have developed an analysis for three kinds of movement of finite main verbs: (i) the V-to-C movement of German matrix V2 clauses, (ii) the V-to-T movement of French matrix clauses, and (iii) the V-to-v movement of English matrix clauses. Furthermore, I have discussed a few concrete examples that demonstrate the ability of the proposed analysis to derive as little functional structure as is necessary to capture the relevant syntactic facts. In the following section, I will widen the scope of the analysis to include clauses where the finite verb is an auxiliary, rather than a main verb. To achieve this, I will extend the Feature Scattering analysis of head movement to finite auxiliary verbs and I will, furthermore, develop an account for the absence of movement to T or C of non-finite verbs.

### **2.3.2 Movement of Finite versus Non-Finite Verbs**

The aim of this section is to extend the analysis of verb movement outlined above from simple clauses containing a single verbal lexical item (i.e. the main verb) to complex clauses that are constructed from multiple verbal lexical items (i.e. the main verb and a number of auxiliaries). The analysis of verb movement in complex clauses needs to account for two facts. Firstly, finite auxiliary verbs display their own patterns of verb movement, which may be similar to the movement of finite main verbs in the same language (e.g. German and French auxiliaries move to C and T respectively, just like main verbs) or different from it (e.g. English auxiliaries move to T, unlike main verbs). Secondly, non-finite main and auxiliary verbs do not display movement to T or C. Rather, in any clause that involves movement to T or C, it is the single finite verb of the clause (either a main or an auxiliary verb) that moves. As we will see in the following discussion, the movement of finite auxiliaries can be captured rather straightforwardly with the theoretical tools developed in the previous sections. On the other hand, the fact that non-finite verbs do not participate in movement to T or C is less straightforward within the analysis of head movement developed here. In a traditional analysis, this fact is attributed to the existence of a single target head (i.e. C or T) that attracts the closest verb to it. However, according to this proposal, the target position is a result, not a cause, of head movement. Rather, head movement is triggered by some property of the moved head itself. Thus, the uniqueness of the target position cannot account for the contrast in verb movement of finite and non-finite verbs. Instead, I will propose an analysis that relies on Feature Bundling (cf. section 1.4.2 of chapter 1) to capture in a principled way the different properties of finite and non-finite verbs in terms of their featural specifications. Accordingly, I will start this section with a brief reminder of the concept of Feature Bundling and other hypotheses related to it. I will then present an analysis of verb movement in complex clauses in German and French. Finally, I will conclude the section with a discussion of some of the peculiarities of verb movement in English, including the different patterns of

movement of finite auxiliaries versus finite main verbs and *do*-support in the case of sentential negation.

Remember that in section 1.4.2 I proposed that the featural specification of a lexical item is determined in the numeration depending on what other lexical items of the same extended projection are included in the same numeration. The intent of this proposal is to ensure that each functional feature appears only once in a given derivation, thus avoiding the potential scattering of redundant functional projections. To achieve this I have proposed three basic hypotheses. Firstly, I suggested that lexical items are, in a sense, underspecified in the lexicon. Specifically, I proposed that they are specified with a single functional feature, which corresponds to the categorial feature that would be ascribed to them in a traditional analysis. A main verb, for example, will be specified with the functional feature V. In the following analysis we will also be discussing auxiliary verbs, so we need to refine the sequence of functional features by adding the feature Aux<sup>10</sup> between T and v. Accordingly, I assume that auxiliaries will be specified with the functional feature Aux in the lexicon. Secondly, I proposed that the numeration contains two things: (i) a number of underspecified lexical items drawn from the lexicon, and (ii) a number of functional features drawn from a functional 'lexicon' that I call the Inventory of Functional Features (IFF). According to this proposal, the IFF contains a number of sequences of functional features, which express various extended projections. For example, any numeration will include the sequence of functional features <C{...}, T{...}, Aux{...}, v{...}, V{...}>, which correspond to the verbal extended projection. Finally, I have formulated a pre-syntactic operation called Feature Bundling, which allocates functional features to the lexical items that are present in a given numeration. Specifically, I have suggested that Feature Bundling maps each functional feature to the lexical item that carries the closest subordinate functional feature in its underspecified featural specification. In more formal terms,

---

<sup>10</sup> In fact, in a more elaborate analysis we would break down Aux to the functional features Perf, Prog and Pass in order to distinguish between perfective, progressive and passive auxiliaries respectively. However, the simplified description in the text suffices for the purpose of this discussion.

I have defined Feature Bundling as in (40) (repeated from section 1.4.2). So, in a numeration containing an auxiliary specified with Aux and a main verb specified with V, the functional features C and T will be mapped to the auxiliary and the functional feature v will be mapped to the main verb. Thus, the resulting featural specification for the auxiliary will be  $\langle C\{\dots\}, T\{\dots\}, Aux\{\dots\} \rangle$  and the specification for the main verb will be  $\langle v\{\dots\}, V\{\dots\} \rangle$ . In comparison, in a numeration that contains only a main verb, all of the functional features C, T, Aux and v will be mapped to the main verb, giving it the featural specification  $\langle C\{\dots\}, T\{\dots\}, Aux\{\dots\}, v\{\dots\}, V\{\dots\} \rangle$ .

(40) Feature Bundling

In a numeration containing a sequence of functional features  $\langle fnF_n\{\dots\}, \dots, fnF_0\{\dots\} \rangle$ , a functional feature  $fnF_x\{\dots\}$  (where  $n \geq x \geq 0$ ) is mapped onto a lexical item bearing a functional feature  $fnF_y$ , iff:

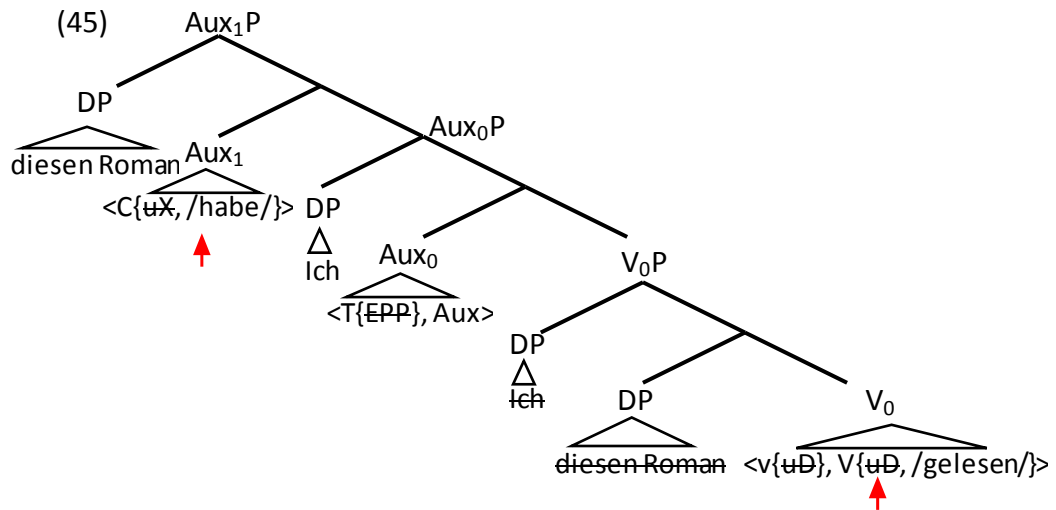
- i.  $x \geq y$ , and
- ii. there is no lexical item bearing a functional feature  $fnF_z$  where  $x \geq z > y$ .

Let us now consider what this understanding of featural specifications adds to the analysis of verb movement developed in the previous section. Remember that I have attributed the V2 position of the finite main verb in the German example in (21b) (repeated in (41) below) to the featural specification in (24) (repeated in (42), with the addition of the newly introduced Aux functional feature). Compare this to an example like (43), where the V2 position is occupied by a finite auxiliary, while the non-finite main verb appears in a clause-final position. As I have mentioned above, after Feature Bundling has taken place the featural specification of the auxiliary will contain the functional features C, T and Aux, while the featural specification of the main verb will contain only v and V. Furthermore, in order to account for the surface position of the auxiliary and the main verb, we have to assume that the phonological features of the former are hosted by the functional feature C, while the phonological features of the latter are hosted by V. Thus, the featural specification of the auxiliary and the main verb must be as in (44a) and (44b) respectively.



- (41) diesen Roman las ich  
 this book read I
- (42) Finite lexical verb (after Feature Bundling)  
 <C{uX, /las/}, T{EPP}, Aux, v{uD}, V{uD}>
- (43) diesen Roman habe Ich gelesen  
 this book have I read
- (44) a. Finite auxiliary verb (after Feature Bundling)  
 <C{uX, /habe/}, T{EPP}, Aux>  
 b. Non-finite lexical verb (after Feature Bundling)  
 <v{uD}, V{uD, /gelesen/}>

Based on the featural specification in (44), the derivation of the clause in (43) will proceed as follows. Firstly, the uninterpretable features of V and v will be eliminated against the object and the subject of the verb projecting  $V_0P$ . At this point the main verb does not have any remaining uninterpretable features, so the derivation will proceed to the auxiliary, which will be merged with  $V_0P$ . Afterwards, the subject will move from the specifier of  $V_0P$  to the specifier of  $Aux_0P$  in order to eliminate the uninterpretable EPP sub-feature of T. However, since the specifier of  $Aux_0P$  is now filled, the object can no longer move to a position that would allow for the uninterpretable uX sub-feature of C to be eliminated. Accordingly, Feature Scattering will move C and its sub-features to  $Aux_1$ , while the object will move to the specifier of  $Aux_1P$  to eliminate the uninterpretable feature uX. The tree in (45) demonstrates that at the end of the derivation the phonological features of the auxiliary are situated in  $Aux_1$  and the phonological features of the main verb are situated in  $V_0$ , which correspond to the V2 and clause-final position respectively.



There is, however, one more issue that needs to be addressed to complete the analysis. So far, I have demonstrated how the featural specifications in (42) and (44) capture the verb movement facts in examples (41) and (43) respectively. Furthermore, I have described how Feature Bundling allocates functional features to underspecified lexical items, with the results shown in (42) and (44). What remains to be done is to formalise the process by which phonological features are placed in the desired position in these featural specifications. I will pursue an approach that attributes the placement of phonological features to the lexical properties of lexical items. Specifically, I will assume that the underspecified featural specification of every lexical item (or, more accurately, of every word form) carries an instruction as to which functional feature must host its phonological features<sup>11</sup>. Thus, in the lexicon, a German finite main verb will be specified as shown

<sup>11</sup> Note, however, that this point highlights a limitation of the analysis of head movement developed in this chapter. Namely, the proposed analysis is not well equipped to account for the fact that, in a given language and structure, lexical items belonging to the same category display a uniform behaviour with respect to head movement (for instance, all German finite verbs move to C and all French finite verbs move to T, while no English finite main verb moves outside the vP). As we have seen throughout this chapter, according to this proposal the surface position of a lexical item depends crucially on the placement of its phonological features within its featural specification, since that factor determines whether those phonological features will be affected by independently triggered Feature Scattering operations. Therefore, the only way to capture generalisations of the kind mentioned above is to assume that the placement of phonological features is uniform across all lexical items of a given category. Ideally, this uniformity should be reduced to some more fundamental property of the linguistic system. However, given the lexicalist approach to

in (46). Note that this featural specification includes, as suggested earlier, the functional feature V that corresponds to the category of the verb. Additionally, (46) includes the phonological features of the verb, which appear as sub-features of C. On the other hand, this featural specification lacks the functional features T, Aux and v, which will be added in the numeration by Feature Bundling. Similarly, a finite auxiliary verb will be specified in the lexicon as shown in (47) and, finally, a non-finite main verb will be specified as in (48).

(46) Finite lexical verb (before Feature Bundling)

<C{..., /las/}, ... , V{...}>

(47) Finite auxiliary verb (before Feature Bundling)

<C{..., /habe/}, ... , Aux>

(48) Non-finite lexical verb (before Feature Bundling)

<..., V{..., /gelesen/}>

The same logic can be applied to the movement of French finite verbs to T. In the previous section I attributed the movement of the finite main verb in example (28) to the featural specification in (29) (repeated below in (49) and (50) respectively). Of course, (50) is the full specification that arises after the application of Feature Bundling. Prior to that, the underspecified finite main verb will have the featural specification in (51), which shows the phonological features of the verb hosted by the functional feature T.

(49) Jean lit souvent des romans

John reads often novels

'John often reads novels'

(50) Finite lexical verb (after Feature Bundling)

<C, T{EPP, /lit/}, Aux, Mod{uA}, v{uD}, V{uD}>

---

the placement of phonological features (in the sense that the relevant information is encoded directly on the lexical items themselves) suggested in the main text, a principled account of this uniformity does not appear to be forthcoming. Nonetheless, even though I acknowledge that this situation is a weakness of the current analysis, I consider it a necessary theoretical trade-off for the exploration of variation in the composition of functional structure, the stated goal of this dissertation.

- (51) Finite lexical verb (before Feature Bundling)  
 <..., T{..., /lit/}, ... , V{...}>

Compare this to a complex clause like (52), where the finite verb is an auxiliary. Here, it is the auxiliary that moves to T (as shown by the fact that it precedes negation), while the non-finite main verb remains inside the vP. In order to account for the movement patterns of these two verbs, it must be the case that the phonological features of the auxiliary are hosted by the functional feature T, while the phonological features of the main verb are hosted by v<sup>12</sup>. Thus, I propose that in the lexicon the finite auxiliary has the underspecified featural specification shown in (53a) and the non-finite main verb has the specification shown in (53b). Furthermore, the two verbs will acquire the full featural specifications shown in (54a-b) after Feature Bundling has taken place (note that in (54a) I am including a functional feature Pol(arity) with an uninterpretable sub-feature uNeg to account for the negative marker appearing in the relevant example).

- (52) Jean n' a pas lu ce livre  
 John prt has Neg read this book  
 'John has not read this book'

- (53) a. Finite auxiliary verb (before Feature Bundling)  
 <..., T{..., /a/}, ... , Aux>  
 b. Non-finite lexical verb (before Feature Bundling)  
 <..., v{..., /lu/}, V{...}>

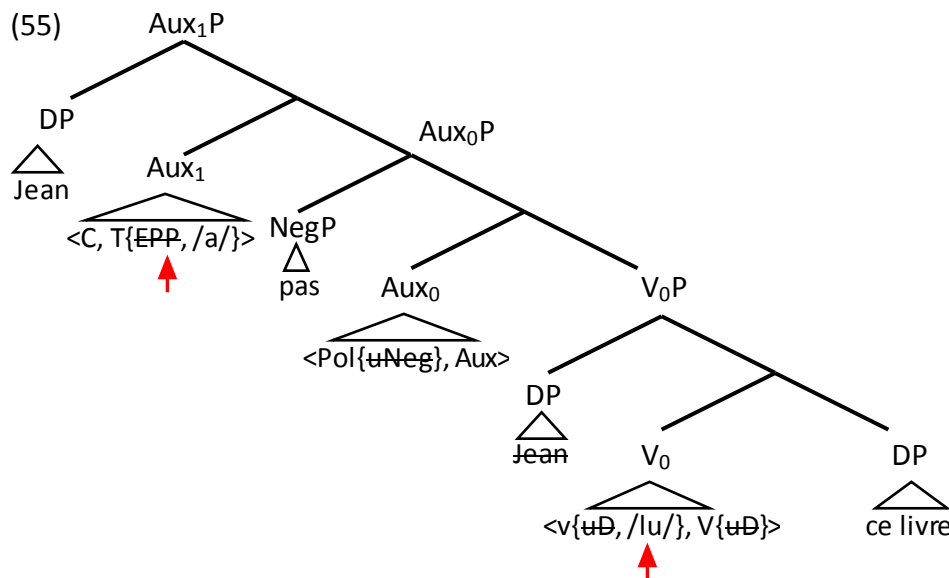
- (54) a. Finite auxiliary verb (after Feature Bundling)  
 <C, T{EPP, /a/}, Pol{uNeg}, Aux>  
 b. Non-finite lexical verb (after Feature Bundling)  
 <v{uD, /lu/}, V{uD}>

According to the featural specifications in (54), the derivation of example (52) will proceed as follows. As expected, the complement and specifier of V<sub>0</sub> will be filled by the object and subject of the verb respectively, thus eliminating the uninterpretable features of V and v. Subsequently, the auxiliary will be introduced in the derivation and merged with V<sub>0</sub>P. The first uninterpretable feature of Aux<sub>0</sub> is

---

<sup>12</sup> Note that the phonological features of the non-finite main verb must be in v, rather than V, to account for the post-verbal position of the objects of a ditransitive verb.

the  $uNeg$  sub-feature of the functional feature  $Pol$ . In order to eliminate this uninterpretable feature, the negative marker will have to appear on the specifier of  $Aux_0P$ . The second (and last) uninterpretable feature of the auxiliary is the  $EPP$  sub-feature of  $T$ . This uninterpretable feature, however, cannot be eliminated at the  $Aux_0$  level since the subject cannot move to its specifier, which is filled by the negative marker. Thus, Feature Scattering moves the functional feature  $T$  and its sub-features to  $Aux_1$ . After Feature Scattering has taken place, the subject will move to the specifier of  $Aux_1P$  and eliminate the uninterpretable  $EPP$  feature. The derivation is represented schematically in (55), which demonstrates that the non-finite main verb surfaces in a position that follows the negative marker (i.e. in  $V_0$ ), while the finite auxiliary has been displaced to a position preceding it (i.e.  $Aux_1$ ).



Finally, let us return to verb movement in English. As we have already seen, English finite main verbs differ from their French counterparts in that they display V-to-v movement, rather than movement to  $T$ . In the previous section, I attributed the surface position of the English verb in example (34) (repeated in (56) below) to the featural specification in (35) (repeated in (57) with the addition of the  $Aux$  functional feature). According to the analysis developed in this section, this featural specification should be the result of Feature Bundling operating on the

underspecified lexical item shown in (58), where the phonological features of the finite main verb appear as sub-features of v.

(56) John often reads novels

(57) Finite lexical verb (after Feature Bundling)  
<C, T{EPP}, Aux, Mod{uA}, v{uD}, /reads/}, V{uD}>

(58) Finite lexical verb (before Feature Bundling)  
<..., v{..., /reads/}, V{...}>

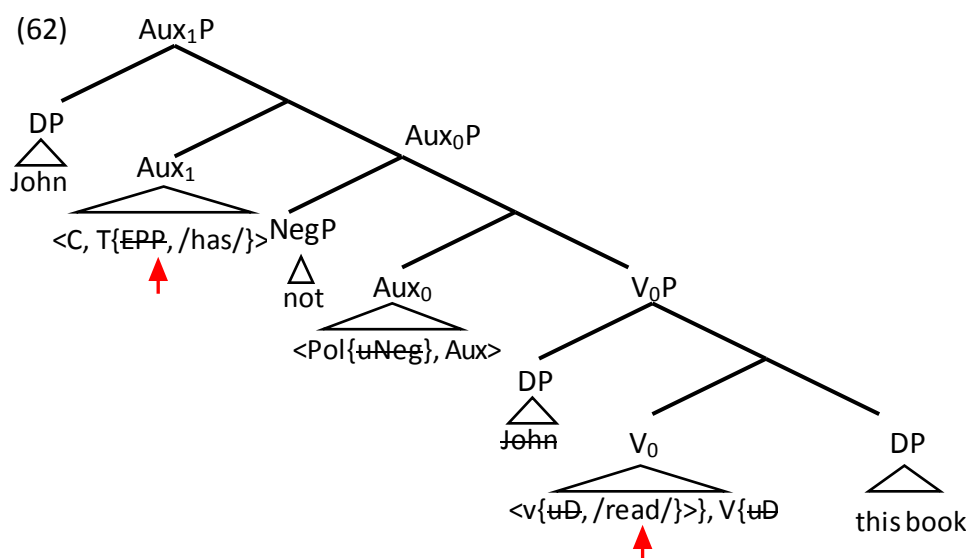
Contrary to English finite main verbs, English finite auxiliaries behave in the same way as their French counterparts. Example (59) demonstrates that English finite auxiliaries appear on the left of the negative marker, which indicates that they move to T. In keeping with the spirit of this analysis, these differences and similarities should be treated as an expression of the lexical properties of finite auxiliaries. Thus, I will assume that the featural specification of an English finite auxiliary in the lexicon is as shown in (60a). Observe that the phonological features of this finite auxiliary appear as sub-features of the functional feature T, unlike the phonological features of English finite main verbs and similar to those of French finite auxiliary and main verbs. Furthermore, since both finite and non-finite English main verbs display V-to-v movement, their underspecified featural specifications will be identical. Thus, in the lexicon, an English non-finite main verb will have a featural specification as shown in (60b), which is the same as the specification of a finite main verb in (58) above. Finally, after the application of Feature Bundling, the finite auxiliary and the non-finite main verb will acquire the full featural specifications in (61a) and (61b) respectively.

(59) John has not read this book

(60) a. Finite auxiliary verb (before Feature Bundling)  
<..., T{..., /has/}, ... , Aux>  
b. Non-finite lexical verb (before Feature Bundling)  
<..., v{..., /read/}, V{...}>

- (61) a. Finite auxiliary verb (after Feature Bundling)  
 <C, T{EPP, /has/}, Pol{uNeg}, Aux>  
 b. Non-finite lexical verb (after Feature Bundling)  
 <v{uD, /read/}, V{uD}>

Note that the featural specifications of the finite auxiliary and the non-finite main verb in (61) are identical to their French counterparts in (54), which is to be expected since the relevant verbs in the two languages display similar movement patterns. Thus, the derivation of example (59) will proceed in the same manner as its French counterpart. The structure in (62) demonstrates the result of this derivation (cf. the discussion of the French derivation in (55) for a detailed description).



In the previous discussion, I sidestepped one peculiarity of English that complicates the analysis of verb movement in this language: namely the phenomenon of *do*-support. Broadly speaking, *do*-support refers to the obligatory use of the dummy auxiliary *do* in various simple tense clauses, including clauses with sentential negation, interrogative clauses, clauses that display negative inversion and clauses with emphatic assertion. I will address the case of sentential negation in the remainder of this section and I will return to the other cases (especially interrogative clauses) in chapter 3. The following examples demonstrate the relevant facts in clauses with sentential negation. Example (63) is a simple tense

affirmative clause where the main verb is finite and there are no additional auxiliary or modal verbs. Compare this to its negative counterpart in (64a). Here, the main verb takes a non-finite form, while a finite form of *do* appears before the negative marker. Note that the dummy auxiliary *do* appears in the same position as other finite auxiliaries (cf. example (59) above). Finally, the ungrammaticality of example (64b) demonstrates that *do*-support is, in fact, obligatory.

(63) John read the book

(64) a. John did not read the book  
b. \*John not read the book

In order to capture these facts, I will make two proposals. Firstly, I will suggest that it is an idiosyncratic lexical property of English main verbs that the functional feature *Pol* in their featural specification cannot carry the sub-feature *uNeg*<sup>13</sup>. The result of this hypothesis is that it becomes impossible to introduce a negative marker in the structure if the only available verbal lexical item in the numeration is the main verb, hence the ungrammaticality of (64b). Thus, in order to derive a negative clause, the numeration must include at least one auxiliary verb. Obviously, if no other auxiliary is compatible with the tense of the clause, then the numeration will have to rely on the dummy auxiliary *do*, as in (64a).

Secondly, I will assume the following featural specifications to account for the derivation of (64a). I will treat the dummy auxiliary *do* as a lexical item with a specification similar to all other auxiliaries. Thus, in the lexicon the dummy auxiliary *do* will be specified as shown in (65a). Furthermore, when included in the numeration, Feature Bundling will ascribe to it the full featural specification shown in (65b). As for the main verb in this example, there is no reason to believe that its specification is any different from other non-finite main verbs. Thus, the main verb

---

<sup>13</sup> Note that it cannot be the case that the main verb is incompatible with the functional feature *Pol*, since I am assuming that functional features are universal. On the other hand, positing cross- or intra-linguistic variation in the sub-features of a functional feature is consistent with the framework developed here.



will be specified as shown in (66a-b), before and after Feature Bundling respectively.

- (65) a. Dummy auxiliary *do* (before Feature Bundling)  
<..., T{..., /did/}, ... , Aux>  
b. Dummy auxiliary *do* (after Feature Bundling)  
<C, T{EPP, /did/}, Pol{uNeg}, Aux>
- (66) a. Non-finite lexical verb (before Feature Bundling)  
<..., v{..., /read/}, V{...}>  
b. Non-finite lexical verb (after Feature Bundling)  
<v{uD, /read/}, V{uD}>

Note that these featural specifications are practically identical to the ones in (60)-(61) above. Therefore, the derivation arising from them will produce the exact same structure as the derivation in (62). As expected, this derivation correctly captures the placement of the dummy auxiliary *do* before the negative marker in example (64a). Thus, this set of assumptions provides an account for the structure, as well as the necessity, of *do*-support in the relevant context.

To summarise, in this section I proposed that the placement of the phonological features of a lexical item in its featural specification is a lexically determined property that may vary from one word form to the other. This hypothesis, in conjunction with the operations Feature Scattering and Feature Bundling proposed in chapter 1, has allowed me to capture (i) the differences in the movement patterns of finite and non-finite main verbs in German and French, (ii) the similarities in the movement patterns of finite auxiliary and finite main verbs in the same languages, and (iii) the differences in the movement patterns of auxiliary and main verbs in English. Thus, this section has concluded the analysis of verb movement in main clauses in the aforementioned languages that was introduced in section 2.3.1. In the following section, I will use the same theoretical tools to further extend the analysis to verb movement in embedded clauses in Germanic languages.

### 2.3.3 Movement in Matrix versus Embedded Clauses

In the previous sections, I have proposed, amongst other things, an analysis of V-to-C movement in German matrix clauses. This analysis carries over rather straightforwardly to other Germanic V2 languages. However, these languages display considerable cross-linguistic variation with respect to verb movement in embedded clauses. In some languages embedded clauses display the same V2 order as matrix clauses, while in others embedded V2 is a limited option. In this section I will provide an account of V-to-C movement (or lack thereof) in embedded clauses across Germanic V2 languages.

I will start this discussion with a brief outline of the relevant empirical facts. We can roughly separate Germanic languages into three categories with respect to the movement of the verb in embedded clauses (see Heycock 2006 for a detailed discussion of the cross-linguistic variation in embedded V2). Firstly, we have languages like Icelandic and Yiddish that always display embedded V2. The following examples from Icelandic demonstrate the point (also note that example (68) demonstrates that the overt complementizer cannot be omitted in Icelandic, which is relevant to the following discussion).

(67) Ég tel að leikarinn sjái áreiðanlega myndina  
I think that the actor saw really the film  
'I think that the actor actually saw the film'

(68) \*Ég tel \_\_\_ leikarinn sjái áreiðanlega myndina  
I think \_\_\_ the actor saw really the film

Vikner (2001)

Unlike Icelandic, Mainland Scandinavian languages do not typically display V2 in embedded clauses. Rather, in these languages embedded V2 is only possible in embedded clauses introduced by a small number of verbs that form a subset of the bridge verbs (with possible variation across languages). The following examples demonstrate the point for Danish. Comparing (69a) with (70a) reveals that embedded V2 is only possible with certain matrix verbs. Furthermore, (70b) shows that even in the environments that allow it, embedded V2 is only an option.

- (69) a. \*Det var en overraskelse at Helge ville gerne læse  
 It was a surprise that Helge would readily read  
 den her bog  
 this here book
- b. Det var en overraskelse at Helge gerne ville læse  
 It was a surprise that Helge readily would read  
 den her bog  
 this here book  
 'It was a surprise that Helge would readily read this here book'
- (70) a. Peter troede at Helge ville gerne læse den her  
 Peter believed that Helge would readily read this here  
 bog  
 book  
 'Peter believed that Helge would readily read this here book'
- b. Peter troede at Helge gerne ville læse den her  
 Peter believed that Helge readily would read this here  
 bog  
 book

Vikner (1995)

Finally, German displays a more restricted version of the pattern seen in Mainland Scandinavian languages. Similarly to those languages, German only allows embedded V2 in embedded clauses introduced by a subset of bridge verbs. However, while the overt complementizer is largely optional in embedded V2 in Mainland Scandinavian, in German the overt complementizer is obligatorily omitted in embedded V2 clauses. This can be seen in examples (71a-b) below. Finally, similarly to Mainland Scandinavian, embedded V2 in German is optional in the relevant environments, as shown in (71c). However, unlike Mainland Scandinavian, embedded clauses that lack V2 require the presence of an overt complementizer, as shown by the ungrammaticality of (71d).

- (71) a. Sie sagte sie wolle keine Bücher kaufen  
 she said she wants no books buy  
 'She said (that) she didn't want to buy any books.'
- b. \*Sie sagte daß sie wolle keine Bücher kaufen  
 she said that she wants no books buy
- c. Sie sagte daß sie keine Bücher kaufen wolle  
 she said that she no books buy wants
- d. \*Sie sagte sie keine Bücher kaufen wolle  
 she said she no books buy wants

Heycock (2006)

These facts suggest that in German the movement of the verb to a V2 position is in complementary distribution with an overt complementizer. This observation is at the heart of the traditional analysis of embedded V2, originally proposed by den Besten (1983), which suggests that the verb in a V2 clause surfaces in the same structural position as an overt complementizer. Thus, the complementary distribution of embedded V2 and an overt complementizer is attributed to the fact that the moved verb and the complementizer are competing for the same position. It is however obvious that this analysis does not carry over to other Germanic languages. As we have already seen, in languages of the Icelandic type the co-occurrence of embedded V2 with an overt complementizer is the norm, while in languages of the Mainland Scandinavian type the co-occurrence of embedded V2 and an overt complementizer is a limited option, but it is an available option nonetheless. The attempts to reconcile den Besten's original analysis with these observations have generally followed one of two paths. The first possibility is to assume that languages that allow (or require) embedded V2 under an overt complementizer have a richer functional structure compared to languages that do not (cf. Vikner 1990, 1995). If, for example, languages of the Icelandic or Mainland Scandinavian type allow for CP-recursion while languages of the German type do not, then the co-occurrence of V2 with an overt complementizer will be possible in the former but not in the latter. The second possibility is to assume that the two language types do not differ in their underlying functional structure but in the placement of the moved verb, as well as the fronted constituent (cf. Diesing 1990, Santorini 1992, Rögnvaldsson and Thráinsson 1990). Specifically, the hypothesis is

that in V2 clauses of the German type the verb moves to C and the fronted constituent moves to the specifier of C, while in V2 clauses of the Icelandic and Mainland Scandinavian type the verb and the fronted constituent move only as high as T and the specifier of T respectively. Thus, it is only in languages of the German type that an embedded verb undergoing V2 and an overt complementizer compete for the same position. In languages of the Icelandic or Mainland Scandinavian type this is simply not the case.

In the following discussion, I will explore how these ideas on embedded V2 in Germanic languages may be formalised within the framework of functional structure developed in this dissertation. In regard to embedded V2 of the German type, I will show that, although it is not entirely transparent, the current framework can maintain the traditional hypothesis that the overt complementizer and the embedded verb compete for the same position. As for embedded V2 of the Icelandic and Mainland Scandinavian type, I will develop an account that incorporates aspects of both of the competing analyses mentioned in the previous paragraph. Roughly speaking, I will propose that finite verbs in embedded V2 clauses in these languages target a different structural position from their counterparts in languages of the German type. Additionally, I will show that due to Feature Scattering the functional structure of embedded V2 clauses is different across the various Germanic languages.

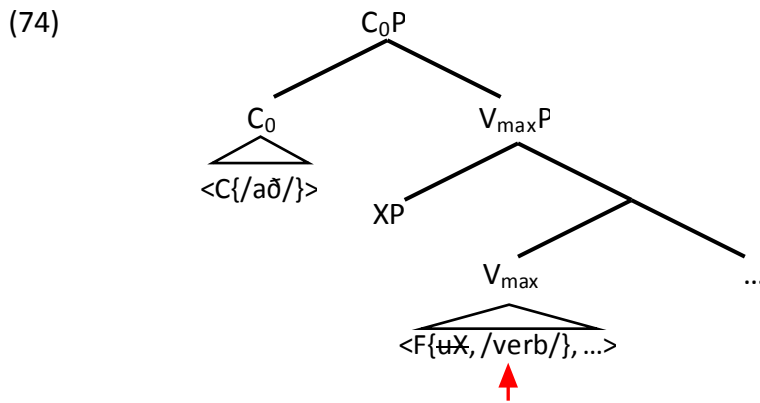
Let us consider the case of Icelandic first. As I have mentioned above, the analysis of embedded V2 in Icelandic will be based on the assumption that the surface position of the moved verb is distinct from the position of the overt complementizer. Of course, according to the analysis of verb movement developed in the previous sections, the surface position of the complementizer and the finite verb will be expressed in terms of the placement of phonological features in the featural specification of the relevant lexical items. Specifically, I will propose that the phonological features of the overt complementizer are situated on the functional feature C. Thus, its featural specification before Feature Bundling will be as shown in (72a). As for the phonological features of a finite main verb (or a finite

auxiliary), I assume that they are situated on some functional feature F, which follows C in the sequence of functional features. The relevant featural specification before Feature Bundling is shown in (72b). Note that the intention behind this hypothesis is to draw a broad sketch of verb movement in Icelandic embedded clauses that will allow for comparisons with embedded clauses in other Germanic languages. Thus, I will leave open the possibility that F corresponds to T or one of the lower heads of the left periphery (e.g. Rizzi's 1997 Fin).

- (72) a. Lexical complementizer (before Feature Bundling)  
 <C{/að/}>  
 b. Finite lexical verb (before Feature Bundling)  
 <..., F{..., /verb/}, ... , V{...}>

In a numeration that contains the complementizer and the finite verb in (72), Feature Bundling will ascribe to these two lexical items the featural specifications shown in (73) below. These specifications will result in a structure where the finite verb will surface in the position of  $V_{\max}$ , the highest head that is associated with its featural specification (i.e. either the highest scattered functional head or the lexical head itself, depending on whether Feature Scattering takes place or not). This is due to the placement of the phonological features of the finite verb in the leftmost functional feature in its specification, namely F. Furthermore, the specifier of the  $V_{\max}$  will be occupied by a fronted constituent, due to the uninterpretable  $uX$  sub-feature that is hosted by the same functional feature F. Finally, the maximal projection  $V_{\max}P$  of the same head will be the complement of the overt complementizer. Thus, the resulting structure will capture the co-occurrence of an overt complementizer with a V2 order that is characteristic of Icelandic. The tree in (74) demonstrates the relevant parts of this structure.

- (73) a. Lexical complementizer (after Feature Bundling)  
 <C{/að/}>  
 b. Finite lexical verb (after Feature Bundling)  
 <F{ $uX$ , /verb/}, ... , v{...}, V{...}>



Note, however, that the analysis presented so far is not sufficient to account for the obligatory presence of an overt complementizer in Icelandic embedded clauses. To clarify, consider the following. The above structure arises from a numeration that contains both an overt complementizer and a verb. However, in principle, it is possible for the complementizer to be absent from the numeration. In that case, Feature Bundling will assign the functional feature C to the featural specification of the verb. Furthermore, this featural specification will culminate to an embedded clause consisting only of the  $V_{\max}P$  in (74). Therefore, the analysis, as it stands, erroneously allows for the omission of the overt complementizer. To address this problem, I will propose a solution that follows the same logic as the analysis of English *do*-support in the previous section. Remember that according to the framework developed in chapter 1 functional features are stored separately from lexical items. The former are drawn from the Inventory of Functional Features while the latter are drawn from the lexicon. It is only during the numeration that the two are combined due to Feature Bundling. Therefore, it is possible that the properties of the former are not always compatible with the properties of the latter. With that in mind, I will suggest that the functional feature C of an embedded clause bears some interpretable sub-feature that distinguishes embedded from matrix clauses. Furthermore, I will suggest that it is a lexical property of Icelandic verbs that they are incompatible with that interpretable sub-feature. Finally, I will assume that the overt complementizer does not have a similar restriction. Therefore, in a numeration that contains an embedded verb but lacks a complementizer, Feature Bundling will fail to assign the functional feature C

of the embedded clause to any lexical item, because there is no appropriate candidate available. Of course, this problem does not arise in a numeration that contains an embedded verb as well as an overt complementizer. Thus, I conclude that the only licit numeration for an embedded clause in Icelandic is the one shown in (73) above and, by extension, the only available derivation for that clause is the one shown in (74).

