UNIVERSITY COLLEGE LONDON

Phrase Structure and Derived Heads

A dissertation submitted in partial fulfilment of the
requirements for the degree of
Doctor of Philosophy in Linguistics

by Dirk Bury

Originally submitted September 2002 and defended
7 March 2003.
This version is from April 2003.

ABSTRACT

This dissertation investigates the theory of phrase structure in the field of generative grammar.

In chapter 1, I propose a new model of phrase structure representations and argue for its conceptual advantages over alternative models. In chapters 2 to 4, I discuss the advantages of this model for the analyses of various empirical phenomena. Chapter 5 is the conclusion.

The proposal is based on two central claims. First, I argue that a moved verb can be the head of its clause, and hence that verb movement can extend a structure (Ackema, Neeleman and Weerman 1993). (This is a rejection of the widely held view that it is always the target of movement that projects; cf. Chomsky 1995.) Since verb movement is not universal, it then follows that syntactic representations are not identical for all languages. This means that the clause structure of a particular language must be learnable, i.e. that the availability of categories without overt realization is highly restricted. In particular, there can be no phonetically empty categories that are not licensed configurationally and that do not receive a semantic interpretation.

Second, I argue that phrase structure should not be thought of in terms of tree diagrams but rather in terms of sets that express dominance relations between categories. This is a radical implementation of the widely held view that linear order is not a property of syntax proper ("the computation of LF") but of the mapping of syntactic structures to the phonological level of representation PF. One result of this approach to phrase structure is that there is no need for categorial projection (cf. Brody 2000).

Chapter 2 analyses the phenomenon of complementizer optionality and its relation to verb movement and adjunction, as well as related problems from English, Italian, and German (e.g. "embedded verb-second").

Chapter 3 investigates subject-verb inversion in verb-second languages and verb-first languages. This involves a discussion of Breton, Modern Welsh, and German, and of the change from verb-second in Middle Welsh to verb-first in Modern Welsh.

Chapter 4 investigates free relative clauses in English and (dialects of) German.

On a personal level, *Freaking Out* is a process whereby an individual casts off outmoded and restricting standards of thinking, dress, and social etiquette in order to express CREATIVELY his relationship to his immediate environment and the social structure as a whole. Less perceptive individuals have referred to us who have chosen this way of thinking and FEELING as "*Freaks*," hence the term: *Freaking Out*. […] We would like to encourage everyone who HEARS this music to join us…become a member of *The United Mutations…FREAK OUT!*

(Frank Zappa, *Freak Out*, 1965)

Table of Contents

ACKNOWLEDGEMENTS

CHAPTER ONE

Outline of the Theory

1 Introduction

The *base component* of a transformational grammar specifies what structures are available for transformations to operate on (cf. Chomsky 1965, McCawley 1968).[1] On the view that transformations cannot introduce structures that could not be generated by the rules of the base component (Emonds 1976), the rules of the base have a wider role than the previous sentence suggests. These rules then determine a set of possible structures that includes all well-formed structures.[2] The rules of the base interact with other linguistic rules, such as conditions on transformations (e.g. locality) or conditions on the licensing of linguistic elements (e.g. case theory), conditions on the mapping between different levels of representation (e.g. shape preservation), and performance constraints (e.g. on processing).

In this dissertation, I outline a model of the base component that embraces a range of fairly conventional assumptions, but that develops them in a less conventional way. The most important among these is the idea that linear order is irrelevant for the computation of syntactic representations. This idea leads to structural representations that can get along without the additional theoretical tool of categorial projection. Moreover, it makes possible the representation of phrase structure as sets that explicitly and exhaustively express hierarchical relations. Chapter 2 explores the interaction of chain formation and selection in the framework built on these ideas. The phenomena covered there include the relation of verb movement and adjunction to complementizer optionality and the (in-)existence of verb-second complements. Chapter 3 proposes a new explanation of verb-second patterns and their relation to verb-initial patterns. Chapter 4 discusses the implications of this framework for the analysis of free relatives. Chapter 5 concludes the dissertation.

In the remainder of this chapter, I introduce the theoretical framework and, in the final section, give a brief overview over empirical scope and results of the dissertation.

---

[1] The background of this dissertation is mainly work in the Principles-and-Parameters approach (e.g. Chomsky 1986b, 1995c), and especially certain ideas proposed in Neeleman and Weerman 1999 and Brody 1997, 2000. However it is possible that the proposals made here (if they make any sense) also have some relevance for other frameworks of generative grammar such as HPSG, LFG, and Word Grammar.

[2] If this is correct, the difference between transformational grammars and transformation-less grammars loses much of its relevance, since, whatever the role of transformations, the structures they generate will have to be compatible with the rules of the base; cf. Koster 1987:101f

2 Moved material can head a structure

2.1 General considerations

Chomsky 1993 assumes that clause structure is derived from individual lexical items by means of two syntactic operations. The first operation, MERGE, combines two objects α and β (lexical items, or complex structures, or one of each) to form a labelled unordered set. In principle, either of the two input objects can provide the label of the set or, in other words, function as the head of the new structure. Thus, merging α and β creates $\{_\alpha \alpha, \beta\}$ or $\{_\beta \alpha, \beta\}$.[3] The sets formed by this operation are usually represented by tree structures. Thus, the two possible sets formed by merge of α and β can be represented by the following trees:[4]

(1) a.　　　α　　　　　b.　　　　β
　　　　　／ \　　　　　　　　／ \
　　　　α　β　　　　　　α　β

On this view, the labelling of a set is a similar process to the X-bar theoretic notion of categorial projection (cf. Chomsky 1970, Jackendoff 1977). That is, the element that provides the label of the output of merge can be said to project.

The second operation, MOVE, also combines two objects to form a labelled unordered set, but here one of the two is a complex element β and the other a simple or complex element α contained within β.[5] Most work in generative syntax follows Chomsky 1995c, where it is assumed that move differs from merge in that only one of the two objects in its input can project, namely the object β originally containing the object to be moved, α:

(2) a.　　　β　　　　　b.　　　　* α
　　　　　／ \　　　　　　　　／ \
　　　　α_i　β　　　　　　α_i　β
　　　　　　△　　　　　　　　　△
　　　　　　t_i　　　　　　　　t_i

---
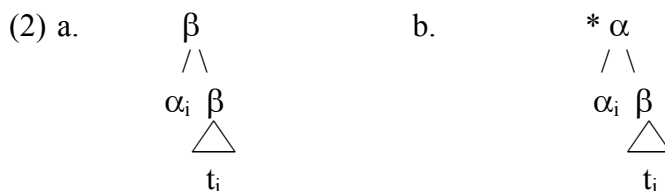
[3] In most cases, the lexical requirements of the input objects and/or the requirements of the elements the merged structure combines with and/or general output conditions will uniquely determine which of the two is the label of the new structure. See Chametzky 2000, Langendoen 2002 for discussion of the formal properties of Chomsky's merge.

[4] These structures are both asymmetric in the sense that in each only α or β projects, not both (Chomsky 1993:246). Uriagereka 1998 seems to suggest that merge is asymmetric in the sense that in principle only one of α and β can ever project.

[5] In contrast to merge, move creates a chain relating the source position and the landing site of the moved element.

This asymmetry between move and merge, that, for move, only one of the two input elements can project, but both can for merge, is often called the *"Target projects" Condition* (e.g. Chomsky 1993:190f, Brody 1998:391ff). However, it is not usually noted that in Chomsky's framework this condition does not follow from any principles of phrase structure, but rather from particular assumptions about movement and its triggers, which are independent of phrase structure. Simplifying somewhat, these assumptions are that movement is driven by the need to check features, and that feature checking is possible only in particular syntactic configurations, namely in a specifier-head relation or under head-adjunction (e.g. Chomsky 1995b:257). Crucially, in any theory that doesn't adopt these additional assumptions, such an asymmetry between move and merge would have to be stipulated.

There is then no principle of phrase structure that bans projection of a moved syntactic object. The assumption of this asymmetry is all the more surprising in the light of recent attempts to derive move as a complex form of merge, or to unify the two operations altogether. Thus, Kitahara 1997 analyses merge as a form of concatenation and move as concatenation plus copying. According to Starke 2001:177, who does not assume this asymmetry, "move is nothing else than an instance of 'merge' operating on two non-adjacent nodes". Similarly the move vs. merge distinction is relaxed in other recent work. Thus, in Chomsky 1999:7, it is argued that "the combination of Agree/Pied-Pipe/Merge is the composite operation Move", and in Chomsky 2001, move is analysed as "external merge". All this suggests that the simplest position is to assume that both merge and move are symmetric operations in the sense that in both operations either of the two objects that the operation combines can project in principle.[6]

Thus, I tentatively maintain Chomsky's view on merge, but adopt a version of move that is not affected by assumptions independent of phrase structure. Substantial revisions to this proposal will be made from section 3 onwards, where a new theory of phrase structure is introduced. Move α, where α originates within β, generates two possible structures, as

---

[6] Of course, the properties of moved material may differ in other ways from those of unmoved material due to independent constraints, e.g. selectional properties are restricted by the Generalised Projection Principle of Brody 1995, 1998; see section 6.

illustrated in (3).[7]

(3) a.   β   b.   α
         $/\$ $/\$
      α_i β α_i β
         △   △
         t_i t_i
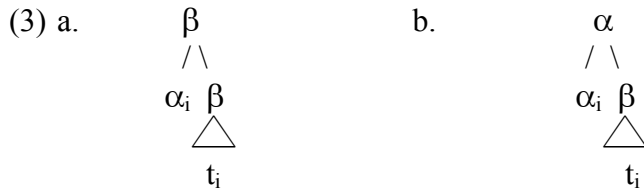
     This conception of move gives rise to two configurations that are not available on the standard view, depending on whether the moved item is a head or a phrase. In the first case, proposed for example by Ackema et al 1993, movement of a head is followed by projection of the moved head. In the second case, movement of a phrase is followed by projection of the moved phrase. The two possibilities are illustrated in (4a) and (4b) respectively:

(4) a.   XP   b.   XP
         $/\$ $/\$
      X_i^0 β XP_i β
         △   △
         t_i t_i
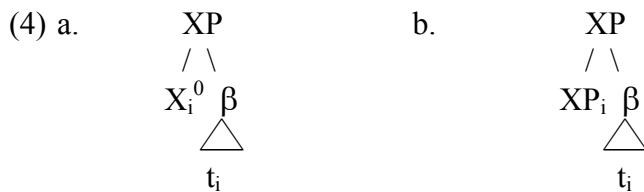
The structure in (4b) created by movement and projection of a phrase resembles a traditional adjunction structure where the constituent β is adjoined to the constituent XP. An analysis of free relatives along these lines is proposed in Bury 1998, which forms the basis of chapter 4, where such structures will be discussed in more detail. In the following section, I will discuss some aspects of structures like (4a) which are created through the projection of a moved head.

2.2 Verb movement without head adjunction

The rejection of the idea that moved material cannot project entails that verb movement can occur beyond its conventional domain of head adjunction structures (see Ackema et al 1993). In particular, what can also happen is that the verb moves from its base position and

---

[7] As with merge, the choice of which element projects after move is determined by the lexical requirements of α and β and/or the requirements of the elements the merged structure combines with and/or general output conditions.

projects a new VP. However, it should be noted that the theory of locality imposes a principled restriction on such structures. Thus, the Head Movement Constraint (HMC; Travis 1984), or some version of Relativised Minimality (Rizzi 1990, 2001), states that a head cannot skip an intervening head position. This means that verb movement followed by projection of the verb is only possible if the moved verb takes the original VP as its complement. In other words, a verb cannot move out of e.g. a CP and project after movement. In (5), the verb moves out of the lower VP and projects a new VP.[8] (6) illustrates a structure that violates the HMC.

(5)

```
            VP
          /    \
        —       V'
              /    \
           V_i      VP
                  /    \
                XP      V'
                      /    \
                    t_i     YP
```

(6)   *

```
            VP
          /    \
        —       V'
              /    \
           V_i      CP
                  /    \
                 C      VP
                      /    \
                    XP      V'
                          /    \
                        t_i     YP
```
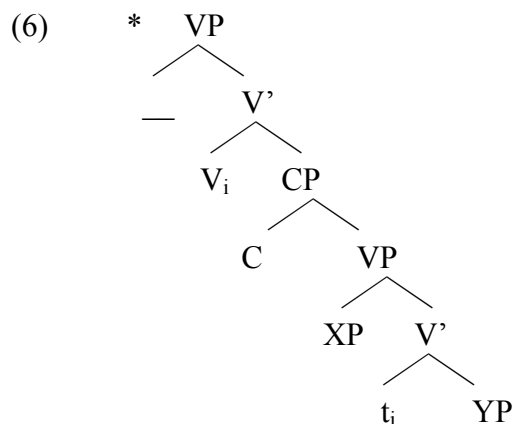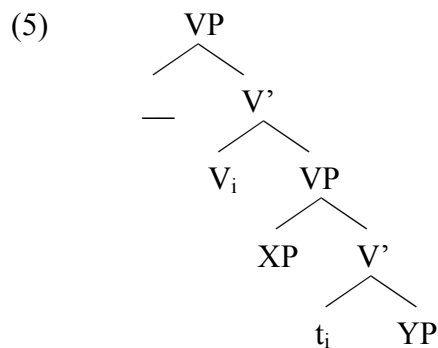
Structures like (5) are sometimes said to involve SELF-ATTACHMENT (cf. Chomsky 1995b:320f). I will refer to them as S-MOVEMENT STRUCTURES. ('s' will be contrasted below with 'p' and is intended as a mnemonic for the notions 'structural' or 'syntactic').

The properties of VPs derived by s-movement, which I will call DERIVED VPS, differ

---

[8] It has been argued by Koeneman 1995, Bury 2000, 2002 that a VP generated in this way always requires a specifier; see section 4 for a new explanation of this specifier requirement, and chapter 3 for detailed discussion.

from base VPs due to a range of independent constraints. For example, given the assumption that s-movement creates a head chain, the Generalised Projection Principle, which states that selection is only possible from the lowest position in a chain (see section 6), entails that the moved verb will not be able to assign θ-roles from its position. This explains why such derived VPs pattern with projections of functional categories rather than with projections of lexical categories in this respect. As noted by Fukui and Speas 1986, Abney 1987 and Grimshaw 1991 and others, the relation between functional heads and a lexical head within an extended projection is different from the thematic selection found between a lexical head and its complement.[9] A further property of functional categories is that they typically take complements of one particular category. As noted above, for derived VPs, this is a direct consequence of the HMC.

A notable property of derived VPs is that their creation cannot be triggered by the features of abstract functional heads, simply because no such heads are involved here (cf. the discussion of head adjunction in the next section). Instead, triggers for verb movement (which may or may not involve the creation of a new position – see section 3.1.3) have to be formulated directly in terms of interface conditions. (See Szendroi 2001 for relevant discussion of focus structures.) In this dissertation, I will discuss two triggers. Building on a proposal by Sproat 1985, and a more recent version of this proposal in Neeleman and Weerman 1999, I will argue in Chapter 3 that verb movement in the Celtic languages takes place in order to license nominative case of the subject. This requirement is formulated as a PF condition. In Chapter 2, I argue that there is a clause-typing parameter that determines whether or not a language must overtly mark main clauses. In some languages, special particles are available to mark main clauses. However not all languages with a positive setting of this parameter have special particles. I argue that in the absence of such particles a language can resort to verb movement to mark main clauses. This function of verb movement results from the interaction of the projection of moved verbs, the particular theory of phrase structure developed in the following section, and a condition on chains that blocks selection after movement (cf. Brody 1995).

While there is currently no theory of triggers on this approach to verb movement, the same is also true for the main alternative, checking theory. In particular, the attempts of

---

[9] For recent discussions of functional categories, see Chametzky 2000:17-32, Fukui and Sakai 2003: section

early checking theory to rely on morphological features to trigger movement (Chomsky 1993) could not be maintained. Instead, it is now widely assumed that movement is triggered by so-called EPP-features, whose sole purpose is to indicate that a (functional) head requires a specifier (cf. Chomsky 1999, Holmberg 2000). So far, it has not been shown that the presence of such a feature on a particular head can be related to independent properties of the head. EPP-features are assumed simply where a moved specifier occurs. While Chomsky 1999:6f suggests that "it is tempting to associate EPP with phi-completeness," this property does not appear to be sufficient to determine whether or not an EPP feature is present, since in some cases such features are optional. This approach at least implicitly acknowledges that at present there is no deep theory of what triggers movement in checking based models. One advantage of the present approach, which doesn't rely on feature checking, is that it makes it more difficult to formulate syntax-internal triggers, which means that triggers typically must be motivated independently as interface conditions. This restriction doesn't hold in an obvious way for checking theory, beyond the trivial requirement that uninterpretable features must be checked by LF or PF. Of course, this requirement merely ensures that features introduced to trigger movement actually do trigger movement.

3 Verb movement and clause structure

Before discussing the implications of the availability of s-movement structures for clause structure, I review some aspects of a different verb movement configuration, namely head adjunction.

3.1 The status of head adjunction

In any theory that assumes some form of the "target projects" condition, that is of the assumption that moved material cannot project, the only way to formalise head movement is in terms of head adjunction.[10] Thus, in most work in Principles-and-Parameters theory, syntactic head movement is taken to be head adjunction (cf. Baker 1988, Chomsky 1993, 1995, Roberts 2001). However, this view is not without problems. In the following tree,

---

2.

[10] An alternative to real head movement would be to analyse all apparent head movement in terms of remnant movement; see Kayne 1998, Koopman and Szabolsci 2000, Nilsen 2001, Müller 2002.

the head H has moved to the head F under head adjunction:

(7)
```
                    FP
                  /    \
                F       HP
               / \     /   \
             Hi  F   AP     H'
                      |    /   \
                      ti        BP
```
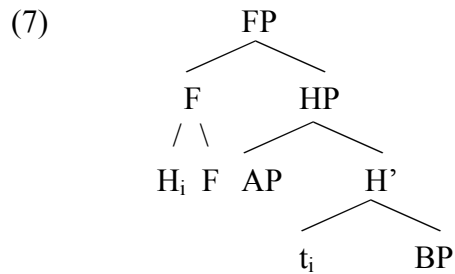
An immediate issue that arises with such structures concerns the c-command requirement on the relation between traces and their antecedents. In a structure like (7), c-command in its traditional sense does not hold, i.e. the first node that dominates the moved H does not dominate its trace (Reinhart 1976). Thus, any theory that assumes head adjunction is forced to assume a more complex definition of the command relation. In a theory without head adjunction, this issue does not need to arise.

Further complications arise from assumptions about the positions to which the head moves under head adjunction. There are two versions of verb movement in this regard. One involves the assumption that verb movement is movement of a bare (uninflected) verb stem to a functional head occupied by an inflectional morpheme. I will refer to this view as the CONCRETE FUNCTIONAL HEAD HYPOTHESIS. On the second view of head adjunction, a verb is inserted fully inflected and moves to an abstract functional head to check (some of) its features. I will refer to this view as the ABSTRACT FUNCTIONAL HEAD HYPOTHESIS.

3.1.1 The Concrete Functional Head view of head adjunction

At the heart of the Concrete Functional Head view is the idea that stems and inflectional morphemes are generated as separate heads within syntax. Complex heads containing both stems and affixes are then the result of syntactic movement, not of processes of a separate morphological component (cf. Chomsky 1957, 1991, Baker 1988). In such a theory, verb movement can be triggered by the STRAY AFFIX FILTER (Lasnik 1981), which requires that certain morphemes not be spelled out on their own, but instead that they must be attached to some other head, e.g. a verb stem. Thus, in a language like French, the verb raises to a Tense morpheme in order to prevent a violation of the Stray Affix Filter. The reverse

happens in English, where the Tense inflection is lowered onto the verb. The two cases are illustrated here:

(8) a.   TP   (before verb movement)   **French**
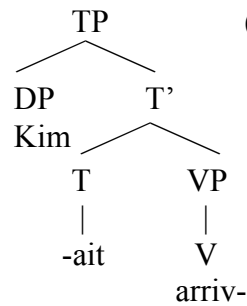
```
              TP
            /    \
          DP      T'
          Kim    /  \
                T    VP
                |    |
              -ait   V
                    arriv-
```

b.   TP   (after verb movement)

```
              TP
            /    \
          DP      T'
          Kim    /  \
                T    VP
               / \   |
          arriv-ᵢ ait V
                      tᵢ
```

(9) a.   TP   (before verb movement)   **English**

```
              TP
            /    \
          DP      T'
          Kim    /  \
                T    VP
                |    |
              -ed    V
                    arriv-
```

b.   TP   (after verb movement)

```
              TP
            /    \
          DP      T'
          Kim    /  \
                T    VP
                |    |
                tᵢ   V
                    / \
              arriv- -edᵢ
```
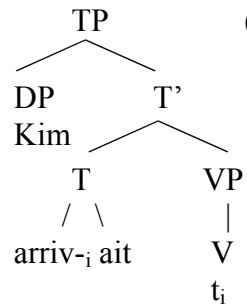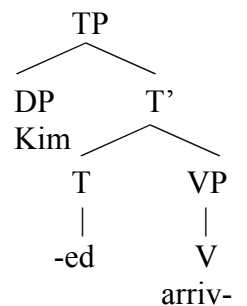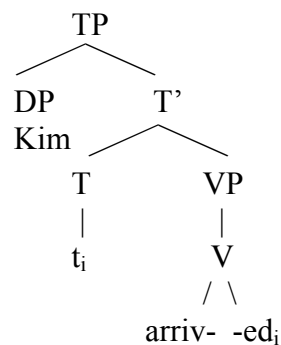
However, it has been observed that a reduction of inflectional morphology to syntax in terms of the Concrete Functional Head view cannot really work (for discussion, see e.g. Ackema 1999). This is so because, in syntax, selectional requirements must always be satisfied at the point of lexical insertion, i.e. before movement takes place (cf. Brody

1995). This conflicts with the Concrete Functional Head idea that it is the morphological requirements of an independently inserted morpheme that trigger movement of the verb. While in syntax proper, selectional requirements cannot be satisfied after movement, the Stray Affix Filter in effect states that selectional requirements, namely the requirement of a bound morpheme to be affixed, trigger movement. A unification of syntax and morphology in this sense cannot be argued to be a conceptual simplification, because the resulting uniform morpho-syntactic component will have to include additional devices that make it possible to capture syntactic generalisations such as the ban on selection after movement that do not hold of those morphological relations that are on this view interpreted as aspects of the theory of syntax.

A further problem with a structure like (9b), where the tense affix lowers onto the verb, is that the trace of the lowered affix is not c-commanded by its antecedent. This is what led many people to reject this approach to verb movement and to adopt a version of the abstract functional head view that will be discussed in the next section.

Finally, it should be noted that one of the motivations for the Concrete Functional Head approach to verb movement is the assumption that it would be desirable to have a single generative component that deals with both traditional syntax and with word formation. The argument is that if inflectional morphology and lexical stems could combine within syntax, this would lead to a simpler theory. For this argument to go through, it is crucial that there are no independent reasons for a separate word formation component. This is so because, once such a component is available, nothing is gained if some aspects of word formation, i.e. the combination of verb stems and inflectional morphemes, can also be accomplished in syntax. In fact, if a word formation component is needed anyway, it could be considered a weakness of the theory if word formation can also be done within syntax. Without discussing the relevant cases, there appear to be certain aspects of derivational morphology and of compounding that can be best captured by a separate word formation component, while it is not clear how they could be captured within syntax (Baker 1988, DiSciullo and Williams 1987, Ackema and Neeleman 2000, 2002). If this intuition is correct, considerations of elegance would favour a theory that can only deal with morphology in a single place, namely the word formation component, and

that is incompatible with a Concrete Functional Head approach to verb movement.[11]

3.1.2 The Abstract Functional Head view of head adjunction

Under the lexicalist hypothesis, verbs are inserted into syntactic structures fully inflected (Chomsky 1970, 1993, Selkirk 1982, Williams 1994). The Abstract Functional Head view of head adjunction is a version of strict lexicalism according to which the inflectional morphology on a verb must be licensed by a functional head without phonetic content and in an appropriate structural configuration. There are different implementations of this idea, but I limit the discussion here to just one. According to Chomsky 1993, the verb must check its features against the matching features of a functional head and this feature checking is possible only in a head-adjunction structure.[12] Consequently, the need to check a feature located on an affix of the verb against some matching features of an abstract functional head higher in the structure can trigger verb movement to create the required head-adjunction structure. On this view, language variation is determined by whether checking takes place overtly or covertly, where only the former gives rise to observable verb movement. In other words, French and English clauses would have the same (LF-)structure namely that in (10b), where [+V] indicates a checked uninterpretable V-feature on T; they would differ in which of the positions of the head movement chain is pronounced. In English, the verb would be pronounced in the position of the trace t_i, while in French the verb would be pronounced in the position of T:

(10)   a.   TP
            ╱╲
         Subj   T'
              ╱╲
         T [+V]   VP
                    |
                 V [+V]

---
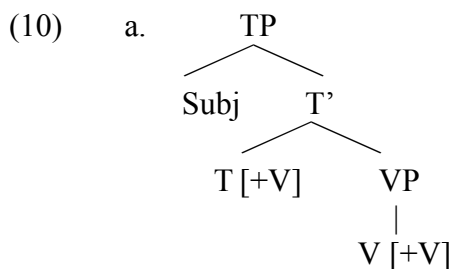
[11] For more discussion of the concrete functional head view, see Ackema 1999:16-21, Ackema and Neeleman 2000, Borer 1998, Chomsky 1993, Koeneman 2000:23-42, Spencer 1991, Stump 2001:17-27.
[12] The idea of feature checking goes back at least to Brody 1985:section 3, whose Case Checking Theory deals with the licensing of case features on nouns.

b.
```
          TP
        /    \
     Subj     T'
            /    \
        T [+V]    VP
        / \        |
    T [ᵥ stem+affix]ᵢ  tᵢ
```
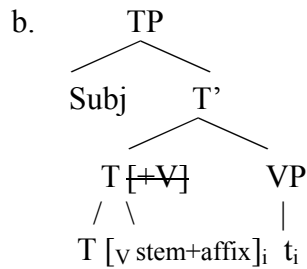
An immediate question is of course how different this approach is from the Concrete Functional Head approach. In particular, the device of feature checking has the same effect under the Abstract Functional Head view as morphological selection does under the Concrete Functional Head view. Attraction of a head by a head with a checking feature is very similar to attraction of a head by a morpheme that wants to be affixed.

Leaving aside this question, the Abstract Functional Head approach and especially its assumption of feature checking introduce a range of new problems. The main problem is the range of redundancies that are intrinsic to the checking operation. The checking operation relies on the assumption that identical features are introduced into the structure on two independent lexical items. In the example in (10a), the functional head T carries an (uninterpretable) [+V] feature, and the verb carries a [+V] feature by virtue of being a verb. The sole purpose of the [+V] feature on T is to trigger movement, and then to delete so complications at LF can be avoided. Since movement-triggering features of abstract functional heads always delete after checking against identical features of a lexical head, it is difficult to imagine what independent motivation there could be for their existence (cf. Brody 1997b). Thus, in addition to being redundant, these features exist only for theory-internal reasons. An approach to movement chains in terms of copies compounds this problem because the copying of heads also copies the heads' uninterpretable features. Since only one of the copies of a head is in the appropriate checking configuration, additional assumptions need to be made to ensure deletion of the uninterpretable features on all copies (cf. Nunes 1995, Cormack and Smith 1997, Gärtner 1998 for some discussion).

A further problematic aspect of feature checking is that it requires an additional statement of the order in which uninterpretable features on a head are checked (cf. Brody 2000: section 2, den Dikken 2002). In particular, in order to account for the mirror

generalisation in terms of feature checking, it is necessary to assume that affixes on the verb are generated in the correct order (and that feature checking takes place from inside to outside). This order must be matched by the order in which abstract functional heads are merged. Thus, to derive the Hungarian word *olvas-hat-om* 'read-PERMISSIVE-1SG-PRESENT', the sequence V-Modal-Tense must be generated in the lexicon, where the complex verb is assembled, and the individual categories V and the abstract functional heads Modal and Tense must be merged into the structure in the correct order (Brody 2000). This means that the sequence of functional heads must be stated as a constraint on morphological operations and in terms of the syntactic distribution of functional heads. Brody compares this duplication to the redundancy of phrase structure rules that state the number and type of arguments of a lexical head, where this information is already available on the lexical entry of the head (2000:37).[13]

3.1.3 Verb movement to functional heads without head adjunction

Brody 2000 proposes an approach to head movement that avoids some of the problems introduced by head adjunction and feature checking. Simplifying somewhat, the idea is that the positions occupied by functional heads in a verb's extended projection (in the sense of Grimshaw 1991) are all positions in which the verb can potentially be pronounced.[14] Verb movement can then be encoded by a feature whose sole purpose is to indicate in which of a range of possible positions the verb is pronounced. This means that no head adjunction structure need be created and there is no mandatory feature duplication as in conventional checking theory. I will call such structures P-MOVEMENT STRUCTURES. ('p' is intended as a mnemonic for notions like 'pronunciation' or 'PF'.)Given (11), p-movement can involve the pronunciation of the verb in the positions of either T or C instead of the pronunciation in the verb's base position V:

---

[13] See Brody 1997a,b, 2000 for further problems of checking theory; see also Abels 2001, Koeneman 2000:43-45, Roberts and Roussou 1997.

[14] I do not adopt a second assumption that is central to Brody's 2000 proposal. Brody's Mirror hypothesis implies that there is a single theoretical module of morphosyntax. In contrast, I assume that there are separate generative systems of morphology and syntax (cf. Ackema and Neeleman 2002: chapter 1).

(11)

```
                    CP
                    |
                   C'
                  /  \
                 C    T'
                     /  \
                   Subj   T'
                         /  \
                        T    VP
                            /  \
                           V    Obj
```
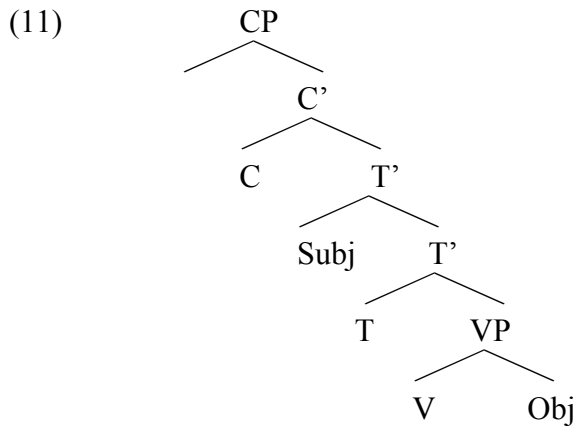
Unlike head adjunction structures (or the s-movement structures discussed above), p-movement structures do not involve syntactic displacement but it amounts to the pronunciation of the verb in a position that is higher than its base position. Since there are no traces and no head adjunction structures are created, no problem arises for the definition of c-command. Since there is no feature checking, no features need to be duplicated. In particular, there is no need for V-features on functional heads to trigger verb movement.[15] Of course, the theory is compatible with an accidental duplication of features in a particular situation, such as the presence of person features on both a subject and an inflected verb. The crucial point is that, unlike Checking Theory, this approach to verb movement doesn't introduce across the board feature duplication by definition.

The effects of the HMC for s-movement follow from general constraints on the mapping between levels of representation. S-movement involves the displacement of phonological features. For concreteness, we can assume that the relevant level of representation for s-movement is PF. Mapping between levels is in general restricted by the requirement that lines associating pairs across two levels may not cross. Furthermore, I assume that each PF head position corresponds to a syntactic head position (and probably vice versa); that is if there are three syntactic head positions, there will be three head positions at PF (cf. Cormack and Smith 2001:221). The syntactic structure in (12a) can

---

[15] There appears to be a residue of redundancy with regards to mirror effects: The sequence of functional heads in clause structure corresponds to the sequence of affixes on the verb. This appearance of a redundancy disappears if syntax and morphology are separate generative components, both interfacing with LF, and the outputs of both must match a universal hierarchy that is stated as a condition on the interface of syntax and morphology with semantics (LF). See Grimshaw 1986, Alsina 1999 for related discussion. There is also no such redundancy in Brody's system, which contains only a single component of morphosyntax.

then be mapped on to the PF in (12b), indicating pronunciation of V in the position of T under s-movement. In contrast, a mapping that would correspond to movement of V to C across intervening T is ruled out because the lines associating the positions of T and V on PF and in syntax cross, as indicated in (12c).[16]

(12) a.　[$_{CP}$ … C [$_{TP}$ … T [$_{VP}$ … V …]]]

　　b.　syntax:　　… C …　　T …　　V …
　　　　　　　　　　　|　　　|　／
　　　　PF:　　　　　H　　　H　　H

　　c.　syntax:　　… C …　　T …　　V …
　　　　　　　　　　　|　／　|
　　　　PF:　　　*　H　　　H　　H

We can summarise then that verb movement always amounts to the pronunciation of the verb in a position different from its base position. This displaced pronunciation can take place in two different structural configurations. Where the verb is pronounced in the position of an independent head I talk about p-movement. Where the verb is pronounced in the position of a copy of the verb, I talk about s-movement.

3.2 Verb movement and abstract functional heads

3.2.1 No need for uninterpretable abstract functional heads in the lexicon

A feature of many proposals adopting the assumption that all verb movement involves head adjunction is that abstract functional heads come in two flavours (cf. Kayne 1994:30). Some heads, e.g. (abstract) Tense, are assumed to be interpretable, that is to have an effect on LF. Other heads, e.g. (abstract) Agreement, are assumed to have no interpretation, and are postulated only to make up for the "paucity of available adjunction sites" (Kayne 1994:30) (or to allow for movement into their specifier position). That is, these uninterpretable heads exclusively function as targets for movement. While this type of functional head has been used to great analytic effect, the assumption of two distinct types of abstract functional heads complicates the theory. Moreover, uninterpretable abstract

---

[16] Since s-movement does not trigger chain formation, there may be cases in which it involves lowering (cf. Abels 2001). A possible example of this is the lowering of the complementizer in Irish discussed by McCloskey 1996b.

functional heads are assumed for purely theory-internal reasons, which means that there can be no independent motivation for such heads (cf. Brody 1997b).[17]

The recognition that there is no structural reason why a moved head shouldn't project, and hence that s-movement structures can exist, makes it possible to simplify the theory. If verb movement is necessary but there is no position to which the verb can move, the verb can simply move and project again. There is no need to assume the existence of uninterpretable heads to make verb movement possible. Since the "paucity of available adjunctions sites" is the only motivation for uninterpretable abstract functional heads, this theoretical device can now be eliminated. That is, the lexicon need not contain such uninterpretable functional heads. (See e.g. Chomsky and Lasnik 1993[1995]:40, Speas 1995 on lexical items without a semantic role.) Compared to a model that assumes such heads, this clearly restricts the range of possible analyses for any particular syntactic structure and is therefore a desirable change.

Maybe it is worth reiterating that this is not a case of one theoretical device being eliminated at the cost of introducing a new, equally complex device. The tool that is eliminated here, namely uninterpretable abstract functional heads, is argued to be unnecessary once a dubious stipulation, namely that moved material cannot project, is given up. Thus, a simplification of the theory, the rejection of the "Target projects" condition, makes available s-movement structures at no extra cost. The availability of these structures makes the device of uninterpretable abstract functional heads redundant.

3.2.2 Covert head movement and triggers of head movement

Under the Abstract Functional Head hypothesis, there is V-to-T movement in all languages. The differences in word order found between English and French are explained in terms of an uninterpretable [+V] feature of T that both languages have. If the feature is strong, verb movement must take place in overt syntax, as in French. If the feature is weak, there is no overt verb movement. However, verb movement must then take place covertly in order to ensure that no uninterpretable [+V] feature remains on T at LF. This is a surprising state of affairs, given the widely held view that head movement does not have an effect on the semantics of a sentence (cf. Chomsky 1995b, Williams 1994, Zwart 2001).

---

[17] This type of head may also be necessary if *all* verb movement is interpreted as s-movement.

Things are different if verb movement is not triggered by feature checking but directly by conditions on the interfaces PF and LF. There is no strong/weak dichotomy in these triggers. If in some language, verb movement is forced by some such condition, it does not follow that in all other languages verbs have to move covertly. Thus, the undesirable prediction that there is covert head movement disappears.

3.2.3 Empty categories and the non-universality of clause structure

Earlier it was argued that the availability of s-movement makes it possible to drop the device of uninterpretable abstract heads from the theoretical repertoire. The availability of s-movement structures indirectly further restricts the availability of empty categories via constraints on language acquisition. This in turn raises the question of whether there can be any abstract heads, and, if yes, under which conditions such heads could be available in a particular language.

In the current proposal, derived VPs may need to be created to represent verb movement, which varies across languages. For example, French and English are usually assumed to differ in their presence vs. lack of verb movement. Indeed, clause structure even varies within a single language. For example, embedded clauses in English in which verb movement exceptionally takes place, e.g. *Jane said that never in her life would$_i$ she* t$_i$ *do that*, have often been analysed as containing two CPs, while regular embedded clauses without verb movement are assumed to contain only a single CP (cf. Culicover 1991, Iatridou and Kroch 1992). Thus, there seem to be good reasons to assume that clause structure is not universal.[18] Since clause structure is not universal, it follows that the structure of a particular (clause in a particular) language must be learnable (Iatridou 1990:553, Thrainsson 1996, Koeneman and Neeleman 2001:190). This means that the learner must have positive evidence for every head in the structure of a clause.[19]

There are different types of evidence for a syntactic head. The most obvious evidence is phonological content. If a sound is pronounced in a sentence, it will need some sort of

---

[18] This is not a necessary consequence. While there is no motivation for covert verb movement intrinsic to the current proposal, the model is, in principle, compatible with covert verb movement and thus with a universal LF. However, the proposal that clause structure is not universal offers a natural solution to the often noted problem that in many current proposals "more than 90% of heads are null" (Starke 2001:161).

[19] It seems that from the perspective of learnability, it would be most advantageous if every head had to be licensed at both interfaces, but this doesn't seem to be a tenable condition, contra Bouchard 1995.

representation within clause structure. Although it is clear that not everything that is pronounced corresponds to an independent syntactic head; e.g. some heads are realised in morphology as affixes.

There are several ways in which a head without phonological content can be licensed. First, in s-movement structures an empty head occurs in the base position of the verb. This illustrates a more general case where phrase structure contains one or more derived heads but where not all of them are pronounced. In this case, the existence of the empty head position comes from whatever triggers the generation of a derived head. An example could be a structure that contains three constituents but only one visible head. If phrase structure is restricted to binary branching, this may make the presence of an extra head detectable.

A second class of empty head results from language-specific deletion rules. For example in English the Doubly Filled Comp Filter (Chomsky and Lasnik 1977, Pesetsky 1998) rules out a CP that contains both a lexical complementizer and a *wh* phrase in its specifier with the result that one of the two must be deleted. Since such rules are restricted to well-defined environments, the rules, and the resulting empty heads, pose no problem for learnability.

A third class of empty head can occur where there is a paradigm containing several overt forms. In this case, the paradigm makes it possible to learn empty forms that realise feature combinations that are not associated with an overt form (Pinker 1984: chapter 5). In the general case, an empty form would correspond to the least marked feature specification. Thus, it is well known that in case paradigms, nominative case is most frequently realised without an overt form. Koeneman (2000:125f) proposes that modal verbs in English form a paradigm that makes it possible to learn an empty syntactic position for clauses without modals.[20]

Finally, since PF can provide evidence for a syntactic head position, an obvious question is whether LF can do the same. Such evidence could come in the form of a particular semantic interpretation of a structure, or as a universal condition on the mapping of syntactic to semantic representations. Given a particular interpretation that can not be related to any visible syntactic head in a structure, there are two possibilities. There may be

---

[20] This proposal receives some support from work on grammaticalisation in English. Krug 2000 argues that gaps in the paradigm of modals are partly responsible for the emergence of new modals, like *hafta* and *wanna*.

an empty syntactic head, or there may be an empty morphological head, i.e. a zero affix, that causes the interpretation.[21] Thus, on its own, a particular semantic interpretation cannot provide unambiguous evidence for a syntactic head. Conditions on the semantic interpretability of syntactic structures provide a different type of evidence. Thus, it is widely assumed that tense has a special role in the interpretation of clauses. For example, Chierchia and McConnel-Ginet (2000:279-282) argue that an elegant mapping between syntax and semantics requires the presence of tense in a position that takes scope over the verb and its arguments, including the subject, within syntactic structure. Similarly, Bouchard (1995:407f) argues that tense must occupy the most prominent head in a sentence.[22] Following this intuition, I propose the following condition:

(13) TENSE CONDITION: Every clause must contain a syntactic head that takes scope over the verb and its arguments and that can be interpreted as the location of tense.

There are two ways in which this condition can be satisfied depending on whether in a language there is an independently available head position with the required properties. On the one hand, such a position may be available in a language that has a special tense particle that dominates the VP, or where there is verb movement across the subject. In the latter case, the position of the moved verb can be interpreted as the location of tense for the purpose of semantic interpretation by virtue of carrying the verb's tense features. On the other hand, no such independently motivated position may be available, which would mean that the Tense Condition must be satisfied in a different way. I propose that in such a situation an abstract tense head must be inserted.

The following examples show the basic structures of French and English. In French, verb movement takes place, which creates a position that has scope over the subject. In English, there is no verb movement, and therefore an abstract T is inserted.[23]

---

[21] Pragmatics may also make available a marked interpretation without overt syntactic or morphological marking. However, such effects are determined by the (non-linguistic) context of an utterance.

[22] See also Wiltschko 2003 for relevant discussion and possibly incompatible proposals.

[23] The underlying difference in verb movement may be related to different properties of inflectional morphology in the two languages. For example, Bouchard (1995:408-411) proposes that tense inflection on the French verb is "strong", which for him means that it can license two syntactic heads, while tense on the verb is weak in English, and only licenses a single projection. (Bouchard does not assume the existence of an

(14) a.   VP   **French**
```
              VP
            /    \
          DP      V'
         Kim_i   /   \
                V     VP
             arrive  /   \
                   t_i    V'
                          |
                        arrive
```

b.   TP   **English**
```
              TP
            /    \
          DP      T'
         Kim_i   /   \
                T      VP
            [-PAST]   /   \
                    t_i    V'
                           |
                         arrives
```
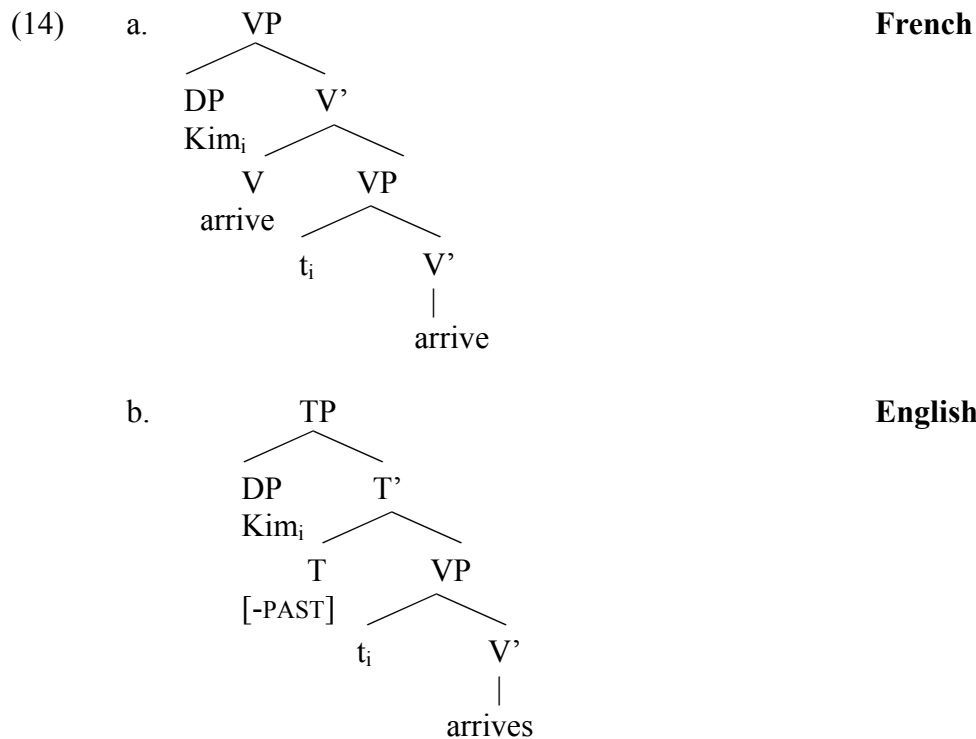
This is then an example of a condition on LF that licenses an abstract head in syntax. (In English T is sometimes realised by auxiliaries.)