Let us now move on to the case of German. The basis of my analysis will be that embedded verbs in German can have one of two possible featural specifications: one that induces V-to-C movement and one that does not. Thus, the four embedded clauses in (71) correspond to the four possible combinations of the two verbal specifications and the presence/absence of an overt complementizer in the numeration. My task then will be to show why some of these numerations converge while others do not. Note that the focus of the following analysis is placed on the structure of embedded clauses and not their distribution. That is, I will develop an account for the obligatory presence of the overt complementizer in embedded clauses with a clause-final verb and the obligatory omission of the complementizer in embedded V2 clauses, but I will not address the fact that the former can be selected by any matrix verb while the latter can only be selected by a small number of verbs.

Similarly to the analysis of V2 in matrix clauses in section 2.3.2 and the analysis of Icelandic embedded V2 in this section, the movement of a verb to a V2 position in a German embedded clause can be attributed to the placement of its phonological features within its featural specification. As long as the phonological features of the verb are situated relatively high in the sequence of functional features, the verb will be pronounced in the position of the highest head associated with its featural specification. However, in the case of German, it is also necessary to keep under consideration the fact that V2 in an embedded clause cannot co-occur with an overt complementizer. I will suggest that the traditional hypothesis that the moved verb and the overt complementizer compete for the same functional position can be expressed in terms of the placement of the



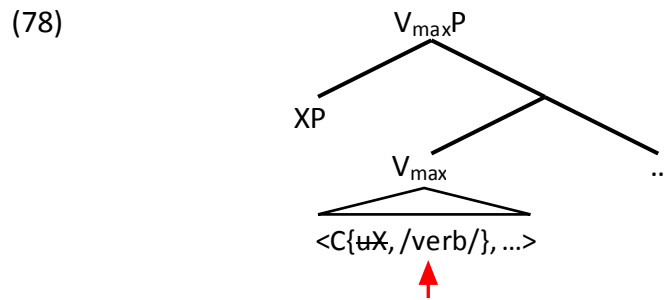
phonological features of these two lexical items in their respective featural specifications. To clarify, consider the following. Let us assume that the overt complementizer in German, similarly to Icelandic, has the featural specification shown in (75) before Feature Bundling. Note that this specification indicates that the phonological features of the complementizer appear on the functional feature C. Let us further assume that before Feature Bundling the featural specification of an embedded verb which surfaces in a V2 position looks like (76), where the phonological features of the verb are again hosted by C.

(75) Lexical complementizer (before Feature Bundling)  
 <C{/daß/}>

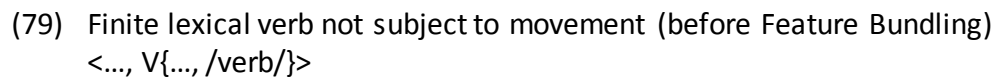
(76) Finite lexical verb subject to movement (before Feature Bundling)  
 <C{..., /verb/}, ... , V{...}>

On these assumptions, a numeration that contains both the verb in (76) and the overt complementizer in (75) is problematic. On one hand Feature Bundling is expected to assign the functional feature C to the complementizer but, on the other hand, the verb also needs a C in its featural specification to host its phonological features. I suggest that Feature Bundling cannot resolve this conflict and, therefore, this numeration is illicit. Consequently, a numeration that includes the verb in (76) must exclude the overt complementizer. Then, Feature Bundling will assign to the embedded verb the full featural specification shown in (77). Furthermore, this numeration will result in an embedded clause that displays V2 (i.e. the verb will surface in the position of the  $V_{max}$ , which will also have the fronted constituent in its specifier) and lacks an overt complementizer. The tree in (78) demonstrates the relevant parts of this structure.

(77) Finite lexical verb subject to movement (after Feature Bundling)  
 <C{uX, /verb/}, T{EPP}, v{uD}, V{uD}>



As I have mentioned above, the featural specification in (76)-(77) is only one of the two possible specifications for embedded verbs in German. The finite verb of an embedded clause that does not display V2 must have a rather different featural specification. In this case the phonological features of the verb must be placed relatively low within its featural specification so that they will not be displaced from the base position of the verb by any potential Feature Scattering operation. Accordingly, I will assume that the specification of such a verb is as shown in (79) before Feature Bundling.

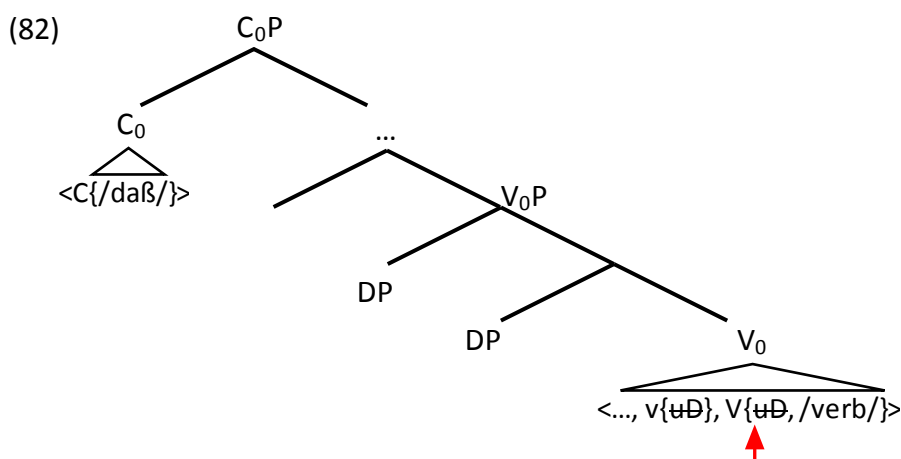


Note that the placement of the phonological features of the verb in this specification not only captures the absence of V-to-C movement but it also allows for the co-existence of the verb and the overt complementizer in the same numeration. As expected, Feature Bundling will assign the functional feature C to the overt complementizer and the functional feature V to the verb. Thus, both the complementizer and the verb will have in their featural specification the appropriate functional feature that ought to host their respective phonological features. The full featural specifications that will arise after Feature Bundling are shown in (80) and (81) below. Furthermore, the structure that will arise from these lexical items is shown in the (partial) tree in (82). What is relevant for the purposes of this discussion is that the embedded clause is headed by an overt complementizer and, furthermore, the embedded verb appears in a clause-final position regardless of how many functional heads may be scattered from it. Thus,

this structure correctly captures the properties of a non-V2 embedded clause in German.

(80) Lexical complementizer (after Feature Bundling)  
 <C{/daß/}>

(81) Finite lexical verb not subject to movement (after Feature Bundling)  
 <T{EPP}, v{uD}, V{uD, /verb/}>



However, this analysis of German non-V2 embedded clauses raises a problem reminiscent of the discussion of embedded clauses in Icelandic. Namely, a numeration that includes the verb in (79) but lacks the overt complementizer in (75) will produce an unattested embedded clause where the verb appears in a clause-final position and the complementizer is absent (cf. the ungrammatical (71)). I will approach this issue in the same fashion that I approached the obligatory complementizer in Icelandic. Specifically, I will assume that German is similar to Icelandic in that matrix and embedded clauses are distinguished by some interpretable sub-feature that is borne by the functional feature C. Furthermore, I will assume that this interpretable feature cannot be assigned to the featural specification of a non-V2 verb like (79). Thus, the non-V2 verb can only be used in a numeration that also includes the overt complementizer, which can bear the problematic interpretable feature. Note that this restriction only applies to the non-V2 featural specification of the embedded verb. Since a V2 verb can appear in an embedded clause that lacks an overt complementizer (and, indeed, it must do so

for independent reasons), it has to be the case that the V2 specification in (76) is compatible with the aforementioned interpretable sub-feature.

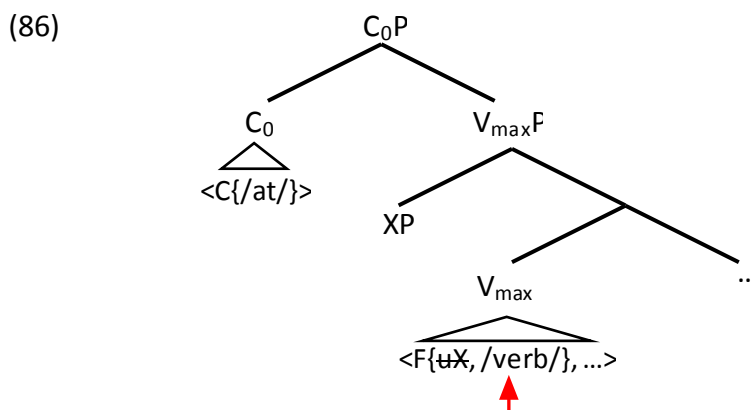
Finally, let us consider the Mainland Scandinavian pattern. The situation in Mainland Scandinavian is similar to the one in German in that embedded V2 is only possible in the complement of select matrix verbs, while the usual case is for embedded clauses to lack V-to-C movement (cf. the Danish examples in (69)-(70)). Thus, similar to the analysis of German, I will assume that embedded verbs in Mainland Scandinavian have two possible featural specifications: one for embedded clauses that involve V-to-C movement and a different one for embedded clauses that do not. I will discuss the derivations that arise from these two specifications in turn, starting from the derivation of embedded V2 clauses. As I have mentioned previously, embedded V2 in Mainland Scandinavian can co-occur with an overt complementizer (cf. example (70a)). In this respect, Mainland Scandinavian is similar to Icelandic and different from German. Remember that this contrast was attributed to the placement of the phonological features of the verb within its featural specification. In German, the phonological features of the verb and the overt complementizer need to be placed on the same functional feature. I have argued that this situation creates a conflict, which prevents Feature Bundling from creating licit featural specifications for the two lexical items. On the other hand, in Icelandic, the phonological features of the verb and the overt complementizer are placed on different functional features, which entails that the previous conflict does not arise. Thus, I will propose that the featural specifications of the verb and the overt complementizer in Mainland Scandinavian are similar to their Icelandic counterparts in this respect. Specifically, I will assume that the overt complementizer is specified as in (83) before Feature Bundling. Furthermore, I will assume that the V2 specification of a finite verb is as in (84) before Feature Bundling.

- (83) Lexical complementizer (before Feature Bundling)  
<C{/at/}>

- (84) Finite lexical verb (before Feature Bundling)  
 <..., F{..., /verb/}, ... , V{...}>

However, the similarities between Mainland Scandinavian and Icelandic stop here. Unlike Icelandic, the overt complementizer in Mainland Scandinavian is not obligatory. Remember that the obligatory presence of the complementizer in Icelandic was attributed to an interpretable sub-feature of the functional feature C that is incompatible with the featural specification of the verb. Consequently, Feature Bundling cannot assign C to an appropriate lexical item in a numeration that lacks an overt complementizer. Obviously, Mainland Scandinavian verbs are not subject to a similar restriction. Thus, the verb in (84) can participate in two, equally well-formed numerations: one that also includes the overt complementizer in (83) and one that does not. In the first case, Feature Bundling will assign to the two lexical items the full featural specification shown in (85). As expected, these specifications will give rise to an embedded V2 clause that is headed by an overt complementizer, as shown in (86). Note that this is identical to the structure of the Icelandic embedded clause in (74) above.

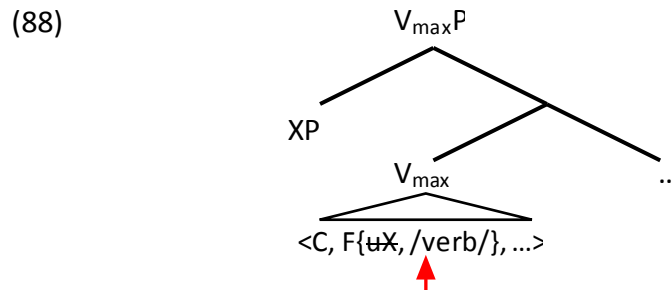
- (85) a. Lexical complementizer (after Feature Bundling)  
 <C{/at/}>  
 b. Finite lexical verb (after Feature Bundling)  
 <F{uX, /verb/}, ... , v{...}, V{...}>



In the second case, Feature Bundling will assign to the verb the featural specification shown in (87). This specification will also produce an embedded V2 clause with V2, but this time without an overt complementizer, as shown in (88).

Note that this structure is comparable to the German embedded clause in (78), with the subtle difference that the  $V_{\max}$  contains the functional feature F as well as C.

(87) Finite lexical verb (after Feature Bundling)  
 $\langle C, F\{uX, /verb/\}, \dots, v\{\dots\}, V\{\dots\}\rangle$

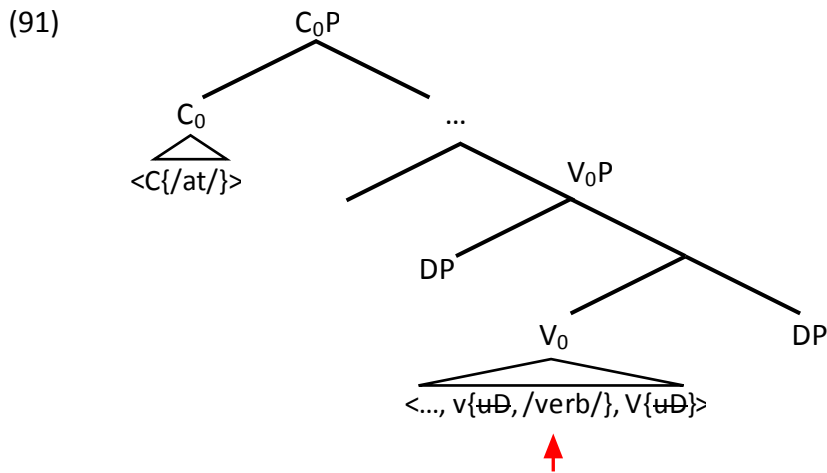


Finally, the non-V2 embedded clauses in Mainland Scandinavian will involve a verb with a different featural specification. Considering that these clauses display an SVO word order, I will assume that in this case the phonological features of the verb are placed on the functional feature  $v$ . Therefore, the featural specification of the relevant verb will be as shown in (89) before Feature Bundling.

(89) Finite lexical verb (before Feature Bundling)  
 $\langle \dots, v\{\dots, /verb/\}, V\{\dots\}\rangle$

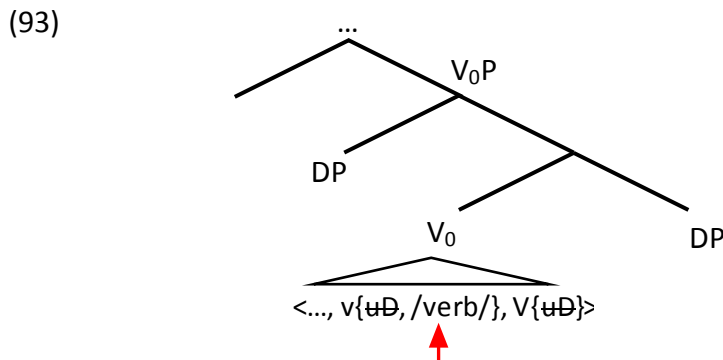
Furthermore, the overt complementizer is optional in non-V2 embedded clauses in Mainland Scandinavian, just like in their V2 counterparts. Thus, I will assume that the verb in (89) does not pose any restriction as to whether it may appear alongside the overt complementizer or on its own in the numeration. In a numeration that contains both the overt complementizer and the verb in (89), Feature Bundling will assign to the two lexical items the featural specifications shown in (90). The structure in (91) demonstrates the resulting clause, which displays an SVO word order and is headed by the overt complementizer. Note that apart from the directionality of the head of  $V_0P$ , this structure is similar to the German non-V2 embedded clause in (82) above.

- (90) a. Lexical complementizer (after Feature Bundling)  
 <C{/at/}>  
 b. Finite lexical verb (after Feature Bundling)  
 <F, ... , v{uD, /verb/}, V{uD}>



Finally, in a numeration that contains the verb in (89) but lacks an overt complementizer, the verb will be assigned the featural specification in (92). As expected, this specification will result in a complementizer-less embedded clause with an SVO word order, as in (93).

- (92) Finite lexical verb (after Feature Bundling)  
 <C, ... , v{uD, /verb/}, V{uD}>



To sum up, I have proposed an analysis of embedded clauses in V2 Germanic languages that involves three components. Firstly, I have provided an account for the cross-linguistic variation in the word order of embedded clauses. The relevant empirical facts are: (i) embedded clauses in languages of the Icelandic type always display V2, (ii) embedded clauses in languages of the German and Mainland

Scandinavian type do not typically display V2, and (iii) embedded V2 is a limited option that is only available in certain contexts in languages of the German and Mainland Scandinavian type. To account for these facts, I suggested that embedded verbs undergoing movement to a V2 position have a distinct featural specification from verbs that surface in their base position. Therefore, the variation in the word order of embedded clauses can be attributed to the availability of one or both of these featural specifications across different languages. Secondly, I have proposed a mechanism to capture the obligatory presence of an overt complementizer in certain contexts. The relevant cases include: (i) embedded clauses in languages of the Icelandic type, and (ii) non-V2 embedded clauses in languages of the German type. I have suggested that the obligatory complementizer in these cases is the result of a failure of Feature Bundling to assign a licit featural specification to a finite embedded verb, unless the numeration contains an overt complementizer. Specifically, I proposed that the functional feature C of an embedded clause bears some interpretable sub-feature that is incompatible with Icelandic verbs and German non-V2 verbs. Accordingly, Feature Bundling will only be able to assign a licit featural specification to these verbs in a numeration where C is assigned to a different lexical item. Thus, the numeration of the relevant clauses must include the overt complementizer. Thirdly, I proposed a mechanism to capture the obligatory absence of the overt complementizer in embedded V2 clauses in languages of the German type. This mechanism is also based on a failure of Feature Bundling to assign a licit featural specification to the relevant verb, but in this case the problem arises in different circumstances and for different underlying reasons. Specifically, I have suggested that the phonological features of an embedded V2 verb in the relevant languages must be placed on the functional feature C. However, C will not be assigned to the verb in a numeration that includes an overt complementizer, leaving the phonological features of the verb stranded. Thus, the numeration of the relevant clause must exclude the overt complementizer to prevent this problem. Finally, I have suggested that the conditions resulting in an obligatory complementizer and those resulting in its obligatory absence are independent of



each other. Thus, when neither of the relevant mechanisms is in play, the overt complementizer is entirely optional. I have argued that this is the case in embedded clauses in languages of the Mainland Scandinavian type.

## **2.4 Conclusions**

In this chapter, I evaluated the merits of re-projective theories of functional structure with respect to the analysis of head movement. In section 2.2.2, I demonstrated that re-projective theories are indeed very successful in capturing the unique properties of head movement that distinguish it from phrasal movement. However, in section 2.2.3, I have argued that treating re-projection as an extension of head movement (as proposed by Ackema et al 1993 or Bury 2003, for example) limits the ability of the theory to capture the properties of functional structure. Accordingly, I adopted the alternative view that head movement is derived by re-projection, which I treat as a fundamental structure-building operation. To capture this hypothesis, I proposed that head movement boils down to the incidental displacement of the phonological features of a lexical item by independently triggered Feature Scattering operations.

In section 2.3.1, I demonstrated how this proposal captures the V-to-C movement of V2 Germanic languages, the V-to-T movement of Romance languages and the V-to-v movement of main verbs in English. In section 2.3.2, I extended the analysis to the movement of finite auxiliary verbs in the same languages. In the same section, I also addressed the fact that only the finite verb of a given clause may display movement to T or C. To account for this, I suggested that the phonological features of a lexical item may be placed in different positions depending on the word form that is used in a given structure. Finally, in section 2.3.3, I addressed the interaction between verb movement and the distribution of the overt complementizer in embedded clauses in V2 Germanic languages. In this respect, I made two fundamental proposals. Firstly, I attributed the obligatory absence of the overt complementizer in certain contexts (i.e. in German V2 embedded clauses) to the hypothesis that the phonological features of both the

complementizer and the verb are situated on the same functional feature. I have suggested that this situation creates a conflict in the numeration (i.e. a single functional feature has to be assigned to two lexical items) that can only be resolved by excluding the overt complementizer. Secondly, I attributed the obligatory presence of the overt complementizer in certain cases (i.e. in Icelandic embedded clauses and German non-V2 embedded clauses) to the hypothesis that the functional feature C bears an interpretable feature that distinguishes embedded from matrix clauses. Furthermore, I have suggested that the relevant interpretable feature is lexically incompatible with the featural specification of an embedded verb. Thus, the numeration has to include the overt complementizer, in order to avoid the illicit assignment of the relevant feature to the featural specification of the verb. Finally, I have suggested that when neither of those factors is in play, the overt complementizer becomes optional (as is the case in Mainland Scandinavian embedded clauses).



### **3.1 Introduction**

In this chapter, I focus on two English structures: *wh*-questions and embedded clauses involving *wh*-extraction. These two structures are some of the most recognisable examples of subject/non-subject asymmetries. The former display an asymmetry with respect to Subject Auxiliary Inversion and *do*-support and the latter display an asymmetry with respect to the distribution of the overt complementizer *that* (i.e. the *that*-trace effect). Since Koopman (1983), it is generally understood that the similarities between these two structures (i.e. they both involve *wh*-movement and they both display subject/non-subject asymmetries) cannot be accidental. Indeed, there is a significant body of research that has pursued a uniform analysis of these two phenomena (amongst others Rizzi 1996, Grimshaw 1997, Pesetsky and Torrego 2001, Roussou 2002, Ishii 2004). In this chapter, I will undertake the same pursuit from the point of view of a re-projective analysis of functional structure. Initially, I will demonstrate that the Feature Scattering analysis of functional structure developed in the previous chapters reaffirms a promising approach to subject/non-subject asymmetries in English *wh*-questions that goes back to George's (1980) Vacuous Movement Hypothesis. In brief, I will show that subject *wh*-questions have a smaller functional structure than their non-subject counterparts, due to the fact that *wh*-subjects do not need to undergo *wh*-movement to appear in a scope-taking position. Therefore, the absence of Subject Auxiliary Inversion and *do*-support in subject *wh*-questions can be

attributed to their reduced functional structure. I will then flesh out this preliminary observation in such a way that the analysis of *wh*-questions can be extended to embedded clauses involving *wh*-extraction. Finally, I will consider if the analysis can be further extended to the anti-*that*-trace effect observed in English (non-*wh*) relative clauses.

This chapter is organised as follows. In section 3.2, I propose an account of (i) the movement of auxiliary verbs to C, and (ii) the phenomenon of *do*-support in English *wh*-questions. Furthermore, I demonstrate that whether the *wh*-word originates in a subject or a non-subject position has an effect on both auxiliary movement and *do*-support. In section 3.3.1, I extend the analysis to the distribution of the overt complementizer *that* in embedded clauses that involve *wh*-extraction (i.e. the *that*-trace effect). In section 3.3.2, I discuss two additional and more complex phenomena that are related to the *that*-trace effect. Those are the amelioration of the *that*-trace effect in embedded clauses that contain a TP adverbial phrase, and the anti-*that*-trace effect observed in English relative clauses.

### **3.2 Subject/Non-Subject Asymmetries in English Wh-Questions**

In this section I will develop an analysis of English *wh*-questions. This empirical domain is relevant to the re-projective theory of functional structure developed in the previous chapters in a number of ways. Firstly, *wh*-questions in English involve movement of the finite auxiliary to C, which is comparable to the V-to-C or Aux-to-C movement of V2 Germanic languages (cf. section 2.3.1). Secondly, simple present and simple past tense *wh*-questions provide another example of *do*-support (cf. the discussion on English sentential negation in section 2.3.2). Finally, English *wh*-questions display a subject/non-subject asymmetry in relation to Aux-to-C movement and *do*-support, which is arguably suggestive of variation in the size of functional structure. Accordingly, one of the main goals of the following discussion is to explore how Feature Scattering may account for this structural asymmetry. The section is organised as follows. I will start the discussion with a brief overview of the relevant empirical data. I will then address the structure of *wh*-questions that

include an auxiliary. At this point, I will focus particularly on the contrast between subject and object wh-questions with regard to Aux-to-C movement. Following that, I will provide an analysis of simple tense wh-questions, including an account of the contrasting pattern of *do*-support between subject and object wh-questions. Finally, I will consider the more complex case of adverb wh-questions, which will require a subtle refinement of the previous analysis.

One of the distinctive properties of non-subject wh-questions in English is that they display Subject Auxiliary Inversion (SAI). The term refers to the inverted order of the auxiliary and the subject in a wh-question in comparison to its declarative counterpart. The contrast between the declarative and interrogative clauses in (1) and (2) demonstrates this point. Under the usual assumption that a declarative clause is a more basic structure, the word order in a wh-question has to be attributed to head movement of the auxiliary to some functional position above the subject. Specifically, SAI is generally treated as an instance of T-to-C movement, in analogy to Germanic V2 (cf. Koopman 1983, Chomsky 1986, Rizzi 1996, Pesetsky and Torrego 2001, amongst many others).

- (1) John is playing chess
- (2) a. What was John playing?  
b. \*What John was playing?

Another characteristic feature of non-subject wh-questions in English is the use of the dummy auxiliary *do* in examples like (3), where the corresponding declarative clause would lack an auxiliary. Note that the order of the auxiliary in relation to the subject in (3a) is the same as in (2a), indicating that non-subject wh-questions with *do*-support also involve T-to-C movement. In fact, it is generally assumed that T-to-C movement plays some role in triggering *do*-support in examples like (3) (see the previous references for various implementations of this broad idea).

- (3) a. What does John play?  
b. \*What John plays?

By comparison, subject *wh*-questions do not display SAI or *do*-support, as shown in the examples in (4)-(5) below. Note that the absence of SAI in (4) does not provide conclusive evidence of absence of T-to-C movement, considering that the subject has to be fronted for independent reasons (i.e. because it is a *wh*-word). However, assuming that in *wh*-questions *do*-support is linked with T-to-C movement, examples like (5) suggest that there is no such movement in subject *wh*-questions. Thus, it is generally assumed that the contrasting pattern of *do*-support in subject versus object *wh*-questions is a symptom of an underlying subject/object asymmetry in the availability of T-to-C movement in the relevant structures.

(4) Who was playing chess?

(5) Who plays chess?

Finally, it is important to note that the examples in (4)-(5) do not present any concrete evidence regarding the final position of the *wh*-subject. Assuming that the auxiliary in these examples surfaces in T, it is unclear whether the *wh*-Subject moves all the way to the specifier of C, or only as far as the specifier of T. By and large, the former approach enjoys wider support in the literature (see, in particular, Koopman 1982, 1983 and Rizzi 1996 for a discussion of the merits of this analysis). As for the latter approach, it has been defended by George (1980), Chomsky (1986) and Ishii (2004) who argue that movement operations without a discernable effect on linear word order ought to be excluded from syntactic theory (the Vacuous Movement Hypothesis). In the following discussion, I will show that the Feature Scattering theory of functional structure developed in this dissertation leads to a similar conclusion. Therefore, I will adopt the view that *wh*-subjects surface in a lower structural position in comparison to *wh*-objects.

After this brief overview of the relevant empirical data, I will proceed to the analysis of the *wh*-questions in (2) and (4), which include an auxiliary. The first thing to consider is the featural specification of the main and the auxiliary verb that participate in the relevant derivations. Assuming the usual sequence of functional

features (i.e. roughly  $\langle C\{\dots\}, T\{\dots\}, Aux\{\dots\}, v\{\dots\}, V\{\dots\} \rangle$ ), Feature Bundling will assign the functional features C, T and Aux to the auxiliary, and the functional features v and V to the main verb. For the most part, the functional features in the specification of the verb and the auxiliary will have the expected sub-features. Namely, V and v will have one uD sub-feature each, and T will have an EPP sub-feature. Furthermore, I will assume that C has a uWh sub-feature, which will be responsible for the fronting of the wh-word.

In addition to uWh, I propose that the functional feature C has an interpretable sub-feature Force<sup>1</sup>. This sub-feature can be marked as declarative (i.e. Force:Decl) or interrogative (i.e. Force:Int). However, I assume that Force is not marked with a Decl or Int value in the numeration. Rather, it acquires an appropriate value in narrow syntax, by entering a relation with its specifier. If the constituent in its specifier is a wh-word, Force will be marked as interrogative. Otherwise, it will be marked as declarative. Furthermore, note that I suggested that Force is an interpretable feature. Therefore, Force can enter a syntactic relation if the appropriate configuration is created by independent reasons, but it cannot trigger Feature Scattering or other syntactic operations on its own. The reasoning behind this proposal will become clearer in the following analysis.

Finally, I will assume that the phonological features of the auxiliary are hosted by the functional feature C, to account for its surface position at the left of the subject. Similarly, I assume that the phonological features of the verb are hosted by v. Therefore, putting all of the above together, the featural specifications of the auxiliary and the main verb (before and after Feature Bundling) will be as shown in (6)-(7) below.

- (6) a. Lexical verb (before Feature Bundling)  
 $\langle \dots, v\{\dots, /verb/\}, V\{\dots\} \rangle$   
 b. Lexical auxiliary verb (before Feature Bundling)  
 $\langle C\{\dots, /aux/\}, \dots, Aux\{\dots\} \rangle$

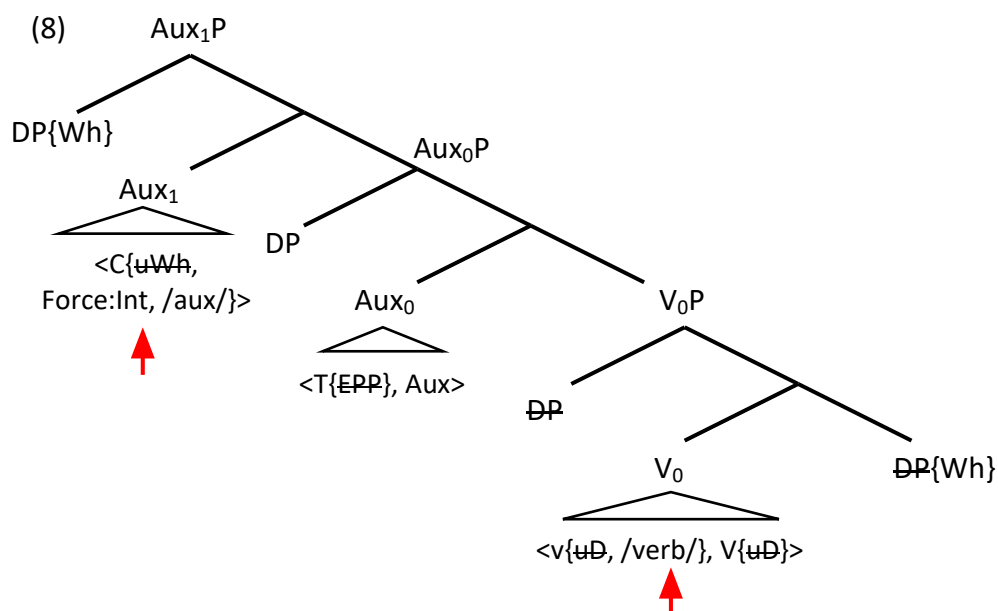
---

<sup>1</sup> See Adger's (2003) clause-type feature for a partly similar proposal.

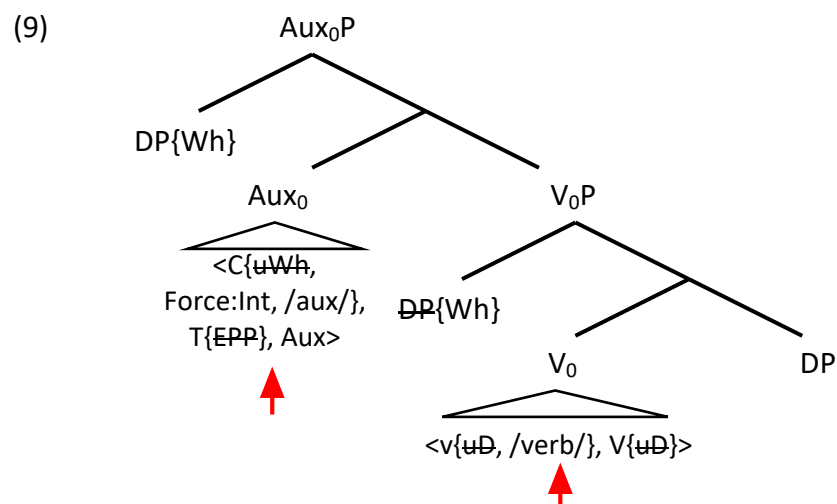


- (7) a. Lexical verb (after Feature Bundling)  
 <v{uD, /verb/}, V{uD}>  
 b. Lexical auxiliary verb (after Feature Bundling)  
 <C{uWh, Force, /aux/}, T{EPP}, Aux>

Let us now consider the derivations that may arise from these featural specifications. The structure in (8) represents the derivation of an object wh-question like (2a). As expected, the derivation will start with the projection of V<sub>0</sub>P. Afterwards, Aux<sub>0</sub> will be merged in the structure and attract the subject of the clause to its specifier, due to the EPP sub-feature of T. Finally, the derivation must eliminate the uninterpretable uWh sub-feature of C. Of course, this operation cannot take place at the Aux<sub>0</sub>P level, because the specifier of Aux<sub>0</sub> is not occupied by a wh-word. Therefore, C and its sub-features will be scattered to Aux<sub>1</sub>. The scattering of Aux<sub>1</sub> will be followed by the movement of the wh-object to the specifier of Aux<sub>1</sub>P. Thus, the uninterpretable sub-feature uWh will be eliminated against the specifier of Aux<sub>1</sub>P and the interpretable sub-feature Force will be marked as Int due to the same Specifier-Head relation. Finally, note that the phonological features of the auxiliary will surface in the position of Aux<sub>1</sub>, which captures the desired word order of the auxiliary in relation to the subject of the clause.



The featural specifications of the auxiliary and the main verb in (7) are also the basis of the derivation of a subject wh-question like (4). The structure in (9) demonstrates the relevant derivation. Note that the operations involved in the projection of  $V_0P$ , the merger of  $Aux_0$  and the movement of the subject to the specifier of  $Aux_0P$  in (9) are identical to the corresponding steps of the derivation of the object wh-question in (8) above. However, the two derivations diverge when it comes to the elimination of the uninterpretable  $uWh$  sub-feature of C. Whereas  $uWh$  had to be scattered to  $Aux_1$  in (8), this is not the case in (9). Rather,  $uWh$  can be eliminated at the  $Aux_0P$  level, since its specifier is already occupied by the wh-word (of course, the same Specifier-Head relation will also mark the interpretable sub-feature Force with the appropriate Int value). Note, further, that the absence of  $Aux_1$  in the structure of a subject wh-question entails that the auxiliary does not undergo movement in this case, in contrast to an object wh-question. Thus, the above proposal suggests that there is a subject/object asymmetry with respect to the head movement of the auxiliary, similarly to traditional analyses of English wh-questions. However, according to this proposal there is a further asymmetry between subject and object wh-questions. Namely, a wh-subject surfaces in a lower structural position than a wh-object. In this regard, the analysis developed here follows in the tradition of the Vacuous Movement Hypothesis approach of George (1980), Chomsky (1986) and Ishii (2004).

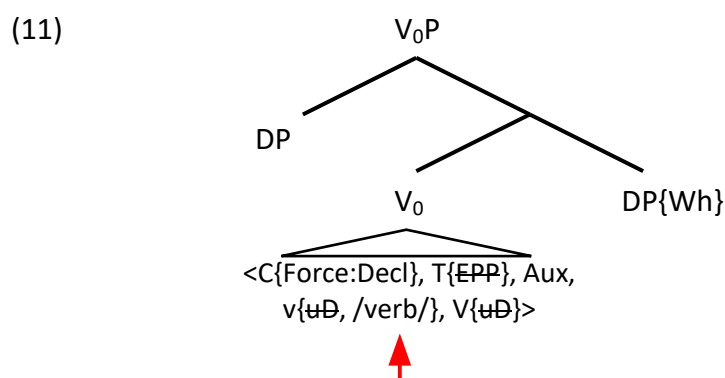


I will now move on to the analysis of the simple tense *wh*-questions in (3) and (5). Remember that the object *wh*-question in (3) involves *do*-support. I have previously discussed *do*-support in the case of English sentential negation in section 2.3.2 of chapter 2. In that section, I proposed that English main verbs are lexically incompatible with the uninterpretable sub-feature *uNeg*, which is responsible for the merger of the negative marker in a negative clause. Therefore, I argued that the numeration of a negative clause must include some auxiliary, which will be able to accommodate *uNeg*. In the case of a simple tense clause, the appropriate choice is the dummy auxiliary *do*. I will pursue a similar approach to account for *do*-support in *wh*-questions. Specifically, I propose that English main verbs are lexically incompatible with the uninterpretable *uWh* sub-feature of *C*. Thus, in a numeration that includes a main verb and a *wh*-word (but no auxiliary verb), Feature Bundling will assign the featural specification in (10) to the verb. Note that, in this case, the functional features *C*, *T* and *Aux* are assigned to the main verb, since the numeration does not include an auxiliary. Note, further, that the functional feature *C* does not have a *uWh* sub-feature (as a result of the hypothesis that this sub-feature is incompatible with a main verb), but it has an interpretable *Force* sub-feature.