3.3 Conclusion: Verb movement configurations

In this section, I discussed the implications of the rejection of the stipulation that moved material cannot project on the analysis of verb movement and on clause structure more generally. I discussed earlier views of verb movement, which involved the assumption that verb movement is always head adjunction. On this view, the adjunction site (the target of verb movement) is either an inflectional affix or an abstract functional head. I argued that the assumption of head adjunction for all cases of verb movement leads to various problems. The Concrete Functional Head view does not succeed in its attempt to reduce inflectional morphology to syntax, because it assumes that morphological selection is possible after movement. This means that a major syntactic generalisation has to be given up. The main problem of the Abstract Functional Head view of verb movement is that it introduces a number of redundancies and feature duplications that cannot be independently

---

abstract T head in English, however.) The relation of so called V-to-T movement and richness of morphology is much debated, and this is not the place to investigate it. See e.g. Koeneman 2000, Bobaljik 2000, Alexiadou and Fanselow 2002 for some discussion.

motivated. Furthermore, the dichotomy of strong/weak features makes the suspect prediction that there is covert head movement. Then, I proposed that verb movement can involve the creation of verbal functional projections headed by copies of the verb. On this view, the occurrence of empty structure is subject to the same constraints as verb movement, and the problems of head adjunction disappear. This view also implies that clause structure is not universal, and hence that it must be learnable. A result of this is that most syntactic heads will have to be overtly realised. Heads licensed by conditions on semantic interpretation may still be empty, and the presence of a head carrying tense with scope over the VP in English finite clauses is a case in point. On this view of clause structure, verb movement always involves the pronunciation of the verb in a "displaced" position α, and two structural verb movement configurations can be distinguished: if the category of α is a copy of the verb, the structure is an s-movement structure. If α corresponds to a different category, the structure is a p-movement structure.[24]

4 Phrase structure

4.1 Categorial projection?

A central assumption in standard theories of phrase structure is that "two nodes are ordered in the left-to-right direction just in case they are not ordered by dominance" (Partee et al 1990:439). In other words, a node A that dominates or is dominated by a node B does not follow or precede B. To implement this assumption it is possible to formulate the following condition on representations of phrase structure:

(15) EXCLUSIVITY CONDITION: In any well-formed tree diagram, for any nodes A and B, A and B stand in the precedence relation if and only if A and B do not stand in the dominance relation. (cf. McCawley 1968:245, Higginbotham 1983:151, Partee et al 1990:440, Speas 1990:23)

---

[24] Note that, where α is a lexical head without phonological content, p-movement is impossible. In such a configuration, pronunciation of the moved category in the higher position would result in a PF identical to that of an s-movement structure in which α corresponds to a copy of the moved category; this would make the presence of α irrecoverable. Thus, there can be no language like English that has an abstract tense head and where the verb is pronounced in the position of T.

The question of what blocks certain overt heads from being targets of p-movement, is the same as the question of what blocks head adjunction to certain heads in a conventional framework. Thus, head adjunction is not possible to heads like English *that* or French *que*. The lack of p-movement to such heads should then receive the same explanation.

Together with the additional assumption that all lexical items must be linearly ordered, this condition directly entails that there must be some device like categorial projection, i.e. that phrase structure representations must contain symbols that do not correspond to word-level categories. A non-terminal node, i.e. a node that dominates other nodes, could not contain a lexical item. So in a structure represented by the following tree, there could be no linear ordering of Y with regard to X or Z, and consequently if X or Z contain a lexical item, Y cannot:

(16)   Y
         / \
        X Z
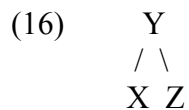
It is then usually assumed that Y inherits its properties from X or Z, i.e. that one of the two projects.

However, the device of categorial projection is not without problems. A first problem is a paradox that concerns selection and the projection level of complements. Since Chomsky 1965, it has usually been assumed that subcategorisation (or selection or theta role assignment) takes place under sisterhood. For example, the verb *read* assigns a theta role to its sister *a book* in (17).

(17)   John [read [a book]]

This leads to the generalisation that "only maximal projections can appear in complement position" (Baltin 1989:(34)) which, as noted by Baltin 1989, is employed in most work on X-bar Theory. This view seems to conflict with another assumption, namely that selection is a head-to-head relation (Muysken 1982:61, Chomsky 1986a:27, Rizzi and Roberts 1989:109, Belletti 2001:503). This view goes back at least to Bresnan's observation that verbs can select for different complement types, i.e. that they do not select a *clausal argument*, but a *clausal argument headed by a particular complementizer* (Bresnan 1970:301 note 5 and the text to which it refers). Thus, while a complement position is usually assumed to be restricted to maximal projections, the selection relation is a relation between a selecting head and the head of its complement, not the complement itself. (See Collins 2003 for a relevant proposal.)

Different questions concern the nature of categorial projection. It has often been claimed that categories of intermediate-level projection are invisible at LF or that they are uninterpretable or that they undergo deletion (Chomsky 1995a, Epstein et al 1998). Similarly, it has been proposed that bar-levels are altogether invisible at LF (Baltin 1991). It has also been proposed that projection levels are not a syntactic primitive at all but that they depend on the structural context of a node, and are read off of the structure, presumably, at LF (Muysken 1982, Speas 1990, Chomsky 1995a, Bury 2002a). This diversity of views suggests that nothing said so far needs to be considered the final word on the issue, and clearly these issues would simply not arise if there were no such thing as categorial projection.[25]

Given the less than solid conceptual and empirical foundations of categorial projection, it seems reasonable to reconsider the status of the Exclusivity Condition in (15), which can only be maintained if something like projection exists, and which, as noted by Partee et al 1990:439f, is not a logical necessity. To this end, a brief look at the history of tree diagrams is illuminating.

The earliest work in generative grammar conceived of sentences as strings of morphemes, that it, as linearised objects. Thus, "[the constituent structure] subcomponent [of a transformational grammar] consists of an ordered set of rewriting rules that generate strings of formatives" (Chomsky 1964:12). For each sentence it is possible to derive a phrase marker based on these phrase structure rules. A phrase structure rule states that a symbol on its left-hand side represents, or should be rewritten as, a string of linearised symbols on its right-hand side. Ultimately, "the set of all those strings which are steps of [a derivation] can serve as a P-marker" (Chomsky 1955/1975:182). To illustrate this point, consider the grammar in (18), the derivation in (19), and the tree in (20), which are all taken from Chomsky 1957:26f:

---

[25] Further problems of categorial projection are discussed in Abels 2001, Brody 1997, 2000. See Carnie 2000

(18)   i.   Sentence → NP + VP
       ii.   NP → T + N[26]
       iii.   VP → Verb + NP
       iv.   T → the
       v.   N → man, ball, etc.
       vi.   Verb → hit, took, etc.

(19)   Sentence
       NP + VP
       T + N + VP
       T + N + Verb + NP
       the + N + Verb + NP
       the + man + Verb + NP
       the + man + hit + NP
       the + man + hit + T + N
       the + man + hit + the + N
       the + man + hit + the + ball

(20)
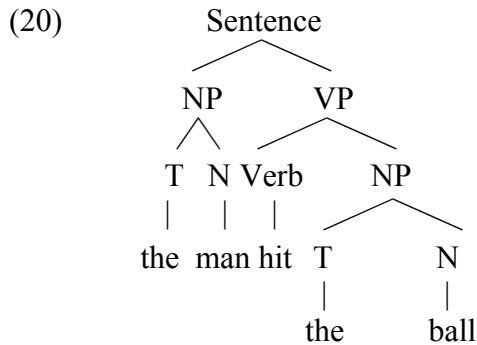
The grammar in (18) generates the strings in (19), which together make up the derivation of the sentence *The man hit the ball*. Thus, line 2 of (19) results from an application of the rule in (18i), line 3 of (19) from the rule in (18ii), and so on. Each new line results from a single rule application. The tree diagram in (20) is then simply a useful representation of (19) that "conveys less information than the derivation [(19)]" but that "retains just what is essential in [(19)] for the determination of the phrase structure (constituent analysis) of the derived sentence *The man hit the ball*" (Chomsky 1957:27f). (20) can be constructed from (19) in the following way. Identical symbols in different lines in (19) (e.g. N in lines 3 and 4) are connected with a vertical line. Similarly, a symbol that occurs on the left-hand side of a rule in (18) is connected with the symbols that occur on the right-hand side of the rule when they are occur in the following line (e.g. NP in line 2 is connected to T and N in line

for a different perspective.

3). Deletion of elements that occur on top of identical elements on the vertical line then yields (20).[27] (See Bach 1964: chapter 3 for useful discussion.)

It is obvious that a tree diagram obtained in this way will meet the Exclusivity Condition. This is so because "in running through the rules one finally produces a line which cannot be further altered by any of the rules because none of its symbols appears on the left hand side of the rule" (Postal 1964:11).

Phrase structure rules turned out to be inadequate because they failed to account for the endocentricity of phrases (Lyons 1968) and duplicated information that is already available in lexical entries. These problems first gave rise to the X-bar schema (Chomsky 1970, Jackendoff 1977) and eventually to the rejection of phrase structure rules in the principles and parameters framework (Chomsky 1981, Stowell 1981, and many others). Moreover, echoing McCawley's 1968 proposals, trees rather than strings of derivations came to be thought of as the true object of syntactic analysis. In spite of these developments, the assumption that representations of phrase structure should meet the Exclusivity Condition continued to be maintained.

The disappearance of phrase structure rules casts further doubts on the Exclusivity Condition in another way. Phrase structure rules were generally taken to produce linearised strings. Thus, Huck 1985 (quoted in Chametzky 1995:170) argues that "phrase structure grammars do not permit rules whose right hand sides are not linearly ordered".[28] However, since at least Stowell 1981, an (often implicit) assumption central to most work in (Chomskyan) generative grammar has been that the relations of precedence and dominance should be kept apart.[29] Phrase structure should only express hierarchical relations among linguistic elements (e.g. Travis 1984, Speas 1990, Chomsky 1993, Chomsky and Lasnik 1993:35, Chametzky 1995, Bobaljik 2002). The idea behind this is that linear order plays no role at the level of logical form, and that it is instead a result of the simple fact that

---

[26] Chomsky's use of *T* here is not to be confused with the contemporary *T*, which usually stands for tense.
[27] Unlike the set of strings in (19), (20) does not indicate in which sequence the rules in (18) have applied. McCawley 1968:246ff notes that this problem is avoided if PS rules are interpreted not as rewrite rules but as node admissibility conditions that define in which context a node is licensed. This insight seems still relevant for the ongoing debate concerning the representational and/or derivational nature of linguistic theory.
[28] But see Gazdar and Pullum 1982.
[29] The separation of order and structure has a long tradition in other frameworks, in particular in dependency grammar (cf. Hudson 1980:192, 1990) but also in GPSG/HPSG and others. Chomsky 1965:123-126 discusses and rejects the possibility of unordered underlying structures. See also the papers collected in Meisel and Pam 1979 for discussion.

humans can pronounce only one sound at a time (Higginbotham 1983:151). Linearisation can then be seen as a function of the mapping of syntactic onto phonological representations. However, once this idea is accepted, there is no reason why two categories that are in a dominance relation in syntax could not be linearly ordered at PF, i.e. why phrase structure should conform to the Exclusivity Condition in (15). So, the nodes X and Z in (16) above can after all be ordered linearly with regard to Y, and hence Y could contain a lexical item. On this view, lexical items can occur in a structure not only as terminal nodes, but also as non-terminals. But if this is the case, there is no longer any need to assume the theoretical construct of categorial projection.

On conceptual grounds independent from those just presented, Brody 1997a, 2000 proposes that there should be no categorial projection, what he calls the TELESCOPE HYPOTHESIS. On this view, a conventional X-bar representation of a category X with a specifier and a complement, like the one in (21), should be represented as (22).[30]

(21)
```
        XP
       / \
    Spec  X'
         / \
        X  Comp
```

(22)
```
        X
       / \
    Spec  Comp
```
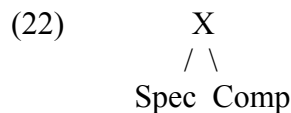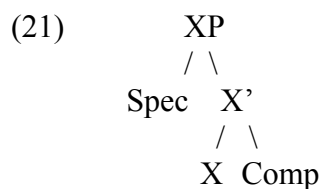
The relation of (21) to (22) can be visualised as a collapsing (or telescoping) of the three nodes of category X in (21) into a single node. X in (22) can then function as both a head and, if it is taken together with the nodes it dominates, as a phrase. However, there are no features or bar-levels in this approach to phrase structure that distinguish word-level and phrase-level categories.[31]

This telescope view of phrase structure is fully compatible with the view of clause structure proposed in the previous section. Given the rejection of categorial projection, the

---

[30] Brody 2000 argues that it may be necessary to add an extra functional head to each lexical category in a telescope structure. On the approach to constituency proposed below this does not seem necessary.
[31] Brody notes that this approach to phrase structure bears some similarity to representations in dependency grammar. For a comparison of the properties of dependency and phrase structure grammars, see Abney 1996, Hudson 1980, Robinson 1970.

notion of "projection after movement" by X must now simply be understood as X's "occurrence as the head of a structure after movement". While in an X-bar theoretic framework a moved verb can project a new VP, the corresponding structure on the telescope view would be one in which the moved copy of a verb dominates the verb's copy in its base position. The clause structure containing derived VPs in (5) above then translates into the telescope structure in (23a), while (23b) illustrates a potential p-movement structure:

(23) a.

```
        V
      /   \
    __      V
          /   \
        X       Y
```

b.

```
        F
      /   \
    __      V
          /   \
        X       Y
```
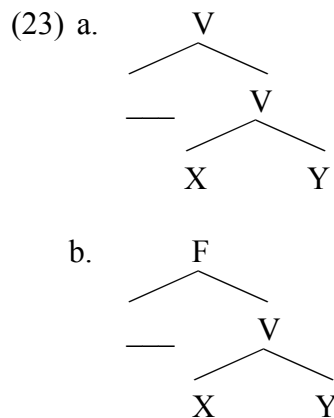
In (23a), the specifier X and the complement Y of the lower copy of the verb are directly dominated by it. In (23b), the verb V is dominated by the head F, and consequently, V can also be pronounced in the position of F.[32]

In the next section, I propose a further modification of the theory of phrase structure. In particular, I argue that, if phrase structure is neutral with regard to precedence, the best way to think of it is in terms of sets of treelets. In addition to leading to a more restrictive theory, this new model has empirical consequences, which fuel the analyses developed throughout this dissertation.

4.2 Phrase markers are sets of treelets

Conventional theories of phrase structure always generate structures that satisfy the Exclusivity Condition, whether they assume some version of X-bar theory (Speas 1990, Kayne 1994) or of "merge as set formation" (Chomsky 1995b). Here, I propose an

---

[32] This of course raises the issue of how specifier and complements can be distinguished; see section 4.3.2.

alternative theory of phrase structure that does not have this property. I assume that a phrase structure is fully determined by a set of linguistic elements called TREELETS[33]. Each treelet consists of a category, called the HEAD CATEGORY, and a statement of which categories the head category dominates, called the DOMINANCE SET.[34]

Syntactic categories contain two types of information, namely categorial features and a set of ATTRIBUTES that states whether the category has any unsatisfied θ-roles (and possibly other grammatically relevant properties to do with binding, case etc.). The Inclusiveness Condition of Chomsky 1995a, Neeleman and van de Koot 2002 constrains the distribution of such information in a structure. Thus, the category of a treelet can correspond to a lexical item inserted in the structure, or it must be inherited from the categories that it immediately dominates,[35] or both.

The dominance set of a treelet is an unordered set that contains all the categories dominated by the treelet's head category.

Each treelet then consists of a set that contains a category and a dominance set. I use angled brackets < >, rather than curly ones { }, to indicate a treelet simply for clarity. Thus, the following represents a treelet of category L $\underline{\theta}_{1\#}\ \theta_2$ whose dominance set is {X,Y}.

(24) $<$L $\underline{\theta}_{1\#}\ \theta_2$, {X,Y}$>$

The categorial features of this treelet are represented as L, and the attributes $\underline{\theta}_{1\#}\ \theta_2$ state that this category has one unsatisfied θ-role θ_2 and one satisfied external θ-role $\underline{\theta}_{1\#}$. (θ-roles are indicated as θ, an external θ-role is underlined, and satisfied θ-roles are marked with a subscripted #.)

---

[33] In Bury 2002b, I call these objects "nodes". This term is somewhat confusing since in conventional usage *node* and *category* are often used interchangeably (modulo the segment/category distinction). In contrast, *treelet* seems less problematic, and in addition it hints at the fact that these objects contain structural information.

[34] This theory is similar to that proposed by Abels 2001, where phrase structures consist of pairs of categories that express immediate dominance relations. Note that while both the present proposal and that of Abels offer a natural home for the rejection of categorial projection, they are in principle independent of it. In contrast, a theory that incorporates certain properties of the two models (namely the dominance pairs of Abels and the reflexive notion of dominance assumed below) may be compatible with projection-free phrase structures only, and hence partially derive the effects of the telescope hypothesis.

[35] This would be the case in s-movement structures, among others.

The head categories of two treelets are identical if they have the same categorial features and the same attributes, regardless of their dominance set. In other words, two categories are distinct if they contain different categorial features, different attributes, or both. Thus, the following categories are all distinct:

(25)   a.   L θ_1 b.   L   c.   L $\underline{\theta}_1$ d.   L $\theta_{1\#}$ e.   K θ_1 f.   L $\theta_1 \theta_2$

Two treelets are identical if they belong to the same category and contain the same dominance set. Conversely, two treelets are distinct if they belong to different categories, contain different dominance sets, or both. The following examples illustrate the range of possibilities:[36]

(26) α = <L, {X,Y}>
         β = <K, {X,Y}>
         γ = <L, {X}>
         δ = <L, {X,Y,Z}>

α is a treelet of category L with the dominance set {X,Y}. The treelet β is distinct from α because it is of a different category. The treelet γ is distinct from α because its dominance set does not contain all of the categories that are contained in the dominance set of α, namely Y is missing. The treelet δ is distinct from α because its dominance set contains a category, namely Z, that is not contained in the dominance set of α.

In contrast to β, γ, and δ, the treelet ε below is identical to α.

(27) ε = <L, {X,Y,Y}>

This is so because ε is of the same category as α and because its dominance set has the same members as the dominance set of α. This last point illustrates a principle of set theory according to which two unordered sets are identical if they contain the same members, regardless of how often they contain a particular member; or in other words "writing the

---

[36] For simplicity, I use simple categories without attributes here; thus, the distinct simple categories K and L here could be replaced by distinct complex categories like L θ_1 and L $\underline{\theta}_1$.
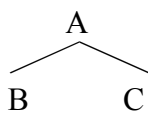
name of a member more than once does not change its membership status" (Partee et al 1990:6). Since this principle will play an important role in the following discussion, it is convenient to give it a name here:

(28) PRINCIPLE OF SET IDENTITY: Two unordered sets are identical if they contain the same members, regardless of how often those members are represented in each set, i.e. {e} = {e,e}.

With treelets the basic building blocks of phrase structure, I take the dominance relation, which holds between categories, to be the basic relation of phrase structure. I follow McCawley 1982, Higginbotham 1983, Partee et al 1990:438, Speas 1990:18-24 and others, and define dominance as a reflexive, transitive and antisymmetric relation. DOMINANCE can then be described by the following axioms (cf. Higginbotham 1983:(2), Speas 1990:20f):

(29) a.   X dominates X (dominance is reflexive)
 b.   if X dominates Y and Y dominates Z, then X dominates Z (dominance is transitive)
 c.   if X dominates Y and Y dominates X, then X = Y (dominance is antisymmetric)
 d.   if X dominates Z and Y dominates Z, then X dominates Y or Y dominates X (or both, if X=Y=Z)

Given these assumptions about categories, treelets and dominance, the conventional tree structure in (30a) should be given formally as the set in (30b):

(30)   a.   A
              B   C

 b.   {<A, {A,B,C}>, <B, {B}>, <C, {C}>}

The first pair, <A,{A,B,C}>, in the set in (30b) states that there is a treelet of category A and that A dominates the categories A, B and C. The second pair states that there is a

treelet of category B and that B dominates the category B. The third pair states that there is a treelet of category C and that C dominates the category C.

In other works, dominance is taken to be a irreflexive relation, that is an item does not dominate itself. Reflexive dominance and irreflexive dominance are equally natural notions. In informal language, the two can be described as follows. Given a set of categories, category α reflexively dominates a category β if α is the *highest* category in the structure containing the categories; hence α can be identical to β. α irreflexively dominates β if α is *higher* than β in the structure containing the categories; here the comparative implies that α must be distinct from β. There also appear to be no formal criteria that favour the definition of dominance in (29) over a definition of dominance as an irreflexive relation.[37] This similarity between the two notions notwithstanding, the choice between the two greatly affects the range of structures a theory describes. An important difference concerns representations that contain a treelet whose head category immediately dominates exclusively an identical category (and, by transitivity, all the categories that this category dominates).

On the view that dominance is reflexive, i.e. that every category dominates itself, the principle of set identity in (28) implies that treelets that contain a category that exclusively dominates an identical category (and all the categories that this category dominates) are identical to treelets that contain only one of the two identical categories. This is so because the dominance set of the two treelets are identical apart from the meaningless double occurrence of a category in the higher of the two treelets, which of course means that formally the two treelets are identical. On the view that dominance is irreflexive, two such treelets would be distinct, because only the dominance set of the higher of the two treelets would contain the shared category. A comparison of the conventional tree in (30a) above to the one in (31) helps to illustrate this point:

---

[37] Partee et al 1990:438 note that reflexive dominance is usually assumed "as a technical convenience".

The predecessor of dominance, the relation ρ in Chomsky 1955/1975:173, which relates strings in a derivation, is irreflexive. In Bach 1964, this relation is reflexive. See also the replies to Oh 1977 in *LI* volume 9. Chomsky 1986a:92 note 11, 1995b:338 assumes irreflexive dominance without discussion. Kayne 1994:134, referring to Chomsky 1986a, takes dominance to be irreflexive. Abels 2001 and Uriagereka 1998

(31)
```
          A
          |
          A
         / \
        B   C
```
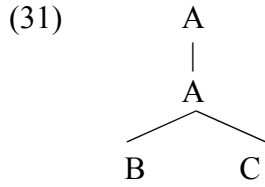
The conversion of the trees to the set notation assumed here would be straightforward if dominance were taken to be irreflexive. The tree in (30a) above would then correspond to the set in (32a), and the tree in (31) to the set in (32b). Crucially, the two distinct trees would correspond to two distinct sets.

(32)   a.   {<A, {B,C}>, <B, {}>, <C, {}>}

        b.   {<A, {A,B,C}>, <A, {B,C}>, <B, {}>, <C, {}>}

However, on the view, adopted here, that dominance is reflexive, this is not so. A simple translation of the trees to phrase structure sets yields the following sets:

(33)   a.   {<A, {A,B,C}>, <B, {B}>, <C, {C}>}

        b.   {<A, {A,A,B,C}>, <A, {A,B,C}>, <B, {B}>, <C, {C}>}

While there is nothing special about the set in (33a), (33b) is more interesting. The first treelet in this set is <A, {A,A,B,C}>. In an attempt to preserve the information expressed by the tree in (31), the dominance set of this treelet contains the category A twice: first, to express the fact that in the corresponding tree the category of the highest node dominates the category A of the second highest node, and second to express the fact that category A of the top node dominates itself. The result of this attempt is an unordered set that contains two occurrences of the same element. However, due to the set identity principle in (28), it is clear that this set is identical to the set containing A only once, i.e. {A,A,B,C} = {A,B,C}. So, instead of {A,A,B,C}, a more accurate representation of the tree in (31) should have {A,B,C} as the dominance set of its first treelet:

---

argue for an irreflexive dominance relation. According to Chametzky 2000:39 note 73, it is "normal practice in formal grammar theory" to take dominance to be reflexive.

(33)   b'.   {<A, {A,B,C}>, <A, {A,B,C}>, <B, {B}>, <C, {C}>}

Now, the phrase structure set itself contains two identical members, namely the two treelets <A, {A,B,C}>, which are meant to represent the highest and second highest nodes in the tree in (31). Again, the set identity principle implies that this set is identical to a set that contains only one of the two identical members. Accordingly, (33b') should be more accurately represented by the following set, which is the same as the set in (33a) that corresponds to the tree in (30a):

(33)   b''. {<A, {A,B,C}>, <B, {B}>, <C, {C}>}   = (33a)

So, ultimately, it turns out that, on the view that dominance is reflexive, the distinct tree structures in (30a) and (31) represent the same phrase structure set and are hence identical.

    Anticipating some of the proposals to be made below, note that this property of the "phrase structures as sets" system means that it is now possible to allow clause structure to be extended freely through the addition of copies of the root category. (This is then somewhat similar in spirit to Speas' 1990, 1995 "project alpha" theory; however, it is not necessary here to constrain the extension of phrase structures by means of an economy condition.) The type of representation assumed here entails that the vacuous addition of an identical category will always collapse to a simpler structure. However, if an identical category is added to the root of a structure and this new category dominates additional material, the representations will not collapse. (Of course the addition of a new category at the root will also change the structure.) The freely available addition of copies at the root plays an important role in the analyses of adjunction and verb movement presented below. Adjunction can involve the creation of a copy if no other adjunction site (specifier position) is available. Verb movement (s-movement) can involve the creation of a copy if no other head position is available. (It should be noted that while there is no constraint on the extension of clause structure through the addition of copies, general constraints on interpretation and interpretability will hold of such representations; see section 6.2.)

    A similar thing to the collapse of the trees in (30) and (31) happens when the following tree is converted to set notation:

(34)
```
        B
      /   \
     A     C
           |
           A
```
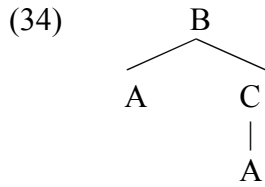
An intuitive translation yields the following set:

(35)   {<B, {B,A,C,A}>, <A, {A}>, <C, {C,A}>, <A, {A}>}

As will be clear by now, the set identity principle in (28) implies that the repeated occurrences of the category A in the dominance set of the first treelet and of the treelet <A, {A}> in the phrase structure set are meaningless. A more appropriate expression of the information contained in this set is then the following:

(36)   {<B, {B,A,C}>, <A, {A}>, <C, {C,A}>}

However, a translation of this set back into tree notation yields not the tree in (34) above, but the following:

(37)
```
     B
     |
     C
     |
     A
```
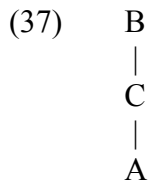
This is then a further case where distinct tree diagrams correspond to a single representation in terms of the sets proposed here.

This discussion shows that some differences in representations that can be expressed in a theory that takes as its primitive relation that of irreflexive dominance cannot be expressed if dominance is taken to be reflexive. This means that assuming dominance to be reflexive leads to a more restrictive theory (of representations of dominance). Similarly, this discussion shows that the view that phrase structures correspond to sets of treelets expressing dominance relations between categories is more restrictive than the conventional view according to which phrase structure corresponds to tree structures. The

trees in (30a) and (31) above are both perfectly reasonable and, crucially, distinct objects. However, the difference that they express does not carry over to the model proposed here.

The greater restrictiveness of the theory proposed here compared to theories that use either traditional tree diagrams or a non-reflexive notion of dominance means that, at this level of discussion, the present proposal is preferable to the alternatives. The major goal of this dissertation is to show that in addition to this conceptual advantage, this theory also leads to analytical advances in a range of empirical domains.[38]

This view of phrase structure also converges with the model of Bare Phrase Structure (Chomsky 1995a) in that both models exclude vacuous projection. That is, a structure containing a category that exclusively dominates an identical category is identical to a structure containing only one of the two categories. Unlike Bare Phrase Structure, this model of course allows for non-branching structures. This is so because here lexical items can occur as non-terminals, a consequence of the rejection of the Exclusivity Condition in (28).

Note that throughout the dissertation I will not always give full representations in terms of phrase structure sets. Instead, to keep the discussion as clear as possible, I attempt to use the simplest possible representation appropriate to make a particular point. Thus, where necessary I will use phrase structure sets, but at times I will use tree structures or labelled brackets, or an appropriate mix of the three.

4.3 Structural relations

4.3.1 Basic definitions

In this theory, phrase structure is fully determined by the phrase structure sets that contain treelets, and by the dominance relations encoded within each treelet. Given these primitives and the notion of dominance in (29) above, the derived relation of immediate dominance can be defined over head categories of treelets as follows.

(38) X immediately dominates Y if X and Y are head categories of distinct treelets and X dominates Y and there is no category Z distinct from X and Y such that X dominates Z

---

[38] Some consequences of the discussion of the trees in (30a) and (31) for the analysis of verb movement and of the discussion of the tree in (34) for "XP" movement will be discussed shortly.

and Z dominates Y.

It should be noted that, since immediate dominance is a derived notion, the phrase structure sets contain no separate statement of immediate dominance relations. Among other things, this means that there can be no two phrase structure sets that are distinguished only by their immediate dominance relations.[39]

Given the derived notion of immediate dominance, the usual structural relations can be defined as follows:

(39) a. X is the DAUGHTER of Y, if Y immediately dominates X.

b. Y is the MOTHER of X, if Y immediately dominates X.

c. X and Y are SISTERS if they are distinct and have the same mother.

Moreover:

(40) a. X is the ROOT of the structure if it is not dominated by a distinct category.

b. X is a TERMINAL if it does not dominate any other categories.

c. X and Y are CONNECTED if X is dominated by Y or Y is dominated by X.

The assumptions of the present proposal require a slight reformulation of the c-command relation (cf. Reinhart 1976). Since the rejection of categorial projection allows lexical heads to dominate other categories, there is no reason to exclude a category from c-commanding categories that it dominates. The relation can then be formulated as follows:

(41) X C-COMMANDS Y iff there is no category Z such that Z immediately dominates X and Z does not dominate Y.

Phrase structure sets can then be constrained by the following conditions (cf. Partee et al 1990:437-441 on properties of tree diagrams).

---

[39] This is the reason why the tree structures in (30) and (31) collapse to a single phrase structure set.

(42) SINGLE ROOT CONDITION: There is exactly one treelet whose dominance set contains all the categories contained in the phrase structure set.

(43) BINARY BRANCHING: Each category has at most two daughters.

Note that the Single Root Condition implies that every category is connected to the root.

It should be noted that while the above notions can all be defined within the model proposed here, it is not necessarily the case that all of them actually have a linguistic reality. The effects of some of them may turn out to be derivable from more primitive notions, while others may turn out to be empirically unfounded. A case of the former is the notion of c-command, which has been derived (or, in one case, eliminated) in different ways in Brody 1997a, 2000, Epstein et al 1998, Neeleman and van de Koot 2002. A case of the latter is the Single Root Condition, which Chametzky 1995 argues against. Since these considerations have no direct bearing on the issues discussed here, I will continue to use these standard notions in order not to detract from innovations in other places.

4.3.2 Specifiers and complements

Now we can turn to the distinction between specifiers and complements.[40] I assume, as in standard X-bar theory, that each head is related to at most two dependents, namely a single specifier/adjunct and a single complement (cf. also Kayne 1994, Cinque 1999, contra Chomsky 1995b). Unlike in conventional approaches, however, it is not possible in the present theory to distinguish *structurally* a specifier from a complement. Thus, while in a conventional structure, the complement is sister to a head and the specifier sister to the first projection of the head, in the model without categorial projection developed here, both specifier and complement are daughters to the head:

---

[40] I am grateful to Klaus Abels for discussion of this issue. He proposed a different approach to the specifier/complement distinction in terms their different status with regard to the functional sequence in classes at the egg school in Novi Sad 2002.

(44)   a.   X-bar structure: XP
                                            / \
                                        Spec   X'
                                               / \
                                              X Comp

        b.   'Telescope' structure   X
                                            / \
                                        Spec Comp

Nevertheless, the complement of a head can be identified through reference to EXTENDED PROJECTIONS (cf. Grimshaw 1991). It has long been recognised that verbs and nouns are related to different functional categories. Thus, e.g. tense and complementizers go with verbs, while determiners go with nouns. A verb is said to form an extended projection with the functional categories to which it is related. In the present proposal, a (non-derived) lexical verb then forms an extended projection with all its derived heads and with associated functional heads such as auxiliaries and complementizers.[41] Intuitively, the complement of a branching category Z is then that daughter of Z that is part of the same extended projection.

5 Derived heads

The simplest phrase structure contains a single category. In standard frameworks, the creation of more complex structures involves an operation called MERGE (e.g. Chomsky 1995). Merge combines two items and forms a set containing the two with a label provided by one of them. This operation is designed to create structures that satisfy the Exclusivity Condition discussed and rejected in section 3.1. This operation is therefore incompatible with the Telescope version of phrase structure adopted here, in which lexical items can appear as non-terminals.

In place of the operation merge, I propose that phrase structure can be extended through the addition of new categories. As noted above, the Single Root Condition requires that each category be connected to the root of the structure. This means that the root must either dominate the new category or that the new category must dominate the root. I assume that only the root is accessible for the extension of phrase structure in this way.

---

[41] The notion of extended projections is formalised in Grimshaw 1991, van Riemsdijk 1998, Ernst 2002:244-246, Matushansky 2002:20f, and it plays a central role in Frank 2002: chapter 2. See Chametzky 2000 for some general discussion.

Given the structure in (45a), extension of the structure through addition of α can then give rise to the structure in (45b,c) but not to (45d):

(45) a. $\{<F, \{F, \beta\}>, <\beta, \{\beta\}>\}$

  b. $\{<F, \{F, \beta, \alpha\}>, <\beta, \{\beta\}>, <\alpha, \{\alpha\}>\}$

  c. $\{<\alpha, \{\alpha, F, \beta\}>, <F, \{F, \beta\}>, <\beta, \{\beta\}>\}$

  d. * $\{<F, \{F, \beta, \alpha\}>, <\beta, \{\beta, \alpha\}>, <\alpha, \{\alpha\}>\}$ (as derived from (45a))

```
a'.  F              b'.  F            c'.  α            d'.  * F
     |                   / \               |                 |
     β                  α   β              F                 β
                                           |                 |
                                           β                 α
```
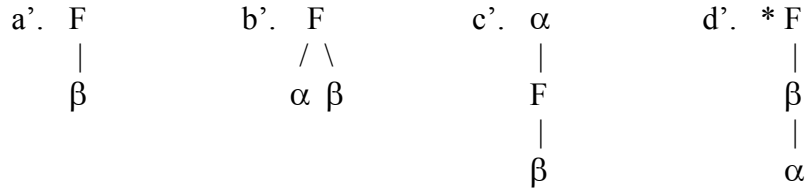
(45d) cannot be derived from (45a) because the addition of α affects not only the root but also the dominance set of the non-root treelet of β.

   Apart from the restriction to occur at the root, the addition of a treelet can apply freely (cf. Speas 1990, 1995). It need not be triggered, and no special constraints on such extensions need be assumed. However, the set identity principle in (28) entails that (45c) is distinct from (45a) only if α is a category distinct from F. As noted in section 4.2, the addition of an identical category is only possible if the new category dominates material that the original category does not dominate. Thus, (45c) is only possible where α corresponds to an additional lexical item.

   However, clause structure can be extended without the insertion of a new lexical item at the root, as well. This is what gives rise to DERIVED HEADS, heads that do not correspond to a lexical item.

5.1 Selection

More complex structures than those in (45) are involved in selection. I assume that if α selects β, the two must be sisters.[42] Selection then involves a configuration of three categories: a selecting category, a selected category, and their mother. The simplest

example of selection is an intransitive verb with an external argument. This configuration can be represented by the following phrase structure set and informal tree:

(46)   a. $\{<V\ \underline{\theta}_{1\#},\ \{V\ \underline{\theta}_{1\#},\ Subj,\ V\ \underline{\theta}_1\}>,\ <Subj,\ \{Subj\}>,\ <V\ \underline{\theta}_1,\ \{V\ \underline{\theta}_1\}>\}$

        b.
$$
\begin{array}{c}
V\ \underline{\theta}_{1\#} \\
\diagup\quad\diagdown \\
Subj\quad V\ \underline{\theta}_1
\end{array}
$$

V $\underline{\theta}_1$ and Subj are sisters. V $\underline{\theta}_1$ represents an intransitive verb with an unsatisfied external θ-role and Subj represents the subject that satisfies it. In V $\underline{\theta}_{1\#}$, the subscripted # indicates that the external θ-role is satisfied. The different status of their θ-roles means that V $\underline{\theta}_{1\#}$ and V $\underline{\theta}_1$ are distinct categories. The root in (46) is then distinct from V $\underline{\theta}_1$, the daughter of which it is a derived head.

    To illustrate further how selection works, representations of transitive and ditransitive verbs are given here:

(47)   a. $\{<V\ \underline{\theta}_{1\#}\theta_{2\#},\ \{V\ \underline{\theta}_{1\#}\theta_{2\#},\ Subj,\ V\ \underline{\theta}_1\theta_{2\#},\ Obj,\ V\ \underline{\theta}_1\ \theta_2\}>,\ <Subj,\ \{Subj\}>,\ <V$
       $\underline{\theta}_1\theta_{2\#},\ \{V\ \underline{\theta}_1\theta_{2\#},\ Obj,\ V\ \underline{\theta}_1\ \theta_2\}>,\ <Obj,\ \{Obj\}>,\ <V\ \underline{\theta}_1\ \theta_2,\ \{V\ \underline{\theta}_1\ \theta_2\}>\}$

        b.
$$
\begin{array}{c}
V\ \underline{\theta}_{1\#}\theta_{2\#} \\
\diagup\quad\diagdown \\
Subj\quad V\ \underline{\theta}_1\theta_{2\#} \\
\diagup\quad\diagdown \\
Obj\quad\ V\ \underline{\theta}_1\ \theta_2
\end{array}
$$

(48)   a. $\{<V\ \underline{\theta}_{1\#}\theta_{2\#}\ \theta_{3\#},\ \{V\ \underline{\theta}_{1\#}\theta_{2\#}\ \theta_{3\#},\ V\ \underline{\theta}_1\theta_{2\#}\ \theta_{3\#}, Obj_b,\ V\ \underline{\theta}_1\ \theta_2\ \theta_{3\#},\ Obj_a,\ V\ \underline{\theta}_1\ \theta_2\ \theta_3\}>,$
       $<Subj,\ \{Subj\}>,\ <V\ \underline{\theta}_1\theta_{2\#}\ \theta_{3\#},\ \{V\ \underline{\theta}_1\theta_{2\#}\ \theta_{3\#},\ Obj_b,\ V\ \underline{\theta}_1\ \theta_2\ \theta_{3\#},\ Obj_a,\ V\ \underline{\theta}_1\ \theta_2\ \theta_3\}>,\ <Obj_b,$
       $\{Obj_b\}>,\ <V\ \underline{\theta}_1\ \theta_2\ \theta_{3\#},\ \{V\ \underline{\theta}_1\ \theta_2\ \theta_{3\#},\ Obj_a,\ V\ \underline{\theta}_1\ \theta_2\ \theta_3\}>,\ <Obj_a,\ \{Obj_a\}>,\ <V\ \underline{\theta}_1\ \theta_2\ \theta_3,\ \{V$
       $\underline{\theta}_1\ \theta_2\ \theta_3\}>\}$

---

[42] Neeleman and van de Koot 2002:531-536 decompose sisterhood in terms of the more primitive relation of immediate dominance in a way that seems compatible with the current proposal.

b.   V $\underline{\theta}_{1\#}\theta_{2\#}\,\theta_{3\#}$
          /   \
      Subj   V $\underline{\theta}_1\theta_{2\#}\,\theta_{3\#}$
              /   \
          Obj   V $\underline{\theta}_1\,\theta_2\,\theta_{3\#}$
                  /   \
              Obj$_a$ V $\underline{\theta}_1\,\theta_2\,\theta_3$
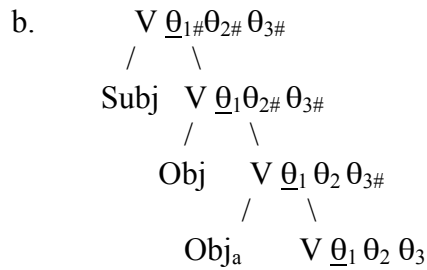
In these structures, selection gives rise to a number of derived heads. In (47), selection by the transitive verb V $\underline{\theta}_1\,\theta_2$ gives rise to the derived heads V $\underline{\theta}_1\theta_{2\#}$ and V $\underline{\theta}_{1\#}\theta_{2\#}$. In (48), selection by the ditransitive verb V $\underline{\theta}_1\,\theta_2\,\theta_3$ gives rise to the derived heads V $\underline{\theta}_1\,\theta_2\,\theta_{3\#}$, V $\underline{\theta}_1\theta_{2\#}\,\theta_{3\#}$, and V $\underline{\theta}_{1\#}\theta_{2\#}\,\theta_{3\#}$.

These structures resemble two different types of structures that have been proposed in the literature, although they differ from both in important ways. As in the light verb structures of the type proposed by Hale and Keyser 1993, Chomsky 1995b and others, several head positions are associated with a single lexeme/verbal stem, and each of the positions is related to a single argument of the verb. However, the structures proposed here differ from standard light verb structures in that the higher head positions associated with the lexeme do not correspond to distinct lexical items. Instead, here, these heads are derived heads licensed configurationally (or derivationally) through selection. Thus, there is for example no lexical item that corresponds to the head V $\underline{\theta}_1\,\theta_{2\#}$ in (47) above. Nevertheless, the properties of these heads mean that they can be ordered as part of the universal sequence of functional categories.

The second similarity is with structures that involve the recursive projection of a single lexical item, where e.g. a single verb can head several separate projections (e.g. Larson 1988, Ackema et al 1993, Bouchard 1995, Koeneman 1995). The structures proposed here differ from such structures in that the derived heads here are categorially distinct from the original head. Structures more similar to the "recursive projection" type will be discussed in the next section.

5.2 Verb movement

The model of phrase structure developed in section 4.2 allows clause structure to be extended freely through the addition of copies (of categories) at the root. I argued in

section 3 that verb movement always involves the pronunciation of the verb in the position of a different head α and that two verb movement configurations could be distinguished depending on whether α corresponds to an independent head (p-movement) or to a copy of the verb (s-movement). In this section I discuss s-movement structures as an example of a structure created through the addition of a copy.

The creation of a copy of the verb on top of the original verb is illustrated informally in the following tree diagrams:

```
(49)    a.   V                  b.       V
                                         |
                                         V
```

However, the tree in (49b) contains two identical categories, and it was already noted earlier that the representation of identical categories is not a trivial matter. Thus, the following trees represent a single phrase structure set:

```
(50)    a.   V                  b.       V
              \                          / \
               V                        X   Y
              / \
             X   Y
```

Here, the higher copy of V in (50a) is supposed to correspond to a moved verb, the lower copy to its trace; X is the specifier of the lower copy and Y its complement. (51) is a set that resembles the tree in (50a) in that the tree's four nodes are matched by four treelets within the phrase structure set.

(51)   {<V, {V,V,X,Y}>, <V, {V,X,Y}>, <X, {X}>, <Y, {Y}>}

The first treelet in the set, <V,{V,V,X,Y}>, corresponds to the highest node in the tree, the second treelet in the set, <V,{V,X,Y}>, corresponds to the second highest node in the tree, and so on. However, following the discussion in section 4.2, it will be clear that not all is well with this set. In particular, the dominance set of the first treelet contains two occurrences of a category V. Given the set identity principle in (28), it is clear that this set

is identical to a set that contains V only once. Thus, the first treelet could be more accurately represented as <V, {V,X,Y}>, which would lead to the following set:

(51)'   {<V, {V,X,Y}>, <V, {V,X,Y}>, <X, {X}>, <Y, {Y}>}

However, this only shifts the problem to the level of the phrase structure set. This set contains two occurrences of the treelet <V, {V,X,Y}>, and obviously the set identity principle implies that this set should rather be represented by the following:

(51)''   {<V, {V,X,Y}>, <X, {X}>, <Y, {Y}>}

Alas, this set seems to correspond most simply to a tree in which there is only one copy of the verb, namely that in (50b).

Thus, this is again the situation in which two distinct tree structures collapse to a single set. Since the sets in (51), (51') and (51'') that are meant to express a potential verb movement structure are identical to a set where no verb movement is possible (because no additional head position is available), the representation of s-movement requires a little more work.