- (10) Lexical verb (after Feature Bundling)  
 <*C*{*Force*}, *T*{*EPP*}, *Aux*, *v*{*uD*, /*verb*/}, *V*{*uD*}>

We can now examine the structures that can be derived from this featural specification. Consider the structure in (11), where the *wh*-word appears in object position. In this structure, the uninterpretable features of *V*, *v* and *T* can all be eliminated at the *V<sub>0</sub>P* level. Furthermore, since the functional feature *C* does not have any uninterpretable features, there is no reason for it to be scattered. Therefore, the interpretable *Force* sub-feature of *C* will be marked as *Decl*, given that the specifier of *V<sub>0</sub>P* is not a *wh*-word. Finally, observe that the *wh*-object remains in-situ throughout the derivation, due to the absence of a *uWh* sub-feature in the featural specification in (10). Thus, the structure in (11) is not a well-formed object *wh*-question (i.e. neither the value of *Force* nor the position of the *wh*-word

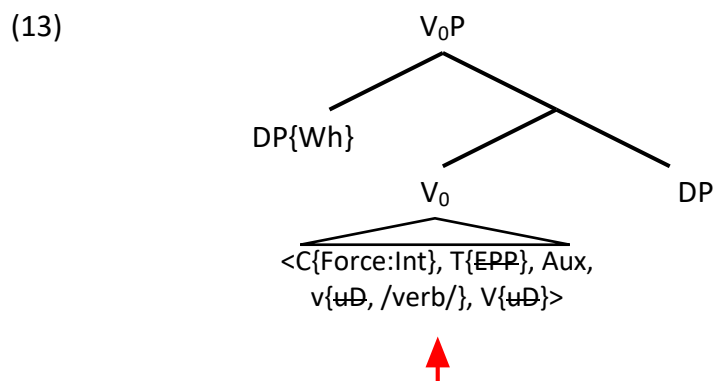
is consistent with an interrogative clause). Note, however, that I am not claiming that this is an illicit derivation. Rather, I assume that (11) is the well-formed structure of an echo question like (12). Nonetheless, the overall conclusion is the same as in the analysis of *do*-support in negative clauses. Namely, a numeration that does not contain any auxiliary verbs cannot derive an object wh-question. Thus, the only way to form a simple tense object wh-question is to include the dummy auxiliary *do* in the numeration. Finally, since I am assuming that *do* is a lexical item like any other auxiliary, I conclude that the structure of all object wh-questions will be as shown in (8) above, regardless of the auxiliary used.



(12) John plays what?

We also have to consider the structure that will be derived from the featural specification in (10) when the wh-word appears in subject position. The relevant structure is shown in (13) below. Note that the derivation in (13) proceeds in exactly the same manner as (11) above. Namely, the uninterpretable sub-features of V, v and T are eliminated at the V<sub>0</sub>P level; the functional feature C remains in V<sub>0</sub> since it does not have any uninterpretable sub-features; and, finally, Force establishes a relation with the specifier of V<sub>0</sub>P. However, in this case, the specifier of V<sub>0</sub>P is occupied by a wh-word. Consequently, Force will receive an Int value. Furthermore, in this structure, the wh-word appears in an initial position, by virtue of the fact that it is the subject of the clause. Therefore, the structure in (13) is, for all intents and purposes, a well-formed subject wh-question. This observation leads us to the following conclusions. Firstly, the fact that (13) successfully derives a wh-question

without resorting to the dummy auxiliary *do*<sup>2</sup>, while (11) fails to do so, accounts for the empirical observation that wh-questions display a subject/object asymmetry with respect to *do*-support (cf. examples (3) and (5) above). Secondly, the auxiliary-less subject wh-question in (13) is similar to the subject wh-question with an auxiliary in (9), and different from the object wh-question in (8), with respect to both head movement and the movement of the wh-word. Thus, the analysis of (13) is consistent with my earlier conclusions regarding the T-to-C asymmetry and the wh-movement asymmetry in subject versus object wh-questions. Finally, note that all of these asymmetries follow seamlessly from the overarching Feature Scattering framework of functional structure. In broad terms, the essence of the analysis is that a well-formed wh-question must involve a Specifier-Head relation between the functional feature C and the wh-word, but the amount of structural space needed to establish this relation varies from one structure to the other depending on the base position of the wh-word.



In the previous paragraphs, I extended the analysis of *do*-support in English negative clauses (cf. section 2.3.2) to *do*-support in object wh-questions. Note, however, that there is a subtle difference between the two analyses. In the case of sentential negation, I argued that the inability of Feature Bundling to assign the

<sup>2</sup> Note that I am not claiming that the numeration must exclude the dummy auxiliary *do* in order to derive a subject wh-question. Rather, I am merely stating that *do* is not necessary in the relevant numeration. Therefore, if the dummy auxiliary *do* is included in the numeration, the derivation will produce a structure like (9). However, (9) is a more complex structure than (13). Thus, I assume that the dummy auxiliary *do* is excluded due to economy considerations, unless a special emphatic interpretation of the wh-question is required.

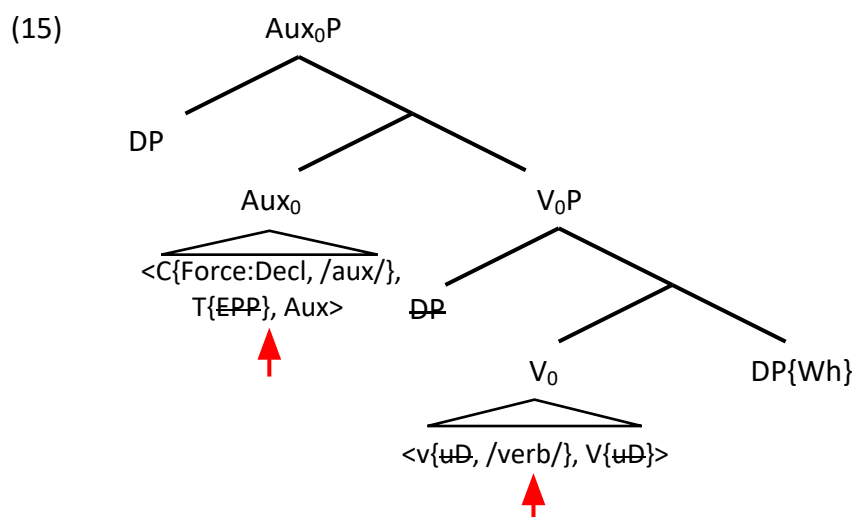
uninterpretable feature *uNeg* to a main verb is fatal for a numeration that does not include an auxiliary. On the contrary, in the case of *wh*-questions, I have tacitly assumed that the uninterpretable feature *uWh* is omitted without any repercussion in numerations that do not include an auxiliary (cf. in particular the featural specification of the main verb in the subject *wh*-question in (13) and the echo question in (11)). I will, however, argue that this conflict is only apparent. Let us assume that a derivation cannot converge until the numeration is exhausted (see Chomsky 1995). With this in mind, consider a numeration that includes a negative marker but lacks *uNeg*. In this scenario, the derivation will not be able to exhaust the numeration, since there is no appropriate uninterpretable feature to trigger the merger of the negative marker in the structure. Thus, this numeration is indeed problematic, as I have assumed in section 2.3.3. Compare this situation with a numeration that includes a *wh*-word but lacks *uWh*. In this case, the absence of *uWh* will not prevent the merger of the *wh*-word, since this uninterpretable feature is only responsible for *wh*-movement. Therefore, there is no reason to expect that this numeration cannot be exhausted. Accordingly, I conclude that the uninterpretable sub-feature *uWh* is truly optional, as I have assumed in the previous discussion.

However, this approach has a further implication. If the uninterpretable sub-feature *uWh* is indeed optional, then it could be omitted even in a numeration that includes an auxiliary. Therefore, the featural specification in (7b) (from which we derived an object and a subject *wh*-question including an auxiliary) should have a counterpart like (14b) below, which lacks a *uWh* sub-feature.

- (14) a. Lexical verb (after Feature Bundling)  
           <*v*{*uD*, /*verb*/}, *V*{*uD*}>  
       b. Lexical auxiliary verb (after Feature Bundling)  
           <*C*{*Force*, /*aux*/}, *T*{*EPP*}, *Aux*>

Let us briefly consider the derivations that may arise from these featural specifications. The structure in (15) represents a derivation where the *wh*-word appears in object position. As expected, the absence of *uWh* in the featural

specification of the auxiliary entails that the wh-object will remain in-situ. Furthermore, the absence of wh-movement entails that the functional feature C, and by extension the phonological features of the auxiliary, will remain in Aux<sub>0</sub>. Finally, considering that C does not establish a Specifier-Head relation with the wh-object, the interpretable sub-feature Force will be marked as Decl. Therefore, the structure in (15) is another example of an echo question (cf. (16) below).

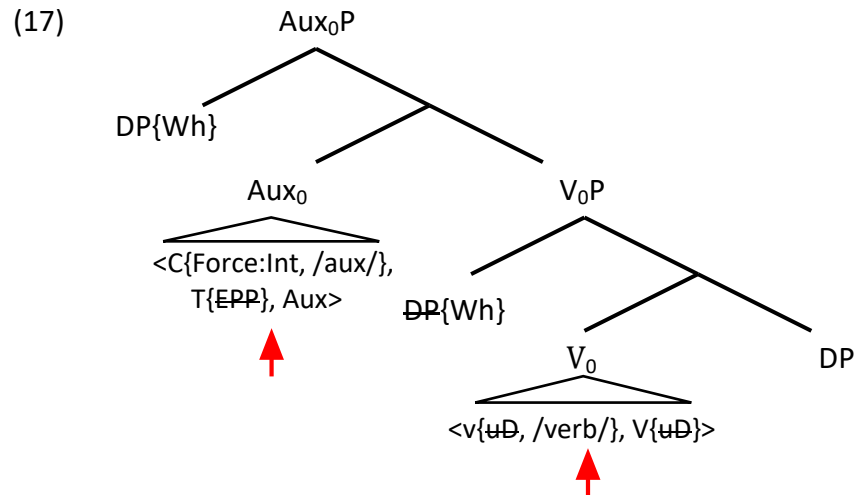


(16) John was playing what?

The structure in (15) contrasts with (17), where the wh-word is the subject of the clause. In this case, the absence of the uninterpretable  $uWh$  sub-feature from the featural specification of the auxiliary proves to be inconsequential. Observe that in this derivation the wh-word moves to the specifier of  $Aux_0P$  for independent reasons (i.e. because of the uninterpretable  $EPP$  sub-feature of  $T$ ). This movement has two consequences. Firstly, the interpretable  $Force$  sub-feature of  $C$  will be marked as  $Int$ . Secondly, the wh-word will surface at the initial position of the clause. These two observations suggest that (17) is a well-formed subject wh-question<sup>3</sup>. Finally, considering that both of the derivations in (15) and (17) produce acceptable structures, I conclude that the optionality of the  $uWh$

<sup>3</sup> Note, further, that this structure does not involve head movement of the auxiliary, as is expected in a subject wh-question.

sub-feature is not problematic for the analysis of English wh-questions developed in this section.



So far I have focused exclusively on the contrast between subject and object wh-questions. However, the situation observed in English wh-questions is best described as a subject/non-subject asymmetry. The following examples demonstrate that questions with an adverbial wh-word pattern with object wh-questions. Namely, adverb wh-questions display SAI (cf. examples (18) and (19)) and *do*-support (cf. examples (20) and (21)). In the remainder of this section I will consider whether the analysis proposed above can account for adverb wh-questions. I will show that this is indeed the case when it comes to questions like (18) and (20), where the wh-word is a VP adverb. However, the analysis of questions like (19) and (21), where the wh-word is a TP adverb, will prove to be problematic. Therefore, I will propose a small modification to the previous analysis.

- (18) a. How was John winning the game?  
 b. \*How John was winning the game?
- (19) a. When was John winning the game?  
 b. \*When John was winning the game?
- (20) a. How did John win the game?  
 b. \*How John won the game?
- (21) a. When did John fix the car?  
 b. \*When John fixed the car?

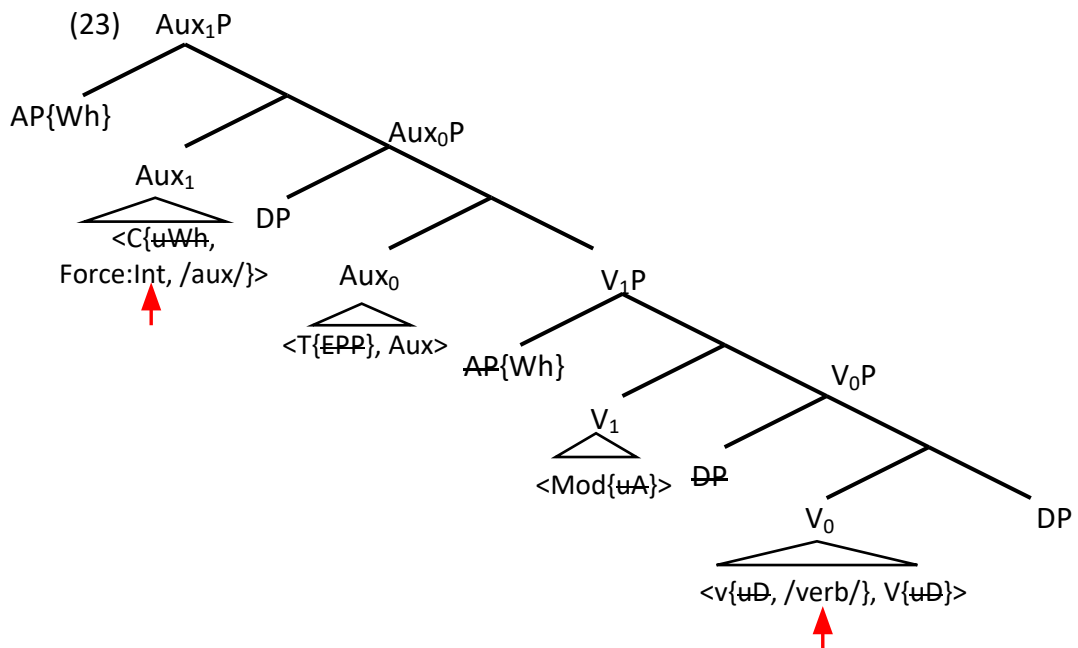


To begin with, I will consider the derivation of (18). That is, a VP adverb wh-question that contains an auxiliary. In this case, the featural specification of the main verb and the auxiliary will be as shown in (22). Note that the specification of the verb contains a functional feature Mod with an uninterpretable sub-feature uA, which is responsible for the merger of the VP adverb in the structure. These featural specifications will derive the structure in (23). This derivation will proceed as follows. As expected, the uninterpretable features of V and v will be eliminated at the V<sub>0</sub>P level. Afterwards, Mod and its uninterpretable sub-feature uA will be scattered to V<sub>1</sub> to enable the merger of the wh-adverb in the specifier of V<sub>1</sub>P. At this point, all of the uninterpretable features in the specification of the main verb are eliminated, so V<sub>1</sub>P will be merged with Aux<sub>0</sub>. Once Aux<sub>0</sub> is introduced into the structure, the subject will move to its specifier due to the uninterpretable EPP sub-feature of T. Considering that the specifier of Aux<sub>0</sub>P is not a wh-word, the uninterpretable uWh sub-feature of C will be scattered to Aux<sub>1</sub>. Of course, this Feature Scattering operation will be accompanied by the movement of the wh-adverb to the specifier of Aux<sub>1</sub>P. Furthermore, the interpretable sub-feature Force will receive an Int value as a result of the Specifier-Head relation between C and the wh-word. Finally, observe that the phonological features of the auxiliary will surface at the position of Aux<sub>1</sub>. All in all, the surface position of the wh-word and the auxiliary, as well as the value of Force, entail that (23) is an accurate representation of the adverb wh-question<sup>4</sup> in (18).

- (22) a. Lexical verb (after Feature Bundling)  
 <Mod{uA}, v{uD, /verb/}, V{uD}>  
 b. Lexical auxiliary verb (after Feature Bundling)  
 <C{uWh, Force, /aux/}, T{EPP}, Aux>

---

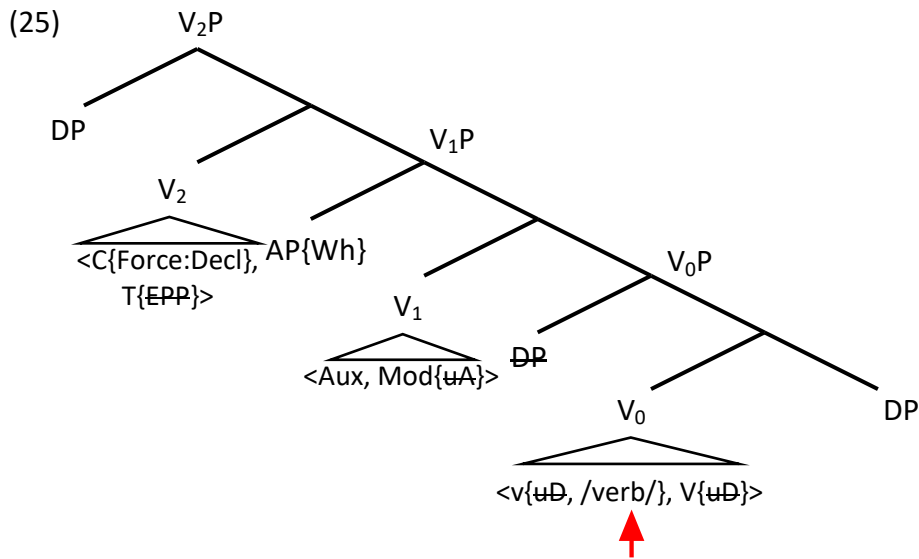
<sup>4</sup> Note that the featural specification of the auxiliary in (22b) has an alternative form that lacks the uninterpretable feature uWh. In the event that uWh is missing, the derivation will differ from (23) in the following ways: (i) there will be no need to move the wh-word from the specifier of V<sub>1</sub>P, (ii) consequently, there will be no need to scatter the functional head Aux<sub>1</sub>P, which entails that the phonological features of the auxiliary will surface in-situ, and (iii) the interpretable sub-feature Force will remain in Aux<sub>0</sub>, where it will receive a Decl value due to the absence of a wh-word in its specifier. Therefore, this alternative derivation will not produce a well-formed wh-question.



Next, I will consider a simple tense question where the wh-word is a VP adverb. As we have seen in example (20), this structure requires *do*-support. Therefore, according to the previous analysis, it has to be the case that a numeration which does not include any auxiliaries fails to derive a well-formed wh-question of this kind. Let us consider if this prediction pans out. In the relevant numeration, the main verb will have the featural specification shown in (24). Further, this specification will produce the structure in (25). Once again, the derivation will begin with the uninterpretable features of V and v, which will be eliminated at the V<sub>0</sub>P level. Then, Mod will be scattered to V<sub>1</sub>, where its uninterpretable sub-feature uA will be eliminated against the wh-adverb in the specifier of V<sub>1</sub>P. The next uninterpretable feature in the specification of the verb is the EPP sub-feature of T, which has to establish a Specifier-Head relation with the subject. To this end, T and its sub-feature will be scattered to V<sub>2</sub> and the subject will move to the specifier of V<sub>2</sub>P. At this point, note that the functional feature C in (24) does not have an uninterpretable uWh sub-feature. Therefore, the wh-adverb will remain in-situ in the specifier of V<sub>1</sub>P and C will remain unscattered in V<sub>2</sub>. Finally, note that the interpretable sub-feature Force will receive a Decl value since its specifier is not a wh-word. Therefore, judging by the value of Force and the surface

position of the *wh*-adverb, (25) is not a well-formed *wh*-question. This is, of course, a desirable outcome which entails that the simple tense *wh*-question under consideration requires the inclusion of *do* in its numeration. Thus, the structure of a VP adverb *wh*-question displaying *do*-support (as in (20a)) will be as shown in (23) above.

(24) Lexical verb (after Feature Bundling)  
 <C{Force}, T{EPP}, Aux, Mod{uA}, v{uD}, /verb/, V{uD}>

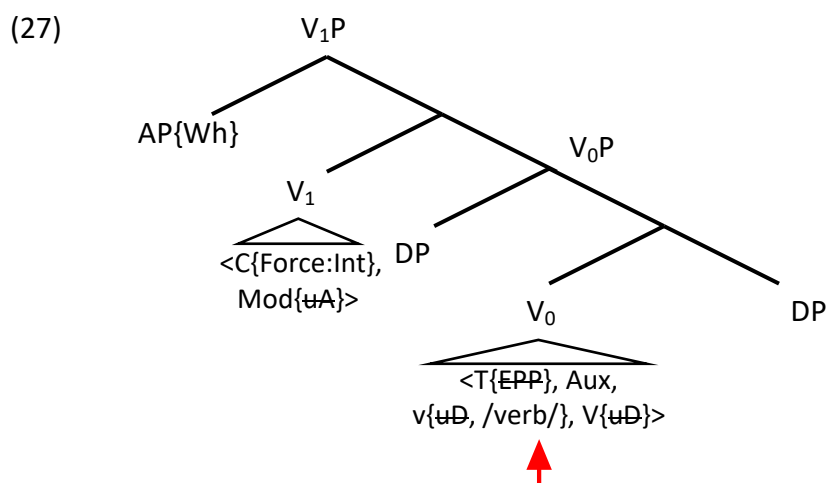


In the previous paragraphs, I demonstrated that the analysis of object *wh*-questions proposed earlier can be extended straightforwardly to the case of adverb *wh*-questions, as long as the *wh*-word is a VP adverb. However, we also have to consider the case of *wh*-words that correspond to a TP adverb. The analysis of adverb *wh*-questions of this kind will prove to be less straightforward. As we will see below, the featural specifications assumed so far fail to account for the obligatory use of *do* in examples like (21) (i.e. a simple tense question where the *wh*-word is a TP adverb). Nonetheless, a careful examination of the relevant derivation will allow me to identify the source of the problem and to propose an appropriate modification of the featural specifications of main and auxiliary verbs that resolves the issue.

In the analysis of object as well as VP adverb *wh*-questions above, I have attributed *do*-support to the inability of an auxiliary-less numeration to derive a

well-formed wh-question. With that in mind, let us consider the derivation of a TP adverb wh-question. In a numeration that lacks an auxiliary, the featural specification of the main verb will be as shown in (26). Note that this specification includes a functional feature Mod with an uninterpretable sub-feature uA, which will account for the merger of the adverb above the maximal projection of T. This featural specification will give rise to the structure in (27). The relevant derivation will proceed as follows. To begin with, the uninterpretable features of V, v and T will be eliminated at the V<sub>0</sub>P level. The next uninterpretable feature in the specification of the verb is the uA sub-feature of Mod. Given that the specifier of V<sub>0</sub> is filled by the subject, Mod and its sub-feature will be scattered to V<sub>1</sub> to allow for the merger of the adverb in the specifier of V<sub>1</sub>P. Observe that, even though the featural specification of the verb lacks an uninterpretable uWh sub-feature, the functional feature C can establish a Specifier-Head relation with the wh-adverb in the V<sub>1</sub>P level. Consequently, the interpretable Force sub-feature of C will receive an Int value. Therefore, the derivation in (27) produces a well-formed wh-question. By extension, it appears that a simple tense question where the wh-word is a TP adverb does not require the inclusion of *do* in its numeration. Thus, we arrive at the erroneous prediction that TP adverb wh-questions pattern with subject wh-questions.

- (26) Lexical verb (after Feature Bundling)  
 <C{Force}, Mod {uA}, T{EPP}, Aux, v{uD, /verb/}, V{uD}>



I suggest that this issue could be resolved if the order of the functional features C and Mod in the featural specification of the verb was reversed. If that were the case, C and its sub-features would not be affected by the scattering of Mod. Therefore, Force would remain in  $V_0$  and receive a Decl value (due to the absence of a wh-word in the specifier of  $V_0P$ ) that is inconsistent with a well-formed wh-question. However, this tentative solution runs counter to the general assumption that sentential adverbs appear between T and C cross-linguistically. Furthermore, it would not be tenable to attribute this abnormal order of C and Mod to some idiosyncratic lexical property of English main verbs, since I have been assuming that the sequence of functional features is universal. There is, however, a more nuanced way of formulating this solution that does not rely on positing a controversial order of C and Mod. Note that what makes this suggestion work, at a purely mechanical level at least, is not the identity of the functional features but the order of their sub-features. Therefore, it would suffice to assume that the set of sub-features that I have hitherto assigned to C (i.e. uWh, Force and the phonological features of auxiliary verbs) are in fact sub-features of some functional feature F that appears between Mod and  $T^5$ . In principle, F could either correspond to one of the lower positions of the left periphery (e.g. Rizzi's 1997 functional head Fin) or, alternatively, it could correspond to one of the higher positions of the middle field (e.g. Pollock's 1989 functional head Agr). However, I will not make any specific claims concerning the identity of F in the following discussion.

Note that this hypothesis necessitates a revision of all the featural specifications assumed so far. For example, the specification in (10) (repeated in (28) below), which represents the main verb in a numeration that lacks an auxiliary, should be revised as in (29). However, observe that the order of all the relevant sub-features is identical between (28) and (29). Therefore, this modification does not alter the previous analysis in any significant respect. The only difference is that

---

<sup>5</sup> Note that this suggestion echoes Rizzi's (1996) analysis of English wh-questions. Rizzi (1996) proposes that interrogative clauses are marked by a +Wh feature, which resides in some functional head between C and T. However, the two analyses differ with respect to the landing site of the wh-word, since Rizzi (1996) assumes that all wh-words move to the specifier of C.

the functional feature C should be replaced with the functional feature F in all of the relevant derivations (i.e. the successful subject wh-question in (13) and the unsuccessful object wh-question in (11), which results in an echo question instead).

(28) Lexical verb (after Feature Bundling) – repeated from (10)  
 <C{Force}, T{EPP}, Aux, v{uD, /verb/}, V{uD}>

(29) Lexical verb (after Feature Bundling) - revised  
 <C, F{Force}, T{EPP}, Aux, v{uD, /verb/}, V{uD}>

As a further example, consider the featural specification of the auxiliary verb in (7b) (repeated in (30b)), which should be revised as in (31). Once again, this revision does not alter the order of the relevant sub-features, so (31) captures the same structures as (7) (i.e. the object and subject wh-questions including an auxiliary in (8) and (9) respectively).

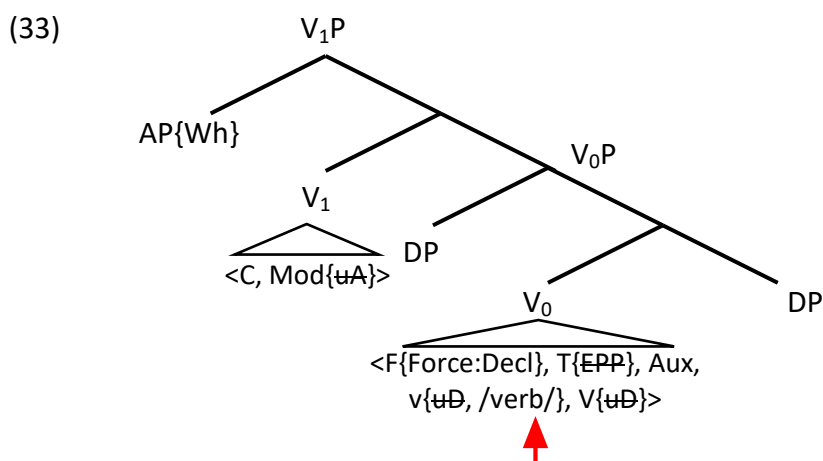
(30) a. Lexical verb (after Feature Bundling) – repeated from (7a)  
 <v{uD, /verb/}, V{uD}>  
 b. Lexical auxiliary verb (after Feature Bundling) – repeated from (7b)  
 <C{uWh, Force, /aux/}, T{EPP}, Aux>

(31) Lexical auxiliary verb (after Feature Bundling) - revised  
 <C, F{uWh, Force, /aux/}, T{EPP}, Aux>

Without any further comment, the exact same reasoning applies to the specifications employed in the analysis of questions where the wh-word is a VP adverb (i.e. (22) and (23)). However, the situation is different when the wh-word corresponds to a TP adverb. Compare the featural specification in (26), which represents the main verb in a numeration that lacks an auxiliary, with the revised version shown in (32). As we can see, the order of the sub-features Force and uA is reversed in the revised specification in (32). Consider, further, the effect of this revision on the corresponding structure. The relevant derivation will proceed as shown in (33). As expected, the uninterpretable features of V, v and T are eliminated at the V<sub>0</sub>P level. The next uninterpretable feature in the specification of the verb is the uA sub-feature of Mod, which will be scattered to V<sub>1</sub>. However, the scattering of Mod leaves the subordinate functional feature F stranded in V<sub>0</sub>.

Therefore, the interpretable Force sub-feature of F will receive a Decl value, given that the specifier of  $V_0P$  is not a wh-word. With this in mind, I suggest that the structure in (33) cannot be interpreted as a wh-question, in spite of the fact that the wh-word appears in a clause-initial position. Thus, we finally arrive at the desirable conclusion that a well-formed question where the wh-word is a TP adverb requires the insertion of the dummy auxiliary *do* in the relevant numeration.

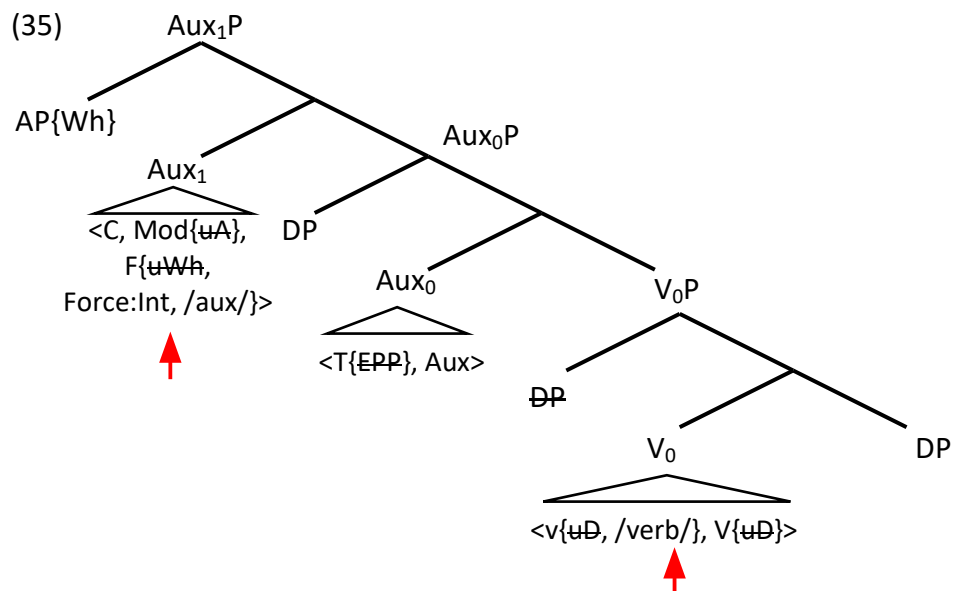
(32) Lexical verb (after Feature Bundling) - revised  
 <C, Mod{uA}, F{Force}, T{EPP}, Aux, v{uD, /verb/}, V{uD}>



Finally, let us consider whether the inclusion of *do* (or any other auxiliary for that matter) in the numeration will result in a successful derivation of a TP adverb wh-question. The featural specifications of the main and the auxiliary verb in the relevant numeration are shown in (34) and the resulting structure is shown in (35). As per usual, the derivation will start with the projection of  $V_0P$ . Once the uninterpretable sub-features in the specification of the verb are eliminated,  $V_0P$  will be merged with  $Aux_0$ . Furthermore, the subject will move to the specifier of  $Aux_0P$  due to the uninterpretable EPP sub-feature of T. At this point, the derivation has to eliminate the uninterpretable  $uWh$  sub-feature of F. Considering that the specifier of  $Aux_0$  is not a wh-word, F will be scattered to  $Aux_1$  to allow for the merger of the wh-adverb in the specifier of  $Aux_1P$ . It is worth noting that this situation is unusual in that  $uWh$  is eliminated against a directly merged (rather than moved) wh-word. However, there is nothing in the overarching framework to preclude this option.

Returning to the derivation, the Specifier-Head relation between F and the wh-adverb entails that the interpretable sub-feature Force will receive an Int value. Finally, the uninterpretable uA sub-feature of Mod will be eliminated against the wh-adverb that has already been merged in the specifier of Aux<sub>1</sub>P. To sum up, this derivation results in a structure where the wh-word appears in a clause-initial position, the auxiliary displays SAI and the interpretable feature Force has received an interrogative value. Thus, the derivation arising from a numeration that contains a main and an auxiliary verb with the featural specifications shown in (34) provides an accurate representation of a TP adverb wh-question<sup>6</sup>.

- (34) a. Lexical verb (after Feature Bundling)  
 <v{uD, /verb/}, V{uD}>  
 b. Lexical auxiliary verb (after Feature Bundling)  
 <C, Mod{uA}, F{uWh, Force, /aux/}, T{EPP}, Aux>



In conclusion, in this section I have developed an account of the subject/non-subject asymmetries observed in English wh-questions. The analysis is based on the overarching hypothesis that a well-formed wh-question must involve a Specifier-

<sup>6</sup> As before, the featural specification of the auxiliary in (34b) has an alternative form that lacks the uninterpretable sub-feature uWh. In the absence of uWh, the functional feature F will not be scattered to Aux<sub>1</sub> and, consequently, the interpretable sub-feature Force will not be marked as interrogative. Therefore, this alternative featural specification will not produce a well-formed wh-question.



Head relation between a wh-word and the functional feature F, which is the locus of an interpretable feature expressing the Force of the clause. Further, I have proposed that Force is paired with an uninterpretable feature uWh, which facilitates the relevant Specifier-Head relation. However, I have suggested that uWh is not compatible with the featural specification of a main verb. Therefore, simple present and past tense wh-questions require the inclusion of the dummy auxiliary *do* in the numeration, unless the necessary Specifier-Head relation between the wh-word and F can be established independently of the uWh feature. Finally, I have demonstrated that wh-subjects are uniquely capable of establishing the relevant Specifier-Head relation, even in a structure that lacks an auxiliary, due to the proximity of F and T (i.e. the functional feature that determines the surface position of the subject). Thus, the unique properties of subject wh-questions are attributed to a structural subject/non-subject asymmetry. In the following section, I will consider whether this analysis can be extended to other subject/non-subject asymmetries relating to the distribution of the overt complementizer *that* in English embedded clauses.

### ***3.3 Subject/Non-Subject Asymmetries in the Distribution of Overt Complementizers***

#### ***3.3.1 The That-Trace Effect***

In the previous section, I focused on the subject/non-subject asymmetries (in relation to SAI, *do*-support and, arguably, wh-movement) that are characteristic of English wh-questions. Another well-known subject/non-subject asymmetry in English is the *that*-trace effect. This asymmetry is observed in structures that involve the extraction of a wh-word from an embedded declarative clause to a matrix interrogative clause. The relevant generalisation is that a wh-subject cannot be extracted over the overt complementizer (cf. the ungrammaticality of (36b)), unlike non-subject wh-words (cf. the grammaticality of (37b), (38b) and (39b)). Since the 1980s, much of the relevant literature has pursued the hypothesis that the subject/non-subject asymmetries in wh-extraction and wh-questions have some

common underlying cause (this approach was originally proposed by Koopman (1983) and was later adopted by Rizzi 1996, Grimshaw 1997, Pesetsky and Torrego 2001, 2004, Roussou 2002, Ishii 2004, and many others). In this section, I will follow this tradition and explore whether the framework developed in this dissertation can provide a uniform analysis of subject/non-subject asymmetries.