The following set differs from those in (51) only in that there is an additional category Z that is dominated by the higher V, which means that the dominance set of the first V treelet also contains Z.

(52) a.   {<V, {V,Z,X,Y}>, <Z, {Z}>, <V, {V,X,Y}>, <X, {X}>, <Y, {Y}>}

   b.   V
         / \
        Z   V
           / \
          X   Y
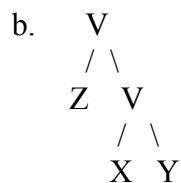
The crucial effect of the additional category Z is that now the treelet that is headed by the higher copy of the verb and the treelet headed by its lower copy have different dominance sets and hence are not identical. Thus, the set in (52a) is a possible representation of a s-

movement structure.

This discussion shows that potential s-movement configurations are exactly those in which a copy of the verb has an additional specifier; intuitively the new specifier is required to distinguish two otherwise non-distinct categories. The theory then implies the following theorem:

(53) Where s-movement takes place, an additional specifier must be created.

In chapter 3, I use this theorem to explain the correlation of verb movement and obligatory specifiers in verb-second contexts.

5.3 Adjunction

Adjunction involves the creation of a new dominance relation between the adjunct and another category. The major structural constraint on adjunction is the Binary Branching condition in (43).

In a structure where no suitable specifier position is available, adjunction may require the creation of a new head position. Schematically, adjunction may then give rise to the following structure:

(54)   a. α b.   {<α, {α, Adj, F, β}, <Adj, {Adj}>, <F, {F, β}>,
                 / \   <β, {β}>}
            Adj   F
                   |
                   β

Here, the category α has been added to the structure to make adjunction to F possible. On the flexible view of clause structure assumed here, not every language will contain a lexical item that corresponds to the category α. Given the Inclusiveness Condition (Chomsky 1995:228, Neeleman and van de Koot 2002), this means that this category must correspond to the category of one of its daughters in languages where no such lexical item is available. It is natural to assume that the category is inherited from its non-adjunct daughter, F above, rather than from the adjunct. (For example, a clause with an adjoined adverb typically still behaves like a clause, not like an adverb.) Unlike insertion of an

argument, insertion of an adjunct does not change the attributes of a category, in particular an adjunct is not selected and does not satisfy an argument role of the category targeted by adjunction. Consequently, the category of the root after adjunction will be the same as the category of the root before adjunction. Thus, in the above structure, α is identical to F, and the treelets of the two cateogries only differ in their dominance sets.[43] A slightly less abstract representation of such a structure would be the following:

(55)   a. $\{<V\ \underline{\theta}_{1\#},\ \{V\ \underline{\theta}_{1\#},\ \text{Adj},\ \beta\},\ <\text{Adj},\ \{\text{Adj}\}>,\ <V\ \underline{\theta}_{1\#},\ \{V\ \underline{\theta}_{1\#},\ \beta\}>,\ <\text{Subj},$
     $\{\text{Subj}\}>,\ <V\ \underline{\theta}_1,\ \{V\ \underline{\theta}_1\}>\}$

     b.   V $\underline{\theta}_{1\#}$
               / \
          Adj   V $\underline{\theta}_{1\#}$
                    / \
               Subj   V $\underline{\theta}_1$
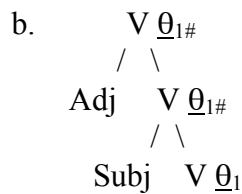
Here, the adjunct Adj is adjoined on top of an intransitive verb phrase. The category dominating the adjunct and the adjunct's sister are identical; the corresponding treelets differ in their dominance sets.

The creation of a new head position to allow for adjunction bears some similarity to Kayne's 1994 analysis of adjunction. In particular, Kayne 1994:30 proposes that certain functional heads may be a consequence of the lack of adjunction sites. While Kayne doesn't discuss the nature of this type of functional head, it should be noted that on many current proposals, such heads are lexical items, a view that clearly contrasts with the present proposal (see also Ackema et al 1993, Nash and Rouveret 1997). Since such heads have no overt realisation (in English) and there appears to be no universal semantic condition that forces their presence, it is not clear what information their lexical entry should contain. This issue does not arise on the view proposed here that clause structure can be freely extended in principle.

It is a striking feature of this analysis that adjunction can involve identical structures to those involved in s-movement. In both structures there is a category immediately

---

[43] There may of course be structures in which adjunction targets the specifier of a lexical category. A relevant example may be Greek, where a fronted focused constituent in an embedded clause may precede the complementizer *oti* (cf. Roussou 2000).

dominating an identical category. The two constructions differ simply in which of the identical categories is pronounced. A further property of these structures will be discussed in the following section.

6 Chains and selection

In much recent work, chains are taken to consist of multiple copies of a single element that stand in a relation of c-command (e.g. Chomsky 1993, Brody 1995, Bobaljik 2002). While the standard conditions on chains, i.e. the head must c-command the tail, locality, etc., carry over, this view is only partially compatible with the present proposal. In this section, I discuss two types of chains, namely specifier chains, which correspond to conventional XP chains, and head chains, and the selection of chains.

6.1 Specifier chains

The copy view of SPECIFIER CHAINS is not compatible in an obvious way with the model of phrase structure developed above. It was noted above that the tree structures in (56), repeated from (34) and (37), cannot be distinguished. If these trees are taken to represent phrase structure sets, both are representations of the set in (57):

(56) a.
```
        B
       / \
      A   C
          |
          A
```
b.
```
      B
      |
      C
      |
      A
```

(57)   {<B, {B,A,C}>, <A, {A}>, <C, {C,A}>}
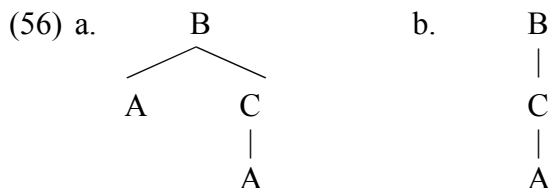
The tree in (56a) resembles a structure containing a "phrasal" copy chain. Thus, the two instances of A can be interpreted as occupying the specifier positions of B and C respectively. Since (56a) collapses with (56b), which would correspond to a structure where A only occurs in the specifier of C, (56a) is not a possible representation of a phrasal movement chain. This means that in the present model, specifier chains cannot be expressed in terms of identical copies.

    Whether this is a desirable consequence or not is not central to the analyses to be

developed here, and I will therefore simply sketch some possible alternative representations of chains.[44]

There are different ways in which such specifier chains can be expressed in the present model. Taking the case of *wh* movement to the specifier of C, the lower position of the *wh* chain can be occupied by a special trace category, as in the simplified structure in (58). Since the two chain positions contain distinct categories, they do not collapse into a single treelet.

(58) a.

$$
\begin{array}{cc}
 & C \\
\textit{wh} & V \\
 & | \\
 & t
\end{array}
$$

b.   {<C, {C, *wh*, V, *t*}>, <*wh*, {*wh*}>, <V, {V, *t*}>, <*t*, {*t*}>}
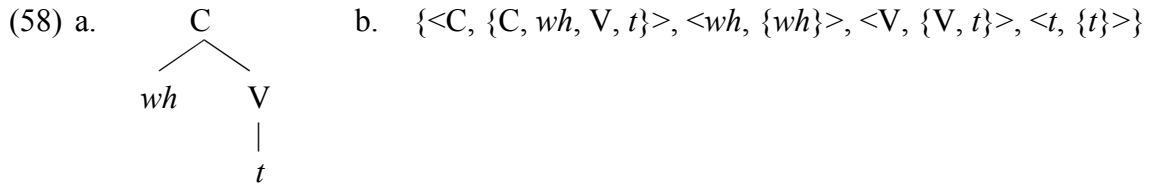
On this approach, chains are not strictly syntactic objects but are formed only at the semantic interface (LF).

A slightly different approach would be to interpret traces as unsatisfied functions that are copied upwards in a structure until they are in the right configuration with an appropriate antecedent (cf. Neeleman and van de Koot 2002 and the slash feature of HPSG). In (59) the trace function is indicated as F (wh). The major difference between such an approach and the previous one is that chain formation is not a matter of LF interpretation, but it is represented directly in syntactic structure.

(59)

$$
\begin{array}{c}
\text{C F (wh)}_{\#} \\
\textit{wh} \qquad \text{V } \underline{\theta}_{1\#}\theta_{2\#}\text{ F (wh)} \\
\diagup \qquad \diagdown \\
\text{Subj } \text{ V } \underline{\theta}_{1}\theta_{2\#}\text{ F (wh)} \\
\diagup \qquad \diagdown \\
\text{F (wh) V } \underline{\theta}_{1}\ \theta_{2}
\end{array}
$$
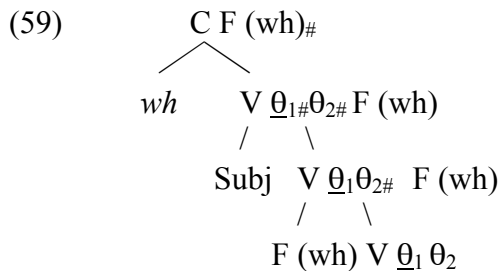
Finally, another possibility would be the insertion of a single category split across different positions within a structure (cf. Cormack and Smith 1997). Thus, a category X with the

---

[44] Note that Cormack and Smith 2002 argue on independent grounds against the copy theory of movement. Some relevant discussion of chains and copies can be found in Brody 1997b, Gärtner 1998, Kracht 2001, Nunes 1995.

features F_1 and F_2 could be split into two parts that contain F_1 and F_2, respectively, and be inserted as in (60). Thus, one part of the category X occurs as the specifier of C, and another, distinct part of X occurs as specifier of V:

(60)

```
            C
          /   \
    X {F₁}     V
             /  \
        X {F₂}    α
```
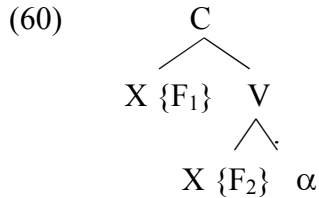
As Cormack and Smith 1997:244 note 40 note, this idea is similar to the copy theory of movement but more restrictive. As under the copy theory of movement, features of a single category can occur in different positions. However, while the copy theory of movement entails that all features occur in all positions of the chain, this is not the case on the split-insertion approach. This means that the latter model does not necessarily require an additional mechanism that deletes duplicated uninterpretable features of multiple copies.

What all these approaches have in common is that the positions of a specifier chain are not occupied by identical elements. Consequently, the fact that usually only one of the positions is pronounced is less surprising than on the copy-view of chains. For the purposes of this dissertation the choice between these options is not relevant, and I will simply represent the lower position in a specifier chain as *t*. What is important however is that the theory contains some mechanism that will allow the two positions to be interpreted as a chain.

6.2 Head chains

As argued in section 4.2, there is no structural constraint on the extension of clause structure through the addition of copies of the root category, which is a feature of adjunction and s-movement structures. This means that while the present model of phrase structure is not compatible with a copy approach to specifier chains, it does allow for the presence of copies of categories.

This raises questions about the interpretation of such structures. While in the theory proposed here all lexical items have some intrinsic semantic content, it is not clear what interpretation the copy of a category could receive. It seems clear that copies cannot be

assigned the same interpretation as the original category. If that was the case, e.g. the copy of a transitive verb would select its own internal argument, just like the original. This suggests that copies of categories cannot receive an interpretation independent from their original. The extension of a structure through the addition of copies can then be seen as the creation of uninterpretable head positions. I assume that structures generated in this way can only satisfy the principle of full interpretation if the copies form a HEAD CHAIN with the original. Chain formation then does two things. First, it ensures that all elements in a representation can receive *some* interpretation, which in the case of copies is an interpretation as part of a head chain. Second, it ensures that copies are assigned the *correct* interpretation, in particular that they do not receive the same interpretation as the original.

Thus, while copies of categories are freely available, their existence gives rise to head chains. This result is relevant for s-movement and adjunction structures. (61) illustrates a structure from a verb-second language, where the verb moves in front of the subject and the specifier of the moved verb contains some other constituent X:

(61)
$$
\begin{array}{c}
V\ \underline{\theta}_{1\#} \\
/\quad\backslash \\
X\quad V\ \underline{\theta}_{1\#} \\
\quad/\quad\backslash \\
\text{Subj}\quad V\ \underline{\theta}_{1}
\end{array}
$$
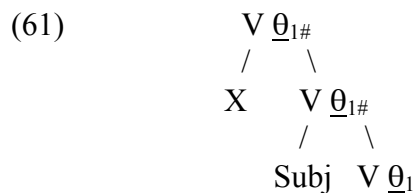
Since X is not selected by the verb, the properties of the category that dominates X are the same as those of X's sister; both are V $\underline{\theta}_{1\#}$; in other words, the higher category is a copy of the lower one. Thus, in such a structure the root category must form a head chain with its daughter.

In the previous section, it was shown that adjunction structures are similar to s-movement structures because they also involve structures extended through the addition of a copy. Thus, the following tree illustrates an English clause with an adverb that is adjoined to the category in whose specifier the subject sits. Since the adverb is not selected the properties of the category dominating it are identical to those of its sister.

(62)
$$
\begin{array}{c}
\text{T}\\
\diagup\quad\diagdown\\
\text{Adv}\qquad \text{T}\\
\diagup\quad\diagdown\\
\text{Subj}_i\qquad \text{V}\ \underline{\theta}_{1\#}\\
\diagup\quad\diagdown\\
t_i\qquad \text{V}\ \underline{\theta}_1
\end{array}
$$
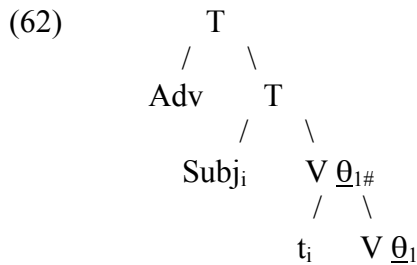
As in the s-movement example, the occurrence of the copy triggers the formation of a head chain which contains the (copied) root category and its (original) daughter.

6.3 The selection of chains

A standard constraint on chains is that only their most deeply embedded member can enter into selectional relations (e.g. Brody 1995, Jackendoff 1997; see chapter 2 for more detailed discussion). This can be formulated as the following condition:

(63)   If X selects Y, both X and Y must be the lowest members of their chains.

This condition can be decomposed into two claims. First, it states that α cannot select β if α is not the lowest member of its chain. Second, it states that α cannot be selected if it is not the lowest member of its chain.

The first statement restricts the distribution of moved elements and other things that occur in a higher chain position. A typical example of this is the ban on movement to θ-positions. Thus, a noun phrase cannot move into a position selected by a higher verb. On the view that there is no categorial projection and that verb movement can create new positions, this part of the condition makes another prediction. Selection is a relation between heads, and on the view that there is no structural distinction between heads and phrases, there can be no way to distinguish a structure in which a phrase is selected and a structure in which a head is selected. In both cases, the relation between the selector and the selectee will be one of sisterhood. Thus, the following structures illustrate "head-phrase" selection with a verb that selects a noun phrase, and "head-head" selection with a verb that selects for a particular complementizer:

(64) a.   V $\underline{\theta}_{1\#}$ b.   V $\theta_{1\#}$
        /   \   /   \
     Subj V $\underline{\theta}_1$ C   V θ_1

This similarity between selection of heads and selection of phrases gains some significance after the discussion of chains in the previous sections, where it was noted that s-movement can only occur in structures in which a head chain is formed, and that adjunction may involve the same type of structure. As there is no categorial projection, there will be no way to distinguish the root of a constituent headed by a moved category and the root of a constituent that itself moved. In both structures, the root categories will not be in the lowest position of their chain. As a result of this, the condition in (63) not only restricts selection of moved constituents but also of constituents whose head is the higher member in a chain. Just like moved constituents, constituents whose head is the higher member in a chain can then not be selected. Since structures involving adjunction to the highest category also involve a head chain, such structures can also not be selected.

The second half of the chain condition above restricts the class of possible selectors. Again, the rejection of categorial projection means that phrases and heads cannot be distinguished for the purpose of selection. Consequently, where a constituent is a possible selector, in the same context, a constituent whose head is the higher member in a chain is not a possible selector. These aspects of the condition in (63) are investigated in chapter 2.

This condition also affects the positions available for adjunction. Adjunction can give rise to head chains. This means that the following structure is not possible:

(65)   a. $\{<$V $\underline{\theta}_{1\#}\,\theta_{2\#}$, $\{$V $\underline{\theta}_{1\#}\,\theta_{2\#}$, Subj, V $\underline{\theta}_1\,\theta_{2\#}$, Adj, V $\underline{\theta}_1\,\theta_{2\#}$, Obj, V $\underline{\theta}_1\,\theta_2\}>$, $<$Subj, $\{$Subj$\}>$, $<$V $\underline{\theta}_1\,\theta_{2\#}$, $\{$V $\underline{\theta}_1\,\theta_{2\#}$, Adj, V $\underline{\theta}_1\,\theta_{2\#}$, Obj, V $\underline{\theta}_1\,\theta_2\}>$, $<$Adj, $\{$Adj $\}>$, $<$V $\underline{\theta}_1\,\theta_{2\#}$, $\{$V $\underline{\theta}_1\,\theta_{2\#}$, Obj, V $\underline{\theta}_1\,\theta_2\}>$, $<$Obj, $\{$Obj$\}>$, $<$V $\underline{\theta}_1\,\theta_2$, $\{$V $\underline{\theta}_1\,\theta_2\}>\}$

        b.   *V $\underline{\theta}_{1\#}\,\theta_{2\#}$
               /   \
           Subj   V $\underline{\theta}_1\,\theta_{2\#}$
                    /   \
               Adj   V $\underline{\theta}_1\,\theta_{2\#}$
                       /   \
                   Obj   V $\underline{\theta}_1\,\theta_2$

Here, the adjunct Adj is adjoined on top of a head of category V $\underline{\theta}_1$ $\theta_{2\#}$, that is to a head with an unsatisfied external argument role. As a result of adjunction, a new node is created and it inherits the category V $\underline{\theta}_1$ $\theta_{2\#}$. These two categories trigger the formation of a chain, and consequently Subj cannot satisfy the remaining unsatisfied role of V $\underline{\theta}_1$ $\theta_{2\#}$ because this would violate the condition in (63). Informally, this means that VP-internal adjunction is impossible. Similarly adjunction to the highest category in a selected or selecting constituent is blocked by (63). However, where an unsatisfied role is inherited by a node of different category, this node will be able to assign the role because it doesn't have to form a chain with the daughter from which it inherits the role.

7 Linearisation and constituent structure

This section briefly addresses the question of how the hierarchical relations provided by phrase structures can be mapped onto a linear order.

The most basic constraint on the mapping of phrase structure to linear order is that linearisation should preserve the constituency of phrase structures. That is given the two constituents $\{\alpha, \beta\}$ $\{\gamma, \delta\}$, linearisation cannot lead to the sequence α-γ-β-δ, where γ appears between two categories of a different constituent, and is hence separate from δ, the other member of its constituent.

The first question then is what constituencies the projection-free phrase structures assumed here define. The following structure is a schematic representation of a verb with a subject S and an object O. The numbered categories correspond to the derived heads of the selecting verb, all of which are positions in which the verb can be potentially pronounced:

(66)
```
          3
         / \
        S   2
           / \
          O   1
```
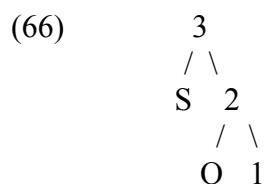
Under the standard definition of constituency, elements dominated by a single category correspond to a constituent. (Essentially, each dominance set corresponds to a constituent.) Thus, the above structure defines the following constituents. Note that 2 in (67a) is itself a complex constituent.

(67)   a.   3   b.   2
                    / \   / \
                   S 2   O 1

These structures still contain some further information. In particular, it should be noted that, within each constituent, there is an asymmetry between mothers and daughters. Thus, in each constituent, there is a single mother and two daughters, which share the same structural relation to their mother. I propose that the asymmetry between mothers and their daughters should be preserved in the mapping to prosodic structure. This can be achieved by means of the following condition:

(68)   In prosodic structure, the position of the mother must separate its daughters.

In addition to capturing the fact that the daughters are structurally distinct from the mother, this condition also captures the fact that both daughters share the same relation to their mother. While in phrase structure the primitive relation is dominance, one of the primitive relations of PF is adjacency. While in phrase structure the two daughters are immediately dominated by the mother, the condition in (68) allows only orders in which both daughters are adjacent to the mother.

The standard definition of constituency together with the above condition then licenses the following linearisations of the structure in (66). The sequences in (69a) give the possible orders for each of the two complex constituents in (66), while (69b) gives the possible strings for the full structure in (66):

(69)   a.   S-3-2   O-2-1
              2-3-S   1-2-O

         b.   S-3- O-2-1; S-3-1-2-O; O-2-1-3-S; 1-2-O-3-S

For clarity, the possible linear sequences for the structure in (66) are repeated here, with an indication of the corresponding major constituent orders: (Recall that the numbered positions correspond to the verb and its derived heads (derived by selection), i.e. the positions in which the verb can be spelled out without s-movement.)

(70)   S-3-O-2-1: S-V-O; S-O-V
        S-3-1-2-O: S-V-O
        O-2-1-3-S: O-V-S   * O-S-V
        1-2-O-3-S: V-O-S   * V-S-O

Thus, out of the six logically possible orders for subject, object and verb, only four can be derived without movement out of the structure in (66). This is an interesting parallel to the possible word orders defined by an X-bar-theoretic representation of a VP. As in the model assumed here, no orders are possible in which the verb and the object are separated by the subject:

(71)   a.   VP   b. [Subj [V Obj]]
               / \   [Subj [Obj V]]
           Subj   V'   [[V Obj] Subj]
                 / \   [[Obj V] Subj]
               V   Obj

8 Outline of the dissertation

The final section of this chapter is a preview of the coming chapters.

Chapter 2 *Selection and head chains*

The assumption that there is no categorial projection interacts with the constraint that blocks selection of non-tail chain positions in the following way. As discussed in section 3, two types verb movement configurations can be distinguished. S-movement structures are those in which the verb is pronounced in the position of a higher copy of the verb. If s-movement involves the highest head in the structure, the structure is headed by a "moved" category, that is a category that is not in the deepest position of its chain. This means that such a clause will be subject to the condition on selection, just like moved categories. This contrasts with the standard view where a moved head always adjoins to a higher head, with the result that the highest projection in a structure is always headed by an unmoved category.

A first prediction of the present proposal then is that a structure in which the highest head is a moved verb cannot be selected. This entails that, if such a structure appears in some embedded context, it should behave differently from regular complements. For example, such a structure can be expected to pattern with adjuncts rather than complements

with regard to extraction. An example comes from (apparently) embedded V2 clauses in German. These clauses appear to be the complements of a matrix verb but their highest head is a moved verb. As predicted, these clauses behave unlike other sentential complements, e.g. with regard to mood marking (den Besten 1983) and also extraction (Reis 1995).

The next prediction concerns optionality of complementizers. As is well known, the complementizer *that* is optional in many contexts. On the view of clause structure maintained here, the absence of *that* (where it is not the result of deletion under the Doubly-Filled-Comp Filter) implies that there is no C level, no CP in standard terms. This entails that in a selected clause with verb movement, the complementizer cannot be omitted. This prediction derives the generalisation observed by Sten Vikner that across the Germanic and Romance languages, optionality of the complementizer is found only in languages that lack "V-to-I" movement (Vikner 2001). Furthermore, the prediction extends also to English where the occurrence of verb movement in embedded clauses blocks omission of the complementizer.

A problem for earlier attempts to deal with complementizer optionality has been the fact that this phenomenon is not limited to selected clauses, but that it occurs in the same way in non-selected clauses, like relative clauses. The Telescope hypothesis offers a way to understand this symmetry. If the relation between a relative clause and the noun that it modifies is taken to be one of selection (by the relative clause), the condition on selection entails that relative clauses pattern with complement clauses with regard to complementizer optionality. More generally, in a selecting clause with verb movement, the complementizer cannot be omitted.

The next prediction concerns the interaction of verb movement and Clause Typing. Some languages, like Welsh and Irish, have special particles to mark matrix clauses. On the view that (some) parameters are related to the features of functional elements, this suggests the existence of the Clause Typing parameter. Work by Cheng 1991 suggests that where some languages have special Clause Typing particles, languages that don't may use movement to clause-type. Thus, the parameter entails that some languages use movement to mark a main clause. The above prediction that a clause headed by a moved verb cannot be an embedded clause also entails its reverse, namely that, if the highest head in a

structure is a moved verb, the structure cannot be a dependent clause. This means that a language without special particles can mark a main clause through verb movement.

I will argue that German is an example of such a language. The claim that verb movement in German serves to mark a main clause naturally accounts for the lack of verb movement in dependent clauses. It also provides further support for the claim that (apparently) embedded V2 clauses are not complements.

These predictions provide strong support for the position on empty heads and the theory of verb movement proposed here and for the Telescope hypothesis. Without the strict constraints on the availability of empty heads implied by the flexible view of clause structure adopted here, it is not clear how these predictions could be derived. In particular, it is no clear how the projection of an abstract empty head on top of an embedded verb movement structure can be ruled out in a principled way once such heads are available. Without the assumption that these cases of verb movement do not involve adjunction to a head inserted higher in the structure, it is not clear how the relevance of the constraint on selection for these data could be explained. Without the abandonment of the phrase/head distinction implied by Telescope, it is also not clear how this condition could be extended in a natural way to structures headed by a moved verb.

Chapter 3 *Verb movement and derived heads*

Chapter 3 deals initially more closely with the theory of verb movement proposed here, and uses a range of data from Welsh, Middle Welsh, Breton, and German to test the theoretical claims.

As already noted, movement of a head that already is the highest head in a structure requires the presence of a specifier (see section 1.3). This condition is a purely structural condition, that is, it does not introduce any restrictions on the category of the constituent in the highest specifier. This implies that, where there are no independent constraints on what occurs in the specifier, verb movement (s-movement) will give rise to a pattern where the moved verb is preceded by a constituent of arbitrary category. That is the verb will appear in the second position. This provides a natural account of the V2 effects found in Breton and German. Furthermore, I will present data from German and Breton to show that the initial constituent in V2 structures does not receive a uniform interpretation. This means

that analyses of these V2 effects in terms of the checking of a semantic or pragmatic feature like Topic or Focus are not tenable. Together with the fact that the initial constituent in V2 structures is of arbitrary category, this provides a strong argument for the structural account of V2, and against checking theoretic accounts of this phenomenon.

Unlike s-movement, p-movement is verb movement that takes place in structures that do not involve copies of the verb. P-movement structures are those in which the verb is pronounced in the position of an independent category. Since here it will not be necessary to create a new specifier, the verb does not have to be preceded by some constituent, i.e. it can be in initial position. However, recall the assumption that the availability of empty heads is tightly constrained. In potential p-movement structures, there must then be some visible head preceding the verb. This derives the typological generalisation that verb-initial languages typically have preverbal particles.

These restrictions on s- and p-movement structures, together with the Head Movement Constraint (the standard assumption that movement of a head cannot skip another head, Travis 1984), imply a further typological prediction. S-movement can only occur when there is no further head that dominates the verb. This implies the absence of preverbal particles in V2 languages.

A pattern relevant in this context can be found in Middle Welsh and Breton. In particular, these V2 languages appear to display preverbal particles on the surface. However, I will argue that in V2 structures, these particles are not represented syntactically as heads dominating the verb.

These considerations further implicate the change of languages from verb-initial to V2. Verb-initial order is only possible where there is a preverbal particle that corresponds to a syntactic head position that dominates the verb, and V2 order only arises if there is no such head. This implies that the change from verb-initial to V2 can only occur in a language that develops preverbal particle heads. Based on work by David Willis (Willis 1998), I show that the change of Welsh from V2 to VSO confirms this prediction. Crucially, the theory proposed here can explain the fact that the development of a new preverbal particle, namely main clause complementizer, and of the emergence of unmarked VSO orders in Welsh occurred at the same time.

The theory of verb movement proposed here, and in particular the analysis of V2

effects, has further implications. The occurrence of V2 patterns is the result of a particular verb movement configuration, but it is in no way related to the *trigger* of this movement. Furthermore, while different triggers may give rise to V2 patterns, the conditions associated with these triggers may lead to different conditions on the V2 structures that they generate. This implies that on the one hand V2 patterns can occur regardless of the trigger of verb movement, and on the other that different triggers of verb movement may give rise to different constraints on V2 structures. A comparison of Breton and German confirms these predictions. Verb movement in German is triggered by the Clause Typing parameter above, while verb movement in Breton is triggered by the requirement that the subject be assigned Case by a preceding verb. Verb movement in German and Breton differs in that verb movement is restricted to root clauses in German but occurs in both root and embedded clauses in Breton. Furthermore, Breton displays adjacency effects on the fronted verb and postverbal subjects while German does not.

Note now that on the standard view, different functional projections are related to different triggers of movement. Given this view, it also follows that V2 patterns can occur in different functional projections in different languages. This means that the occurrence of V2 patterns in different languages is not a valid argument in support of a universal clause structure. Furthermore, it suggests that at least not all verb movements in all V2 structures in all languages should be analysed in terms of the checking of some feature on an abstract functional head C.

Chapter 4 *Free relatives and derived heads*

Chapters 1 to 3 are concerned mainly with movement of the verb, that is movement of a traditional head or X^0 category. However, the telescope hypothesis implies that there is no syntactic difference between X^0 or XP (phrasal) categories. Thus, it is expected that if a moved "head" can be the head of a structure, so can a "phrase".

One structure created in this way, namely free relatives, is the subject of chapter 4. The (old) problem with free relatives is that while they look like clauses, their distribution is more similar to that of noun phrases modified by a relative clause. This has lead to proposals that assume that a free relative is in fact structurally identical to a noun plus relative clause structure, but that the head noun of the free relative is phonologically

empty. Such proposals lead to complications concerning the licensing of empty nouns, especially in languages like English or German that do not usually allow for null nominals. In the model of clause structure assumed here an alternative analysis becomes available. I propose that free relatives involve two copies of the *wh*-phrase, one of which occupies the position that corresponds to that of the head noun in headed relative clause constructions and the other occupies the highest specifier inside the relative clause.

This analysis straightforwardly derives two well-known facts about free relatives: First, a free relative always has the same categorial properties as the *wh*-phrase that introduces it. Second, conflicting case requirements in matrix clause and internal to the relative may affect the well-formedness of a free relative.

Furthermore the proposed analysis restricts the class of possible *wh*-phrases that can introduce a free relative. In particular, this analysis implies that a *wh*-phrase that lacks a semantic restriction cannot introduce a free relative. This explains why in English free relatives cannot be introduced by *which*, *that* or no complementizer (the last two cases would both involve an empty operator). Similar restrictions from German and a dialect of German will be discussed. The final section of the chapter deals with the question of which of the two copies of the operator is pronounced.

Chapter 5 *Conclusion*
Chapter 5 is the conclusion.

CHAPTER TWO

Selection and head chains

1 Introduction

This chapter deals with the effects of a constraint on selection on the shape and the distribution of embedded clauses. This is an area where the predictions of a theory without categorial projection in which moved elements can head structures differ in a clear way from those made by a more conventional theory. In this first section, I introduce the constraint on selection, recapitulate some of the proposals made in chapter 1, and outline the structure of this chapter.

One of the earliest insights that emerged in the field of generative-transformational grammar is that certain grammatical relations can only hold of two linguistic items if neither of the two has been involved in a transformational operation. A strong form of this view is implied in the following passage from Katz and Postal 1964:[1]

> It appears that in the formally motivated underlying P-markers provided by the simplest transformational grammar there is associated with each grammatical relation a unique subconfiguration of constituents that can be taken as the formal basis for these relations. But in derived P-markers no such unique correlation between grammatical relations and configurations of constituents can be found. This is the most important sense in which derived P-markers provide only a superficial account of grammatical structure, with the 'deeper' facts represented only in underlying P-markers. (1964:39)

In fact, Chomsky (1965:117, 1966) argues that this generalisation predates the history of modern generative grammar by a couple of hundred years. He attributes the following view of language to the work of the 17th century French philosophers who developed the Port Royal Grammar:

---

[1] See also Harris 1957:290 for relevant remarks.

The surface structure resulting from […] transformations does not directly express the meaning relations of the words, of course, except in the simplest cases. It is the deep structure underlying the actual utterance […] that conveys the semantic content of the sentence. (1966:35)

More recently, this generalisation has been discussed explicitly in relation to the thematic properties of chain positions, especially in the work of Brody 1987:(3), 1993:(9).[2] I will refer to the type of relation that follows this generalisation as SELECTIONAL RELATIONS, or as instances of SELECTION, and I assume that selection takes place under sisterhood.[3] The generalisation can then be formalised as follows:[4]

(1) If α selects β, both α and β must occupy the deepest position of their chains.

The following examples illustrate the empirical basis of this generalisation.

(2) a.   * John$_i$ hit t_i (Brody 1995:15)

    b.   * John$_i$ believes t_i to seem that S   (Brody 1993:(8))

The examples in (2) illustrate the ban on movement to a selected position. There appear to be no transitive verbs like *hit* that do not assign case to their object. Similarly, there are no verbs that are like *believe* except that they do not assign case to the embedded subject. If the subject *John* forms a chain with the trace in object position in (2a), and with the trace in the embedded subject position in (2b), these structures can be ruled out by the constraint in (1), which bans θ-role assignment to the moved noun *John*, and therefore entails that the

[2] Brody 1995:12 calls (1) the MAIN THEMATIC CONDITION, and later discusses it as one aspect of his GENERALISED PROJECTION PRINCIPLE. Chomsky 1995b:312 uses the more catchy name CHAIN CONDITION for (1), which I may sometimes use below. For relevant discussion of the generalisation, see also Chomsky 1986b:131-144, Chomsky and Lasnik 1993:46, Chomsky 1995b: 312-316, Jackendoff 1997:101-103.
[3] For an approach to grammatical relations that does not invoke the sisterhood relation, which is compatible with the theory developed here, see Neeleman and van de Koot 2002: section 2.
[4] There have been some proposals according to which this generalisation does not hold, at least not in the strongest form (e.g. Bošković 1994, Cormack 1995:253ff, Hornstein 1999, Manzini and Roussou 2000). Nevertheless, I assume that (1) does hold, without attempting to explain it. Critical discussion of the arguments in Hornstein 1999 can be found in Culicover and Jackendoff 2001. Relevant discussion of why (1) should hold, and different approaches to how it can be derived, can be found in Chomsky 1981, 1986b, Brody 1995, 1998, Neeleman and van de Koot 2002, among others.

subject θ-role of *hit* and *believe,* respectively, cannot be assigned.

The example in (3a) illustrates the reverse of those in (2), namely that an element that is not in the lowest position of its chain cannot assign a θ-role.

(3) a. * [Den alten Mann]$_i$ traf$_j$ gestern ein Freund von sich$_i$ t_j

        The old man   met yesterday a   friend   of   himself

    b. * [Den alten Mann]$_i$ traf$_j$ gestern ein Freund   von sich t_i t_j

        The old man   met yesterday a   friend   of   himself

It is generally assumed that binding is only possible from A-positions. If *traf* could assign its internal θ-role to *den alten Mann* after movement, *den alten Mann* would sit in an A-position, and should be able to bind the reflexive *sich*. On the other hand, if, as (1) implies, θ-role assignment is only possible from the lowest position in a chain, the only way for *den alten Mann* to receive a θ-role is if it is moved out of the object position of *traf,* as in (3b). However, in (3b) *den alten Mann* occupies an A-bar position and hence cannot bind the reflexive. (The trace of *den alten Mann* cannot bind *sich* because it doesn't c-command it.)

While the generalisation in (1) is mainly discussed with regard to the relation between a verb and its arguments, it seems to hold more generally. Thus, Brody 1997a:53 note 4 suggests that (1) extends also to adverbial modification. Moreover, Chomsky and Lasnik 1993:45 argue following Williams 1980 that "[a]n argument may also receive a semantic role […] by predication by an XP […], possibly an open sentence", like a relative clause. The fact that they take the question of whether or not this semantic role should be considered a θ-role as a theory-internal matter implies that empirically modification by a relative clause behaves similar to θ-role assignment, and that it can plausibly be expected to conform to (1). I know of no explicit theory of grammatical relations that defines which relations exactly behave according to (1), and I will not attempt to develop such a theory.[5] The aim of this chapter is simply to show how a range of apparently unrelated patterns can be understood in the context of (1).

---

[5] I assume that the relation between a functional head like the complementizer *that* or the auxiliary *have* and their complements is not of the same type as relations restricted by (1). See Grimshaw 1991, van Riemsdijk 1998, Chametzky 2000, among others. See also chapter 1.

The patterns to be discussed involve two types of selectional relations, namely θ-role assignment and the relation between a restrictive relative clause and the head that it modifies. Moreover, the analysis of selectional relations is also helpful for the evaluation of the different predictions made by alternative theories of phrase structure. Selection as a relation between sisters and the ban on selection from non-chain-tail positions are affected by the proposals of chapter 1 in several ways.

Recall that according to the Telescope hypothesis (Brody 1997a, 2000) adopted here, there is no categorial projection. This means that the conventional X-bar structure in (4a) should be represented as in (4b).[6]

(4) a.   XP   b.   X
            /   \   /   \
      Specifier   X'   Specifier Complement
                /   \
            X Complement
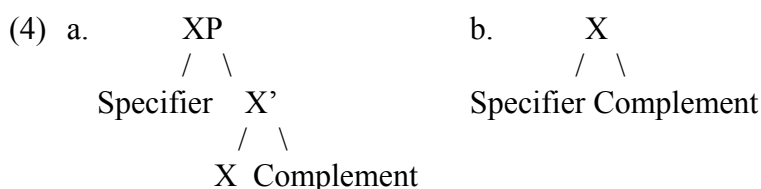
Consequently, while in conventional theories selection in principle involves sisterhood between a head and the projection of a head (or in the case of a complex selector sisterhood between two projections), here selection always involves sisterhood between two heads.

Furthermore, the assumption that material that does not occupy the deepest position within its chain can head a structure (Ackema et al 1993), combined with the Telescope hypothesis, may give rise to structures whose head is a member of a (complex) chain. Since selection is based on sisterhood, selection of (or by) such structures is expected to behave in line with the generalisation in (1). Thus, even though the structures themselves may not be part of a complex chain, it is predicted that they cannot enter into selectional relations. These points are discussed in greater detail throughout the chapter, but the line of argumentation will be the following. The theory assumed here implies that there are structures that cannot select or be selected. Where feasible such a structure may be turned into a possible selector or selectee through insertion of an extra head, typically a

---

[6] See chapter 1: section 4 for fuller discussion.
    Note that for the purposes of this chapter nothing hinges on the claim of chapter 1 that phrase structure should be represented in terms of phrase structure sets and that tree diagrams are merely convenient ways to represent these sets. To keep the discussion as simple as possible, I therefore only use tree structures here.

complementizer. This has the effect that the head of the structure is no longer part of a complex chain. Where this is not possible, such structures simply cannot partake in selectional relations.

The next section deals with the relation of the generalisation in (1) and verb movement structures that involve the pronunciation of the verb in the position of a higher copy of the verb, i.e. s-movement, and in particular with the effect of verb movement on the distribution of complementizers in complement clauses. In section 3, it is argued that like verb movement, adjunction may give rise to a head chain. Consequently, it is expected that adjunction interacts in the same way with (1) as verb movement, and hence has the same effect on the distribution of complementizers. Section 4 extends the analysis of complementizer optionality to relative clauses. The claim is that while complement clauses conform to (1) because they are selected, relative clauses do so because they themselves select. Section 5 discusses the implications of these proposals for the analysis of apparently embedded V2 clauses in German, and section 6 shows how verb movement through its interaction with (1) can play a role in clause-typing and offers a new account of the complementarity of verb movement and complementizers in German.

2 Verb movement and complementizer optionality in complement clauses

As discussed in chapter 1, the creation of a new head position can be triggered by the need to express head movement. Since in such a context, the higher head position doesn't correspond to an additional lexical item, it must inherit its properties from its daughter. Given the structure in (5a), verb movement can give rise to the structure in (5b), where the categorial information of the lower node V is inherited by the new root node in (5b). Thus, s-movement structures can involve the presence of a copy of the verb:[7]

(5) a.   V   b.   V
            / \   / \
           A   B   C   V
                                               / \
                                              A   B

---

[7] For reasons not relevant here, such verb movement is only possible if the new head has its own specifier; see chapters 1and 3 for discussion.

As proposed in chapter 1, the existence of copies in a structure triggers the creation of a head chain that contains the copies and the original category in order to ensure interpretability of the structure. The formation of a chain containing the two categories means that the head of the structure in (5b) is not in the lowest position of its chain. Since selection takes place under sisterhood between two heads, it follows that such a structure will be subject to the chain condition in (1) above. This case must be distinguished from a situation where a new head position is created to make possible the satisfaction of an unsatisfied selectional requirement. Take the case where a new head is created to allow the head V θ_1 to be in a selection configuration with a head N that satisfies its θ-requirement. In that case, the categorial information of V θ_1 has to be copied up to the additional head position, but since in this position the requirement for a θ_1 is no longer active, the properties of the new head differ from those of its daughter:

(6) a.   V θ_1 b.   V $\theta_{1\#}$
                                                                 / \
                                                                N   V θ_1

Thus, the categories V θ_1 and V $\theta_{1\#}$ are distinct, and therefore there is no head chain. Since no complex chain is formed, a structure like (6b) conforms to the chain condition in (1) by definition when it is selected. (If there is no complex chain, every position is the lowest position in its chain.)

Some relevant structures are illustrated here with examples from English and French:[8]

(7) a.   'Jean cuit des pâtes.'   b.   'John cooked pasta.'

              cuit $\underline{\theta}_\# \theta_{2\#}$ T
              /   \   /   \
        Jean$_i$ cuit $\underline{\theta}_\# \theta_{2\#}$ John$_i$ cooked $\underline{\theta}_{1\#} \theta_{2\#}$
              /   \   /   \
          t$_i$ cuit $\underline{\theta}_1 \theta_{2\#}$ t$_i$ cooked $\underline{\theta}_1 \theta_{2\#}$
                 /   \   /   \
        des pâtes   cuit $\underline{\theta}_1 \theta_2$ qpasta   cooked $\underline{\theta}_1 \theta_2$

---

[8] See chapter 1: section 3 for some discussion of the difference between languages with and without verb movement.

In French, a language where verb movement involves a new head position that contains a copy of the verb, the root node and its (non-subject) daughter form a head chain. This is so because the verb's subject-role is satisfied by *Jean* before it moves, and consequently the two heads share the same categorial features, i.e. they are copies. In (7a), both heads are of category *cuit* $\underline{\theta}_\# \theta_{2\#}$.[9] In contrast, in English, a language without verb movement, the two highest clausal heads do not form a chain since they are not copies.

The membership in a complex chain of the root node in s-movement structures affects the structural contexts in which they can occur. In particular, the condition in (1) implies that a structure like (7a) is not selectable, and therefore that it cannot sit in a position that satisfies a predicate's θ-role. In contrast, selection of a structure like in (7b) is no problem for the condition because the root node is in the lowest position of its (one-member) chain. These two situations are illustrated here in a schematic way; with verb movement in (8a), without in (8b):

(8) a. $*V_{matrix}\ \underline{\theta}_1\ \theta_{2\#}$ b. $V_{matrix}\ \underline{\theta}_1\ \theta_{2\#}$

                  /   \   /   \

        $V\ \underline{\theta}_\#\ \theta_{2\#}$ $V_{matrix}\ \underline{\theta}_1\ \theta_2$ T $V_{matrix}\ \underline{\theta}_1\ \theta_2$

         /   \   /   \

      $Subj_i$ $V\ \underline{\theta}_\#\ \theta_{2\#}$ $Subj_i$ $V\ \underline{\theta}_1\ \theta_{2\#}$

                /   \   /   \

           t_i $V\ \underline{\theta}_1\ \theta_{2\#}$ t_i $V\ \underline{\theta}_1\ \theta_{2\#}$

                   /   \   /   \

          Obj $V\underline{\theta}_1\ \theta_2$ Obj $V\ \underline{\theta}_1\ \theta_2$
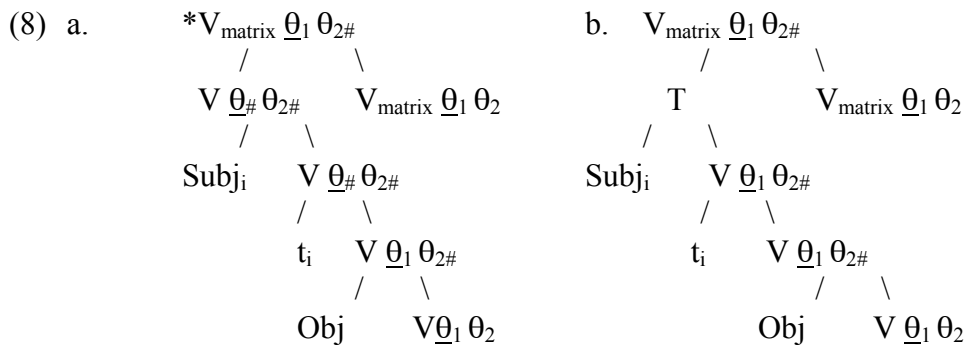
On this view of verb movement and chain formation, it is a consequence of the chain condition in (1) that a clause headed by a moved verb cannot occur in a selected position.[10] For future reference, I give this result its own number:

(9) A clause headed by a moved verb cannot be selected.