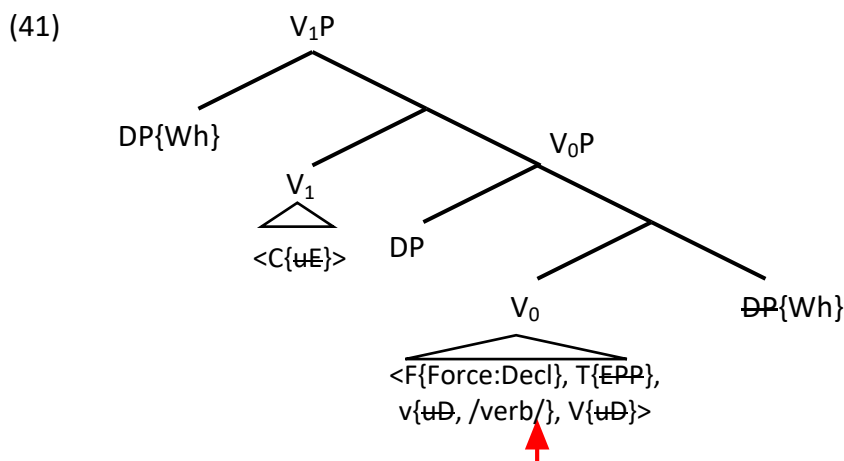
- (36) a. Who do you think [ \_ likes john? ]  
b. \*Who do you think [ that \_ likes john? ]
- (37) a. Who do you think [ John likes \_ ? ]  
b. Who do you think [ that John likes \_ ? ]
- (38) a. How do you think [ John \_ fixed the car? ]  
b. How do you think [ that John \_ fixed the car? ]
- (39) a. When did you say [ \_ John fixed the car? ]  
b. When did you say [ that \_ John fixed the car? ]

I will begin this discussion by addressing the various cases of wh-extraction from embedded clauses that do not contain an overt complementizer (cf. the (a) version of the examples above). The usual assumption is that wh-extraction is a two-step process, where the moved wh-word has to pass through the edge of the embedded clause before it can reach its surface position in the matrix clause. Of course, due to the variable nature of functional structure within this analysis, there are certain cases where the wh-word will appear at the edge of the embedded clause independently of wh-movement. For instance, the base position of the wh-subject in (36a) and the TP adverb wh-word in (39a) correspond to the edge of the (truncated) embedded clause. Therefore, in these cases, wh-extraction can take place in a single step. On the contrary, the wh-object in (37a) and the VP adverb wh-word in (38a) do not appear at the edge of the embedded clause by default. Therefore, they have to move from their base position to the specifier of some derived functional head above the (surface) position of the embedded subject. However, remember that I have proposed that main verbs are lexically incompatible with the uninterpretable sub-feature *uWh*, which is responsible for the movement of a wh-word in a matrix question. Therefore, it has to be the case

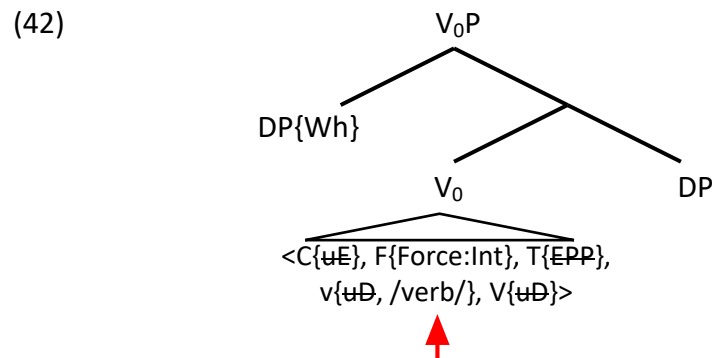
that the featural specification of the embedded verb contains some other uninterpretable feature that causes the movement of the wh-word. I will assume that the relevant feature is an ‘edge feature’ (cf. Chomsky 2005) that is distinct from the uWh of a wh-question. I will annotate this feature as uE. Furthermore, I will assume that uE is hosted by the functional feature C (rather than F), given that it has to mark the edge of the embedded clause. Therefore, the featural specification of the embedded verb, in a numeration that does not contain an overt complementizer, will be as in (40).

- (40) Lexical verb (after Feature Bundling)  
 <C{uE}, F{Force}, T{EPP}, v{uD, /verb/}, V{uD}>

Let us consider some of the structures that will arise from this featural specification. The structure in (41) represents the derivation of (37a) (i.e. the case of extraction of a wh-object from an embedded clause without an overt complementizer). In brief, this derivation will proceed as follows. The uninterpretable sub-features of V, v and T are all eliminated at the V<sub>0</sub>P level. Observe that the functional feature F does not have any uninterpretable features, so it will necessarily remain unscattered in V<sub>0</sub>. Observe, further, that the interpretable Force sub-feature of F will receive a Decl value, since the specifier of V<sub>0</sub>P is not a wh-word. The remaining uninterpretable feature in the specification of the verb is the uE sub-feature of C. Considering that uE cannot establish a Specifier-Head relation with the wh-object at the V<sub>0</sub>P level, C will be scattered to V<sub>1</sub>. Finally, the wh-word will move to the specifier of V<sub>1</sub>P, which concludes the derivation of the embedded clause.



Let us also consider the derivation of (36a) (i.e. the case of extraction of a wh-subject from an embedded clause lacking an overt complementizer). The resulting structure is shown in (42) below. In this case, all of the uninterpretable features of the verb, including the uE sub-feature of C, will be eliminated at the V<sub>0</sub>P level. What is of interest in this derivation is that the interpretable Force sub-feature of F will be marked as Int, due to the presence of the wh-subject in the specifier of V<sub>0</sub>P. Obviously, the Int value of Force is at odds with the declarative interpretation of the embedded clause in (36a). To resolve this quandary, I will make the following proposal: while uninterpretable features are eliminated once and for all when they enter a syntactic relation, I suggest that the value of interpretable features is constantly re-evaluated at every step of the derivation. Therefore, the Int value of Force in (42) is, in effect, temporary. At the end of the derivation, the wh-subject will move from the specifier of the embedded V<sub>0</sub>P to an appropriate position in the matrix clause. At that point, the interpretable sub-feature Force of the embedded verb will revert to a Decl value, which is consistent with the declarative interpretation of the embedded clause.



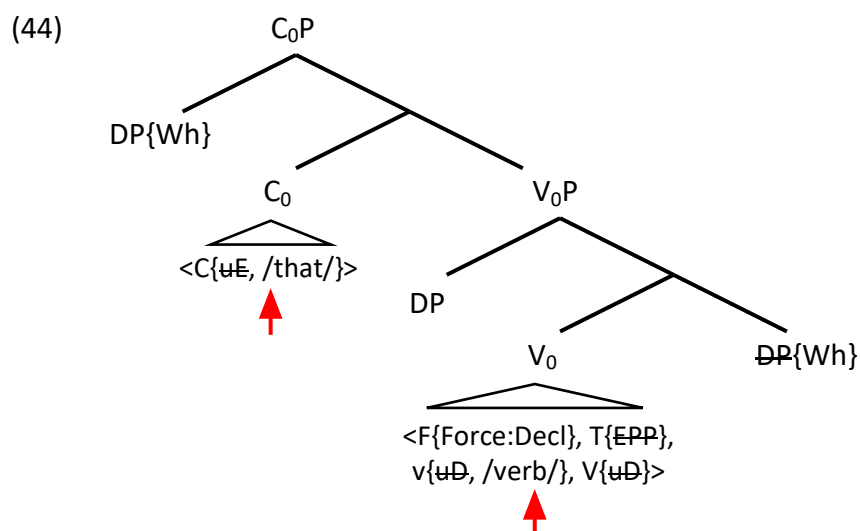
I will now turn my attention to wh-extraction over the overt complementizer *that* (cf. the (b) version in examples (36)-(39) above). To begin with, we need to consider the featural specification of the overt complementizer. I will simply assume that the specification of the complementizer consists of the single functional feature C, which also hosts its phonological features. Of course, when the embedded clause involves wh-extraction, the complementizer will be assigned the uninterpretable sub-feature uE as well. Therefore, in a numeration that contains

both an embedded verb and an overt complementizer, the featural specification of the two lexical items will be as in (43).

- (43) a. Lexical verb (after Feature Bundling)  
<F{Force}, T{EPP}, v{uD, /verb/}, V{uD}>  
b. Lexical complementizer (after Feature Bundling)  
<C{uE, /that/}>

As I have mentioned above, these featural specifications will give rise to the (b) versions of the examples in (36)-(39). Note that this includes the ungrammatical (36b) (i.e. the extraction of a *wh*-subject over the overt complementizer). In the following discussion, I will raise two questions. Firstly, I will consider whether the base position of a subject versus a non-subject *wh*-word is responsible for any differences between the derivation of a structure involving subject extraction and a structure involving extraction of a non-subject. Secondly, I will consider whether any such difference may be relevant to the ungrammaticality of (36b). To explore these points, I will focus specifically on the contrast between the extraction of a *wh*-subject versus a *wh*-object.

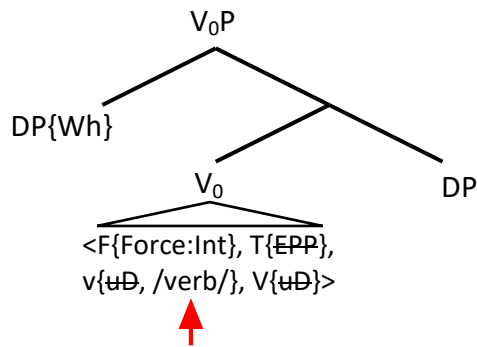
I will address the grammatical case of object extraction in (37b) first. The relevant derivation is shown in (44). The projection of the  $V_0P$  is rather predictable. The uninterpretable features of V, v and T are all eliminated in the  $V_0P$  level. One point of interest is that the interpretable Force sub-feature of F receives a Decl value, given that the specifier of  $V_0P$  is not a *wh*-word. Once the uninterpretable features in the specification of the verb are eliminated,  $C_0$  is introduced in the structure. Finally, the *wh*-object will move to the specifier of  $C_0P$  to eliminate the uninterpretable uE sub-feature of C, thus concluding the derivation of the embedded clause.



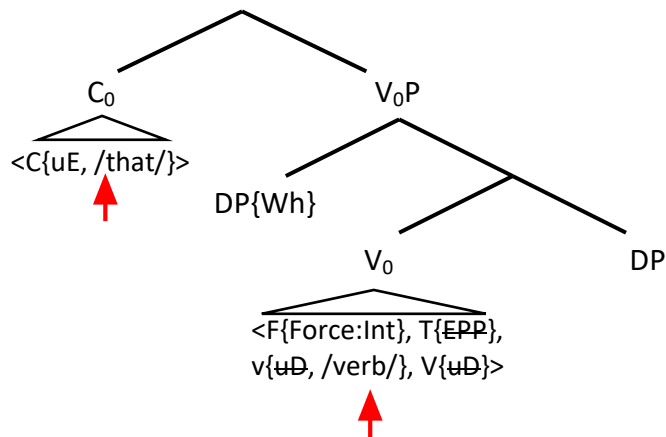
Let us now consider the (ungrammatical) case of subject extraction in (36b). The relevant derivation is shown in (45) below. Note that in this case I demonstrate some of the crucial steps of the derivation individually. (45a) represents the state of the derivation at the point of completion of the  $V_0P$ . At this stage, all of the uninterpretable sub-features of the verb have been eliminated and the interpretable sub-feature Force bears the value Int, due to the presence of the wh-subject in the specifier of  $V_0P$ . If the derivation proceeds unimpeded, we expect  $C_0$  to be merged with  $V_0P$ , as shown in (45b). Then, the wh-subject will have to move to the specifier of  $C_0P$  to eliminate the  $uE$  sub-feature of  $C$ . Finally, as the wh-subject evacuates the specifier of  $V_0P$ , the value of Force will revert to Decl. Thus, the final structure of the embedded clause will be as in (45c). Observe that the base position of the wh-word in (44) and (45) (i.e. a wh-object and a wh-subject respectively) does not have an effect on the end result of the two derivations, but it does have a temporary effect on the featural specification of the verb between the completion of the  $V_0P$  and the completion of  $C_0P$ . Therefore, the ungrammaticality of subject extraction over the overt complementizer has to be linked to the intermediate steps (45a-b) in the derivation of the embedded clause. In broad terms, I will suggest that the overt complementizer *that* cannot head an interrogative embedded clause. However, due to the transient nature of the value of Force, I will formulate this hypothesis in much narrower terms. Specifically, I

propose that the overt complementizer is prohibited from taking a complement headed by a lexical or functional head containing a Force:Int sub-feature. Therefore, the derivation of a structure involving subject extraction will not be able to proceed beyond the step shown in (45a) and, consequently, it will crash<sup>7</sup>.

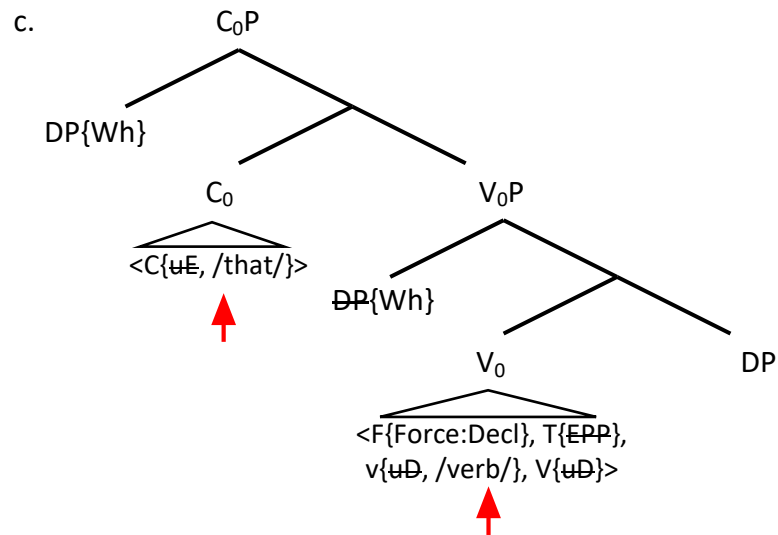
(45) a.



b.



<sup>7</sup> Note that the extraction of a wh-adverb will pattern with the extraction of a wh-object. This is because the functional feature F does not contain any uninterpretable sub-features. Thus, F will necessarily surface on the same functional head as T. Furthermore, due to the EPP sub-feature of T, the specifier of the relevant functional head will necessarily be occupied by the subject. Consequently, the (temporary) Int value of Force will only come into play when the subject is a wh-word.



In conclusion, in this section I have argued that subject/non-subject asymmetries in English *wh*-questions and embedded clauses involving *wh*-extraction overlap in a significant way. I have proposed that both phenomena boil down to the fact that a *wh*-subject appears in a privileged structural position that allows it to assign an interrogative value to the interpretable sub-feature *Force*. Broadly speaking, this situation entails that a *wh*-subject allows for a smaller functional structure. This translates to absence of *wh*-movement, auxiliary movement and *do*-support in the case of *wh*-questions and absence of the overt complementizer in the case of *wh*-extraction. However, while auxiliaries are optional in subject *wh*-questions, the complementizer is entirely unacceptable in structures involving extraction of a *wh*-subject. To account for this difference, I relied on the observation that an interrogative value of *Force* is desired in *wh*-questions but problematic in embedded clauses with *wh*-extraction. Specifically, I proposed that the overt complementizer cannot be merged with the projection of the embedded verb if its head bears a *Force:Int* sub-feature. In the following section, I will extend this analysis to two additional structures that are, in different ways, related to the *that*-trace effect (namely, the adverb amelioration effect and the anti-*that*-trace effect of English relative clauses).



### 3.3.2 Amelioration and Anti-That-Trace Effects

In the previous discussion, I focused on the analysis of prototypical examples of the *that*-trace effect. However, the distribution of the overt complementizer in the context of *wh*-extraction is a larger issue. In this section, I will provide a tentative analysis of two additional phenomena that relate to this issue. Firstly, I will address the adverb amelioration effect, which is directly relevant to the analysis of the *that*-trace effect. Secondly, I will discuss the distribution of the overt complementizer in relative clauses, which provides an example of an anti-*that*-trace effect.

I will begin this section with a discussion of the adverb amelioration effect. Contrary to what we have seen so far, the extraction of a *wh*-subject over an overt complementizer appears to be acceptable if the embedded clause also contains a TP adverbial phrase (cf. Bresnan 1997, Culicover 1993, Browning 1996, Grimshaw 1997, Rizzi 1997, Pesetsky and Torrego 2001, Ishii 2004, amongst others). The ‘ameliorating’ effect of the adverbial phrase can be seen in example (46), which would otherwise be a prime example of the *that*-trace effect.

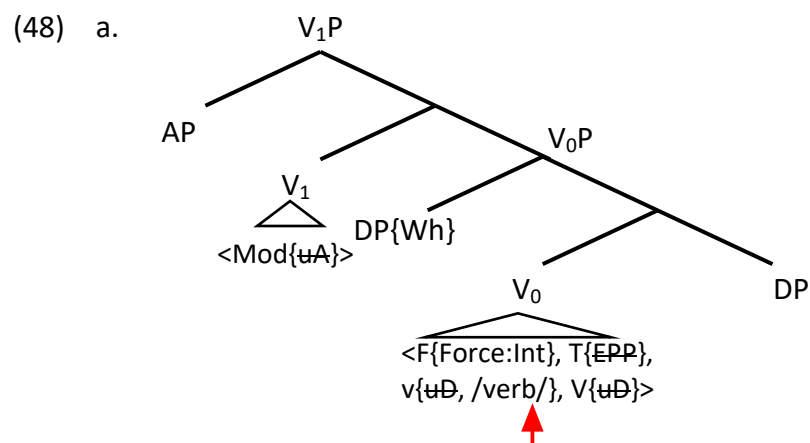
- (46) Sue met the man who Mary is claiming that [for all intents and purposes] \_\_ was the mayor of the city.

Pesetsky and Torrego (2001)

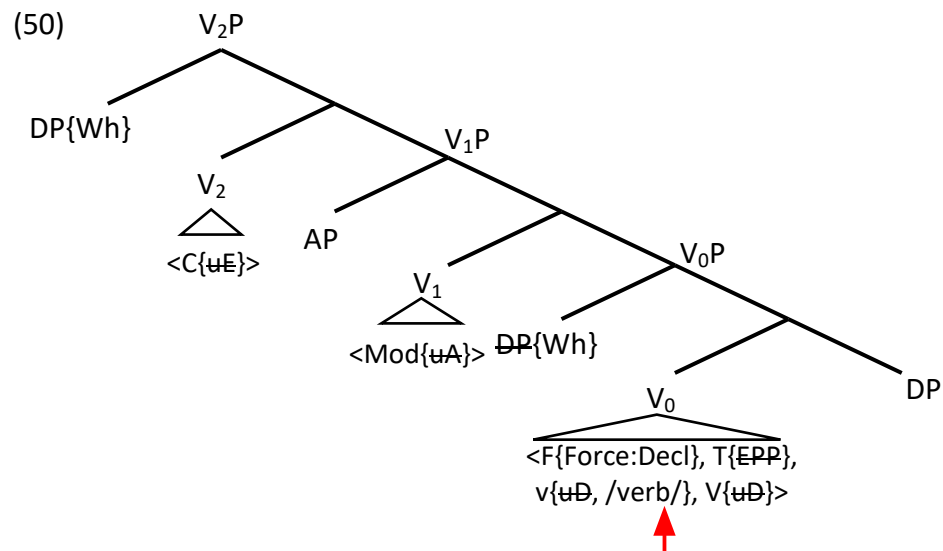
I will argue that the grammaticality of (46) follows straightforwardly from the analysis of the *that*-trace effect developed in the previous section. Consider the following. All things being equal, the embedded clause in this example ought to be the product of the featural specifications in (47). Note that the specification of the embedded verb in (47a) contains a  $\text{Mod}\{uA\}$  feature situated above *F*, which will be responsible for the merger of the adverbial phrase. The derivation arising from these featural specifications is represented incrementally in (48a-b). (48a) demonstrates the intermediate structure that will be projected from the featural specification of the embedded verb. Similarly to (45a) above, we can see that the uninterpretable sub-features of *V*, *v* and *T* are eliminated at the  $V_0P$  level. Furthermore, at this stage the interpretable Force sub-feature of *F* will be marked

as Int, since the specifier of  $V_0P$  is a wh-word. However, in this case, the embedded verb has one more uninterpretable feature, i.e. the uA sub-feature of Mod. Of course, this sub-feature cannot be eliminated at the  $V_0P$  level, so it will be scattered to  $V_1$ . Finally, the adverbial phrase will be merged in the specifier of  $V_1P$ . The question now is whether the overt complementizer can take the projection of the embedded verb as its complement. Observe that the scattering of  $V_1$  and the merger of the adverbial phrase do not alter the interrogative value of Force, considering that the position of the wh-subject is unaffected by these operations. However, the scattering of  $V_1$  entails that the head of the prospective complement of the overt complementizer does not bear the offending Force:Int sub-feature. Therefore, assuming that the merger of  $C_0$  with  $V_1P$  is a strictly local operation (i.e. an operation that only evaluates the feature content of the closest head to  $C_0$ , namely  $V_1$ ), the derivation will successfully proceed from (48a) to (48b). Subsequently, the wh-subject will move to the specifier of  $C_0P$  in order to eliminate the uninterpretable uE sub-feature of C. Note, finally, that at this point the interpretable Force sub-feature of F will revert to a Decl value. Thus, at the end of the derivation the structure of the embedded clause will be as shown in (48b).

- (47) a. Lexical verb (after Feature Bundling)  
 <Mod{uA}, F{Force}, T{EPP}, v{uD}, /verb/}, V{uD}>  
 b. Lexical complementizer (after Feature Bundling)  
 <C{uE}, /that/}>







Therefore, this analysis predicts that the adverb effect is a true amelioration effect, where the presence of an adverbial phrase repairs the ungrammaticality of subject extraction over an overt complementizer but does not affect the grammaticality of subject extraction from an embedded clause that is not introduced by an overt complementizer. Note that this prediction is not particularly common in the relevant literature. Rather, most analyses of the *that*-trace effect (Grimshaw (1997), Pesetsky and Torrego (2001), Ishii (2003), amongst others) arrive at the prediction that the presence of an adverbial phrase reverses the grammaticality pattern of subject extraction; i.e. the overt complementizer becomes not only acceptable, but necessary in the context of subject extraction. Unfortunately, the relevant grammaticality judgements are far from conclusive. For instance, Pesetsky and Torrego (2001: p40; note 37) acknowledge that examples like (51) are not strongly unacceptable. Pesetsky and Torrego (2001) attribute examples of this kind to an alternative, parenthetical parse. By comparison, the analysis developed here suggests that the limited acceptability of (51) should be attributed to independent factors. Tentatively, this situation could be linked to the fact that an embedded TP adverbial phrase renders the overt complementizer nearly unacceptable even in the absence of *wh*-extraction, as demonstrated by the contrast between examples (52a) and (52b). However, at this point, I do not have an account for the pattern in (52).

- (51) ?Sue met the man who Mary is claiming [for all intents and purposes] \_\_\_ was the mayor of the city.
- (52) a. Mary is claiming that [for all intents and purposes] John is the mayor of the city.  
 b. ??Mary is claiming [for all intents and purposes] John is the mayor of the city.

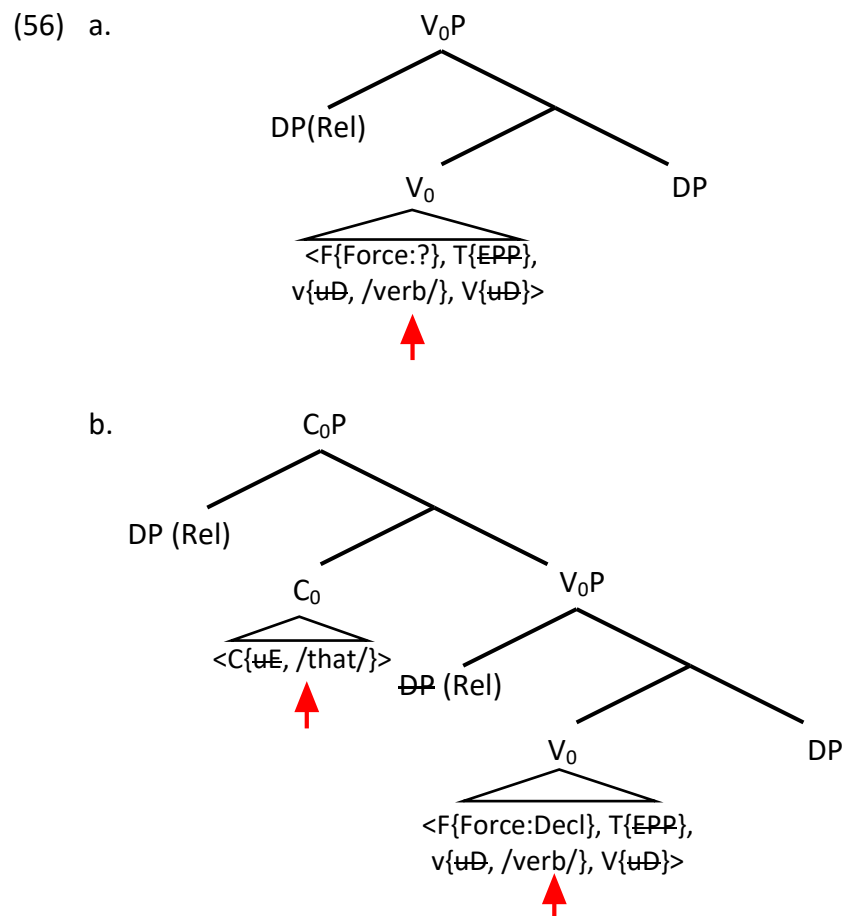
The second issue I will address in this section is the distribution of the overt complementizer in English (non-wh) relative clauses. The following examples demonstrate that relative clauses display an anti-*that*-trace effect. The contrast between (53a) and (53b) shows that the overt complementizer *that* is necessary in structures where the relativized nominal corresponds to the subject of the relative clause. Furthermore, the examples in (54) demonstrate that the overt complementizer is optional in structures where the relativized nominal is the object of the relative clause. Thus, relative clauses are similar to wh-extraction in so far as they display a subject/non-subject asymmetry, but the two structures display the inverse pattern of grammaticality with respect to the use of the overt complementizer.

- (53) a. \*This is the girl \_ likes John  
 b. This is the girl that \_ likes John
- (54) a. This is the girl John likes \_  
 b. This is the girl that John likes \_

I will suggest that the contrast between relative clauses and wh-extraction is not the result of some structural difference but, rather, the result of a difference in the feature content of a relativized nominal as opposed to a wh-word. To clarify, consider the following. Suppose that the featural specifications of the verb and the complementizer in the subject relative in (53b) are identical to the corresponding specifications of an embedded clause displaying subject extraction. Then, the derivation of the subject relative will follow the exact same logic as the derivation of a structure involving subject extraction. The relevant featural specifications are shown in (55) and the resulting derivation is shown in (56), which are modelled

after (43) and (45) respectively. Note that (56a) represents the state of the derivation after the projection of the  $V_0P$ . At this point, the value of the interpretable sub-feature Force is determined by the relativized nominal in the specifier of  $V_0P$ . If we simply assume that the relativized nominal does not bear a Wh feature, it follows that Force will not be assigned an Int value and, consequently, the  $V_0P$  will be a licit complement for the overt complementizer. Therefore, the derivation will successfully produce the relative clause in (56b).

- (55) a. Lexical verb (after Feature Bundling)  
 <F{Force}, T{EPP}, v{uD}, /verb/}, V{uD}>  
 b. Lexical complementizer (after Feature Bundling)  
 <C{uE}, /that/}>



However, capturing the ungrammaticality of a subject relative without an overt complementizer as in (53a) is less straightforward. Following the same reasoning as above, the featural specification of the relevant verb will be as shown



put forward the tentative suggestion that the relativized nominal assigns some other value to the interpretable sub-feature Force, such that the projection of the verb is rendered an illicit complement for the head noun.

### **3.4 Conclusion**

In this chapter, I proposed an analysis of subject/non-subject asymmetries in English *wh*-questions and embedded clauses involving *wh*-extraction. In section 3.2, I demonstrated that the Feature Scattering analysis of functional structure predicts that the structure of subject *wh*-questions is smaller than the structure of their non-subject counterparts. I have shown that this contrast suffices to account for the absence of Subject Auxiliary Inversion in subject *wh*-questions. Furthermore, I proposed two additional hypotheses to account for the absence of *do*-support in the same environment. Firstly, I attributed *wh*-movement to an uninterpretable *uWh* sub-feature, which is present in the featural specification of an auxiliary verb but not in the specification of a main verb. Consequently, a numeration that does not include an auxiliary can derive a subject *wh*-question (due to the fact that a *wh*-subject appears in a scope-taking position for independent reasons), but not an object or a VP adverb *wh*-question. Secondly, I posited an interpretable sub-feature Force, which is hosted by a functional feature F situated between T and C. Additionally, I suggested that Force receives an interrogative value when it establishes a Specifier-Head relation with a *wh*-word, which I claimed to be a prerequisite for a well-formed *wh*-question. The proximity of Force to T entails that *wh*-subjects can establish the required Specifier-Head relation, while *wh*-words that appear in the position of a TP adverb cannot. Thus, I concluded that the position of the subject is privileged compared to the position of the object or various adverbial phrases.

In section 3.3.1, I have extended this analysis to the *that*-trace effect. As I have shown, the observations regarding the structural asymmetry between the position of subjects and non-subjects carry over from *wh*-questions to embedded clauses involving *wh*-extraction. In order to link this asymmetry to the distribution



of the overt complementizer, I proposed that the complementizer cannot take as its complement a constituent that is headed by a functional head marked as interrogative. Consequently, the unique ability of a *wh*-subject to assign an interrogative value to the interpretable sub-feature Force becomes detrimental for the derivation of an embedded clause that includes an overt complementizer.

Furthermore, in section 3.3.2, I demonstrated that this analysis also accounts for the observation that the ungrammaticality of subject extraction over an overt complementizer is repaired when the embedded clause contains a TP adverbial phrase (i.e. the adverb amelioration effect). This is the case because the presence of the adverbial phrase triggers the scattering of an additional functional head that intervenes between the complementizer and the head hosting Force. The situation is less clear with respect to the extraction of a *wh*-subject from a complementizer-less embedded clause that contains a TP adverbial phrase. I arrive at the prediction that this case should be grammatical as well. However, the relevant examples are neither fully grammatical nor fully ungrammatical. This suggests that, at the very least, there are independent factors constraining the presence of TP adverbial phrases in embedded clauses. It remains an open question whether those factors can be incorporated in the overall analysis of the *that*-trace effect developed in this chapter.

Finally, I outlined a tentative analysis of the anti-*that*-trace effect in English relative clauses. I suggested that the contrast between the *that*-trace and the anti-*that*-trace effect is not of a structural nature, but rather it relates to some featural difference between *wh*-words and relativized nominals. Specifically, I suggested that the Specifier-Head relation between the interpretable sub-feature Force and the relativized nominal results in some value that (i) allows for the merger of the overt complementizer with the projection of the embedded verb, but (ii) disallows the merger of the head noun with the projection of the embedded verb. Thus, the overt complementizer is not just permissible but in fact necessary for the derivation of a subject relative. However, this discussion focused on the mechanical aspect of the analysis and it did not provide any suggestion with respect

to the identity of the relevant feature.



#### **4.1 Introduction**

In this chapter, I will address two additional cases of head movement: namely, the movement of finite verbs and the movement of nouns in Semitic and Celtic languages. What separates these examples of head movement from the ones discussed in chapter 2, is that they target the absolute initial position in their respective phrase resulting in verb-initial (specifically, VSO) clauses and noun-initial nominal phrases. I will argue that head movement of this kind challenges the hypothesis that the sole purpose of re-projection is to create (non-vacuous) specifier positions (Giorgi and Pianesi 1997, Bury 2003). In particular, I will demonstrate that the movement of the noun in Semitic Construct State nominals provides a compelling argument against this view. Therefore, I will defend a less restrictive version of re-projection. Specifically, I will propose that there are two types of uninterpretable features that are able to trigger Feature Scattering and, by extension, head movement. Type-1 uninterpretable features can only be eliminated by appearing in a configuration where they c-command a matching interpretable feature. Conversely, type-2 uninterpretable feature are eliminated in a configuration where they are c-commanded by a matching interpretable feature. The second type of feature was employed in the analysis of V2, residual V2 and SVO word orders in the previous chapters, where verb movement is associated with various Specifier-Head relations. In this chapter, I will employ the first type of

feature to the analysis of Semitic and Celtic verb-initial clauses and noun-initial nominal phrases, where head movement targets an absolute initial position.

The chapter is organised as follows. Section 4.2 provides an overview of the relevant empirical data, as well as a detailed discussion of the arguments for and against the view that re-projection serves the purpose of establishing Specifier-Head relations. In section 4.3.1, I propose an alternative view on the purpose of Feature Scattering. On the basis of this proposal, in section 4.3.2, I develop an analysis of noun-initial nominal phrases in Semitic and Celtic languages, including Construct State and Free State nominals. Finally, in section 4.3.2, I extend this analysis to the VSO word order observed in the same languages.

#### ***4.2 Head Movement to First Position and Re-Projection***

In chapter 2, I have discussed head movement through the lens of V-to-C movement in V2 languages (e.g. most Germanic languages) and V-to-T movement in SVO languages (e.g. Romance languages). These two cases of head movement have a subtle similarity. In both cases, the moved head enters a relation with a specifier in its target position. In broad terms, re-projective theories can provide a principled account of structures where head movement coincides with a Specifier-Head relation. Remember that the basic premise of the relevant theories is that functional structure is not predetermined, but rather it emerges dynamically during the derivation. Specifically, structural positions are created only if they are material to the syntactic relations that need to be established during a given derivation. By extension, considering that each instance of re-projection creates an additional functional head and a corresponding specifier, it stands to reason that the relation between these two structural positions is a potentially significant aspect of re-projection. However, Giorgi and Pianesi (1997) and Bury (2003) independently arrive at a stronger conclusion. They propose that re-projection (and re-projective head movement) takes place solely for the purpose of creating specifier positions. In this section, I will argue that this view of re-projection is too restrictive, especially when considering the parallel cases of verb movement in languages with verb-initial

clauses and noun movement in languages with noun-initial nominal phrases. Thus, I will conclude that re-projection must be formulated in such a way that the link between Specifier-Head relations and re-projective head movement is weakened. I will start the discussion in this section with a brief overview of the theoretical and empirical facts that motivate a head movement analysis of verb-initial and noun-initial constructions. I will then provide an outline of the theory internal considerations underpinning Giorgi and Pianesi's (1997) and Bury's (2003) conclusions on the role of Specifier-Head relations in a re-projective theory of functional structure and head movement. Finally, I will evaluate this view of re-projection from the perspective of verb-initial and noun-initial constructions.

Let us begin with a discussion of VSO languages. The view that VSO is derived from a different underlying word order grew out of research in Celtic languages in the late 1970s and the 1980s (Jones and Thomas 1977, Emonds 1980, McCloskey 1983, Sproat 1985, amongst others). In theoretical terms, the advantage of such an approach is that it reconciles the linear order of verb, subject and object with the hierarchical structure of the VP. Assuming that the VP has a universal underlying  $[_{VP} S [_{V'} V O]]$  structure (where the verb and its object form a constituent excluding the subject), a linear order where the  $[V O]$  constituent is separated by the subject can only be captured if one of those elements undergoes movement. Empirically, some of the earliest evidence in support of the existence of a  $[V O]$  constituent in Celtic languages comes from the behaviour of non-finite verbs (McCloskey 1983, Sproat 1985). Unlike their finite counterparts, non-finite verbs in most Celtic languages typically appear in a SVO order<sup>1</sup>. Furthermore, it is also relevant that the non-finite verb and the object undergo clefting as a constituent (see McCloskey 1983 for Irish and Sproat 1985 for Welsh; see also the introduction of Carnie and Guilfoyle 2000:4ff for an overview of additional evidence beyond Celtic).

---

<sup>1</sup> With the exception of non-finite verbs in some varieties of Irish and Scottish Gaelic, which appear in SOV rather than SVO order. See Chung and McCloskey (1987), McCloskey and Sells (1988), Carnie (1995), Bobaljik and Carnie 1996, Adger (1996), for a discussion of relevant structures.

These considerations have given rise to numerous analyses that treat VSO as a derived, rather than base, word order. VSO has been variously attributed to movement of the subject from a VP-external base position to a position inside the VP (Choe 1987, Shlonsky 1987, Chung 1990), to rightward movement of the object from an underlying VOS order (England 1991), and, more recently, to remnant VP movement (where both the subject and the object are moved outside the VP prior to the fronting of the VP remnant; Lee 2000, Massam 2000, Rackowski and Travis 2000). However, the most influential proposal has undoubtedly been the verb-fronting analysis originally proposed by Emonds (1980) and subsequently refined by many others (Sproat 1985, Hendrick 1988, 1991, Stowell 1989, Guilfoyle 1990, Rouveret 1990, 1991, Duffield 1991, Koopman and Sportiche 1991, McCloskey 1991, 1996a, 1996b, 2017, Bobaljik and Carnie 1996; also Mohammad 1989 and Fassi Fehri 1989, 1993 for Arabic). The basic claim of this analysis is that the verb moves from a base position inside a VP with the familiar hierarchical structure, to some functional head higher in the structure, thus surfacing in an initial position in the clause. In spite of the apparent simplicity of this hypothesis, the details of the analysis have proven to be a very fruitful area of research. The major points of interest (and contention) have been the precise landing site of the moved verb, as well as, the surface position of the subject. I will discuss these issues in more detail in section 4.3.3.

Similarly to the VSO literature, research on VOS languages has also been concerned with the question of whether VOS is a base or derived word order. In this case, however, researchers face a rather different set of theoretical and empirical challenges. First of all, note that the constituency of the universal underlying structure of the VP is a non-issue in the case of VOS languages: it is entirely possible to derive a VOS word order from a VP with the familiar hierarchical structure, simply by assuming that the specifier of the verb is linearized to the right (i.e. [<sub>VP</sub> [<sub>V'</sub> V O ] S ]; England 1991, Aissen 1992, Guilfoyle, Hung and Travis 1992, Chung 1998, and others). There are, however, other theoretical considerations that have motivated the analysis of VOS as a derived word order. One such influence has been

Kayne's (1994) proposal for an antisymmetric syntactic structure, which excludes rightward specifiers and thus entails that SVO is the universal underlying word order. Remnant VP movement analyses of VOS (where the subject is removed from the VP prior to the movement of the VP remnant; Lee 2000, Massam 2000, Rackowski and Travis 2000, Coon 2010, and others) have been developed on the basis of this premise. Another theoretical benefit of the remnant VP movement analysis of VOS that has been highlighted in the relevant literature is that it can provide a uniform analysis of VSO and VOS clauses, both cross-linguistically (i.e. across languages with strict VSO word order and languages with strict VOS word order) and intra-linguistically (i.e. across VSO and VOS clauses in languages with alternating VSO/VOS word order). According to this approach, the difference between VSO and VOS boils down to whether the object of the verb has been removed from the VP prior to the movement of the remnant, or not.

Nonetheless, Carnie, Dooley and Harley (2005), Chung (2006) and, more recently, Clemens and Polinsky (2018) argue that empirical facts do not justify a uniform analysis of verb-initial word orders (not only of VSO and VOS as a whole, but also of VOS independently of VSO and vice versa). Rather, these works suggest that it is more likely that there are multiple paths for the derivation of verb-initial word orders (including verb movement and remnant VP movement) across languages and structures. Under this light, it is worth mentioning that verb movement analysis have been proposed not only for VSO languages as we saw above, but also for languages with alternating VSO/VOS word order (see Richards 2000, Rackowski 2002, Rackowski and Richards 2005, Clemens and Coon 2018). In this case, VSO clauses are derived from an underlying SVO word order by verb movement (as we have seen above for Celtic and Semitic languages) and VOS clauses are derived by a combination of verb movement and phrasal movement (either leftward movement of the object to a intermediate position between the base position of the subject and the surface position of the verb, or leftward movement of the subject).



For the purposes of this discussion, I will not adopt any position on the matter of a uniform analysis of VSO languages, VOS languages, or verb-initial languages as a whole. The main concern of this chapter is the implications of verb movement to a clause-initial position for the proposed theory of verb movement. As such, I will focus on those verb-initial languages where verb movement is the least controversial; namely the VSO languages of the Celtic and Semitic families. Of course, the arguments I will present do carry over to any other VSO or VOS language that can be reasonably assumed to involve verb movement, but I will remain impartial as to how common or uncommon verb movement is in verb-initial languages at large. Finally, a quick note on the choice of terminology: when using the term ‘verb-initial languages’ (with head movement), I do not wish to imply that the verb movement analysis of Celtic and Semitic languages should be extended to other VSO/VOS languages. Rather, I choose this terminology to highlight the similarities between head movement of a verb to a clause-initial position and head movement of a noun to an initial position within the nominal phrase (which is, indeed, another characteristic property of Celtic and Semitic languages).

Let us now shift our attention from verb movement to the movement of a head Noun in the nominal domain. The most robust evidence for head movement in nominals comes from languages with noun-initial nominal phrases. Here, I will mostly focus on Semitic and Celtic languages because these languages have both noun-initial nominals and verb-initial clauses, which encourages certain comparisons between nominal and clausal structure. The motivation for positing head movement in these languages comes, once again, from a conflict between surface word order and the commonly held view of the underlying structure of nominal phrases. In brief, nominal structure is believed to be a close parallel of clausal structure<sup>2</sup>: the lexical head (i.e. the noun) with its arguments constitute the most deeply embedded projection of the nominal phrase (similar to a verb and its arguments in the clausal structure), the determiner constitutes the highest functional head of the nominal phrase (comparable to the role of a complementizer

---

<sup>2</sup> A view that essentially originates with the DP Hypothesis, originally proposed by Abney (1987), and subsequently supported and elaborated by numerous linguists.

in the clause), while genitives, adjectives, numerals, etc., populate various intermediate functional projections. Therefore, similar to verb-initial word orders, we can take noun-initial nominal phrases as an indication of movement of the head noun.

Let us then have a closer look at one prominent example of a noun-initial nominal, namely the Semitic Construct State. Semitic languages can express possessives in two ways whose use depends on stylistic, semantic and other factors. The first available structure, sometimes called the Free State or Free Genitive, displays a word order as shown in (1). The head Noun is preceded by an optional proclitic determiner (Semitic languages usually have an overt definite determiner, for example the Hebrew *ha-*, while indefiniteness is expressed with a null determiner) and followed by a prepositional phrase expressing the possessor. Any adjectives modifying the head noun appear between the noun and the PP. The second available structure, known as the Construct State, displays a word order as shown in (2). Here, the head noun appears in absolute initial position and the determiner is necessarily absent regardless of the in/definite interpretation of the nominal. The head noun is then immediately followed by a genitive noun (which, of course, lacks the usual possessive preposition). Note that in the Semitic Construct State, the head noun and the possessor cannot be separated: any adjectives modifying the head noun will appear at the end of the nominal phrase. One final feature of the Construct State is a kind of in/definiteness harmony between the head noun and the possessor noun: i.e. the in/definite interpretation of the head noun is not signified by the presence/absence of a determiner, but instead it depends on the in/definiteness of the possessor noun which, if appropriate, will be marked for definiteness by an overt determiner (for a more detailed view on Semitic nominals see, amongst others, Mohammad 1988, Ritter 1988, Fassi Fehri 1989, 1993, Siloni 1991, 1996, 1997; see also Roberts 2001 for an overview of the literature on the Construct State).

(1) Free State (Semitic):                    D N (Adj) PP

(2) Construct State (Semitic):            N Possessor (Adj)

Consider first the order seen in the Construct State. In the relevant literature, it is usually assumed that the genitive noun expressing the possessor is base generated at the specifier of the head noun (and potentially moved to the specifier of a higher functional head, depending on the particulars of various analyses). This is based on the thematic relation that holds between the head noun and the possessor which, especially for deverbal nouns, is parallel to the thematic relation between a verb and its external argument (cf. Siloni 1996, amongst others). Additionally, it is generally assumed that specifiers are linearized to the left of the head noun in Semitic languages, based on the directionality of specifiers in the clausal domain (cf. Ritter 1988; Siloni 1996 presents additional arguments based on the ordering of the external and internal argument of the noun and the asymmetric c-command relation that holds between them). Therefore, it is safe to conclude that the surface position of the head noun on the left of the genitive in the Construct State can only be derived from the underlying structure of the nominal via head movement. Finally, it is worth noting that the word order of the Free State construction might also be seen as evidence for noun movement under some views. If one adopts a theoretical model that excludes rightward adjunction (as proposed by Kayne 1994), then the position of the noun to the left of the adjective can only be attributed to head movement.

Noun-initial nominal phrases are also found in Celtic languages. As a matter of fact, possessive structures in Celtic languages are very similar to the Semitic possessives discussed above. Once again we find two possible possessive constructions: one where the head noun is preceded by a proclitic determiner and followed by a PP possessor and one where the head noun appears in absolute initial position (with a necessarily absent determiner) and followed by a genitive noun possessor. However, Celtic possessive constructions differ from their Semitic counterpart in two respects: Firstly, in the Celtic counterpart of the Construct State, adjectives intervene between the noun and the genitive, in contrast to Semitic. So the Celtic Construct State displays a word order as shown in (3). Secondly, the in/definite harmony of the Semitic Construct State is not found in the Celtic

counterpart of that structure<sup>3</sup>. Nonetheless, the relative position of the head noun and the genitive in the Celtic Construct State, as well as the overall similarities of the possessive constructions in Celtic and Semitic languages, provide suggestive evidence for a noun movement analysis of Celtic nominals (see Duffield 1996 for a comparison between Semitic and Celtic nominals; see also Guilfoyle 1988).

(3) Construct State (Celtic):                    N (Adj) Possessor

In the remainder of this section, I will consider how a re-projective theory of functional structure may account for head movement to an initial position. As I have mentioned earlier, this is not a straightforward task. This type of movement unavoidably undermines the idea that Specifier-Head relations are a significant (and, according to some, even indispensable) part of the re-projective mechanism. The question, then, is how to broaden the scope of this mechanism so as to encompass head movement to an initial position, while maintaining the relevance of Specifier-Head relations where appropriate. To address this issue, I will start with a brief outline of the strongest position regarding the role of Specifier-Head relations within a re-projective theory (i.e. Giorgi and Pianesi's 1997 and Bury's 2003 view that the very purpose of re-projection is to enable Specifier-Head relations), and then I will consider what are the theoretically available options to weaken that position.

Remember that Giorgi and Pianesi (1997) formulate re-projection in terms of a Feature Scattering operation. Furthermore, they suggest that the application of Feature Scattering during a given derivation is subject to an economy condition. Specifically, they adopt a formulation of this condition proposed by Chomsky (1995), according to which, if a given numeration produces more than one convergent derivation, the preferred one is the one with the least number of operations. This implies that Feature Scattering is a costly operation that should be

---

<sup>3</sup> It seems that the correlation between the in/definiteness of the possessor and the in/definiteness of the head noun in Welsh and Irish is limited, but not entirely absent, in comparison to the Semitic in/definiteness harmony. Specifically, a definite possessor forces a definite interpretation to the head noun, while an indefinite possessor allows for either a definite or an indefinite interpretation of the head noun (Duffield 1996, Roberts 2001).

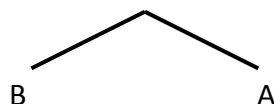
avoided unless necessary. Giorgi and Pianesi (1997) propose that the cost of Feature Scattering is offset only when the additional structural space created by the operation (i.e. the additional specifier) is necessary to accommodate some constituent. Thus, the process of scattering a functional head is intrinsically linked with the creation of the corresponding specifier.

In fact, Giorgi and Pianesi's (1997) version of Feature Scattering is even more restrictive than that. Specifically, they suggest that functional heads cannot be scattered in order to accommodate a moved phrase in their specifier. Only a phrase that is directly merged in the relevant specifier justifies the scattering of a functional head. At first sight, this suggestion seems to entail that phrasal movement can only target the specifier of a non-scattered functional head (i.e. a functional head that is drawn from the lexicon). However, Giorgi and Pianesi (1997) propose the following addendum to their theory. They propose that there are two versions of the Feature Scattering operation. The first one, called Feature Scatter A, is a syntactic operation with the properties discussed above and in section 1.3.1 of chapter 1. The second version, called Feature Scatter B, is an operation that takes place in the numeration. Thus, while Feature Scatter A affects the syntactic structure, Feature Scatter B essentially alters the numeration. Consequently, a numeration where a given syncretic category has been split to separate functional heads due to Feature Scatter B is essentially different from a numeration where the same syncretic category is unaffected by Feature Scatter B. This conclusion entails that the derivations that arise from these two numerations are not in competition for the purposes of the economy condition. Thus, while Feature Scatter A cannot take place to accommodate phrasal movement, Feature Scatter B is not restricted in the same way. Note that, although Giorgi and Pianesi (1997) do not discuss head movement to an initial position, the division between Feature Scatter A and Feature Scatter B may be relevant for the analysis of this type of movement. I will return to this point below. However, before doing so, I will briefly consider Bury's (2003) version of re-projection, which is similar to Giorgi and Pianesi (1997) in so far as it

predicts that the scattering of a functional head is necessarily accompanied by the creation of a specifier.

Bury (2003) proposes a system that replaces traditional syntactic trees with a set-theoretic representation of the dominance relations that hold between the heads in a given syntactic structure. Simplifying somewhat, Bury introduces the concept of a 'treelet', which consists of the categorial feature of a head H and a set of all the heads dominated by H (i.e. the Dominance Set in his terminology). Furthermore, he proposes that a syntactic structure can be represented as the sum of the treelets that correspond to the various heads of the relevant structure. So, for example, a simple structure like (4) is captured by the treelets in (5). Note that the Dominance set of each treelet in (5) is represented inside the curly brackets '{...}'. Note, further, that Bury (2003) adopts a reflexive view of Dominance, which entails that each of the heads in (4) dominates itself. Thus, the treelets in (5) state that the head H dominates the heads A, B and itself, while the heads A and B dominate only themselves. Finally, note that Bury adopts Brody's (1997) Telescoped view of syntactic structure which eliminates intermediate and maximal projections and, therefore, removes the distinction between terminal and non-terminal nodes. Thus, all the nodes in (4) are treated as heads.

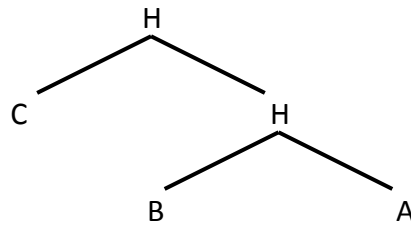
(4)



- (5) a. H {H, A, B}  
b. A {A}  
c. B {B}

Furthermore, Bury (2003) assumes that a syntactic structure can be extended by re-merging a head at the root of the structure, a common hypothesis in the re-projective literature. So, the structure in (4) may be extended as shown in (6), which is a representation of the set of treelets in (7). Therefore, the system has the means to derive functional heads via re-projective head movement.

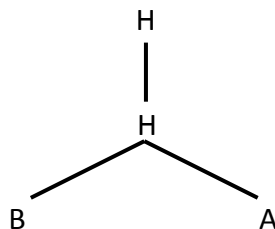
(6)



- (7) a. H {H, A, B, C}  
b. H {H, A, B}  
c. A {A}  
d. B {B}  
e. C {C}

Finally, Bury (2003) observes that expressing dominance relations through Set Theory has an interesting implication. One basic principle of Set Theory is that two sets that have the same members are identical, regardless of the cardinality of each member (e.g. the sets {X} and {X, X} are in fact identical). Suppose then that we extend the structure in (4) as in (8), where the head H is moved and re-merged with the structure but its specifier remains empty. This structure should correspond to the set of treelets in (9). Note, however, that the set in (9a) is identical to the set in (9b) (i.e. the second H in (9a) is not relevant for the identity of the set). By extension, the set of treelets in (9) is identical to the set of treelets in (5) (i.e. the treelet in (9a), which is the same as the treelet in (9b) is not relevant for the identity of the set). Therefore, it is impossible to distinguish between the structures in (4) and (8) in set-theoretic terms. On the basis of this observation, Bury (2003) concludes that it is impossible to extend a syntactic structure via re-merger of a head, if its specifier is not filled. Of course, this conclusion is very reminiscent of Giorgi and Pianesi's (1997) version of Feature Scatter A.

(8)



- (9) a. H {H, H, A, B}  
b. H {H, A, B}  
c. A {A}  
d. B {B}

Let us now return to the issue of head movement to an initial position. As I have mentioned previously, the structure of verb-initial clauses and noun-initial nominals in Celtic and Semitic languages involves head movement to a functional head whose specifier remains empty. Obviously, the relevant functional head (and the movement of a verb or noun to it) cannot be the product of re-projection under Giorgi and Pianesi (1997) or Bury's (2003) version of this operation. Broadly speaking, there are two possible solutions to this problem. The first option is to develop an account of structures involving head movement to an initial position that does not rely on re-projection. Note that this implies that both the origin of the target functional head (i.e. whether it is a lexical item, a scattered head, or some other option) and head movement itself must receive an alternative analysis in the relevant cases. The second option is to extend re-projection to all functional heads and all cases of head movement. This approach requires a less restrictive version of the theory that allows for the re-projection of a functional head regardless of the presence or absence of a specifier, pace Giorgi and Pianesi (1997) and Bury (2003). In the following paragraphs I will consider the first of these two approaches, as defended by Bury (2003). I will show that, although Bury's (2003) analysis of verb-initial clauses is successful, it does not carry over to noun-initial nominals in Celtic and Semitic languages. Thus, I will conclude that it is more profitable to pursue the second alternative.

As I have mentioned above, Bury (2003, 2005, 2007) opts for a non-re-projective account of verb movement in verb-initial languages. His analysis rests on the observation that the moved verb in the relevant languages is typically preceded by an (optional) preverbal particle (Carnie and Guilfoyle 2000, Carstairs-McCarthy 1999). Note that these particles are commonly considered to be heads in the relevant literature (c.f. Hendrick 1988, Chung and McCloskey 1987, McCloskey 1996a, Duffield 1990, amongst others). Bury (2003, 2005, 2007), in



particular, treats them as lexical complementizers. Furthermore, Bury assumes that when one of these particles is present in a given structure, it provides a landing site for verb movement. Of course, the movement of one lexical head to another (i.e. the lexical verb and the lexical complementizer respectively) cannot be attributed to re-projection by definition. Therefore, Bury suggests that, in addition to re-projective head movement, there is a second head movement mechanism that functions similarly to a traditional movement operation (albeit modified to fit within Bury's theory of syntactic structure). However, this analysis is not restricted to clauses that include one of the optional preverbal particles. Bury (2003) adopts Koenenman's (2000) view that null functional heads can be acquired by a learner as long as they are part of a paradigm of overt functional heads. According to this, the existence of a paradigm of overt complementizers (i.e. the preverbal particles) justifies the availability of a null complementizer in the lexicon of verb-initial languages. Thus, a clause that lacks the optional preverbal particle can also be analysed as an instance of (non-re-projective) head movement to a lexical complementizer.

In conclusion, Bury (2003, 2005, 2007) posits two types of null functional heads and two distinct head movement operations. In chapter 1, I have argued against an approach like this, on the grounds that it incurs an unnecessary cost to the theory (c.f. section 1.3.2). In the following, I will show that such an approach is not only costly, but ultimately unsuccessful. Let us consider how Bury's (2003) analysis of verb-initial languages may be extended to noun-initial nominal phrases in Semitic and Celtic languages. Following Koenenman's (2000) argument regarding the acquireability of null non-derived functional heads, it is reasonable to assume that the existence of one overt determiner facilitates the acquisition of a null determiner with a contrasting semantic interpretation. This may indeed be the case with the pair of the overt definite determiner and the null indefinite determiner that is used in the Free State construction in Semitic and Celtic languages. However, as we have seen previously, the situation is different in the Construct State. There, the overt determiner is obligatorily absent, regardless of the in/definite

interpretation of the NP. Therefore, the learner is expected to acquire a second definite determiner, even though this determiner is null and, additionally, it is used exclusively in a construction where it never contrasts with some overt counterpart. I believe that this hypothesis is rather unsatisfactory, especially within a system where the learner has an alternative way of analysing the relevant structures (i.e. by treating noun movement in the Construct State as re-projective movement). Therefore, I conclude that, within a re-projective framework, it is preferable to pursue an analysis of the Construct State as an instance of head movement to a functional position that is derived by re-projection, rather than movement to a non-derived functional head drawn from the lexicon.

If this reasoning is on the right track, the Construct State presents us with a case where the functional structure of the nominal phrase is extended via re-projective head movement, even though the specifier of the derived head remains radically empty. This entails that we need to abandon the claim that re-projection is tied to the creation of a specifier, at least in its stronger form proposed by Bury (2003). Furthermore, the rejection of this hypothesis undermines Bury's (2003) analysis of verb-initial languages, since it removes the motivation for positing two distinct head movement operations. Rather, it is more reasonable to posit a single head movement operation (i.e. some version of re-projective head movement), which is triggered by different factors in different structures: Specifier-Head relations still seem to be relevant in certain cases (e.g. V2 languages with V-to-C movement and SVO languages with V-to-T movement), while head-initial structures seem to require a different trigger. Note, of course, that this approach is relevant to both verb-initial clauses and noun-initial nominal phrases. In the following section I will examine what the relevant triggers may be in both of these cases of head movement.

Before concluding this section, I will add a brief comment regarding Giorgi and Pianesi's (1997) version of a re-projective theory. As I have mentioned above, Giorgi and Pianesi do not discuss verb-initial languages or noun-initial nominal phrases. However, their analysis is similar to that of Bury (2003) in so far as they assume that

the scattering of a functional head is necessarily accompanied by the creation of a specifier. Therefore, in cases of head movement to an initial position, the target functional head ought to be attributed to something other than Feature Scatter A. According to their analysis, the relevant functional head could either be drawn from the lexicon or it could be the product of Feature Scatter B (i.e. the version of the Feature Scattering operation that applies in the numeration). However, it is hard to see what could motivate Feature Scatter B in the head-initial structures under consideration. Of course, syntactic factors should not be relevant to the motivation of Feature Scatter B, since it is a pre-syntactic operation. Alternatively, Feature Scatter B could be motivated by morphological factors. But, in that case, the same argument that challenges the existence of a null lexical Determiner in the Construct State applies to the proposition that this null Determiner is the result of Feature Scatter B. Thus, I return to my previous conclusion that the Construct State points towards a re-projective analysis of the functional structure of head-initial structures.

To summarise, in this section I have considered the possibility that structures involving head movement to an initial position do not fall within the scope of the mechanism of re-projection. I have shown that this approach requires the inclusion of two types of null functional heads and two distinct head movement operations in the theory. Furthermore, I have argued that this is both a costly and unsatisfactory solution for at least some of the relevant structures (i.e. the Semitic and Celtic Construct State). Thus, I have concluded that it is necessary to modify the theory of re-projection so as to incorporate the analysis of head movement to an initial position. I will pursue this approach in section 4.3.

### ***4.3 A Feature Scattering Analysis of Head Movement to First Position***

#### ***4.3.1 Two Triggers for Feature Scattering***

In the previous section I concluded that verb-initial clauses and noun-initial nominal phrases present certain complications for a re-projective analysis of head movement. Furthermore, I have tentatively suggested that the solutions may lie in

the factors that trigger the scattering of a functional head across different structures. Throughout the discussion in sections 4.3.1 to 4.3.3, I will explore what the relevant factors may be. In section 4.3.1, I will consider how one could distinguish between structures involving head movement to an initial position (i.e. verb-initial clauses and noun-initial nominals) and other types of head movement (e.g. V-to-C movement in V2 languages and V-to-T movement in SVO languages) in theoretical terms. I will then flesh out the analysis of noun-initial nominal phrases in section 4.3.2 and the analysis of verb-initial clauses in section 4.3.3.

In chapter 1, I have proposed that lexical items have complex featural specifications that consist of a sequence of functional features and their un/interpretable sub-features. Furthermore, I have suggested that the sequence of functional features captures the distribution of the (derived and non-derived) heads that make up the functional structure, while the un/interpretable sub-features capture the syntactic relations that hold between heads and phrases. The overarching theme of this analysis is that the blueprint of a given syntactic structure is encoded in the featural specification of the lexical items that are involved in its derivation. With that in mind, we would expect that the featural specification of a head H should be able to capture two configurations involving itself and a phrase XP: the first is a syntactic structure where the head H c-commands the XP and the second is a structure where the head H is c-commanded by XP. In order to achieve this, I propose that uninterpretable features belong in one of two categories, depending on the syntactic environment in which they can be eliminated:

- (10) a. Type-1 uninterpretable features:  
An uninterpretable feature  $uF_1$  of type-1 that is part of the featural specification of a (derived or non-derived) head H can be eliminated iff H c-commands an XP with a matching interpretable feature.
- b. Type-2 uninterpretable features:  
An uninterpretable feature  $uF_2$  of type-2 that is part of the featural specification of a (derived or non-derived) head H can be eliminated iff H is c-commanded by an XP with a matching interpretable feature.

Interestingly, due to the formulation of Feature Scattering developed in chapter 1, features of type-2 will always result in a Specifier-Head relation without

any need to explicitly add any constraint to this effect in (10b). This is the case because of the hypothesis that the sequence of functional features dictates the order of syntactic operations, as well as the featural specification of derived functional heads (because of the ancillary hypothesis that Feature Scattering does not displace individual features, but an entire array of functional features and their respective sub-features). Therefore, at any given point during a derivation, any uninterpretable feature that is eligible to trigger a syntactic operation will always be situated at the highest head of the relevant intermediate structure. Obviously, during that same step of the derivation, the only structural position c-commanding the head is the specifier of that head. Thus, features of type-2 are necessarily eliminated in a Specifier-Head configuration. Consequently, although the definition of a type-2 feature in (10a) is not as narrow as the usual definition of a strong feature, it has the exact same result. Finally, let me clarify that the analyses proposed in chapter 2 (i.e. the V-to-C movement of Germanic V2 languages, the V-to-T movement of Romance languages, the Aux-to-T movement of English auxiliaries and V-to-v movement of English main verbs) rely exclusively on uninterpretable features of type-2, since they all involve Specifier-Head relations.

Let us now consider what structures may arise from type-1 uninterpretable features. What is of particular relevance to the topic of this chapter is that type-1 features will result in head-initial constructions under certain circumstances. Consider a derivation involving an uninterpretable feature of type-1. As mentioned above, at the specific step of the derivation when this feature will be considered for elimination, it will be situated at the highest head of the structure that has been derived up to this point. Suppose, further, that the only available constituent with a matching feature at this time happens to be the specifier of this head. Then, Feature Scattering will create a new functional head at the top of the structure to host the uninterpretable type-1 feature. Finally, if the relevant type-1 feature is high enough within the feature specification, then this operation will take place during the last step of the derivation. Therefore, the resulting functional head will

end up occupying the initial position of the structure<sup>4</sup>. In the following sections, I will consider how this proposal can be applied to the analysis of noun-initial nominals and verb-initial clauses in Semitic and Celtic languages<sup>5</sup>.

### ***4.3.2 A Feature Scattering Analysis of Semitic and Celtic Nominal Phrases***

Having established the theoretical tools necessary to capture head movement to an initial position within the re-projective framework proposed in chapter 1, we can now take a closer look at the relevant structures. In this section, I will discuss the structure of nominal phrases in Semitic and Celtic languages. I will begin this discussion with a more detailed overview of the contrasts between the Free and the Construct State, as well as the contrasts between the Semitic and the Celtic Construct State. After that, I will propose an analysis of these structures.

Both Semitic and Celtic nominals express definiteness with a prenominal definite determiner. The Free State nominals in (11a) and (12a) demonstrate this determiner for Hebrew and Irish respectively. On the other hand, both languages lack an indefinite determiner as can be seen in the indefinite Free State nominals in (11b) and (12b).

---

<sup>4</sup> The other possible scenario is that a head bearing a type-1 uninterpretable feature already c-commands some constituent with a matching feature. In that case, there is no need to scatter this type-1 uninterpretable feature across the specifier of the existing head. Therefore, the relevant feature will not appear in initial position, unless the specifier of the head was empty for independent reasons.

<sup>5</sup> Type-1 uninterpretable features might also be relevant to the analysis of yes/no questions that involve verb or auxiliary movement to a clause-initial position, as in the case of English. There is, however, an obstacle for such an analysis. The relevant type-1 uninterpretable feature, whichever that might be, must only be present in the featural specification of an auxiliary when it appears in a yes/no question. When the same auxiliary appears in a wh-question, the relevant feature must be absent. Otherwise, we will arrive at the erroneous prediction that auxiliaries move in subject wh-questions, as well as non-subject wh-questions and yes/no questions. At this point, it is not clear to me how the desired distribution of the relevant type-1 uninterpretable feature could be achieved. Therefore, I will leave the analysis of English yes/no questions for future research.

- (11) a. ha-bayit shel ha-'ish  
 the-house of the-man  
 'the house of the man'  
 b. bayit shel ha-'ish  
 house of the-man  
 'a house of the man'

Siloni (1996)

- (12) a. an pictiur den fhear  
 the picture of + the man  
 'the picture of the man'  
 b. pictiur den fhear  
 picture of + the man  
 'a picture of the man'

Duffield (1996)

The distribution of the definite determiner is the first obvious difference between the Free and the Construct State. Unlike the Free State, the definite determiner is necessarily absent in Construct State nominals with a definite interpretation. This is demonstrated in (13) for Hebrew and (14) for Irish.

- (13) (\*ha-) beyt ha-more  
 (\*the) home the-teacher  
 'the man's home'

Ritter (1988)

- (14) (\*an) teach an fhir  
 (\*the) house the man.Gen  
 'the man's house'

Duffield (1996)

Another thing to observe from the previous examples is that the possessor is expressed differently in the Free versus the Construct State. In the Free State the possessor is expressed by a prepositional phrase, while in the Construct State it is expressed by a genitive noun. Note also that both the PP in the Free State and the genitive noun in the Construct State appear after the head noun of the nominal. As we will see below, the order of the head noun and the possessor proves to be significant for the analysis of both the Free and the Construct State. I will address the two cases in turn in the following paragraphs, starting from the Construct State.

The relative position of the head and the possessor noun in the Construct State constitutes the basis for the argument that this construction involves movement of the head noun. In brief, the usual assumption is that the underlying position of the possessor is at the specifier of the head noun and that this specifier is linearized at the left of the head. Therefore, the underlying structure of the NP is  $[_{NP} \text{ Poss } [_{N'} \text{ N}]]$ . Consequently, the surface  $[\text{N Poss}]$  word order can only be the result of movement of the head noun across the possessor in its specifier to a higher functional projection. As for the hypothesized underlying structure of the NP, evidence for it comes from the asymmetric c-command relations that are observed when a head noun has multiple arguments. Siloni (1996) demonstrates the relevant c-command relations by investigating violations of Condition A and C of Binding Theory and Weak Crossover effects. To elaborate the reasoning of this argument, consider the following example which revolves around Condition A. Note that in example (15) the head noun has two arguments (an Agent and a Patient). Note also that it is possible for the Patient to be a reflexive pronoun bound by the Agent (as in (15a)), but it is not possible for the Agent to be a reflexive bound by the Patient (as in (15b)). Given that reflexive pronouns need to be c-commanded by their antecedent according to Principle A of Binding Theory, the pattern in (15) suggests that the Agent asymmetrically c-commands the Patient.

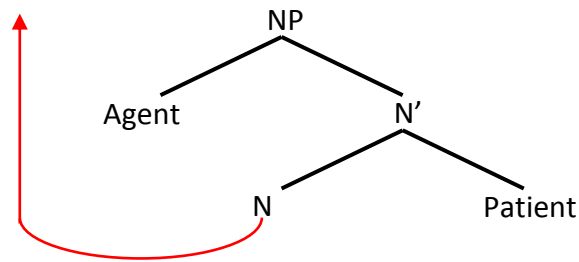
- (15) a. harisat      ha-cava    'et    'acmo  
           destruction the-army Acc    itself  
           'the army's destruction of itself'  
       b. \*harisat      'acmo    'et    ha-cava  
           destruction    itself    Acc    the-army

Siloni (1996)

Therefore, in order to account for the observed c-command relations, as well as the linear order displayed in (15) (i.e. the Agent precedes the Patient), we have to assume a structure where the Agent is a leftward specifier situated higher than the Patient. Thus, we conclude that (at least as a rough first approximation) the structure of the Construct State is as in (16) below.



(16)



Let us now turn to the relative order of the PP and the head noun in the Free State. As we have already seen in (11), the PP in the Free State also appears post-nominally. Ritter (1988) attributes the position of the PP simply to rightward adjunction of the PP with the DP. This simple proposal has the additional benefit that it accounts for the difference in the placement of adjectives in the Free State versus the Construct State. However, before going into more detail in the distribution of adjectives, let us consider an alternative view on the position of the PP proposed by Siloni (1996). Siloni (1996) observes that in the Free State the PP necessarily precedes an argument of the head noun that bears the Patient thematic role (cf. the contrast between (17a) and (17b)). Furthermore, (18) demonstrates that the PP and the Patient in the Free State display the same asymmetric c-command relation that we saw in the Construct State. Therefore, Siloni (1996) assumes that the Free State has the same underlying structure as the Construct State. In the following, I will adopt Siloni's (1996) view, as the evidence in (17) and (18) point strongly against a rightward adjunction analysis for the PP in the Free State.

- (17) a. ha-harisa            shel ha-cava 'et ha-'ir  
      the-destruction of the-army Acc the-city  
      'the army's destruction of the city'
- b. \*ha-harisa            'et ha-'ir shel ha-cava  
      the-destruction Acc the-city of the-army  
      'the army's destruction of the city'

- (18) a. ha-harisa            shel ha-cava    'et    'acmo  
           the-destruction of    the-army Acc    itself  
           'the army's destruction of itself'
- b. \*ha-harisa            shel    'acmo    'et    ha-cava  
           the-destruction of    itself    Acc    the-army  
           'the army's destruction of itself'

Siloni (1996)

In the previous discussion I alluded to the significance of the positioning of the adjective in relation to the other syntactic material in the nominal phrase. Indeed, the position of the adjective is distinct not only between the Semitic Free State and the Semitic Construct State, but also between the Semitic Construct State and its Celtic Counterpart. Generally speaking, both Semitic and Celtic adjectives are typically post-nominal. More specifically, in the Free State, in both Semitic and Celtic, the adjective appears immediately after the head noun, while the PP appears at the final position. Example (19) demonstrates the relevant order in the Free State for Hebrew.

- (19) ha-bayit    ha-gadol    shel ha-'isha  
       the-house the-big    of    the-woman  
       'the woman's big house'

Siloni (1997)

Unlike the Free State, in the Semitic Construct State, it is the possessor that must appear immediately after the head noun, while the adjective appears in the final position. However, this is not the case for Celtic languages. In the Celtic Construct State, the order is similar to the order of the Free State with the adjective following the head noun and the possessor at final position. The examples in (20a) and (20b) demonstrate this contrast. (Note that the adjective in the Hebrew example may modify either the head noun or the possessor. Conversely, in the Irish example, the adjective unambiguously modifies the head noun. An adjective modifying the possessor would appear in the final position).

- (20) a. beyt ha-'isha ha-gadol  
house the-woman the-big  
'the woman's big house/the big woman's house'

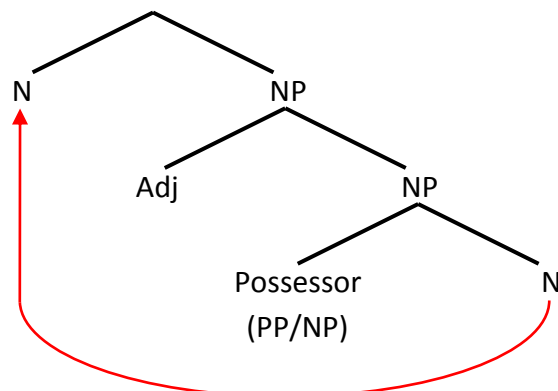
Siloni (1997)

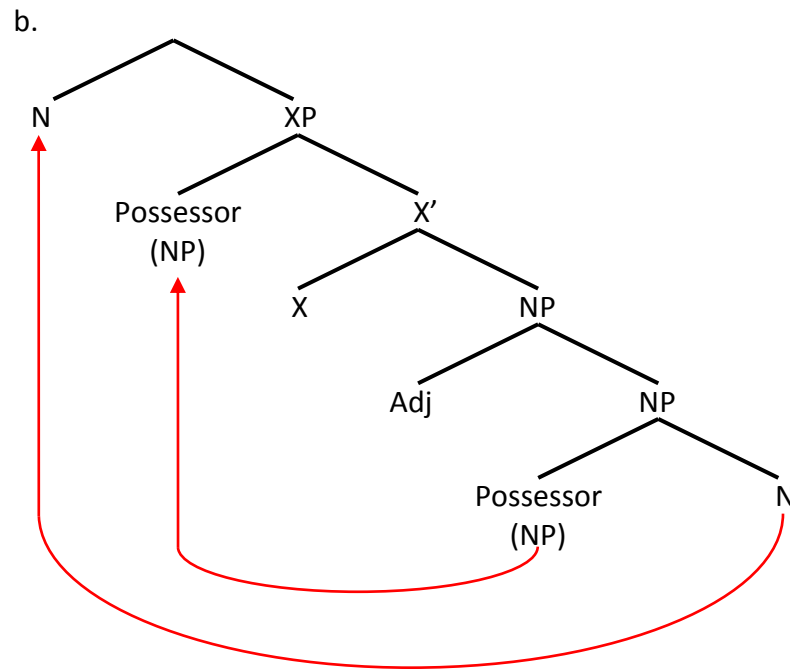
- b. guth laidir an tsagairt  
voice strong the priest.Gen  
'the priest's powerful voice'

Duffield (1996)

As I have mentioned earlier, Siloni (1996) argues that the Free State and the Construct State have the same underlying structure (whereby the base position of both PPs and genitive nouns is at the specifier of the head noun). The question then is how to capture the different position of the adjective in the two structures. Siloni's (1996) solution has two components: Firstly, the adjective is assumed to be left-adjoined to the NP in both structures. Secondly, while the PP in the Free State is assumed to stay in-situ (and therefore to the right of the adjective), the possessor in the Construct State is hypothesized to move to a higher structural position to the left of the adjective. Duffield (1996) uses the same device to account for the different position of the adjective in the Semitic Construct State versus the Celtic Construct State. Therefore, we can capture the observed variation with just two basic structures. (21a) is an approximation of the structure of a Free State nominal, as well as a Celtic Construct State, while (21b) is the structure of the Semitic Construct State.