---

[9] By assumption, the trigger of verb movement, possibly a condition on case assignment or properties of agreement morphology, does not affect the attributes of the verb. This idea may be related to the complementarity of feature checking and θ-theory assumed in Chomsky 1995:322.

[10] It would be more accurate to say that such a clause cannot be assigned a θ-role. However, on the assumption that all of a predicate's θ-roles must be assigned, some other element would have to sit in the selected position. This would in turn keep the verb movement structure from occurring in that position.

This prediction makes for an interesting difference between the present proposal and more conventional approaches to clause structure. On the standard view, verb movement is analysed as adjunction of the moved verb to some functional head. Thus, in the X-bar tree in (10) the matrix verb selects a clause in which V-to-I movement has taken place. While there is adjunction of a moved item, namely the verb, to I, it is clear that the head of the selected IP, namely the functional head I, does not itself move:

(10)
```
              V'
          ┌───────┐
          V       IP
              ┌────────┐
            Subj      I'
                  ┌──────────┐
                  I         VP
                 / \      ┌─────┐
               Vᵢ  I    tᵢ    Obj
```
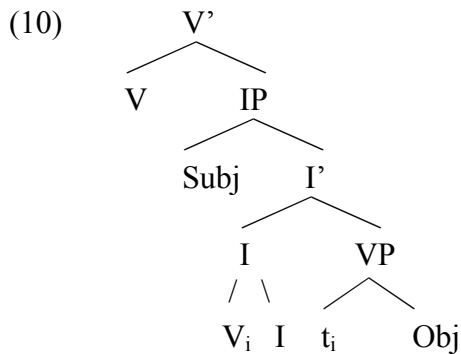
Leaving aside the possibility of a novel device that makes I invisible in selected contexts, there is no obvious way in which the condition in (1) could be extended to rule out selection of such a structure. Consequently, on such an approach to verb movement, nothing like the prediction in (9) follows.[11] Similarly, in a theory without categorial projection that takes verb movement to create a new head position rather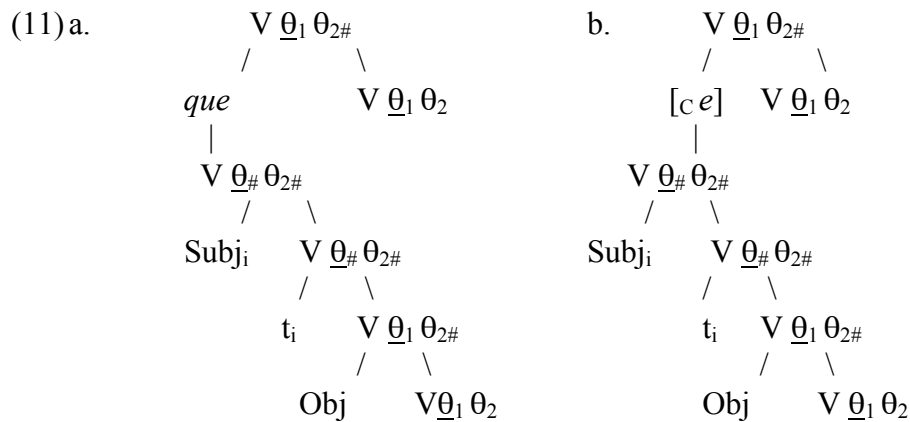 than as head adjunction, but that assumes that complement clauses without visible *that* differ from those with a visible *that* not structurally but simply in that the structural position of *that* contains an empty element, prediction (9) does not follow. On such a view, the structures with and without (overt) complementizer in a verb movement language would be the following:

---

[11] In fact, an early attempt to derive such a prediction by Rizzi and Roberts 1989, assumes that verb movement is head adjunction. To rule out I-to-C movement in selected contexts, Rizzi and Roberts introduce the concept of radically empty heads, i.e. head positions that contain no features at all. Rizzi and Roberts' substitution into an empty head position can be taken to be a precursor of more recent self-attachment proposals (e.g. Ackema et al 1993) and the device of proxy categories of Nash and Rouveret 1997. The major difference to the current proposal would then be the assumption of categorial projection and the distinction between movement to I and movement to C. Rizzi and Roberts' proposal that C is radically empty while I is

(11) a.
$$V\ \underline{\theta}_1\ \theta_{2\#}$$

```
            V θ₁ θ₂#
           /        \
        que          V θ₁ θ₂
         |
       V θ# θ₂#
       /      \
    Subjᵢ      V θ# θ₂#
              /      \
            tᵢ        V θ₁ θ₂#
                     /      \
                   Obj       Vθ₁ θ₂
```

b.
```
            V θ₁ θ₂#
           /        \
       [c e]         V θ₁ θ₂
         |
       V θ# θ₂#
       /      \
    Subjᵢ      V θ# θ₂#
              /      \
            tᵢ        V θ₁ θ₂#
                     /      \
                   Obj       V θ₁ θ₂
```

In both cases, the highest head of the embedded clause is not the moved verb but an unmoved complementizer, which happens to be empty in (11b). It follows that there is no difference between these structures from the perspective of the condition in (1), and hence selection of both should be equally good or bad. The empirical validity of (9) then seems to be a useful help in the choice between these approaches to head movement and clause structure.

An immediate question raised by the prediction in (9) is how clausal complements can be selected at all in languages with verb movement. The answer is obvious if we consider the following example from French, a language with verb movement:

(12) a.   Jean pense [que [Pierre dort [Pierre dort]]]

b.
```
            pense
           /     \
        que       pense
         |
        dort
        /   \
    Pierreᵢ   dort
             /   \
           tᵢ     dort
```
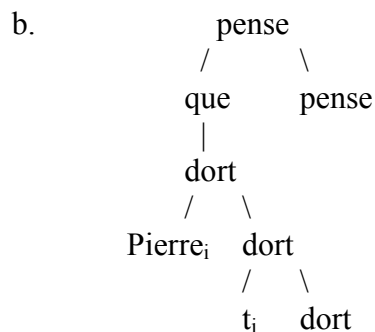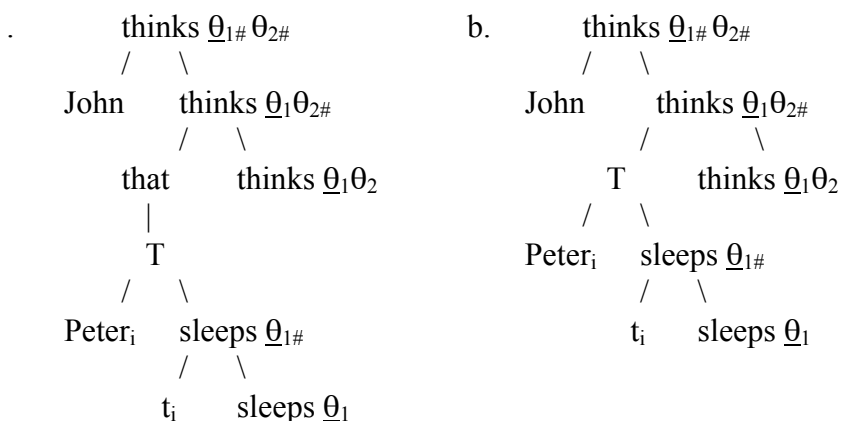
The embedded clause is headed by the complementizer *que*. As the simplified tree in (12b) illustrates, the sister of the selecting head *pense* is the complementizer, which did not move. Since the complementizer is the head of the embedded clause, rather than the moved

---

not means that movement to the two positions should give rise to different effects in embedded contexts.

verb, the fact that verb movement has taken place in this clause does not affect its selectability.[12]

The role of the complementizer in making a clause in a verb movement language available for selection implies that in languages with verb movement, complementizers should be obligatory. This contrasts with languages without verb movement. Since here an embedded clause is not headed by a moved verb, its selection is unproblematic. This implies that unless there are additional constraints, complementizers could in principle be optional. Indeed, the complementizer *that* in English, a language without verb movement, is usually optional.[13] Since there is no verb movement, the condition in (1) is irrelevant to the selection of embedded clauses. The following trees illustrate clauses with and clauses without a complementizer. In both, the sister of the selecting head *thinks* is a head that doesn't move, namely the complementizer *that* or the head T:

(13)   'John thinks (that) Peter sleeps.'

```
      .      thinks θ̲₁# θ₂#              b.         thinks θ̲₁# θ₂#
            /    \                                   /    \
      John      thinks θ̲₁θ₂#               John         thinks θ̲₁θ₂#
                /    \                                    /       \
          that       thinks θ̲₁θ₂                  T           thinks θ̲₁θ₂
            |                                     /    \
            T                                 Peterᵢ    sleeps θ̲₁#
          /    \                                        /    \
    Peterᵢ      sleeps θ̲₁#                         tᵢ       sleeps θ̲₁
            /    \
          tᵢ      sleeps θ̲₁
```

This model predicts that while complementizers should in principle be optional in languages without verb movement, they should be obligatory in languages with verb

---

However, the arguments presented in this chapter suggest that a generalisation might have been missed there.

[12] Again, the relation between the complementizer and its daughter is not one of selection; see note 5 above.

[13] The complementizer is only optional if the clause sits in its base position. Grimshaw 1997:411n17 points out that speaker judgements vary in such examples with respect to complementizer optionality (cf. also McCloskey 1992:23). In contrast, examples like (i) where the complement clause doesn't occupy its base position, are ungrammatical for all speakers. This suggests that different constraints on the deletion of *that* may play a role. (i) *(That) he left so early shows that he was tired*.

movement.[14]

This indeed appears to be the case. Sten Vikner recently observed that this pattern holds across the Germanic and Romance languages. Vikner's generalisation is stated in (14) and illustrated in the following examples from Vikner 2001. In Danish and English, which don't have (obligatory) verb movement in embedded clauses, complementizers are optional, while in French and Icelandic, which have verb movement, complementizers are obligatory.[15]

(14) Complementizers in embedded clauses can only be optionally omitted in languages without overt (V-to-I) verb movement. In languages with verb movement, complementizers cannot be omitted.

(15) a. Jeg tror at skuespilleren virkelig så filmen.   Danish

   b. I think that the actor   actually saw the film.   English

   c. Jeg tror   skuespilleren virkelig så filmen.   Danish

   d. I think   the actor   actually saw the film.   English

(16) a. Ég tel að   leikarinn sjái áreiðanlega myndina.   Icelandic

   b. Je crois que   l'acteur voit vraitment   le film.   French

   c. *Ég tel   leikarinn sjái áreiðanlega myndina.   Icelandic

   d. *Je crois   l'acteur voit vraitment   le film.   French

---

[14] This issue is independent of the question of what determines the choice of +/- complementizer in cases where the complementizer is optional. Thompson and Mulac 1991 discuss this choice in English based on conversational discourse. They note that presence of complementizer is preferred in structures where the subjects and main verbs of the matrix clause "behave very much like single epistemic morphemes in other languages" (p.239). Thus, the lighter (in the sense of semantically bleached) the matrix and the more "assertion" the embedded clause, the more likely is omission of the complementizer. The lightness of the matrix *viz.* the likelihood of *that* omission, is increased by (a combination of) the presence of "first and second person subjects, the verbs *think* and *guess*, pronominal complement subjects, and auxiliaries, indirect objects, and adverbs" (p.249). It seems then that syntax doesn't have to say much about when an optional complementizer is used. However, syntax constrains the range of constructions where a complementizer can potentially be left out.

[15] Sten Vikner (personal communication) points out that Italian subjunctives appear to be a counterexample to his generalisation since Italian has verb movement but complementizers are optional in certain subjunctive contexts; see section 5: note 33 for some comments.

  A second type of complementizer optionality (CO) occurs in Florentine Italian (Cocchi and Poletto 2001). Unlike CO with subjunctives, this type of CO is not restricted to complements of bridge verbs. However, it can take place only where the embedded verb is preceded by a preverbal subject or object clitic, by negation, or by an auxiliary. In the present framework, this pattern could receive a natural account if these preverbal elements can be analysed as heads that license the pronunciation of the verb in their position. These heads would then function like complementizers in that they allow an embedded verb movement structure to escape the effect of the chain condition in (1). Further work is needed to see if this idea is tenable.

In fact, the pattern observed by Vikner also extends to English. While in general English doesn't display verb movement, there are certain contexts in which verb movement does occur:

(17) a. All too seldom did$_i$ he t$_i$ bring her flowers.

   b. Not even that small consideration did$_i$ he t$_i$ ever bestow on his partner.

   c. Never in her life would$_i$ she t$_i$ accept this solution.

   d. Never in his life had$_i$ he t$_i$ seen such a book.

   e. Under no circumstances would$_i$ he t$_i$ do it.

   f. Beyond the next hill stood$_i$ a large fortress t$_i$.[16]

   g. "Go to hell!" shouted$_i$ the witness t$_i$ at the judge.

Unlike regular English clauses, these constructions are headed by an auxiliary or main verb that moves in front of the subject. This means that if such a construction occurs in a selected position and there is no complementizer, a selected clause will be headed by a moved verb. According to prediction (9), such structures should be ungrammatical. The following examples show that this prediction is correct:[17]

(18)   a. I know that all too seldom does he bring her flowers. (Bolinger 1977:515)

    b. You can well imagine that not even that small consideration did he ever bestow on his partner. (Bolinger 1977:519)

    c. She swore/insisted/thought that never in her life would she accept this solution. (Grimshaw 1997:(44a))

    d. The publisher told us that never in his life had he seen such a book. (Grimshaw 1979:(46))

    e. He said that under no circumstances would he do it. (Rizzi and Roberts

---

[16] While locative inversion patterns with more obvious cases of movement with regard to the distribution of *that*, it is not obvious that the correct analysis of locative inversion involves movement of the verb. For some discussion, see Bresnan 1994.

[17] The following examples show that the matrix verbs in these examples are usually compatible with *that*-less complements: *I know (that) you're only joking, You can well imagine (that) he would be late, She swore/insisted/thought (that) she didn't do it, The publisher told us (that) he liked the book, The scouts reported (that) they had discovered a large fortress, The judge was shocked (that) he hadn't been informed earlier*.

78

1989:(42))

    f.   The scout reported that beyond the next hill stood a large fortress. (Hooper and Thompson 1973:(48))

    g.   ?*The judge was shocked that "Go to hell!" shouted the witness at him.

(19)   a.   *I know all too seldom does he bring her flowers.

    b'.   *You can well imagine not even that small consideration did he ever bestow on his partner.

    c'.   *She swore/insisted/thought never in her life would she accept this solution. (Grimshaw 1997:(44b))

    d'.   *The publisher told us never in his life had he seen such a book.

    e'.   ?* He said under no circumstances would he do it. (Rizzi and Roberts 1989:(43))

    f'.   *The scout reported beyond the next hill stood a large fortress.

    g'.   *The judge was shocked "Go to hell!" shouted the witness at him.

Thus, the condition on chains in (1) together with the present view of phrase structure correctly predicts one aspect of the distribution of complementizers, namely the correlation of verb movement and optionality of complementizers in selected clauses.[18] In contrast, this prediction doesn't follow in an obvious way on standard views (even given some condition like (1)).[19]

The next section extends this analysis to similar patterns related to adjunction.

3 Complementizer optionality and topicalisation/adjunction in complement clauses

As discussed in chapter 1, structures headed by the copy of a lower head not only feature in the representation of verb (s-)movement. On the assumption that no abstract functional heads are available whose specifier could contain adverbs or topicalised phrases and that there is no categorial projection, the need to place an adjunct can create structures very

---

[18] Some speakers seem to marginally allow embedded inversion into the highest position in certain contexts, as in (i) from McCloskey 1992:(81). (i) ?*Does the chair often know which of the promotions did the Dean support?*

[19] This story implies the diachronic prediction that English should have acquired complementizer optionality at the same time as it lost "V-to-I" movement. See Arnold 1995, 1997 for relevant discussion.

similar to s-movement structures. Given the assumption that phrase structure is at most binary branching, the presence of an adjunct can trigger the creation a new head position if all otherwise suitable adjunction sites already have two daughters. As in the case of verb movement, this additional head position, if it doesn't correspond to an additional lexical item, must inherit its properties from its daughter. Given the structure in (20a), adjunction can give rise to the structure in (20b), where the new root node inherits the categorial information of the lower node V:

```
(20) a.      V                  b.              V
           /  \                              /    \
          A   B                          Adjunct   V
                                                  / \
                                                 A   B
```

This structure of course parallels that of verb movement discussed in the previous section. Since the two highest nodes in (20b) are copies, the must form a chain. The formation of a chain means that this structure, which is dominated by head that is not in the root position of its chain, will be subject to the chain condition in (1) above. This case must be distinguished from a situation where a head that doesn't already have two daughters is available.[20] In such a case, it may be possible that the adjunct can be added as a daughter to this head, if no other constraint would be violated. The relevant structures are given in (21). In (21a), the verb V is dominated by a non-branching head F. As shown in (21b), placement of an adjunct in such a structure is possible without the creation of a new head position. In such a structure, the two highest nodes belong to different categories, and consequently no chain formation is necessary. The chain condition will not apply to a structure like (21b). Structures like (21b) can be found for example in Greek where a focused constituent in an embedded clause may precede complementizers like *oti* (Roussou 2000). This is illustrated in (21c).

```
(21) a.      F                  b.              F
             |                                /   \
             V                            Adjunct  V
           /  \                                   / \
          A   B                                  A   B
```
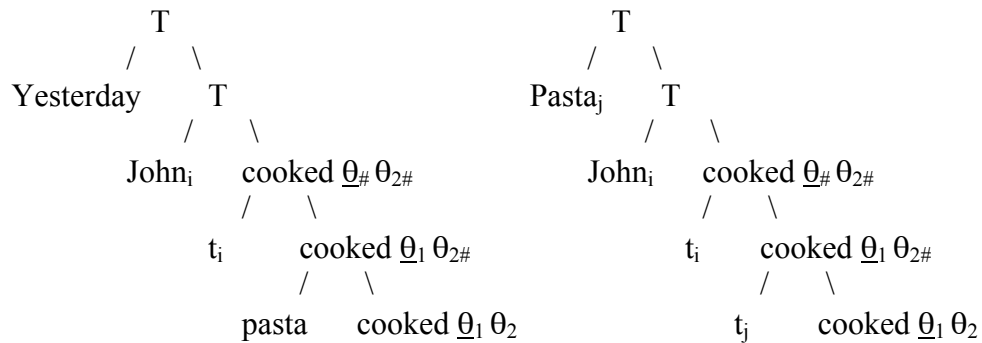
---

[20] This may be the case of languages that have special particles in adjunction structures, and it is the default in theories that allow abstract functional heads.
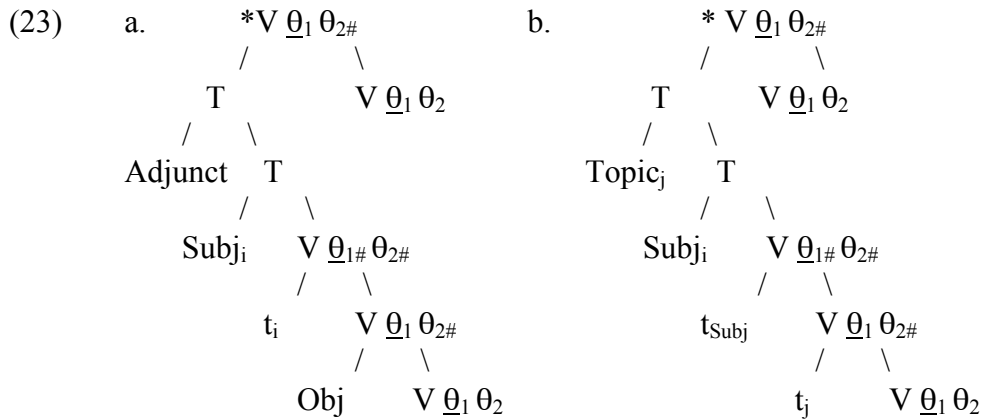
c. Fovame [$_C$ o   YANNIS oti tha   me   apatisi]
   fear-1SG   the-NOM Yannis that will   me-ACC cheat-3SG
   'I fear that Yannis will cheat on me.' (Marika Lekakou, p.c.)