- (21) a.





The final unique property of Construct State nominals is that the definite or indefinite interpretation of the head noun seems to be linked to the definiteness of the possessor noun. Specifically in Semitic languages we observe a kind of in/definiteness harmony where the in/definiteness of the head is entirely dependent on the possessor. Given the absence of a definite determiner in the Construct State, the in/definiteness of the head noun can be diagnosed by looking at the in/definiteness of adjectives modifying it. In (22a) the head noun is interpreted as definite, as indicated by the definiteness marking on the adjective (i.e. the proclitic definite determiner). Of course, the adjective may modify either the head or the possessor noun, but here we are interested in the first possibility. On the other hand, the head noun in (22b) is indefinite as indicated by the indefinite adjective. Thus, in both cases the in/definiteness interpretation of the head noun corresponds to the in/definiteness of the possessor.

- (22) a. beyt ha-ish ha-gadol  
house the-man the big  
'the big man's house/the man's big house'  
b. beyt ish gadol  
house man big  
'a big man's house/a man's big house'

Roberts (2001)

The situation is much less clear in the Celtic Construct State. Roberts (2001), claims that definite possessors in Celtic have the same effect as in Semitic, in that they force a definite interpretation of the head noun, while indefinite possessors do not pose any restriction on the in/definite interpretation of the head noun. Example (23) demonstrates the first case: the head noun of a Construct State nominal with a definite possessor can only be understood as definite (note that adjectives in Celtic languages are not marked for definiteness, so the argument is based on interpretative evidence alone). Example (24) demonstrates the second case: a Construct State nominal with an indefinite possessor is compatible with both a definite and an indefinite interpretation of the head noun.

- (23) pictiur an fhir  
picture the man.Gen  
'the/\*a picture of the man'

Irish; Duffield (1996)

- (24) mab brenin  
son king  
'the son of a king/a son of a king'

Welsh; Rouveret (1994)

In the previous paragraphs, I provided a fairly broad overview of Semitic and Celtic nominal phrases. In so doing, I glossed over some important details that complicate the analysis of the relevant structures (see the appendix for discussion). Thus, the following discussion is meant to be a jumping-off point for the understanding of Semitic and Celtic nominal phrases, rather than a comprehensive analysis. Before moving on to that discussion, let us summarise some of the key points of the previous overview. Firstly, I will adopt the view that the underlying structures of the Free and Construct State are similar (Siloni 1990). For the purposes

of the following discussion, the two most important aspects of that structure are the placement of the possessor and the adjective, which are highlighted in (25a-b) respectively.

- (25) a. The base position of the possessor (the PP in the Free State or the genitive noun in the Construct State) is at the specifier of N.
- b. Adjectives are left-adjoined to the NP.

Secondly, the examination of the surface word order of the head noun, the possessor and the adjective indicates that the relevant structures involve two movement operations: head movement of the head noun and phrasal movement of the possessor. The former is observed in both the Construct and the Free State in Semitic and Celtic languages (cf. (26a)). The latter is observed in the Semitic Construct State only (cf. (26b)).

- (26) a. The surface order of the head noun at the left of the possessor (either a PP or a genitive noun) suggests that the head noun undergoes movement to a higher functional position in both the Free and the Construct State.
- b. In the Semitic Construct State the position of the possessor at the left of the adjective suggests that the possessor moves across the adjective to the specifier of some intermediate functional projection between the surface and base position of the head noun. In the Celtic Construct State, as well as the Free State, the possessor surfaces in final position at the right of the adjective suggesting that in these structures the possessor remains in-situ.

Thirdly, the overt determiner displays different distribution in the Free and Construct State (cf. (27a-b) respectively). This difference is observed in both Semitic and Celtic languages.

- (27) a. Free State nominals have a prenominal definite determiner, while they lack an indefinite determiner.
- b. Construct State nominals lack a determiner regardless of the in/definite interpretation of the head noun. Consequently, the head noun appears in initial position.

Finally, in the construct State, the in/definite interpretation of the head noun is (fully or partly) determined by the in/definiteness of the possessor. Due to this,

the Construct State displays a form of in/definiteness harmony. In particular, the Semitic Construct State displays a fully-fledged in/definiteness harmony (cf. (28a)), while the Celtic Construct State displays a more limited version of the same effect (cf. (28b)).

- (28) a. In the Semitic Construct State the in/definiteness of the possessor determines the in/definite interpretation of the head noun.
- b. In the Celtic Construct State a definite possessor (arguably) forces a definite interpretation for the head noun, while an indefinite possessor does not constrain the in/definite interpretation of the head noun.

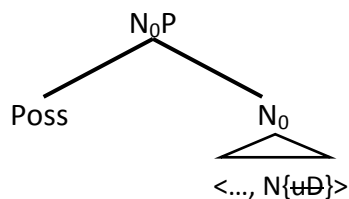
In the remainder of this section I will develop an analysis of Free and Construct State nominals that is based on the theoretical tools developed in the previous chapters and in section 4.3.1. The discussion is organised as follows. I will firstly present a preliminary analysis of the Semitic Construct State, the Celtic Construct State and the Free State, in that order. I will then compare the structure of the relevant nominal phrases, on the basis of this preliminary analysis. Finally, this comparison will lead to certain conclusions that will allow me to flesh out the final version of the analysis of noun-initial nominal phrases in Semitic and Celtic languages.

As I have already mentioned, I adopt the view that the internal structure of a nominal phrase is a parallel of the internal structure of a clause. Therefore, my analysis of the nominal phrase will follow the same general direction as the analysis of the clause in the previous chapter. First of all, I will assume that a noun is stored in the lexicon with an underspecified featural specification, which consists solely of the functional feature N. Furthermore, I will assume that N is the lowest member of a sequence of functional features that captures the structure of a nominal extended projection. Of course, due to Feature Bundling, the members of this sequence will be assigned to the lexical nominal heads (i.e. the noun, the determiner, etc.) that happen to be present in a given numeration. I will elaborate on the details of the nominal sequence of functional features and the role of Feature Bundling in the following paragraphs. However, this very rudimentary description of the system

suffices to capture the base position of the genitive noun in relation to the head noun in the Construct State (as per (25a)). Thus, I will address this point first. Similarly to the analysis of a clause, I will assume that during the numeration the functional feature N is assigned one or more uninterpretable sub-features which are responsible for selecting the arguments of the head noun. In the Construct State, the relevant feature is a uD uninterpretable feature of type-2. Accordingly, a first partial representation of the featural specification of a noun will be as in (29) below. This featural specification will give us the very first step in the derivation of the Construct State (shown in (30)), which roughly corresponds to the traditional maximal projection of the lexical head of the nominal (i.e. the NP).

- (29) a. Semitic Construct State noun (before Feature Bundling)  
 <..., N{...}>  
 b. Semitic Construct State noun (after Feature Bundling)  
 <..., N{uD}>

(30)

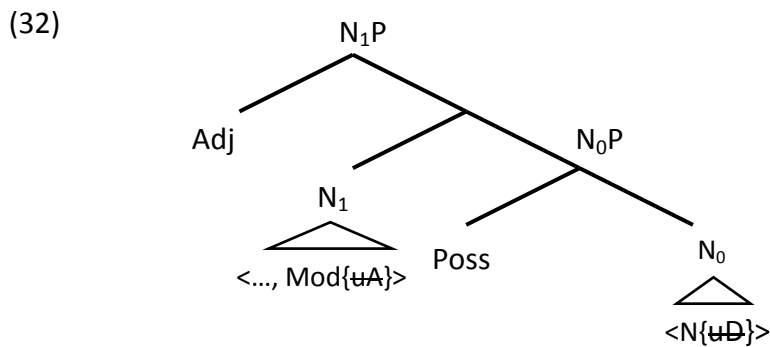


The second point I will address is the position of an adjective in the Construct State. Broadly speaking, I will follow Siloni's (1996) view that adjectives are adjoined to the NP (as per (25b)). However, in the framework developed in the previous chapters, I have adopted Cinque's (1999) analysis of adverbials as specifiers of dedicated functional heads. Under this analysis of modifiers, Siloni's (1996) proposal translates to the hypothesis that adjectives are the leftward specifier of a functional head that immediately dominates the NP. Furthermore, in the current framework the relevant functional head will be attributed to the (optional) Feature Scattering of an appropriate functional feature from the featural specification of the head noun. Specifically, I will assume that the nominal sequence of functional features contains a feature Mod(ifier), which appears immediately to the left of the feature N (i.e. <..., Mod{...}, N{...}>). Furthermore, I will assume that Feature



Bundling will assign the functional feature Mod to the specification of the noun. Thus, I will amend the featural specification in (29) as shown in (31). Note that the functional feature Mod in (31) bears an uninterpretable sub-feature  $uA$ , which will be responsible for the merger of an adjective in the structure. This sub-feature is placed in parentheses to indicate that it is optional. It will only be present in a numeration that includes an adjective. Finally, the derivation of a Construct State nominal that includes an adjective will proceed as shown in (32). This structure shows that the uninterpretable sub-feature  $uA$  cannot be eliminated at the  $N_0P$  level, thus causing the Feature Scattering of the functional feature Mod to  $N_1P$ .

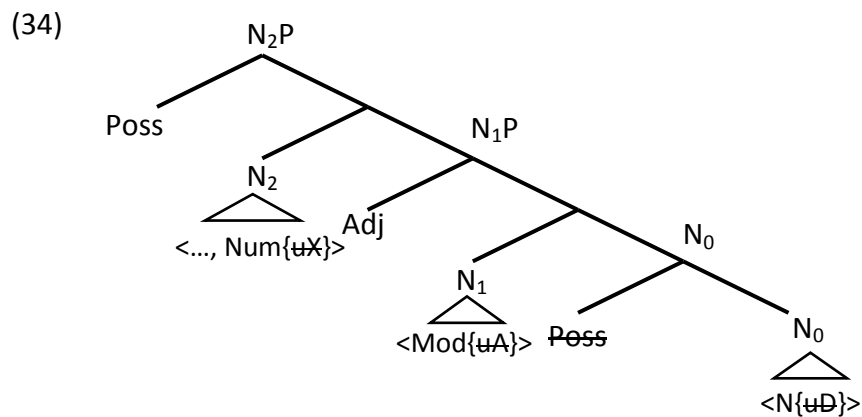
(31) Semitic Construct State noun (after Feature Bundling) - amended  
 <..., Mod{ $uA$ }, N{ $uD$ }>



While (31)-(32) captures the base position of the adjective, we still need to account for the surface order of the adjective in relation to the possessor in the Semitic Construct State. As already stated, the position of the possessor at the left of the adjective can be attributed to a movement of the possessor across the adjective to some intermediate functional projection (cf. (26b)). I will follow Ritter (1991) in assuming that the relevant functional projection is Num(ber), but, of course, I will analyse it in terms of Feature Bundling and Feature Scattering. Therefore, I will further refine the nominal sequence of functional features to include a feature Num immediately to the left of Mod (i.e. <..., Num{...}, Mod{...}, N{...}>). Once again, Feature Bundling will assign the functional feature Num to the specification of the noun. Furthermore, I propose that Num bears some

uninterpretable sub-feature  $uX^6$  of type-2 (cf. (10b)), which needs to be checked against the possessor in a Specifier-Head relation. Thus, I will further amend the featural specification in (31) as shown in (33). Accordingly, the derivation in (32) will proceed as shown in (34). Briefly, the specifier of  $N_1$  is occupied by the adjective, so the type-2 sub-feature  $uX$  cannot be eliminated. Therefore, the functional feature Num is scattered to  $N_2P$ .

(33) Semitic Construct State noun (after Feature Bundling) - amended  
 <..., Num{ $uX$ }, Mod{(uA)}, N{uD}>



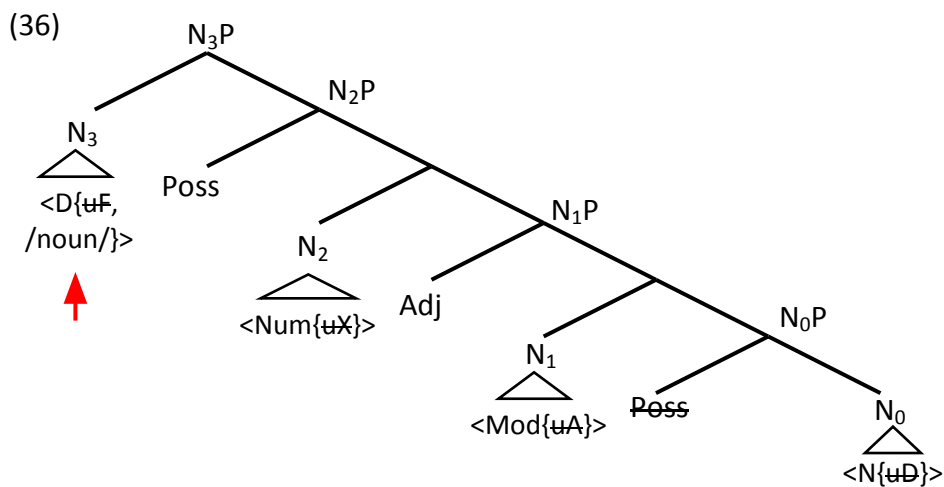
The last point to address is the movement of the head noun across the possessor to the initial position (as per (26a)). Following Ritter (1988) and Siloni (1996) (amongst others), I assume that the functional projection where the noun surfaces corresponds to D. Of course, according to the discussion in section 4.2, I will assume that this functional projection is headed by a scattered functional head, rather than a null determiner drawn from the lexicon. To achieve this, I will assume the following: (i) the highest feature in the nominal sequence of functional features is the feature D (i.e. <D{...}, Num{...}, Mod{...}, N{...}>), (ii) the functional feature D bears some uninterpretable sub-feature  $uF$  of type-1 (cf. (10a), which needs to be scattered to a position c-commanding the possessor in order to be eliminated, and (iii) the phonological features of the noun are also hosted by the functional feature

---

<sup>6</sup> Note that in the literature on the Semitic Construct State there is no consensus regarding the syntactic relation that triggers the movement of the possessor to Num. Thus, I will not make any claim about the category of X in Num{ $uX$ }.

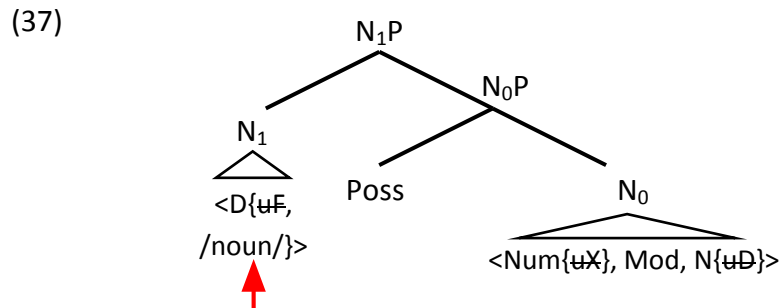
D. Thus, the featural specification of the noun in (33) should be amended as in (35b). Note that I have previously proposed that the placement of the phonological features of a lexical item is determined in the lexicon rather than the numeration, so it is also necessary to amend the underspecified featural specification of the noun in (29a) as shown in (35a). Finally, the derivation in (34) will proceed as shown in (36). What we see in this structure is that the type-1 uninterpretable feature  $uF$  cannot be eliminated at  $N_2$ , because it does not c-command the possessor. Therefore, Feature Scattering will move  $D$  and its sub-features, including the offending  $uF$  and the phonological features of the noun, to  $N_3$ . Consequently, the noun will surface at the initial position of the nominal phrase (i.e. the position of  $N_3$ ).

- (35) a. Semitic Construct State noun (before Feature Bundling) - amended  
 <C{..., /noun/}, ... , N{...}>  
 b. Semitic Construct State noun (after Feature Bundling) - amended  
 <C{ $uF$ , /noun/}, Num{ $uX$ }, Mod{{ $uA$ }}, N{ $uD$ }>



To be thorough, let us also consider the derivation of a Semitic Construct State nominal that does not include any adjectives. The relevant derivation will be based again in the featural specification in (35), with the only difference that the functional feature  $Mod$  will not have an uninterpretable sub-feature. As the structure in (37) demonstrates, both of the type-2 uninterpretable features in the specification of the noun (i.e. the  $uD$  sub-feature of  $N$  and the  $uX$  sub-feature of  $Num$ ) will be eliminated at  $N_0$  against the possessor at the specifier of  $N_0P$ .

However,  $uF$ , the type-1 uninterpretable sub-feature of  $D$ , cannot be eliminated at the same point, since  $N_0$  does not c-command the possessor. Rather,  $D$  and its sub-features have to be scattered to  $N_1$ . Therefore, the phonological features of the noun will surface in the initial position as expected.

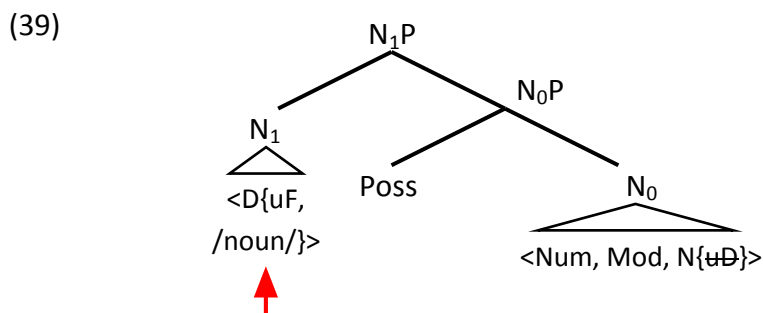


In the discussion above, I have developed a preliminary analysis of the Semitic Construct State. What remains to be done is to identify the category of the uninterpretable feature  $uF$  that triggers the scattering of the functional feature  $D$ , which hosts the phonological features of the noun, to an initial position. To achieve this, it is necessary to compare the Semitic Construct State with its Celtic counterpart, as well as with the Free State construction in both Semitic and Celtic languages. Therefore, in the following paragraphs, I will present a preliminary analysis of Celtic Construct State nominals and Free State nominals, before returning to the issue of the identity of  $uF$ .

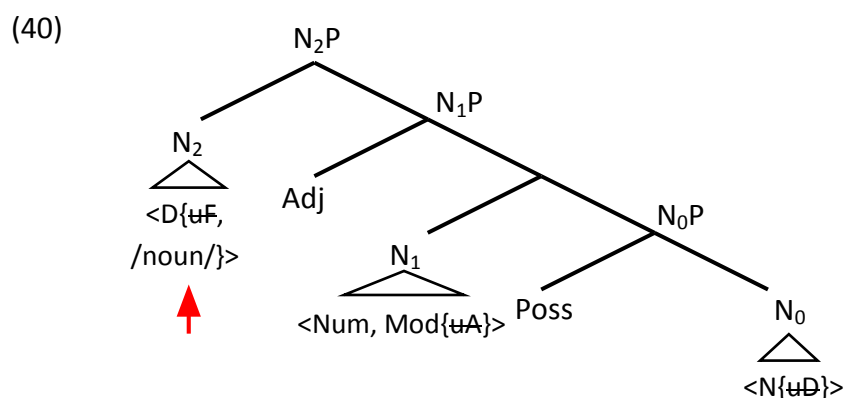
As mentioned in (26b), one of the main differences between the Semitic and the Celtic Construct State is the position of the adjective in relation to the possessor. While the adjective follows the possessor in the Semitic Construct State, in the Celtic Construct State the adjective precedes the possessor. Following Duffield (1996), I will assume that what differentiates the Celtic from the Semitic Construct State is that while the possessor moves to  $Num$  in Semitic, in Celtic it remains in-situ. Therefore, I propose that the functional feature  $Num$  in Celtic languages lacks the  $uX$  sub-feature seen in the specification of the Semitic noun in (35b). The proposed featural specification is shown in (38) below.

- (38) Celtic Construct State noun (after Feature Bundling)  
 <D{ $uF$ , /noun/}, Num, Mod{{ $uA$ }}, N{ $D$ }>

In the absence of an adjective, the contrast between (35b) and (38) is practically inconsequential. The derivation of a Celtic Construct State (without an adjective) will include one less operation compared to the derivation of the Semitic Construct State, due to the absence of the *uX* sub-feature on Num. However, the resulting structures will be identical since the uninterpretable sub-feature *uX* does not trigger any Feature Scattering operation in a Semitic Construct State nominal without an adjective. For comparison, the structure of a Celtic Construct State lacking an adjective is shown in (39) below.

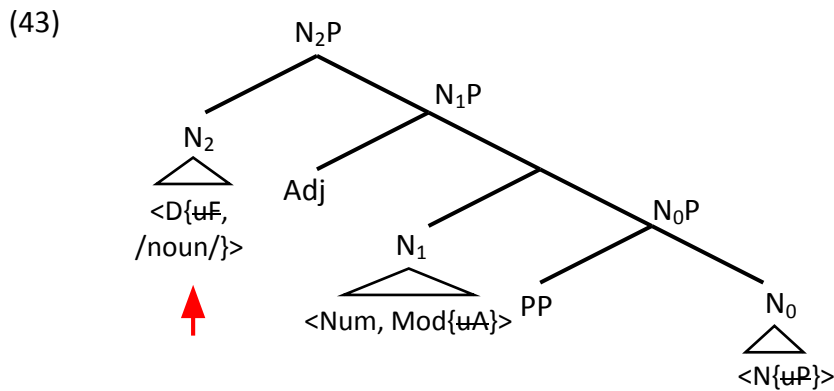
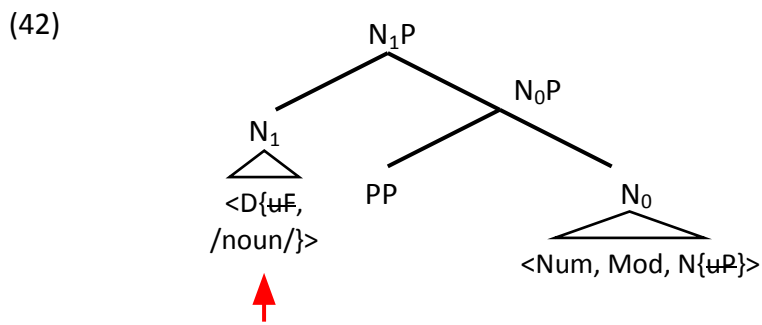


While the featural specification in (38) leads to a rather straightforward derivation for a Celtic Construct State nominal lacking any adjectives, the inclusion of an adjective in the numeration presents us with an unexpected problem. Based on the order of the adjective and the possessor in the Celtic Construct State (cf. (26b)), the intended derivation should look like (40). As we would expect, in (40) the uninterpretable sub-feature of  $N\{uD\}$  is eliminated at the  $N_0P$  level and the uninterpretable sub-feature of  $Mod\{uA\}$  is scattered and eliminated at the  $N_1P$  level. Additionally, the absence of *uX* on Num entails that the possessor remains in-situ as desired. Finally, in order to capture the initial position of the noun, we expect  $D\{uF, /noun/\}$  to be scattered to  $N_2$ . Note however that at the  $N_1P$  level,  $D\{uF\}$  already c-commands the possessor. Therefore, the uninterpretable sub-feature *uF* should be able to be checked at the  $N_1P$  level without resorting to Feature Scattering. This is of course an undesirable conclusion. Thus, the structure in (40) suggests that *uF* needs to be scattered to a position where it c-command the adjective, as well as the possessor.



Let us also consider the structure of Free State nominal phrases in Semitic and Celtic languages. Remember that according to Siloni (1996) there is evidence that Free State and Construct State nominals have similar underlying structures (cf. (25)). Remember also that the surface order of adjective and possessor in the Free State is the same as in the Celtic Construct State, indicating that in both of these constructions the possessor remains in-situ unlike the Semitic Construct State (cf. (26b)). Therefore, we can conclude that the featural specification of a noun in a Free State nominal (in both Semitic and Celtic languages) is fundamentally similar to the specification of the noun in the Celtic Construct State shown in (38). The only notable difference is that the uninterpretable sub-feature of N should be uP, rather than uD, since the argument of the noun in the Free State is a PP. Accordingly, I assume that the relevant featural specification is the one shown in (41). Furthermore, the derivation of the Free State will be essentially identical to the derivation of a Celtic Construct State nominal. The one significant difference between these two constructions is the presence of a determiner in a definite Free State nominal, a topic that I will address later. For the moment, I will focus on the derivation of an indefinite Free State nominal, which will arise from a numeration that does not include a determiner. For reference, (42) represents the structure of a Free State nominal lacking an adjective, while (43) is the structure of a Free State nominal with an adjective. Note that these structures are modelled on the structure of the Celtic Construct State in (39)-(40) above.

- (41) Free State noun (after Feature Bundling)  
 <D{uF, /noun/}, Num, Mod{{uA}}, N{uP}>



Although the derivation of the Free State in (42) and (43) is largely a straightforward extension of the previous analysis of the Construct State, there are two important points to note. Firstly, the derivation of a Free State nominal phrase including an adjective in (43) presents the same problem as the Celtic Construct State counterpart that includes an adjective. Namely, we need to motivate the scattering of  $D\{uF\}$  from  $N_1$  to  $N_2$  across the adjective. Secondly, in a Free State nominal without an adjective as in (42), the surface position of the noun to the left of the PP raises a similar problem. Namely, we also need to motivate the scattering of  $D\{uF\}$  across a PP.

To sum up our findings so far,  $D\{uF\}$  is scattered to a position where it c-commands (i) a genitive noun in the Semitic Construct State (with or without an adjective) and in the Celtic Construct State (as long as there is no adjective), (ii) a possessor PP in the Free State (again in the absence of an adjective), and (iii) an adjective in the event that a Free State nominal or a Celtic Construct State nominal involve an adjective. Consider also the following: All of these c-command relations are established in order to eliminate the uninterpretable sub-feature  $uF$ . If it was

not for the disparate nature of genitive nouns, PPs and adjectives, we could simply assume that all of those syntactic objects have some feature in common that matches uF. Potentially, that could even help us identify the category of uF. However, the disparity of the relevant elements suggests that a more elaborate analysis might be in order. One possibility that is worth considering is that we are not dealing with a single feature but a collection of two or more type-1 features that all contribute to the same result: i.e. the scattering of D and its sub-features (including the phonological features of the noun) to an initial position in the various structures that we are considering. Specifically, I will argue that a uniform analysis of the scattering of D across a genitive noun and across a PP is possible, but I will propose that the scattering of D across an adjective is an independent (although fundamentally similar) operation.

Let us then consider what the genitive noun and the PP might have in common. More specifically, we are interested in features of the genitive noun and the PP that can be reasonably assumed to show some relation with the head noun or its functional structure. In other words, features like number, person and gender are orthogonal to this particular issue, since neither the genitive noun nor the (NP inside the) PP agree with the head-noun for any of these features. Definiteness is another option that initially seems promising given the definiteness harmony in the Semitic Construct State, but must be eventually discarded because the definiteness of PPs is independent from the definiteness of the head noun in the Free State. This leaves us with one last option: case. The possessor in the Construct State is obviously marked for case. Arguably, this is also true for the possessor in the Free State, under the assumption that the preposition is in fact a case marker (as suggested by Ritter (1991) for Hebrew). It would then be reasonable to assume that D establishes a relation with the external argument of the noun (i.e. the possessor) for the purpose of case assignment in both the Construct and the Free State. Therefore, I propose that the feature D in the featural specification of the noun bears a uCase sub-feature with an appropriate value (i.e. uCase:Gen in the Construct State and uCase:Prep) in the Free State), that needs to be eliminated



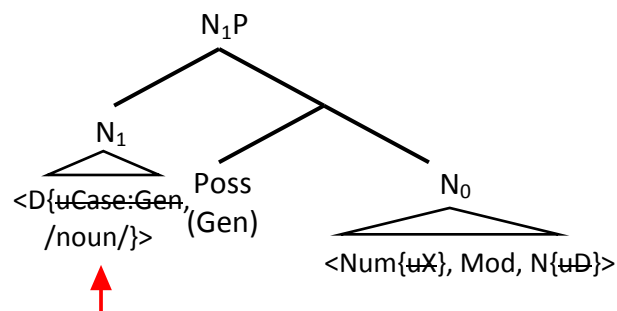
against a possessor bearing a matching feature. Furthermore, in order to ensure the scattering of D to a position where it c-commands the possessor, I assume that uCase is a type-1 feature. Thus, we can amend the earlier featural specifications by replacing uF with uCase, as shown in (44)-(46). These specifications should be able to capture the derivation of Semitic Construct State, Celtic Construct State and Free State nominal phrases that lack an adjective. The relevant derivations are rather predictable as they follow the exact same reasoning as (37), (39) and (42) respectively. For reference, I repeat these structures with the appropriate modifications in (47)-(49) below.

(44) Semitic Construct State noun (after Feature Bundling) – amended  
 <D{uCase:Gen, /noun/}, Num{uX}, Mod{(uA)}, N{uD}>

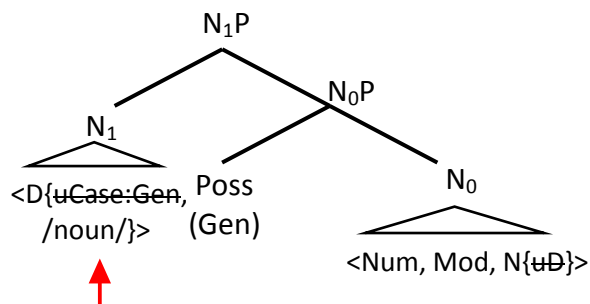
(45) Celtic Construct State noun (after Feature Bundling) – amended  
 <D{uCase:Gen, /noun/}, Num, Mod{(uA)}, N{uD}>

(46) Free State noun (after Feature Bundling) – amended  
 <D{uCase:Prep, /noun/}, Num, Mod{(uA)}, N{uD}>

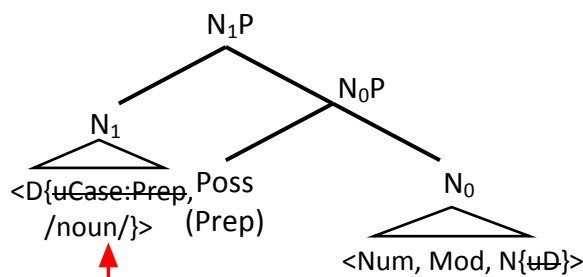
(47) Semitic Construct State (without adjectives)



(48) Celtic Construct State (without adjectives)



(49) Semitic and Celtic Free State (without adjectives)



While the previous analysis offers a uniform account for the scattering of D in Free and Construct State nominals lacking an adjective, it is doubtful that it could be extended to the scattering of D across adjectives, considering that adjectives are not case marked in either Semitic or Celtic languages. Therefore, the scattering of D across an adjective has to be attributed to some different feature. Adjectives in both Semitic and Celtic languages typically agree with the noun they modify for person, number and gender<sup>7</sup>. In principle the scattering of D could be attributed to any of those features, but for clarity, I will assume that the relevant feature is number. Therefore, I propose that the featural specification of the noun contains an uninterpretable *uNum* sub-feature, which is eliminated against a matching feature on the adjective. Of course, I assume that *uNum* is a type-1 feature that needs to be scattered to a position c-commanding the adjective. Furthermore, I assume that this sub-feature is situated on the functional feature *Num*. As a result, the scattering of *Num* will also affect D, due to their order in the sequence of functional features. Finally, with the addition of the *uNum* sub-feature in the specifications in (44)-(46) above, we arrive at the final featural specification of the noun in Semitic and Celtic nominal phrases. The relevant specifications are shown in (50)-(52) below.

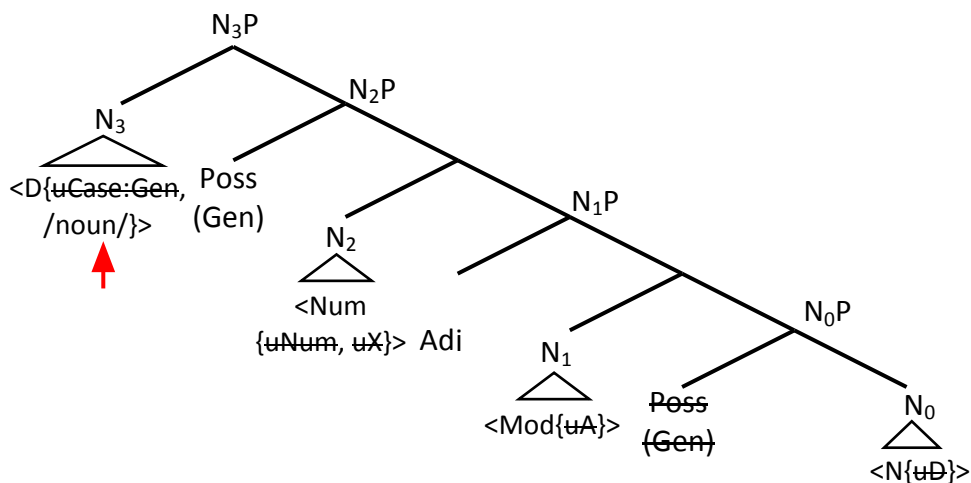
(50) Semitic Construct State noun (after Feature Bundling) – final version  
 $\langle D\{uCase:Gen, /noun/\}, Num\{uX, uNum\}, Mod\{(uA)\}, N\{uD\} \rangle$

<sup>7</sup> Semitic adjectives also agree with their noun for definiteness. However, this is not the case for Celtic adjectives. So for the sake of consistency between the analysis of Semitic and Celtic languages, I will not consider definiteness as a strong candidate.

- (51) Celtic Construct State noun (after Feature Bundling) – final version  
 <D{uCase:Gen, /noun/}, Num{uNum}, Mod{(uA)}, N{uD}>
- (52) Free State noun (after Feature Bundling) – final version  
 <D{uCase:Prep, /noun/}, Num{uNum}, Mod{(uA)}, N{uD}>

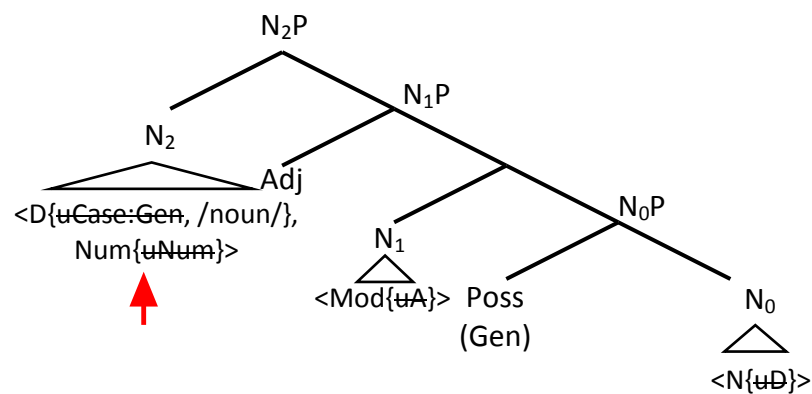
Let us then consider the interaction of D{uCase} and Num{uNum} in the derivations of nominal phrases that include an adjective. The structure in (53) represents the derivation of a Semitic Construct State nominal. What is interesting here is the Feature Scattering operation from N<sub>1</sub> to N<sub>2</sub>. Note that the feature Num has two uninterpretable sub-features, neither of which can be eliminated at the N<sub>1</sub>P level: uNum needs to be moved to a position where it c-commands the adjective, while uX needs to move to a position with an empty specifier that can provide a landing site for the movement of the possessor. Therefore these two sub-features simultaneously trigger the scattering of Num to N<sub>2</sub>P. Furthermore, D{uCase} cannot be checked at the N<sub>2</sub>P level, since the possessor has moved to the specifier of N<sub>2</sub> at this stage. Thus, D is scattered to N<sub>3</sub>, which concludes the derivation. Note that in this derivation the final position of the phonological features of the noun essentially depends on D{uCase}, while Num{uX} and (redundantly) Num{uNum} are responsible for the intermediate steps of the derivation.

(53) Semitic Construct State (with an adjective)



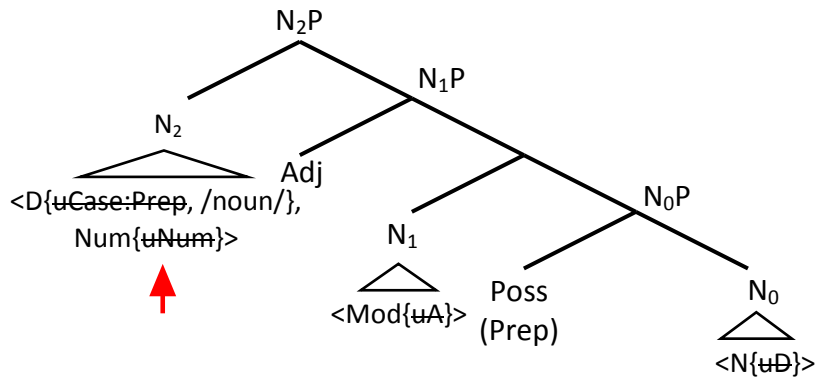
Let us also consider the derivation of a Celtic Construct State nominal phrase that includes an adjective. The structure in (54) demonstrates the relevant derivation. Note that the functional feature Num does not have an uninterpretable uX sub-feature in this structure, in contrast to the Semitic Construct State in (53) above. Nonetheless, Num will undergo Feature Scattering from N<sub>1</sub> to N<sub>2</sub>, because of the uninterpretable uNum sub-feature that needs to c-command the adjective. Note also that there is no need for any further scattering operations in this structure, since D{uCase} already c-commands the possessor at the N<sub>2</sub>P level. Therefore, the final position of the phonological features of the noun in the Celtic Construct State is practically determined by the intermediate feature Num{uNum}.

(54) Celtic Construct State (with an adjective)



Finally, the derivation of a Semitic or Celtic Free State nominal phrase that includes an adjective will be identical to the derivation of the Celtic Construct State above. For reference, I demonstrate the resulting structure in (55). Note that this structure corresponds to an indefinite Free State nominal, which does not have a determiner. I will briefly address the derivation of a definite Free State nominal phrase below.

(55) Semitic and Celtic Free State (with an adjective)



To sum up the above analysis, the derivations in (47)-(49) and (53)-(55) demonstrate that the scattering of  $D\{uCase\}$  (which has to appear in a position c-commanding the possessor) and  $Num\{uNum\}$  (which has to appear in a position c-commanding the adjective) have a cumulative effect on the surface position of the phonological features of the noun. The effects of the former are demonstrated most clearly in nominal phrases that lack an adjective (i.e. (47)-(49)). Similarly, the effects of the latter are most obvious in nominal phrases where the noun is immediately followed by the adjective (i.e. the Celtic Construct State in (54) and the Free State in both Semitic and Celtic languages in (55)). Finally, in a structure with a [N Poss Adj] word order (i.e. a Semitic Construct State nominal phrase with an adjective as in (53)), the phonological features of the noun are displaced in two steps, firstly by the scattering of  $Num\{uNum\}$  and subsequently by the scattering of  $D\{uCase\}$ .

Note, however, that all of the structures in (47)-(49) and (53)-(55) have been derived from numerations that do not include a lexical determiner. Thus, the above analysis does not provide any insight on the distribution of determiners in Semitic and Celtic nominal phrases. I will remedy this omission in the following paragraphs.

To begin with, remember that the distribution of the lexical determiner within the Free State correlates with the in/definite interpretation of the nominal phrase (cf. (27a)). Specifically, a definite nominal requires the determiner to be present, while an indefinite nominal requires it to be absent. The second part of this generalisation can be captured rather straightforwardly by the assumption that the

determiner is lexically specified as definite. Therefore, the lexical determiner is incompatible with an indefinite nominal phrase. However, the first part of this generalisation requires a more nuanced approach. Note that the hypothesis that the lexical determiner is inherently definite only takes us part of the way: it ensures that a definite nominal phrase that includes a lexical determiner is grammatical, but it does not explain why a definite nominal phrase lacking a lexical determiner is ungrammatical. To address this, consider the following. In Semitic and Celtic languages, in/definiteness is not marked on the noun itself. Therefore, if it were possible for a definite nominal phrase to lack a determiner, it would be indistinguishable from its indefinite counterpart. This suggests that the presence of the definite determiner assists with the parsing of the semantic interpretation of the nominal. I will, in fact, assume that this effect has been grammaticalized. Broadly speaking, I will suggest that nouns in Semitic and Celtic languages are effectively lexically specified as indefinite. To be more precise, remember that I have proposed that functional features and their sub-features are assigned to the featural specification of lexical items in the numeration. Therefore, the numeration of a Free State nominal will include a functional feature D which has an interpretable Def:+/- sub-feature. If this sub-feature is marked with a positive value (i.e. Def:+) the numeration will produce a definite nominal phrase and, conversely, if the sub-feature is marked with a negative value (i.e. Def:-) the numeration will produce an indefinite nominal phrase. Furthermore, I propose that the lexical properties of nouns and determiners prevent the former from acquiring a Def:+ value (at least in the Free State) and the latter from acquiring a Def:- value. Consequently, a licit numeration for a definite Free State nominal has to include the lexical determiner, because a functional feature D with a Def:+ sub-feature cannot be assigned to the noun. Similarly, a licit numeration for an indefinite Free State nominal must not include a determiner, since a functional Feature D with a Def:- sub-feature cannot be assigned to the determiner.

Let us also compare the distribution of the lexical determiner in Construct State versus Free State nominal phrases. In the Construct State, the lexical

determiner is necessarily absent regardless of the semantic interpretation of the nominal (cf. (27b)). Thus, it seems unlikely that the analysis of the Free State could be extended to Construct State nominal phrases. Instead, I will approach this case in the same manner as the obligatory absence of an overt complementizer in German embedded V2 clauses (cf. section 2.3.3 in chapter 2). Remember that the absence of a complementizer in that context was attributed to the hypothesis that its phonological features and the phonological features of the verb have to be hosted by the same functional feature, which leads to an insurmountable conflict in a numeration that contains both of these lexical items. Specifically, such a numeration cannot produce a licit featural specification for both the complementizer and the verb, because the aforementioned functional feature can only be assigned to one of them. Therefore, the only acceptable numeration under these circumstances is the one that excludes the overt complementizer. The obligatory absence of a lexical determiner in the Construct State can be captured in the same manner, if we assume that the phonological features of the determiner and the noun are hosted by the same functional feature. However, note that the placement of the phonological features of the noun has to be different in the Construct State (where the lexical determiner is always absent) versus the Free State (which is, in principle, compatible with a lexical determiner). To achieve this, it is necessary to posit that the nominal sequence of functional features contains two features above Num. Let us tentatively call them  $D_1$  and  $D_2$ , where  $D_2$  is the highest feature in the sequence. Then, we can assume that  $D_2$  hosts the phonological features of a lexical determiner and a Construct State noun, while  $D_1$  hosts the phonological feature of a Free State noun. Consequently, the lexical determiner is (in principle) compatible with the numeration of a Free State nominal phrase, but not with the numeration of a Construct State nominal.

Before concluding this section, I will provide a brief sketch of an analysis of the in/definiteness harmony observed in the Semitic Construct State and (partially) the Celtic Construct State (cf. (28)). In the previous discussion, I have attributed the movement of the noun to uCase (amongst other features). If this proposal is on the

right track, then the in/definite interpretation of the noun in the Construct State is not directly linked to its movement. I believe that this is not a negative result considering that the differences in the in/definite interpretation of Semitic versus Celtic Construct State nominal phrases do not correlate with any differences in the movement of the noun in the same structures. However, it should be abundantly clear that the proposed analysis of the Semitic and Celtic nominals predicts that there is a strong link between D (which is usually assumed to be the host of in/definiteness features) and the possessor. So far, I have argued that this link relates to case assignment, but it would not be far-fetched to assume that the in/definite features of D and the possessor in the Construct State are also linked, independently of uCase.

Building on the discussion on the distribution of the lexical determiner, definiteness harmony in the Semitic Construct State could be captured by assuming that D has an interpretable Def sub-feature, which receives an appropriate positive or negative value by establishing a relation with the possessor in narrow syntax. Conversely, the absence of any in/definiteness harmony in the Free State could be captured by the assumption that, in this case, the interpretable sub-feature Def receives its value in the numeration. Finally, the partial in/definiteness harmony of the Celtic Construct State could be captured as a combination of the previous two cases. Specifically, I suggest that the interpretable sub-feature Def in the Celtic Construct State will either receive a positive value or no value at all, during the numeration. A negative value can only arise when Def establishes a relation with an indefinite possessor in narrow syntax. Therefore, there can be four possible combinations in the numeration: (i) a definite possessor and an unmarked noun, (ii) a definite possessor and a definite noun, (iii) an indefinite possessor and an unmarked noun, and (iv) an indefinite possessor and a definite noun. In the first case, the interpretation of the noun will be determined by the definiteness of the possessor. In the second case, the interpretation of the noun is independent from, but nonetheless harmonic with, the definiteness of the possessor. In the third case, the interpretation of the noun will be determined by the indefiniteness of the



possessor. In the fourth case, the interpretation of the noun is independent from, and additionally not harmonic with, the indefiniteness of the possessor. Thus, a definite possessor will always be coupled with a definite noun (cf. (i-ii)), while an indefinite possessor may appear either with a definite or an indefinite noun (cf. (iii-iv)).

To summarise, in this section I have developed an analysis of Semitic and Celtic nominal phrases. Broadly speaking, I have proposed that the movement of the noun to an initial position in the relevant structures is the result of one or more Feature Scattering operations that are triggered by type-1 uninterpretable sub-features (i.e. uninterpretable features that must c-command a matching interpretable feature in order to be eliminated). To be more precise, I have attributed the aforementioned Feature Scattering operations to the uninterpretable sub-features *uNum* and *uCase*, which are responsible for the displacement of the phonological features of the noun across adjectives and (genitive or prepositional) possessors respectively. Additionally, I have proposed an account for the distribution of the lexical determiner in Semitic and Celtic nominal phrases and I have provided a tentative analysis of the (full or partial) in/definiteness harmony observed in the Semitic and Celtic Construct State. In the following section, I will extend the analysis of head movement in noun-initial nominal phrases to the comparable case of verb-initial clauses in the same language families.

#### ***4.3.3 A Feature Scattering Analysis of Verb-Initial Languages***

One of the fundamental assumptions in this dissertation is that Feature Scattering takes place to provide an appropriate configuration for the elimination of uninterpretable sub-features. In section 4.3.1, I proposed that there are two types of uninterpretable features. Type-1 features are eliminated in a configuration where they c-command a matching interpretable feature, while type-2 features are eliminated when they are c-commanded by a matching interpretable feature. Furthermore, in the same section I suggested that type-1 features can account for head movement to an initial position, under specific circumstances. In section 4.3.2,

I discussed the implementation of that proposal in the case of Semitic and Celtic nominal phrases. In this section, I will return to the clausal domain, and I will explore how this theory fares in the analysis of the VSO word order of Semitic and Celtic languages. I will start the section with a brief overview of the literature on Semitic and Celtic VSO. In particular, I will focus on the issue of the surface position of the verb and the subject in the relevant structures. After that introductory discussion, I will present a Feature Scattering analysis of verb movement in VSO clauses in Semitic and Celtic languages.

As mentioned in section 4.2.1, the verb movement analysis of VSO enjoys relatively widespread acceptance. Nonetheless, the details of the relevant syntactic structure remain controversial. In particular, there are two main points of contention: the surface position of the verb and the surface position of the subject. As we will see in the following discussion, these two questions intersect with each other in a significant way.

Regarding verb movement, the earlier literature in VSO languages was divided in two camps. Namely, the movement of the verb was treated either as an instance of V-to-C movement (Emonds 1980, Stowell 1989, Doherty 1996, amongst others) or as an instance of V-to-T movement (Sproat 1985, Guilfoyle 1990, McCloskey 1991, Koopman and Sportiche 1991; also Mohammad 1989 and Fassi Fehri 1989 for Arabic). A major point of debate between these two views is the extent to which verb-initial languages can be compared to V2 languages. Insofar as both language groups seem to have some structural position that is reserved for the finite verb of the clause, the comparison between them is justified and the V-to-C hypothesis seems rather appealing. On the other hand, it would appear that the comparison falls apart when looking at embedded clauses: Celtic languages do not display any restriction on embedded VSO that could be comparable to the restricted distribution of V2 in embedded clauses in most (but not all) Germanic languages. Since the restricted distribution of embedded V2 has traditionally been tied to V-to-C movement, extending the V-to-C analysis to verb-initial languages seems spurious. However, both of these arguments are undermined by the contemporary

understanding of the complexity of embedded V2. For example, the case of Icelandic, where embedded V2 co-occurs with an overt complementizer, suggests that a unitary V-to-C analysis is insufficient to capture the range of V2 phenomena across the Germanic languages. Therefore, comparing V2 and verb-initial languages in terms of the landing site of the verb does not appear to be a particularly productive exercise.

A more substantial source of evidence for the movement operations involved in the derivation of VSO in Celtic languages comes from the linearization of the verb and the subject in relation to various kinds of adverbs. McCloskey (1996a) observes that verbs in Irish appear on the right of TP adjoined adverbs, as shown in (56) below. This piece of evidence suggests that the verb does not move any higher than the TP, contrary to the V-to-C hypothesis. Additionally, McCloskey (1996b) demonstrates that the subject in Irish appears on the left of VP adjoined adverbs (cf. example (57)). Similarly, Koopman and Sportiche (1991) observe that subjects in Welsh appear to the left of the negative marker (cf. example (58)). Examples (57) and (58) suggest that the subject moves outside of the VP and (presumably) surfaces on the specifier of T, which goes against the V-to-T hypothesis.

- (56) Deiridis            an chead   Nollaig   eile   go   dtiocfadh  
 they-used-to-say   the first   Christmas   other   prt   would-come  
 se anios.  
 he up  
 'They used to say that next Christmas he would come up'  
Irish; McCloskey (1996a)

- (57) Deireann   siad   i gconai   paidir   roimh   am   lui.  
 say            they always   prayer   before   time   lie  
 'They always say a prayer before bed-time'  
Irish; McCloskey (1996b)

- (58) Welodd   Emrys   ddim   draig.  
 saw        Emrys   Neg   dragon  
 'Emrys didn't see a dragon'  
Welsh; Koopman and Sportiche (1991)

Apart from challenging both the V-to-C and the V-to-T hypotheses, the examples in (56) and (57)-(58) have a deeper implication. They suggest that an

analysis where there is a single functional head in the structural space between the CP and the VP is insufficient. Rather, it seems that the structure of Celtic VSO languages requires at a minimum two functional heads in the middle field: the lower of the two accommodates the moved subject in its specifier and the higher of the two serves as the landing site for the moved verb itself. Different versions of this approach, usually making very different assumptions about the category of the relevant functional heads, have been pursued by Hendrick (1991), Rouveret (1991), Bobaljik and Carnie (1996) and McCloskey (1996b, 2017) (amongst others). In the following discussion, I will examine this general approach from the perspective of a re-projective theory of head movement. For the most part, I will be interested in the mechanics of the two relevant movement operations and I will not add anything significant on the topic of the category of the functional heads that host the moved verb and subject. For concreteness, I will adopt the proposal of Rouveret (1991) and Bobaljik and Carnie (1996), who suggest that the subject moves to the specifier of T, while the verb moves to an Agr (or Agr<sub>s</sub>) head above T.

Of course, in the current framework, the hierarchical relations of the verb, the subject and the functional heads that are associated with their movement, are all to be expressed in terms of the sequence of functional features. Specifically, I will adopt the verbal sequence in (59), where the functional feature Agr is situated between C and T.

- (59) Verbal sequence of functional features:  
 <C{...}, Agr{...}, T{...}, Neg{...}, Mod{...}, v{...}, V{...}>

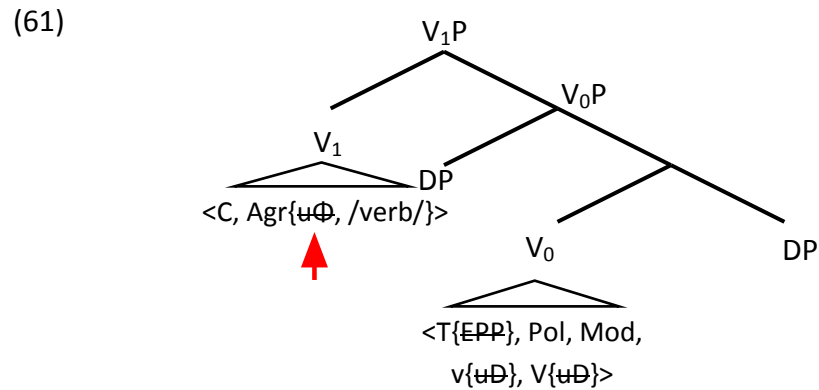
Furthermore, I will assume that the functional features T and Agr are endowed with appropriate sub-features that trigger the movement of the subject to the specifier of T and the movement (or scattering) of the verb to Agr. The movement of the subject can be attributed to the usual EPP sub-feature of T. Note that, according to the analysis developed in this chapter, EPP is a type-2 feature that must be c-commanded by the subject in order to be eliminated. Conversely, the scattering of Agr (and the movement of the verb to it) should be attributed to a type-1 uninterpretable sub-feature that needs to appear in a configuration where it

c-commands a matching interpretable feature. This is a necessary conclusion, considering that the specifier of Agr remains empty in Celtic and Semitic VSO clauses. Moreover, we have to assume that Agr's sub-feature has to be eliminated against the subject of the clause. If this sub-feature could be eliminated against any other constituent lower down in the structure, then the scattering of Agr would not be necessary and, therefore, it would not be possible to derive the desired verb-initial order. This is a significant observation when considering the identity of the relevant uninterpretable sub-feature. What we are looking for is a syntactic relation that holds between a verbal functional head (in the middle field) and the subject. Two possibilities (that are in fact not mutually exclusive) immediately come to mind: subject-verb agreement and case-assignment. For concreteness, I will assume that the uninterpretable sub-feature in question is related to agreement and I will annotate it as Agr{uΦ}, although this is not to deny the possibility that the same derived functional head is also involved in case-assignment (cf. Rouveret 1991). Finally, I will assume that the phonological features of the verb also appear as sub-features of Agr, so that the surface position of the verb will be dictated by the Feature Scattering operation triggered by uΦ. Putting everything together, I propose that the featural specification of verbs in Celtic and Semitic languages is as shown in (60) below.

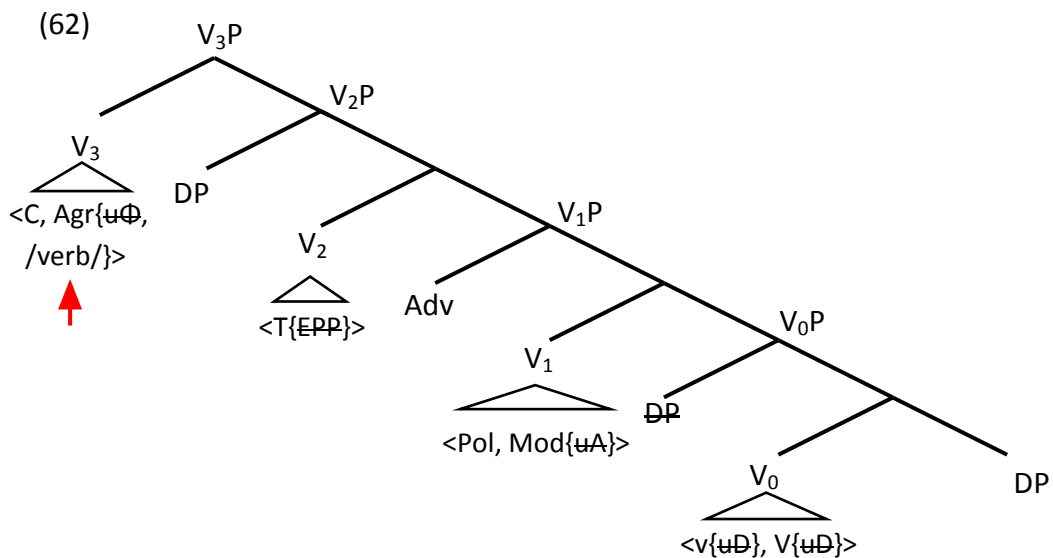
- (60) a. Lexical verb (before Feature Bundling)  
 <Agr{..., /verb/}, ... , V{...}>  
 b. Lexical verb (after Feature Bundling)  
 <C, Agr{uΦ, /verb/}, T{EPP}, Pol{(uNeg)}, Mod{(uA)}, v{uD}, V{uD}>

Note that the featural specification in (60) can derive more than one structure, depending on the numeration. The simplest case is a numeration that does not include a negative marker or adverbial phrases. The structure in (61) represents the result of this derivation. As expected, the object and subject of the clause are merged with V<sub>0</sub>, in order to eliminate the uninterpretable sub-features of V and v respectively. Furthermore, the uninterpretable EPP sub-feature of T can also be eliminated at the V<sub>0</sub>P level, since the subject is already in its specifier.

Subsequently, the type-1 uninterpretable  $u\Phi$  sub-feature of Agr will trigger the scattering of the functional head  $V_1$ , since  $u\Phi$  does not c-command the subject at the  $V_0P$  level. Therefore, the phonological features of the verb, which are also hosted by Agr, will surface at the position of  $V_1$  (i.e. in a clause-initial position).



By comparison, a derivation that includes an adverbial phrase or negative marker will consist of a few additional steps. For example, consider the structure in (62), which includes an adverb. In this case, the uninterpretable  $uA$  sub-feature of Mod needs to be eliminated before EPP. This will result in the scattering of  $V_1$ , since  $V_0P$  cannot accommodate the adverb. However, at the  $V_1P$  level, the Specifier-Head relation between EPP and the subject has been disturbed. Therefore, Feature Scattering will take place for a second time, creating the functional head  $V_2$ . Furthermore, the scattering of  $V_2$  will be accompanied by the phrasal movement of the subject to the specifier of  $V_2P$ . At this point, the derivation of (62) converges with the derivation of (61) above. Considering that the uninterpretable  $u\Phi$  sub-feature of Agr is located in  $V_2$  during this intermediate step of the derivation, it does not c-command the subject in its specifier. Thus, the functional feature Agr and its sub-features (including the phonological features of the verb) will be scattered to a third functional head  $V_3$ . Consequently, the verb will surface, once again, in a clause-initial position.



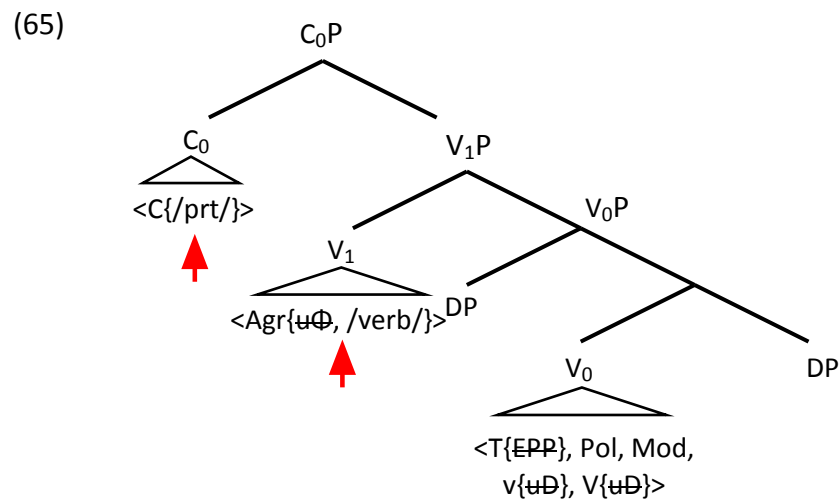
To conclude the discussion of Celtic and Semitic VSO, we also need to address the issue of the preverbal particles that are commonplace in verb-initial languages. The relevant literature presents a vast array of proposals about the category of those particles. For example, Hendrick (1988), Willis (1998) and Bury (2003, 2005) take preverbal particles to be complementizers, Chung and McCloskey (1987), McCloskey (1996a) and Sadler (1988) treat them as a combined C and Infl head, while Duffield (1990) takes it a step farther and suggests that different particles are realizations of separate C, T, Agr and Neg heads. The common thread amongst these proposals is that preverbal particles are treated as independent functional heads that attract the finite verb of the clause<sup>8</sup>. For the purposes of this analysis, I will adopt the simplest amongst these proposals. Namely, I will assume that preverbal particles are lexical complementizers. However, I will not adopt the view that the verb moves to the position of the preverbal particle. As I have argued in section 4.2, I consider head movement of one lexical head to another to be superfluous in, or even incompatible with, a re-projective theory of functional structure. Rather, I will assume that the verb surfaces in a derived functional head

<sup>8</sup> It is worth mentioning, however, that there are also proposals that treat preverbal particles as a part of a complex verb that is constructed prior to narrow syntax. Timm (1988) and Stump (1988) pursue this approach for Breton. Stump (1988) in particular, claims that this is the case only for affirmative particles, while he treats negative particles as complementizers.

that appears immediately below the position of the preverbal particle in the functional structure. Specifically, I will propose that preverbal particles are specified as shown in (63), when in the lexicon. (63) states that Feature Bundling must assign the functional feature C to the preverbal particle during the numeration. Furthermore, C must host the phonological features of the particle. Therefore, in a numeration that includes a verb as well as a preverbal particle, Feature Bundling will assign the featural specifications in (64) to these two lexical items. Finally, the specifications in (64) will derive the structure in (65). This structure demonstrates that the preverbal particle is directly merged with a V<sub>1</sub>P, which will be derived in exactly the same manner as in (61) above. Thus, the verb will surface in a clause-initial position and immediately preceded by the particle.

(63) Lexical complementizer (before Feature Bundling)  
 <C{..., /prt/}>

(64) a. Lexical complementizer (after Feature Bundling)  
 <C{/prt/}>  
 b. Lexical verb (after Feature Bundling)  
 <Agr{uΦ, /verb/}, T{EPP}, Pol{(uNeg)}, Mod{(uA)}, v{uD}, V{uD}>





To summarise, in this section I have proposed an analysis of verb movement to a clause-initial position in Semitic and Celtic VSO languages. This analysis is based on the hypothesis that, in the relevant structures, Feature Scattering takes place in order to create a configuration where an appropriate uninterpretable feature (namely, the  $u\Phi$  sub-feature of Agr) c-commands the subject. Thus, verb-initial clauses and noun-initial nominal phrases in Semitic and Celtic languages are derived in a fundamentally similar way. Furthermore, this approach succeeds in incorporating the analysis of verb movement to a clause-initial position within the overarching Feature Scattering theory of functional structure.

#### **4.4 Conclusion**

In this chapter, I argued against the view that the sole function of re-projection is the creation of specifier positions (Giorgi and Pianesi 1997, Bury 2003). Specifically, in section 4.2, I suggested that the structure of Semitic and Celtic VSO clauses and, especially, noun-initial nominal phrases offers compelling evidence in favour of a less restrictive version of re-projection. In section 4.3.1, I proposed an amendment to the concept of Feature Scattering developed in chapter 1 to achieve this goal. Namely, I suggested that Feature Scattering may be triggered either to create a configuration where an uninterpretable feature c-commands a matching interpretable feature (i.e. a type-1 feature), or a configuration where the uninterpretable feature is c-commanded by a matching interpretable feature (i.e. a type-2 feature).

In section 4.3.2, I demonstrated that this hypothesis provides a principled account of the movement of a noun to an initial position within Semitic and Celtic nominal phrases. However, due to the intricacies of the distribution of the head noun, the possessor noun and the (optional) adjective, I have proposed that noun movement in the relevant structures is the cumulative effect of not one but two type-1 uninterpretable features (i.e.  $uCase$  and  $uNum$ ). Furthermore, in section 4.3.3, I have demonstrated that this analysis can be straightforwardly extended to the movement of the verb to a clause-initial position in the same languages. Thus, I

have arrived at a uniform analysis of head movement in the clausal and the nominal domain in Semitic and Celtic languages.



In this dissertation, I developed a new framework of functional structure that dynamically determines how many, or how few, functional heads must be projected in any given derivation to capture the syntactic relations that are relevant to it. This proposal builds on a strand of research that (partly) attributes covert functional structure to the re-projection of lexical or functional heads (e.g. Ackema et al 1993, Giorgi and Pianesi 1997, Nash and Rouveret 1997, Koenenman 2000, Bury 2003, and others). However, this framework is more radical than its predecessors in so far as it derives the entirety of covert functional structure via re-projection, in order to completely eliminate null and syntactically inert functional heads from any given structure.

The fundamental mechanism of this framework is based on two hypotheses. Firstly, I suggested that lexical items have more expansive and more structured featural specifications than previously assumed. Specifically, I proposed that lexical items consist of a sequence of functional features  $fnF$  that each bears one or more un/interpretable sub-features  $(u)F$ , as shown in (1).

(1) Lexical head:  $X_0 \langle fnF_n \{(u)F_z, \dots\}, \dots, fnF_1 \{(u)F_y, \dots\}, fnF_0 \{(u)F_x, \dots\} \rangle$

Additionally, I suggested that functional features and their sub-features serve a different purpose in the derivation. Specifically, I proposed that syntactic operations are triggered by the uninterpretability of sub-features (cf. Chomsky 1995), while the order of those operations is determined by the sequence of functional features (cf. Giorgi and Pianesi's 1997 Universal Ordering Constraint).

Secondly, I adopted a modified version of Giorgi and Pianesi's (1997) Feature Scattering. Following Giorgi and Pianesi, I define Feature Scattering as a last resort operation that creates new functional heads by displacing uninterpretable features that cannot establish an appropriate syntactic relation in their original position. In addition, I propose that Feature Scattering does not affect individual features, but entire segments of the featural specification of lexical items, as described in (2).

## (2) Feature Scattering

Assume that at a given (intermediate) step of a derivation, an uninterpretable feature  $uG$  of a head  $X_0$  - where  $X_0$  has the featural specification in (i) - cannot establish the necessary relation with a matching interpretable feature.

$$i. X_0 < fnF_n\{(u)F_z, \dots\}, \dots, fnF_m\{uG, \dots\}, \dots, fnF_1\{(u)F_y, \dots\}, fnF_0\{(u)F_x, \dots\} >$$

Then Feature Scattering will strip a segment of the featural specification of  $X_0$ , leaving  $X_0$  with the altered featural specification in (ii) and creating a new functional head  $X_1$  with the featural specification in (iii) that merges with  $X_0P$ .

$$ii. X_0 < \dots, fnF_1\{(u)F_y, \dots\}, fnF_0\{(u)F_x, \dots\} >$$

$$iii. X_1 < fnF_n\{(u)F_z, \dots\}, \dots, fnF_m\{uG, \dots\} >$$

The interaction of (1) and (2) entails that Feature Scattering takes place in a recursive manner that can potentially derive multiple functional heads (which will abide to a presumably universal hierarchical order) from a single lexical item. Therefore, covert functional structure can be captured entirely by this mechanism.

In addition, I have argued that null functional heads ought to be attributed to Feature Scattering, regardless of whether their specifier is empty or not (pace Giorgi and Pianesi 1997 and Bury 2003). To this end, I proposed that Feature Scattering may be triggered by two types of uninterpretable features that have to appear in different configurations to establish the necessary syntactic relations, as described in (3). According to this hypothesis, only type-2 features establish a Specifier-Head relation, while type-1 features establish a relation with some constituent inside their complement.

- (3) a. Type-1 uninterpretable features:  
An uninterpretable feature  $uF_1$  of type-1 that is part of the featural specification of a (derived or non-derived) head H can be eliminated iff H c-commands an XP with a matching interpretable feature.
- b. Type-2 uninterpretable features:  
An uninterpretable feature  $uF_2$  of type-2 that is part of the featural specification of a (derived or non-derived) head H can be eliminated iff H is c-commanded by an XP with a matching interpretable feature.

Finally, I have argued that Feature Scattering is a fundamental structure-building operation that subsumes head movement (pace Ackema et al 1993 and Bury 2003, who assume the opposite relation between these two operations). To incorporate head movement in the overall analysis of functional structure, I have simply proposed that the phonological features of a lexical item are also part of its featural specification. This hypothesis entails that independently triggered Feature Scattering operations may inadvertently displace the phonological features of a lexical item.

This proposal, in conjunction with (1)-(3), has provided an account of a wide variety of structures involving head movement. In chapter 2, I focused on instances of head movement that appear to be linked to a Specifier-Head relation, which I attributed to the scattering of type-2 features. Amongst other cases, I addressed the movement of the verb to C, T or v in Germanic V2 languages, Romance SVO languages and English respectively. In chapter 4, I discussed some instances of head movement to an absolute initial position, which I attributed to the scattering of type-1 features. Specifically, I examined the movement of the verb in verb-initial clauses and the movement of the noun in noun-initial nominal phrases in Semitic and Celtic languages.

In chapter 3, I discussed the separate issue of subject/non-subject asymmetries in English wh-questions and embedded clauses involving wh-extraction. I demonstrated that the subject version of those clauses is predicted to have a smaller functional structure than their non-subject counterparts. Therefore, I suggested that it is this structural asymmetry that affects Subject Auxiliary Inversion

and *do*-support in *wh*-questions and the distribution of the overt complementizer *that* in embedded clauses.

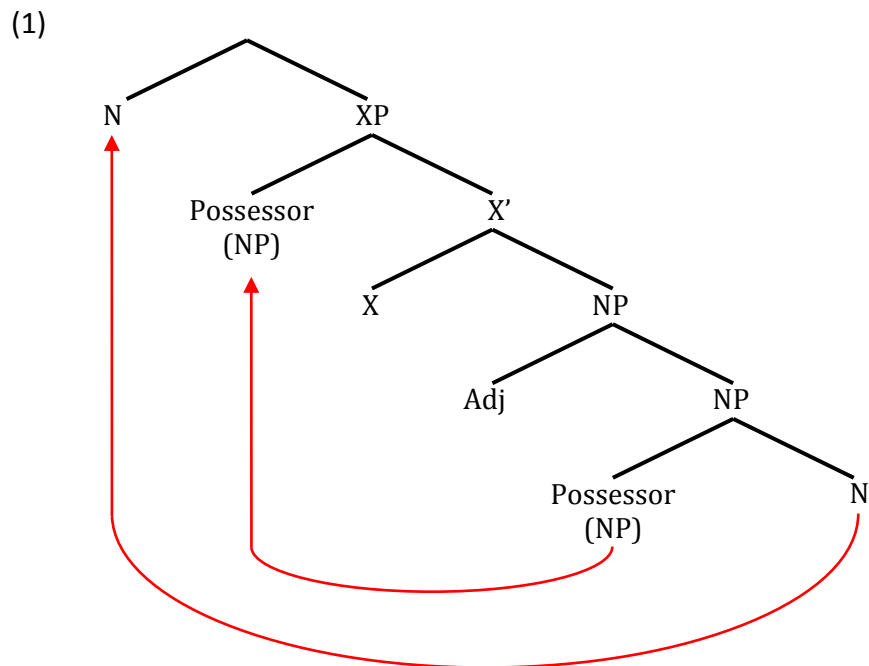
In conclusion, the diverse syntactic phenomena I addressed throughout this dissertation demonstrate that the proposals in (1)-(3) provide a theory of functional structure that can capture the potentially complex and non-local syntactic relations of any given structure, while completely eliminating any covert functional head that does not contribute to the pertinent syntactic relations.

## Appendix

---

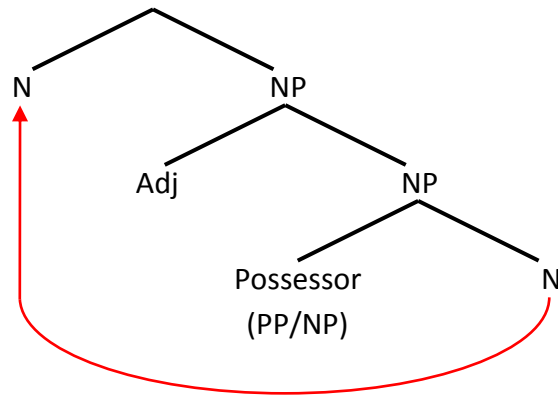
In this appendix, I outline two problematic aspects of the analysis presented in section 4.3.2 of chapter 4, the resolution of which I will leave for future work. The first has to do with the validity of a uniform analysis of Semitic and Celtic nominal phrases. The second has to do with the validity of a head movement analysis of nominal phrases more broadly.

In section 4.3.2, I assumed that the basic structure of Semitic Construct State nominals is (1), while the basic structure of Semitic Free State nominals, as well as all Celtic nominal phrases, is (2) (both repeated from chapter 4).





(2)



However, there is good reason to believe that the (surface) position of a genitive possessor in Celtic nominal phrases is much higher than what is suggested in (2). Consider the following. According to Adger (2013), the genitive possessor in Gaelic appears to the left of the universal quantifier *uile* ‘all’, an element that can be independently shown to occupy a relatively high position in the nominal phrase (i.e. above demonstratives; see Adger 2013 for details). Example (3) illustrates this point. Under most analyses (including the one developed in chapter 4), this distribution suggests that the genitive possessor is structurally higher than *uile*.

(3) Cheannaich mi dealbhan Sheumais uile de Mhàiri.  
buy.past I pictures Seumas.gen all of Mhàiri  
‘I bought all Seumas’s pictures of Mhàiri.’

Furthermore, Adger (2013) demonstrates that Gaelic genitive possessors can bind into a PP complement of the noun, but not vice versa, indicating that the genitive possessor asymmetrically c-commands the PP complement (as we also saw in Semitic). To demonstrate, compare (4a) (where the reciprocal in the PP complement is bound by the genitive possessor) with (4b) (where the reciprocal genitive cannot be bound by the PP complement).

- (4) a. dealbhan nan caileagan de a cheile  
 pictures the.gen.pl girls of each-other  
 'the girls pictures of each other'  
 b. \*dealbhan a cheile dhe na caileagan  
 pictures each-other of the.pl girls  
 'each other's pictures of the girls'

This observation is not particularly instructive on its own. However, Adger (2013) argues that 'complement' PPs in Celtic languages do not surface in the position of the complement of the noun, but in some structural position outside the NP<sup>1</sup>. This can be seen in examples involving coordination, as in (5). Here, adjectives and numerals appear inside the two conjuncts, while the PP complement appears outside of them.

- (5) Cheannaich e [còig dealbhan mòra] agus [trì iomhaighean  
 buy.past he [five pictures big] and [three images  
 beaga] de Mhàiri.  
 small] of Màiri  
 'He bought five big pictures and three small images of Màiri.'

Therefore, the combination of (4) and (5) provides another piece of evidence for the position of the genitive possessor being much higher than what I assumed in (2). Thus far, these observations are not especially problematic for the analysis proposed in chapter 4. The surface position of the genitive possessor in Celtic languages could be attributed to a movement operation that places it in some appropriate structural position outside the NP, but below the surface position of the noun. However, remember that possessors in Celtic languages appear on the right of adjectives. Therefore, if the surface position of Celtic genitive possessors is as high as Adger (2013) suggests (i.e. above demonstratives and the universal quantifier), phrasal movement of the genitive possessor does not suffice to account for its position in relation to adjectives (which under any reasonable assumption, do not appear higher than demonstratives). At the very least, we would have to

---

<sup>1</sup> In fact, Adger (2013) arrives at a much stronger conclusion. He suggests that the PPs that are traditionally treated as complements of the noun are universally base merged outside the NP. I will briefly return to this point towards the end of this appendix.

assume that adjectives branch to the right, as suggested by Willis (2006) for Welsh. Finally, note that, although a solution to this problem within the confines of the analysis proposed in chapter 4 is not inconceivable, this situation suggests that the analysis of Celtic nominals is much less similar to the analysis of Semitic nominals than what I have assumed. At this point, it is not clear to me how these discrepancies could be resolved. I will leave further discussion of this issue for future research.

Let us now move on to another issue that merits discussion. In section 4.3.2 of chapter 4, I focused on a head movement analysis of nominal phrases in Semitic and Celtic languages. There are, however, some important alternatives that should be acknowledged. What is of interest here is Cinque's (2005) analysis of Greenberg's (1963) Universal 20 on the distribution of demonstratives, numerals and adjectives within the nominal phrase. According to Greenberg, these elements appear in that order when they are prenominal, while they appear either in the same or in the reverse order when they are postnominal. Similar observations have been made with respect to the distribution of various classes of adjectives (e.g. adjectives of quality, size, shape, colour, origin, etc) in relation to each other. There is a certain degree of regularity in the order of prenominal adjectives, which is either kept or reversed when adjectives are postnominal (Sproat and Shih 1991, Cinque 1994, Scott 2002, Laenzlinger 2005). It should be noted that Semitic languages display the 'reverse' order both in the distribution of determiners, numerals and adjectives, and in the distribution of classes of adjectives (Fassi Fehri 1999, Cinque 2000, Shlonsky 2004). With respect to the former, although demonstratives and numerals are not uniformly postnominal in the various Semitic languages and dialects, when they are postnominal they appear in the reverse order, as shown in the Standard Arabic example in (6). With respect to the latter, Semitic languages systematically display the reverse order, as shown in the Hebrew examples in (7) where the order of the Hebrew adjectives contrasts with the order of the adjectives in the English translation. Finally, Celtic nominal phrases also display the reverse order of demonstratives and adjectives, both of which appear

postnominally, as shown in the Gaelic example in (8). Numerals, however, appear prenominally (Duffield 1991, Rouveret 1994)<sup>2</sup>.

- (6)  $\text{\textasciixu}\text{-u}$                        $\text{l-jadiidat-u}$      $\text{\textasciixu}\text{-\textasciixu}\text{-u}$                        $\text{haa}\text{\textasciixu}$   
the-newspapers-nom    the-new-nom    the-three-nom    these  
‘these three new newspapers’

Fassi Fehri (1999)

- (7) a.  $\text{para}$      $\text{\textsc{s}vecarit}$      $\text{xuma}$   
cow    swiss    brown  
‘a brown swiss cow’  
b.  $\text{ha-\textsc{s}ulxan}$      $\text{ha-\textsc{s}axor}$      $\text{ha-\textsc{a}rox}$   
table            black            long  
‘the black long table’  
c.  $\text{naknikiya}$     ‘ostrit             $\text{k\textsc{s}era}$   
sausage            Austrian            Kosher  
‘a Kosher Austrian sausage’

Shlonsky (2004)

- (8)  $\text{na}$      $\text{dealbhan}$      $\text{snog}$      $\text{ud}$   
the    pictures    nice    that  
‘those nice pictures’

Adger (2013)

As the previous description suggests, there is an asymmetry in the distribution of determiners, numerals and adjectives, as well as in the distribution of various classes of adjectives. Namely, the order of those elements is more restricted in a prenominal than a postnominal position. As a matter of fact, Cinque (2005) has demonstrated that the disparity between prenominal and postnominal orders is even bigger than what Universal 20, in its original formulation, suggests. Cinque shows that the first clause of Universal 20 is correct (i.e. the order demonstrative, numeral, adjective is in deed the only attested option in prenominal position), but the second clause is too strong (i.e. apart from the demonstrative, numeral, adjective order and its reverse, there are several other attested permutations in

---

<sup>2</sup> Note, however, that the comparison between Semitic and Celtic nominal phrases breaks down when it comes to the distribution of classes of adjectives. Willis (2006) observes that some Welsh adjectives (e.g. adjectives of size, colour and origin) appear in the order of prenominal adjective languages, while others (e.g. adjectives of quality and age) appear in the reverse order.

postnominal position). Note further that demonstratives, numerals and adjectives do not all have to appear on the same side of the noun, which gives rise to even more possible combinations. These facts create a rather complex empirical puzzle (i.e. there are twenty-four possible orders, out of which fourteen are attested and ten are not, according to Cinque 2005).

Nonetheless, Cinque (1996, 2000, 2005) has proposed an analysis involving roll-up phrasal movement, which successfully derives all attested orders while excluding all unattested orders. In brief, Cinque makes the following assumptions. Firstly, he suggests that demonstratives, numerals and adjectives are the left-ward specifiers of dedicated functional heads, which are part of a universal functional hierarchy. The hierarchy is such that the position of demonstratives is the higher and the position of adjectives the lower of the three elements. Secondly, he suggests that the movement operation may only apply to phrases, not heads. Finally, he suggests that, in the nominal domain, all movement operations must apply to some constituent that contains the noun. For the purposes of this discussion, it is not necessary to demonstrate how these assumptions derive the full range of cross-linguistically attested orders of demonstrative, numeral and adjective. However, let us consider the derivation of the 'reverse' postnominal order that we see in Semitic nominal phrases, to demonstrate the roll-up nature of phrasal movement according to this analysis. Following Cinque (2005), I will call the functional head that accommodates the demonstrative in its specifier  $W$ , the functional head that accommodates the numeral  $X$ , and the functional head that accommodates the adjective  $Y$ . Furthermore, Cinque assumes that above each one of those heads there is an associated Agr head (i.e. an  $Agr_W$ , an  $Agr_X$  and an  $Agr_Y$ ). With that in mind, the suggestion is that the NP moves to the specifier of  $Agr_Y$ , above and to the left of the adjective. Then,  $Agr_Y P$  (which contains the noun and the adjective, in that order) moves to the specifier of  $Agr_X$ , above and to the left of the numeral. Finally,  $Agr_X P$  (which contains the noun, the adjective and the numeral, in that order) moves to the specifier of  $Agr_W$ , above and to the left of the demonstrative. The resulting structure (shown in (9)) displays the order noun,

adjective, numeral, demonstrative (for a detailed discussion of this analysis see Cinque 2005; see also Cinque 2000 and Shlonsky 2004 for an analysis of Semitic nominal phrases under this approach).

- (9) [Agr<sub>w</sub>P [Agr<sub>x</sub>P [Agr<sub>y</sub>P NP [Agr<sub>y</sub> [YP AP [Y tNP]]]]] [Agr<sub>x</sub> [XP NumP [X tAgr<sub>y</sub>P]]]] [Agr<sub>w</sub> [WP DemP [W tAgr<sub>x</sub>P]]]]

As I have mentioned above, this analysis partly relies on the assumption that movement in the nominal phrase always affects phrases, rather than heads. Therefore, to the extent that this approach provides a successful account of the (revised) Universal 20, it poses a serious challenge to a head movement analysis of noun-initial nominal phrases. There is however one residual problem with this approach. Note that apart from excluding head movement, the analysis also excludes remnant movement (i.e. if movement must always apply to a constituent that contains the noun, it is not possible to move the complement or specifier of the noun, thus creating a remnant NP). As pointed out by Fassi Fehri (1999), Pereltsvaig (1996) and Willis (1996), this entails that a PP complement of a noun should be immediately adjacent to it. This prediction is not borne out in either Semitic or Celtic languages, where PP complements usually appear at the end of the nominal phrase (see example (10) from Standard Arabic and (11) from Gaelic, where the PP follows an adjective). Finally, note that removing or weakening the requirement that movement applies to a constituent containing the noun is not an available option, because this requirement plays a large part in constraining the orders of demonstrative, numeral and adjective that can be derived by this analysis (see Abels and Neeleman 2012 for relevant discussion).

- (10) muḥaarabat-u l-ḥukuumat-l l-muntaḍarat-u  
 Fighting-nom the-government-gen the-expected-nom  
 li-l-irtiṣaa?-i  
 of-the-corruption  
 'the expected fighting of the corruption by the government  
 Fassi Fehri (1999)

- (11) an dealbh mòr brèagha de Mhàiri  
The picture big beautiful of Màiri  
'The big beautiful picture of Mhàiri'

Adger (2013)

Another approach to the structure of the nominal phrase that is of interest in connection to the aforementioned problem is that of Adger (2013) (Adger discusses Gaelic in depth, but also draws from a wide range of languages; Kane 2015 extends Adger's analysis to Irish; Sadler 2000 is an earlier analysis of Welsh that reaches partly similar conclusions). Simplifying somewhat, Adger (2013) proposes that the base position of 'complement' PPs, as well as PP and genitive possessors, is not inside the NP. Rather, he proposes that they appear in a functional position outside the NP and above the base position of adjectives. This hypothesis, he argues, captures a cross-linguistic generalisation that he calls the PP Peripherality: in the nominal phrase, PPs and adjectives that appear on the same side of the noun are ordered such that adjectives are closer to the noun than the PP. Interestingly, the underlying structure proposed by Adger (2013) provides a solution to the exact residual problem of Cinque's (2005) analysis of nominal phrases via roll-up phrasal movement. Therefore, a combination of the two seems to be a promising avenue of investigation. Note, however, that the nominal phrase structure proposed by Adger (2013) represents a significant departure from the traditional view regarding the parallelism of clausal and nominal structure. This is an issue that goes beyond the scope of this dissertation. As such, I will not adopt any position on the matter.

## References

---

- Abels, K. and Neeleman, A. (2012). Linear asymmetries and LCA. *Syntax*, 15(1), 25-74.
- Abney, S. (1987). *The English noun phrase in its sentential aspect*. Doctoral dissertation, Massachusetts Institute of Technology.
- Ackema, P., Neeleman, A. and Weerman, F. (1993). Deriving functional projections. In A. Schafer (ed.), *Proceedings of NELS 23* (pp. 17-31). Amherst: GSLA.
- Adger, D. (1996). Aspect, agreement and measure phrases in Scottish Gaelic. In R.D. Borsley and I. Roberts (eds.), *The syntax of the Celtic languages: A comparative perspective* (pp. 200-222). Cambridge: Cambridge University Press.
- Adger, D. (2003). *Core syntax: A minimalist approach*. Oxford: Oxford University Press.
- Adger, D. (2013). *A syntax of substance*. Cambridge, Mass.: MIT Press.
- Aissen, J. (1992). Topic and focus in Mayan. *Language*, 68(1), 43-80.
- Baker, M. (1985). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry*, 16(3), 373-415.
- Baker, M. (1988). *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Belletti, A. (1990). *Generalized verb movement: Aspects of verb syntax*. Torino: Rosenberg & Sellier.
- Besten, H. den. (1983). On the interaction of root transformations and lexical deletive rules. In W. Abraham (ed.), *On the formal syntax of the Westgermania* (pp. 47-131). Amsterdam: John Benjamins.
- Bobaljik, J. (1995). *Morphosyntax: The syntax of verbal inflection*. Doctoral dissertation, Massachusetts Institute of Technology.
- Bobaljik, J. and Carnie, A. (1996). A minimalist approach to some problems of Irish word order. In R.D. Borsley and I. Roberts (eds.), *The syntax of the Celtic languages: A comparative perspective* (pp. 223-240). Cambridge: Cambridge University Press.
- Bobaljik, J. and Thráinsson, H. (1998). Two heads aren't always better than one. *Syntax*, 1(1), 37-71.
- Boeckx, C. and Stjepanović, S. (2001). Head-ing toward PF. *Linguistic Inquiry*, 32(2), 345-355.
- Bresnan, J. (1970). On complementizers: Toward a syntactic theory of complement types. *Foundations of Language*, 6(3), 297-321.
- Bresnan, J. (1977). Variables in the theory of transformations. In P.W. Culicover, T.



- Wasow and A. Akmajian (eds.), *Formal syntax* (pp. 157-196). New York: Academic Press.
- Brody, M. (1997). Mirror theory. Ms., University College London.
- Brody, M. (2000). Mirror theory: Syntactic representations in perfect syntax. *Linguistic Inquiry*, 31(1), 29-56.
- Browning, M.A. (1996). CP recursion and *that-t* effects. *Linguistic Inquiry*, 27(2), 237-255.
- Bury, D. (2003). *Phrase structure and derived heads*. Doctoral dissertation, University College London.
- Bury, D. (2005). Preverbal particles in verb-initial languages. In A. Carnie, S.A. Dooley and H. Harley (eds.), *Verb first: On the syntax of verb-initial languages* (pp. 135-154). Amsterdam: John Benjamins.
- Bury, D. (2007). Verb movement and VSO-VOS alternations. *UCL Working Papers in Linguistics*, 19, 77-91.
- Carnie, A. (1995). *Non-verbal predication and head-movement*. Doctoral dissertation, Massachusetts Institute of Technology.
- Carnie, A., Dooley, S.A. and Harley, H. (2005). *Verb first: On the syntax of verb-initial languages*. Amsterdam: John Benjamins.
- Carnie, A. and Guilfoyle, E. (2000). *The syntax of verb initial languages*. Oxford: Oxford University Press.
- Carstairs-McCarthy, A. (1999). *The origins of complex language: An inquiry into the evolutionary beginnings of sentences, syllables, and truth*. Oxford: Oxford University Press.
- Choe, J.W. (1987). LF movement and pied-piping. *Linguistic Inquiry*, 18(2), 348-353.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1970). Remarks on nominalization. In R.A. Jacobs and P.S. Rosenbaum (eds.), *Readings in English transformational grammar* (pp. 184-221). Waltham, Mass.: Ginn.
- Chomsky, N. (1986). *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1991). Some notes on economy of derivation and representation. In R. Freidin (ed.), *Principles and parameters in comparative grammar* (pp. 417-454). Cambridge, Mass.: MIT Press.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and S.J. Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 1-52). Cambridge, Mass.: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT press.
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (ed.), *Ken Hale: A life in language* (pp. 1-52). Cambridge, Mass.: MIT Press.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1-22.

- Chung, S. (1990). VP's and verb movement in Chamorro. *Natural Language and Linguistic Theory*, 8(4), 559-619.
- Chung, S. (1998). *The design of agreement: Evidence from Chamorro*. Chicago: University of Chicago Press.
- Chung, S. (2006). Properties of VOS languages. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax, volume III* (pp. 685-720). Oxford: Blackwell.
- Chung, S. and McCloskey, J. (1987). Government, barriers and small clauses in Modern Irish. *Linguistic Inquiry*, 18(2), 173-237.
- Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock and R. Zanuttini (eds.), *Paths towards universal grammar: Studies in honor of Richard S. Kayne* (pp. 85-110). Washington, D.C.: Georgetown University Press.
- Cinque, G. (1996). The 'antisymmetric' programme: Theoretical and typological implications. *Journal of Linguistics*, 32(2), 447-464.
- Cinque, G. (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Cinque, G. (2000). On Greenberg's universal 20 and the Semitic DP. *University of Venice Working Papers in Linguistics*, 10(2), 45-61.
- Cinque, G. (2002). *Functional structure in DP and IP: The cartography of syntactic structures, volume 1*. Oxford: Oxford University Press.
- Cinque, G. (2005). Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry*, 36(3), 325-332.
- Cinque, G. (2006). *Restructuring and functional heads: The cartography of syntactic structures, volume 4*. Oxford: Oxford University Press.
- Clemens, L. and Coon, J. (2018). Deriving verb-initial word order in Mayan. *Language*, 94(2), 237-280.
- Clemens, L. and Polinsky, M. (2017). Verb-initial word orders (primarily in Austronesian and Mayan languages). In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax, volume VIII*, 2<sup>nd</sup> edition (pp. 4747-4796). Hoboken, NJ: Wiley Blackwell.
- Coon, J. (2010). VOS as predicate fronting in Chol Mayan. *Lingua*, 120(2), 354-378.
- Culicover, P.W. (1993). Evidence against ECP accounts of the *that-t* effect. *Linguistic Inquiry*, 24(3), 557-561.
- Diesing, M. (1990). Verb movement and the subject position in Yiddish. *Natural Language and Linguistic Theory*, 8(1), 41-79.
- Doherty, C. (1996). Clausal structure and the Modern Irish copula. *Natural Language and Linguistic Theory*, 14(1), 1-46.
- Donati, C. (2006). On *wh*-head movement. In L.L.-S. Cheng and N. Corver (eds.), *Wh-movement: Moving on* (pp. 21-46). Cambridge, Mass.: MIT Press.

- Duffield, N. (1990). Movement and mutation in Modern Irish. *MIT Working Papers in Linguistics*, 12, 31-45.
- Duffield, N. (1991). *Particles and projections*. Doctoral dissertation, University of Southern California.
- Duffield, N. (1996). On structural invariance and lexical diversity in VSO languages: Arguments from Irish noun phrases. In R.D. Borsley and I. Roberts (eds), *The syntax of the Celtic languages: A comparative perspective* (pp. 314-340). Cambridge: Cambridge University Press.
- Emonds, J. (1970). *Root and structure-preserving transformations*. Doctoral dissertation, Massachusetts Institute of Technology.
- Emonds, J. (1976). *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. New York: Academic Press.
- Emonds, J. (1980). Word order in generative grammar. *Journal of Linguistic Research*, 1, 33-54.
- England, N.C. (1991). Changes in basic word order in Mayan languages. *International Journal of American Linguistics*, 57(4), 446-486.
- Fassi Fehri, A. (1989). Generalized IP structure, case and VS word order. *MIT Working Papers in Linguistics*, 10, 75-113.
- Fassi Fehri, A. (1993). *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer.
- Fassi Fehri, A. (1999). Arabic modifying adjectives and DP structures. *Studia Linguistica*, 53(2), 105-154.
- George, L.M. (1980). *Analogical generalization in natural language syntax*. Doctoral dissertation, Massachusetts Institute of Technology.
- Giorgi, A. and Pianesi, F. (1997). *Tense and aspect: From semantics to morphosyntax*. Oxford: Oxford University Press.
- Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (ed.), *Universals of language* (pp. 73-113). Cambridge, Mass.: MIT Press.
- Grimshaw, J. (1991). Extended projections. Ms., Brandeis University.
- Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry*, 28(3), 373-422.
- Grimshaw, J. (2000). Locality and extended projection. In P. Coopmans, M. Everaert and J. Grimshaw (eds.), *Lexical specification and insertion* (pp. 115-133). Amsterdam: John Benjamins.
- Guilfoyle, E. (1988). Parameters and functional projection. In J. Blevins and J. Carter (eds.), *Proceedings of NELS 18* (pp. 193-207). Amherst: GSLA.
- Guilfoyle, E. (1990). *Functional categories and phrase structure parameters*. Doctoral dissertation, McGill University.

- Guilfoyle, E., Hung, H., and Travis, L. (1992). Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory*, 10(3), 375-414.
- Harley, H. (2004). Merge, conflation, and head movement: The first sister principle revisited. In K. Moulton and M. Wolf (eds.), *Proceedings of NELS 34* (pp. 239-254). Amherst: GSLA.
- Hendrick, R. (1988). *Anaphora in Celtic and universal grammar*. Dordrecht: Kluwer.
- Hendrick, R. (1991). The morphosyntax of aspect. *Lingua*, 85(2-3), 171-210.
- Heycock, C. (2006). Embedded root phenomena. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax, volume II* (pp. 174-209). Oxford: Blackwell.
- Holmberg, A. (2015). Verb second. In T. Kiss and A. Alexiadou (eds.), *Syntax - Theory and analysis: An international handbook* (pp. 342-383). Berlin: Mouton de Gruyter.
- Iatridou, S. (1990). About Agr(P). *Linguistic Inquiry*, 21(4), 551-577.
- Ishii, T. (2004). The phase impenetrability condition, the vacuous movement hypothesis, and *that-t* effects. *Lingua*, 114(2), 183-215.
- Jackendoff, R. (1977). *X̄ syntax: A study of phrase structure*. Cambridge, Mass.: MIT Press.
- Jones, M. and Thomas, A.R. (1977). *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press for the Schools Council.
- Kane, F. (2015). The fine structure of the Irish NO. Doctoral dissertation, Ulster University.
- Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Kayne, R. and Pollock, J.-Y. (2001). New thoughts on stylistic inversion. In A. Hulk and J.-Y. Pollock (eds.), *Subject inversion in Romance and the theory of universal grammar* (pp. 107-162). Oxford: Oxford University Press.
- Koenenman, O. (2000). *The flexible nature of verb movement*. Doctoral dissertation, Utrecht University.
- Koopman, H. (1982). Theoretical implications of the distribution of *quoi*. In J. Pustejovsky and P. Sells (eds.), *Proceedings of NELS 12* (pp. 153-162). Amherst: GSLA.
- Koopman, H. (1983). ECP effects in main clauses. *Linguistic Inquiry*, 14(2), 346-350.
- Koopman, H. (1984). *The syntax of verbs: From verb movement rules in the Kru languages to universal grammar*. Dordrecht: Foris.
- Koopman, H. and Sportiche, D. (1991). The position of subjects. *Lingua*, 85(2-3), 211-258.
- Koopman, H. and Szabolcsi, A. (2000). *Verbal complexes*. Cambridge, Mass.: MIT Press.

- Laenzlinger, C. (2005). French adjective ordering: Perspectives on DP-internal movement types. *Lingua*, 115(5), 645-689.
- Larson, R. (1988). On the double object construction. *Linguistic Inquiry*, 19(3), 335-391.
- Lechner, W. (2006). An interpretive effect of head movement. In M. Frascarelli (ed.), *Phases of interpretation* (pp. 45-69). Berlin: Mouton de Gruyter.
- Lee, F. (2000). VP remnant movement and VSO in Quiavini Zapotec. In A. Carnie and E. Guilfoyle (eds.), *The syntax of verb initial languages* (pp. 143-162). Oxford: Oxford University Press.
- Mahajan, A. (2000). Eliminating head movement. *GLOW Newsletter*, 44, 44-45.
- Mahajan, A. (2003). Word order and (remnant) VP movement. In S. Karimi (ed.), *Word order and scrambling* (pp. 217-237). Oxford: Blackwell.
- Massam, D. (2000). VSO and VOS: Aspects of Niuean word order. In A. Carnie and E. Guilfoyle (eds.), *The syntax of verb initial languages* (pp. 97-116). Oxford: Oxford University Press.
- Matushansky, O. (2006). Head movement in linguistic theory. *Linguistic Inquiry*, 37(1), 69-109.
- McCloskey, J. (1983). A VP in a VSO language? In G. Gazdar, E. Klein and G.K. Pullum (eds.), *Order, concord and constituency* (pp. 9-55). Dordrecht: Foris.
- McCloskey, J. (1991). Clause structure, ellipsis and proper government in Irish. *Lingua*, 85(2-3), 259-302.
- McCloskey, J. (1996a). On the scope of verb movement in Irish. *Natural Language and Linguistic Theory*, 14(1), 47-104.
- McCloskey, J. (1996b). Subjects and subject positions in Irish. In R.D. Borsley and I. Roberts (eds.), *The syntax of the Celtic languages: A comparative perspective* (pp. 241-283). Cambridge: Cambridge University Press.
- McCloskey, J. (2017). Ellipsis, polarity, and the cartography of verb-initial orders in Irish. In E. Aboh, E. Haeberli, G. Puskás and M. Schönenberger (eds.), *Elements of comparative syntax: Theory and description* (pp. 99-152). Boston, Mass.: Mouton de Gruyter.
- McCloskey, J. and Sells, P. (1988). Control and A-chains in Modern Irish. *Natural Language and Linguistic Theory*, 6(2), 143-189.
- Mohammad, M., 1988. On the parallelism between IP and DP. In H. Borer (ed.), *Proceedings of WCCFL 7* (pp. 241-254). Stanford, Calif.: CSLI, Stanford University.
- Mohammad, M. (1989). *The sentential structure of Arabic*. Doctoral dissertation, University of Southern California.
- Müller, G. (2004). Verb-second as vP-first. *The Journal of Comparative Germanic Linguistics*, 7(3), 179-234.

- Munaro, N., Poletto, C. and Pollock, J.-Y. (2001). Eppur si muove! On comparing French and Bellunese Wh-movement. In P. Pica and J. Rooryck (eds.), *Linguistic variation yearbook 1* (pp. 147-180). Amsterdam: John Benjamins.
- Nash, L. and Rouveret, A. (1997). Proxy categories in phrase-structure theory. In K. Kusumoto (ed.), *Proceedings of NELS 27* (pp. 287-304). Amherst: GSLA.
- Nash, L. and Rouveret, A. (2002). Cliticization as unselective attract. *Catalan Journal of Linguistics*, 1, 157-199.
- Nilsen, Ø. (2003). *Eliminating positions: Syntax and semantics of sentence modification*. Doctoral dissertation, Utrecht University.
- Ouhalla, J. (1988). *The syntax of head movement: A study of Berber*. Doctoral dissertation, University College London.
- Ouhalla, J. (1991). *Functional categories and parametric variation*. London: Routledge.
- Ouhalla, J. (1994). Verb movement and word order in Arabic. In D. Lightfoot and N. Hornstein (eds.), *Verb movement* (pp. 41-72). Cambridge: Cambridge University Press.
- Pereltsvaig, A. (2006). Head movement in Hebrew nominals: A reply to Shlonsky. *Lingua*, 116(8), A1-A40.
- Pesetsky, D. and Torrego, E. (2001). T-to-C movement: Causes and consequences. In M. Kenstowicz (ed.), *Ken Hale: A life in language* (pp. 355-426). Cambridge, Mass.: MIT Press.
- Pesetsky, D. and Torrego, E. (2004). Tense, case, and the nature of syntactic categories. In J. Guéron and J. Lecarme (eds.), *The syntax of time* (pp. 495-537). Cambridge, Mass.: MIT Press.
- Poletto, C. and Pollock, J.-Y. (2004). On the left periphery of some Romance wh-questions. In L. Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures, volume 2* (pp. 251-296). Oxford: Oxford University Press.
- Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20(3), 365-424.
- Pollock, J.-Y. (2006). Subject clitics and complex inversion. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell companion to syntax, volume IV* (pp. 601-659). Oxford: Blackwell.
- Rackowski, A. (2002). *The structure of Tagalog: Specificity, voice, and the distribution of arguments*. Doctoral dissertation, Massachusetts Institute of Technology.
- Rackowski, A. and Richards, N. (2005). Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry*, 36(4), 565-599.
- Rackowski, A. and Travis, L. (2000). V-initial languages: X or XP movement and adverbial placement. In A. Carnie and E. Guilfoyle (eds.), *The syntax of verb initial languages* (pp. 117-141). Oxford: Oxford University Press.

- Ramchand, G. and Svenonius, P. (2014). Deriving the functional hierarchy. *Language Sciences*, 46, 152-174.
- Reinhart, T. (1981). Definite NP anaphora and c-command domains. *Linguistic Inquiry*, 12(4), 605-635.
- Richards, N. (2000). Another look at subjects in Tagalog. In I. Paul, V. Philips and L. Travis (eds.), *Formal issues in Austronesian linguistics* (pp. 105-116). Dordrecht: Kluwer.
- Ritter, E. (1988). A head-movement approach to construct-state noun phrases. *Linguistics*, 26(6), 909-929.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In S. Rothstein (ed.), *Perspectives on phrase structure: Heads and licencing* (pp. 37-62). San Diego: Academic Press.
- Rizzi, L. (1996). Residual verb second and the wh-criterion. In A. Belletti and L. Rizzi (eds.), *Parameters and functional heads: Essays in comparative syntax* (pp. 63-90). Oxford: Oxford University Press.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar: Handbook in generative syntax* (pp. 281-337). Dordrecht: kluwer.
- Rizzi, L. (2004). *The structure of CP and IP: The cartography of syntactic structures, volume 2*. Oxford: Oxford University Press.
- Roberts, I. (1991). Excorporation and minimality. *Linguistic Inquiry*, 22(1), 209-218.
- Roberts, I. (2001). Head movement. In M. Baltin and C. Collins (eds.), *The handbook of contemporary syntactic theory* (pp. 113-147). Oxford: Blackwell.
- Roberts, I. (2010). *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, Mass.: MIT Press.
- Roberts, I. (2011). Head movement and the minimalist program. In C. Boeckx (ed.), *The Oxford handbook of linguistic minimalism* (pp. 195-219). Oxford: Oxford University Press.
- Rögnvaldsson, E. and Thráinsson, H. (1990). On Icelandic word order once more. In J. Maling and A. Zaenen (eds.), *Modern Icelandic syntax* (pp. 3-40). San Diego: Academic Press.
- Roussou, A. (2002). C, T, and the subject: *That-t* phenomena revisited. *Lingua*, 112(1), 13-52.
- Rouveret, A. (1990). X-bar theory, minimality and barrierhood in Welsh. In R. Hendrick (ed.), *The syntax of modern Celtic languages* (pp. 27-79). San Diego: Academic Press.
- Rouveret, A. (1991). Functional categories and agreement. *The Linguistic Review*, 8(2-4), 353-387.
- Rouveret, A. (1994). *Syntaxe du gallois: Principes généraux et typologie*. Paris: CNRS Éditions.

- Sadler, L. (1988). *Welsh syntax: A government-binding approach*. London: Croom Helm.
- Sadler, L. (2000). Noun phrase structure in Welsh. In M. Butt and T.H. King (eds.), *Argument Realization* (pp. 73-110). Stanford, Calif.: CSLI Publications.
- Santorini, B. (1992). Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory*, 10(4), 595-640.
- Scott, G. (2002). *The syntax and semantics of adjectival modification*. Doctoral dissertation, University of London.
- Shlonsky, U. (1987). *Null and displaced subjects*. Doctoral dissertation, Massachusetts Institute of Technology.
- Shlonsky, U. (2004). The form of Semitic noun phrases. *Lingua*, 114(12), 1465-1526.
- Siloni, T. (1991). Noun raising and the structure of noun phrases. *MIT Working Papers in Linguistics*, 14, 255-70.
- Siloni, T. (1996). Hebrew noun phrases: Generalized noun raising. In A. Belletti and L. Rizzi (eds.), *Parameters and functional heads: Essays in comparative syntax* (pp. 239-267). Oxford: Oxford University Press.
- Siloni, T. (1997). *Noun phrases and nominalizations: The syntax of DPs*. Dordrecht: Kluwer.
- Sproat, R. (1985). Welsh syntax and VSO structure. *Natural Language and Linguistic Theory*, 3(2), 173-216.
- Sproat, R. and Shih, C. (1991). The cross-linguistic distribution of adjective ordering restrictions. In C. Georgopoulos and R. Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda* (pp. 565-593). Dordrecht: Kluwer.
- Starke, M. (2001). *Move dissolves into merge: A theory of locality*. Doctoral dissertation, University of Geneva.
- Starke, M. (2004). On the inexistence of specifiers and the nature of heads. In A. Belletti (ed.), *Structures and beyond: The cartography of syntactic structures, volume 3* (pp. 252-268). Oxford: Oxford University Press.
- Starke, M. (2009). Nanosyntax: A short primer to a new approach to language. *Nordlyd*, 36(1), 1-6.
- Stowell, T. (1989). Raising in Irish and the projection principle. *Natural Language and Linguistic Theory*, 7(3), 317-359.
- Stump, G.T. (1989). Further remarks on Breton agreement. *Natural Language and Linguistic Theory*, 7(3), 429-471.
- Surányi, B. (2005). Head movement and reprojection. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica*, 26, 313-342.
- Szabolcsi, A. (1983). The possessor that run away from home. *The Linguistic Review*, 3(1), 89-102.



- Szabolcsi, A. (1994). The noun phrase. In F. Kiefer and K. Kiss (eds.), *The syntactic structure of Hungarian* (179-274). San Diego: Academic Press.
- Thráinsson, H. (1993). On the structure of infinitival complements. *Harvard Working Papers in Linguistics*, 3, 181-213.
- Thráinsson, H. (1996). On the (non-)universality of functional categories. In W. Abraham, S.D. Epstein, H. Thráinsson and J.-W. Zwart (eds.), *Minimal ideas: Syntactic studies in the minimalist framework* (pp. 253-281). Amsterdam: John Benjamins.
- Timm, L.A. (1988). Relative clause formation in Breton. *Word*, 39(2), 79-107.
- Travis, L. (1984). *Parameters and effects of word order variation*. Doctoral dissertation, Massachusetts Institute of Technology.
- Vikner, S. (1990). *Verb movement and the licensing of NP-positions in the Germanic languages*. Doctoral dissertation, University of Geneva.
- Vikner, S. (1995). *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.
- Vikner, S. (2001). V-to-I movement and *do*-insertion in optimality theory. In J. Grimshaw, G. Legendre and S. Vikner (eds.), *Optimality-theoretic syntax* (pp. 427-464). Cambridge, Mass.: MIT Press.
- Weerman, F. (1989). *The V2 conspiracy: A synchronic and a diachronic analysis of verbal positions in Germanic languages*. Dordrecht: Foris.
- Williams, E. (2013). Generative semantics, generative morphosyntax. *Syntax*, 16(1), 77-108.
- Willis, D. (1998). *Syntactic change in Welsh: A study of the loss of verb-second*. Oxford: Clarendon Press.
- Willis, D. (2006). Against N-raising and NP-raising analyses of Welsh noun phrases. *Lingua*, 116(11), 1807-1839.
- Zwart, J.-W. (1993). *Dutch syntax: A minimalist approach*. Doctoral dissertation, University of Groningen.