Examples that correspond to the structure where a new head is created in (20b) are illustrated here. In the example in (22a), the adverb *yesterday* and, in (22b), the topicalised object *pasta* are adjoined to the top of the clause. Since neither the adverb nor the fronted object enter into a selectional relation with the root node, this node inherits the category from its (clausal) daughter unchanged:

(22) a. 'Yesterday John cooked pasta.'   b. 'Pasta, John cooked.'

```
              T                                    T
            /   \                                /   \
      Yesterday   T                         Pastaⱼ    T
                 /  \                               /  \
           Johnᵢ    cooked θ# θ2#            Johnᵢ    cooked θ# θ2#
                   /   \                             /   \
               tᵢ     cooked θ1 θ2#              tᵢ     cooked θ1 θ2#
                      /   \                             /   \
                  pasta   cooked θ1 θ2              tⱼ     cooked θ1 θ2
```
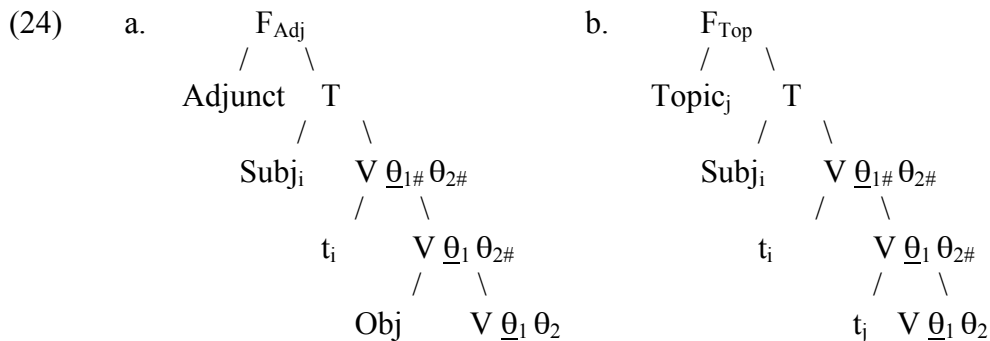
As in the case of verb movement discussed in the last section, the root node and its daughter must form a head chain because they are copies. And again, the membership in a complex chain of the root node of these structures affects the positions in which they can occur. In particular, the chain condition in (1) implies that these structures are not selectable, and therefore that they cannot sit in a position to which a θ-role is assigned. These two situations are illustrated here in a schematic way, adjunction in (23a), topicalisation in (23b):

(23) a.   *V $\underline{\theta}_1\,\theta_{2\#}$ b.   * V $\underline{\theta}_1\,\theta_{2\#}$
                /   \   /   \
           T   V $\underline{\theta}_1\,\theta_2$ T   V $\underline{\theta}_1\,\theta_2$
         /   \   /   \
    Adjunct   T   Topic$_j$ T
            /   \   /   \
      Subj$_i$ V $\underline{\theta}_{1\#}\,\theta_{2\#}$ Subj$_i$ V $\underline{\theta}_{1\#}\,\theta_{2\#}$
              /   \   /   \
         t$_i$ V $\underline{\theta}_1\,\theta_{2\#}$ t$_{Subj}$ V $\underline{\theta}_1\,\theta_{2\#}$
                 /   \   /   \
            Obj   V $\underline{\theta}_1\,\theta_2$ t$_j$ V $\underline{\theta}_1\,\theta_2$

It is then expected that even in a language without verb movement like English, complementizers should become obligatory if adjunction to the highest head in a complement clause occurs.

There are of course very popular alternatives to such analyses of topicalisation and adjunction. In particular, Rizzi 1997 and Cinque 1999, respectively, argue that topicalised constituents and adverbials and other adjuncts occur in the specifier positions of dedicated functional heads, not in adjoined positions. Simplified structures in a projection-free representation are given here:

(24)   a.   F$_{Adj}$ b.   F$_{Top}$
                  /   \   /   \
           Adjunct   T   Topic$_j$ T
                  /   \   /   \
             Subj$_i$ V $\underline{\theta}_{1\#}\,\theta_{2\#}$ Subj$_i$ V $\underline{\theta}_{1\#}\,\theta_{2\#}$
                     /   \   /   \
                t$_i$ V $\underline{\theta}_1\,\theta_{2\#}$ t$_i$ V $\underline{\theta}_1\,\theta_{2\#}$
                        /   \   /   \
                   Obj   V $\underline{\theta}_1\,\theta_2$ t$_j$ V $\underline{\theta}_1\,\theta_2$

While these proposals of course differ from the present one in their assumptions about clause structure and functional heads, the most important thing for the present argument is the fact that in these analyses, the root node does not form a chain with its daughter. That is, the heads F$_{Adj}$ in (24a) and F$_{Top}$ in (24b) have different categories from their daughters and consequently no head chain needs to be formed. Since there is no head chain, the chain condition in (1) does not block selection of such structures. This means that on such a view there is no obvious reason why selection of a clause with an initial adjunct should be

impossible, and consequently this view doesn't entail that there should be a restriction on the omission of an otherwise optional complementizer.

The argument is then the same as in the previous section. The present proposal predicts that adjunction and topicalisation should block omission of an otherwise optional complementizer in an embedded clause. The standard alternative makes no such prediction.[21] The empirical accuracy of this prediction is then again a good indication of the relative merit of the different theories.

Again, the present proposal gets it right. Where topicalisation or initial adverbs occur in a complement clause *that* cannot be omitted:[22]

(25)   a.   I hope that this book you will read. (Doherty 1997:3a)

       b.   This proves that Joyce he'd read but that Yeats he hadn't. (Doherty 1997:3a)

       c.   She prayed that next Wednesday the check would arrive. (Doherty 1997:12a)

       d.   We maintain that in London a nice flat is hard to find (Doherty 1997:12c)

       e.   She swore/insisted/thought that (,) most of the time (,) they accepted this solution. (Grimshaw 1997:(43a))

(26)   a.   *I hope this book you will read. (Doherty 1997:7a)

       b.   *This proves Joyce he'd read but that Yeats he hadn't. (Doherty 1997:7a)

       c.   *She prayed next Wednesday the check would arrive. (Doherty 1997:15a)

       d.   *We maintain in London a nice flat is hard to find (Doherty 1997:12c) (Doherty 1997:15c)

       e.   *She swore/insisted/thought (,) most of the time (,) they accepted this solution. (Grimshaw 1997:(43b))

While the present proposal predicts the effect of adjunction and topicalisation in a straightforward way, it is not clear how this effect could be captured in the alternative

---

[21] And again, an alternative that posits a covert complementizer in place of a missing *that*/*que* would pattern with the standard proposal.

[22] There are similar examples where a complementizer is not obligatory: (i)*I'm sure the lecture by Kayne he wouldn't miss.* (Watanabe 1993:145n23 (attributed to Noam Chomsky)) (ii) *He said no matter what stand he takes it would be misconstrued that he was sympathetic to one or the other of the Republicans.* (Elsness

approaches.

In fact, this point is even more serious in view of the result of the previous section, where the effects of verb movement on complementizer optionality were discussed. In both cases, the present proposal outperforms the alternatives. But beyond that, this proposal also can account for the fact that verb movement and adjunction and topicalisation affect complementizer optionality in the same way. Since verb movement and adjunction (through movement or base generation) give rise to the same structural configuration, their identical behaviour with regard to selection and especially the distribution of *that* is expected. It is once more not at all clear how this generalisation could be captured in different proposals.

The similar behaviour of movement of the verb to the highest head in a complement clause and adjunction to the top projection in a complement clause has already been noted by McCloskey (1992:12n9,19) (see also Kayne 1994:28). However, McCloskey's concern is not with the distribution of the complementizer *that*, and his CP recursion analysis is not designed to provide an analysis of it. Furthermore, while the parallel restrictions on the availability of embedded inversion and embedded adjunction are both related to whether or not the projection targeted by these operations is s-selected, they are ruled out by independent conditions in McCloskey's proposal. Following Rizzi and Roberts 1989:107, embedded inversion is ruled out by the assumption that I-to-C movement is substitution and substitution into a selected position violates the Projection Principle. Embedded adjunction is ruled out by a condition called Adjunction Prohibition formulated specifically to rule out adjunction to selected constituents.[23]

In Grimshaw 1997, the similarity of inversion in embedded clauses and adjunction in embedded clauses follows from a single constraint called Pure-EP:

(27)   No adjunction takes place to the highest node in a subordinate extended projection; and no movement takes place into the highest head of a subordinate extended projection. (Grimshaw 1997:394)

1984:524). I guess here the 'matrix clause' must be a parenthetical (as *I guess* in the sentence immediately preceding this parenthesis must be, actually).

A unified account of the two phenomena is of course desirable, and this constraint is clearly designed to capture this generalisation. However unless this constraint can be derived from independently motivated principles of the grammar it amounts to no more than a restatement of the facts. As Grimshaw 1997:395 points out, this formulation of Pure-EP suggests that there could in fact be two separate constraints, one on adjunction and one on verb movement. Whether this would be a more accurate analysis is an empirical question, and, as Grimshaw notes, at least for English, there seems to be no motivation for a separation of the two constraints. This can be taken as a further argument against Pure-EP, whose formulation invites such a separation.

In the next section, it will be discussed how this approach to complementizer optionality can be extended from complement clauses to a type of clause that is not selected, namely relative clauses.

4 Complementizer optionality in relative clauses
### 4.1 The distribution of the complementizer

So far, only one half of the chain condition in (1) has been discussed, namely its restriction of the selection of material that doesn't occupy the lowest position within its chain. However, as already noted, the implications of the chain condition go further. In particular, it does not only rule out selection *of* moved material, but also selection *by* moved material. This case is illustrated by the examples in (3) at the beginning of this chapter. As for the discussion of selection of higher chain members, the rejection of categorial projection is relevant here. Since there is no categorial projection, a structure whose head is not the lowest member of its chain will resemble a structure that is itself not the lowest member of its chain. Both structures are headed by a category that is not the lowest member of its chain. Given that selection is a relation between heads, this means that the chain condition not only rules out selection by a category that is not the lowest member of its chain, but also selection by a structure whose head is not the lowest member of its chain:

(28) A structure whose head is not in the lowest position in its chain cannot select.

---

[23] See also Chomsky 1986:6, 16, Rochemont 1989. McCloskey 1992:11 suggests that the Adjunction
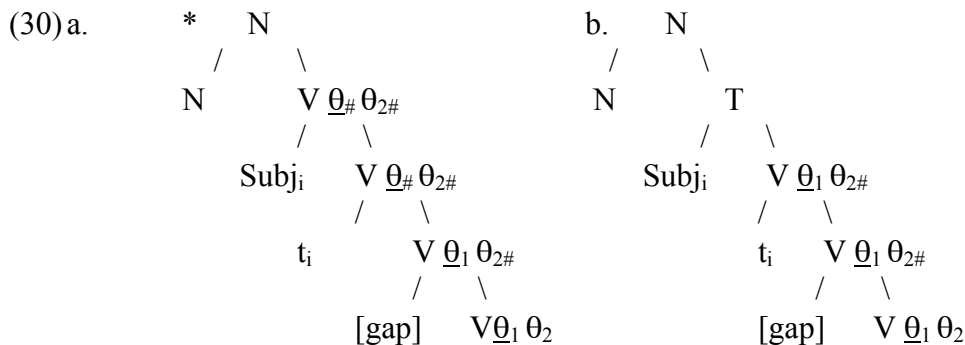
As discussed in section 1, the relation between a relative clause and the head that it modifies is similar to the selectional relation between a verb and an argument to which it assigns a θ-role. Thus, the chain condition and its correlate in (28) are expected to restrict the former relation, as well.

Given the assumption that selection takes place under sisterhood, a relative clause construction has the following structure. The relative clause constituent labelled RC selects its sister, the constituent labelled N.

(29)

```
            N
          /   \
        N       RC
        △       △
```
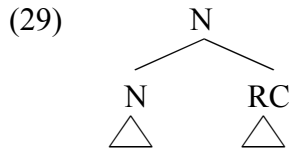
On the view, adopted here, that clauses with and without the complementizer *that* have different structures, in particular that only those with *that* have a complementizer level, this suggests that Vikner's generalisation should not be restricted to complement clauses, but also to relative clauses. It is expected that relative clauses behave in the same way as complement clauses with regard to the correlation of complementizer optionality and verb movement and adjunction/topicalisation.

In relative clauses in languages without verb movement, complementizers are predicted to be optional, whereas in relative clauses in languages with verb movement, complementizers are predicted to be obligatory. This is illustrated here in a schematic way with structures where the gap is in the object position of the relative clause. (30a) is a structure with verb movement and without a complementizer, while (30b) is a structure without verb movement or complementizer:

(30) a.   * N   b.   N
          /   \   /   \
        N   V $\underline{\theta_\#}\,\theta_{2\#}$ N   T
                /   \   /   \
            $Subj_i$ V $\underline{\theta_\#}\,\theta_{2\#}$ $Subj_i$ V $\underline{\theta_1}\,\theta_{2\#}$
                    /   \   /   \
                  t_i V $\underline{\theta_1}\,\theta_{2\#}$ t_i V $\underline{\theta_1}\,\theta_{2\#}$
                      /   \   /   \
                  [gap]   V$\underline{\theta_1}\,\theta_2$ [gap]   V $\underline{\theta_1}\,\theta_2$

Prohibition may be related to the Projection Principle, but he doesn't investigate this possibility.
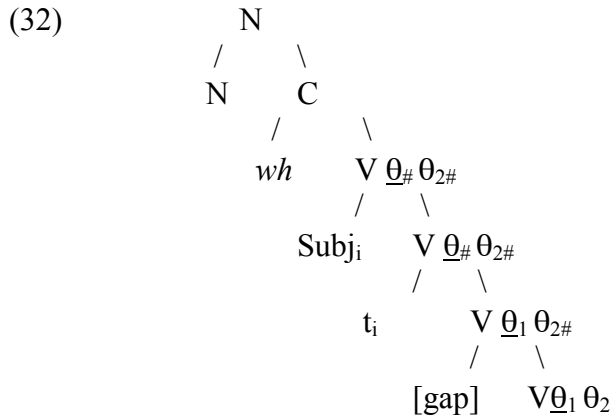
In the example in (30a) with verb movement, the two highest clausal heads in the relative clause form a head chain because they are copies. This means that selection by this relative clause should be impossible. In the example in (30b) without verb movement, the two highest categories in the relative clause do not form a head chain; hence selection by this relative clause should be possible.[24] In the corresponding structures with a complementizer, selection by the relative clause is ok with and without verb movement. (31a) is a structure with verb movement and with a complementizer, while (31b) is a structure without verb movement with a complementizer:

(31) a.
```
              N
             / \
          N     C
                |
              V θ# θ2#
              /   \
        Subji     V θ# θ2#
                  /   \
              tI      V θ1 θ2#
                      /   \
                  [gap]    Vθ1 θ2
```

b.
```
              N
             / \
          N     C
                |
                T
               / \
         Subji     V θ1 θ2#
                   /   \
               ti      V θ1 θ2#
                       /   \
                   [gap]    V θ1 θ2
```
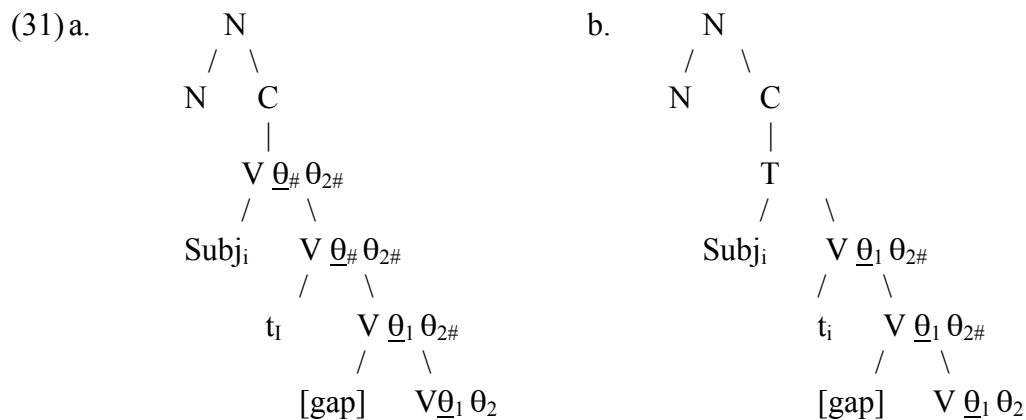
One more time, there is a well-known standard analysis that does not make these predictions. It has been widely assumed since Chomsky and Lasnik 1977 that all relative clauses are CPs and that the absence of an overt complementizer is the result of an optional deletion rule.[25] Thus, all relative clauses would have the structure in (32), and vary only with regard to how *wh* and C are pronounced. The standard analysis of relative clauses is illustrated in the following projection-free structure:

---

[24] For the moment, it is enough to assume that the gap in the object position turns the relative clause into an open predicate. These structures will be discussed in more detail below.

[25] Although there have also been several proposals according to which not all types of relative clauses have the same structure; cf. Weisler 1980, Doherty 1993, Grimshaw 1997, Bakovic and Keer 2001, Lee 2001 among others. For more discussion see below.

(32)

```
                    N
                  /   \
                N      C
              /   \
            wh      V θ_# θ_{2#}
                  /   \
              Subj_i   V θ_# θ_{2#}
                     /   \
                   t_i    V θ_1 θ_{2#}
                        /   \
                    [gap]    V θ_1 θ_2
```

Since on this view, all relative clauses are headed by a complementizer that does not form a chain with its daughter, no interaction of selection with verb movement is expected. This is true if verb movement is analysed as proposed here, but even more so if it is analysed in terms of head adjunction, where the head of the relative clause would not have moved regardless of whether there is a complementizer or not.

Again, the present model is more successful. In languages without verb movement complementizers may be optional in relative clauses, while in languages with verb movement, complementizers are obligatory:

(33) **Languages without verb movement: English, Danish**

    a.   I know a film (that) Frank directed.

    b.   Jeg kender en bog (som) denne lingvist har skrevet.

           I   know   a book   that   this   linguist has written (Vikner 1991:(4))

(34) **Languages with verb movement: Italian, French**

    a.   L'uomo *(che) ti vuole e la.

           'The man that wants you is there.' (Cinque 1981-82: (1))

    b.   L'homme *(que) je connais

           The man   that   I   know (Pesetsky 1998:(12d))

As with complementizer optionality in complement clauses, there are contexts where this pattern carries over to English. Where inversion, an initial adjunct, or a topicalised

constituent occur in a relative clause in English, the complementizer cannot be omitted.[26] In the following examples, the standard CP analysis of relatives is indicated to highlight the fact that under such an analysis the grammaticality pattern appears rather random:

(35) a.   the film [which$_i$ [$_C$ 0]   [under no circumstances did I want see t$_i$]]

   b.   the film [OP$_i$ [$_C$ that] [under no circumstances did I want see t$_i$]]

   c.   * the film [OP$_i$ [$_C$ 0]   [under no circumstances did I want see t$_i$]]

(36) a.   *The people when you get home who want to talk to you right away… (McCloskey 1992:(32))

   b.   This is the woman *(who) most of the time John likes. (Doherty 1993:63)

   c.   There's the man *(who) just this morning I met in the shop. (Doherty 1993:63)

(37) a.   this is the book [which$_i$ [$_C$ 0] years ago [everyone was talking about t$_i$]]

   b.   this is the book [OP$_i$ [$_C$ that] years ago [everyone was talking about t$_i$]]

   c.   * this is the book [OP$_i$ [$_C$ 0] years ago [everyone was talking about t$_i$]]

Just as in the case of complement clauses, the basic principles of the present proposal suffice to account for the observed pattern.[27] In contrast, the standard analysis would have to invoke a novel condition to capture the data.[28]

---

[26] While it seems clear that relative clauses "strongly disallow inversions" (Rizzi and Roberts 1989:114 note 21), inversion appears to be marginally possible for some speakers. In examples (i) to (iii) the relatives are presumably non-restrictive, and these differ from restrictive relatives also in other ways (see McCawley 1988, Fabb 1990, Borsley 1992). However, (iv) and (v) seem to involve restrictive relatives, and are hence potential problems: (i) *These are the people none of whom had I ever seen* (Culicover 1991:16); (ii) *Here are the results, none of which were we expecting* (Hudson 1995); (iii) *A formal wedding invitation should come in a squarish envelope, inside which should be several increasingly small envelopes accompanied by some sheets of what appears to be Soviet Union toilet paper* (Green 1996:(6f)); (iv) *Places where, upon mentioning the name of an habitue friend, might be obtained strange whiskey and fresh gin in many of their ramifications* (Green 1996:(6g)); (v) *These are the people none of whom had I ever seen* (Culicover 1991:16).
[27] The following example from McCloskey 1992:(33) shows that adjunction to non-relative clause adjuncts is also impossible: *I graduated while at college without having really learned anything*. This may suggest that the relation between non-relative clause adjuncts and their hosts may also be subject to the chain condition.
[28] In contrast to adjunction to clausal heads adjunction to nominal heads does not seem to trigger chain formation. If adjunction to N triggered the formation of a head chain, no more than one (restrictive) modifier should be possible. However, adjunction to nominals does not appear to be restricted in such a way (e.g. *the big book about politics by Chomsky that everyone read*.) At this point I can only speculate that the reason for this clause/noun asymmetry may have to do with differences in the functional structure in the two domains. A more serious investigation of this issue remains a subject of future research.

4.2 The internal structure of non-*wh* relative clauses

I consider now in a little more detail the structure of relatives without a complementizer in English. First, there is the question of what is the exact structure of a relative clause without *that* or *wh*. According to the standard analysis, such relative clauses contain a null operator that moves from the gap position to the highest specifier within the relative clause. However, in the present system such an analysis would be problematic. This is due to the assumption that there are no abstract empty heads, and hence that if there is no visible complementizer, there is no covert C-position either.[29] The following is a relevant structure:

(38)
```
        *    N
            /   \
        N         γ
                /   \
            OP        T
                    /   \
               Subj_i    V θ_{1#} θ_{2#}
                        /   \
                   t_i        V θ_1 θ_{2#}
                            /   \
                       [gap]      V θ_1 θ_2
```
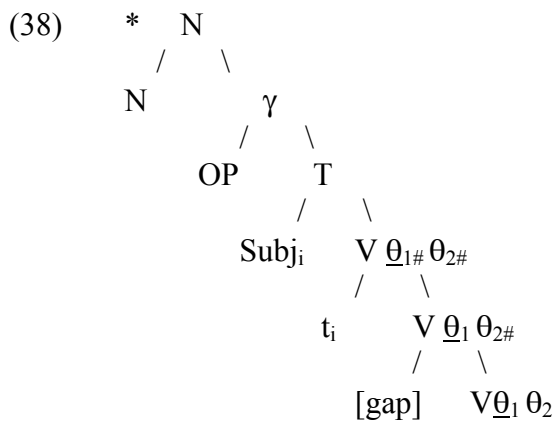
This structure raises the question of what could be the category of the highest head in the relative clause, i.e. the position indicated as γ above. Since it cannot be an abstract functional head, the only possibility is that it inherits its category from its daughter T. The head dominating OP and its daughter, would then have to form a head chain, because both would be of category T, i.e. copies. However, the relation between the relative clause and the head that it modifies is one of selection. The existence of a head chain containing the highest head in the relative clause would then mean that this clause cannot modify its sister due to the chain condition. Since there is no other position that a null operator could occupy in a structure without an abstract C-head, it follows that an analysis of *that*-less relative clause in terms of null operator movement is not compatible with the present

---

[29] Of course an overt *that* may be subject to deletion in well-defined, learnable contexts. This is the case of *that*-deletion in accordance with the Doubly-filled-Comp Filter; cf. Chomsky and Lasnik 1977, Pesetsky 1998.
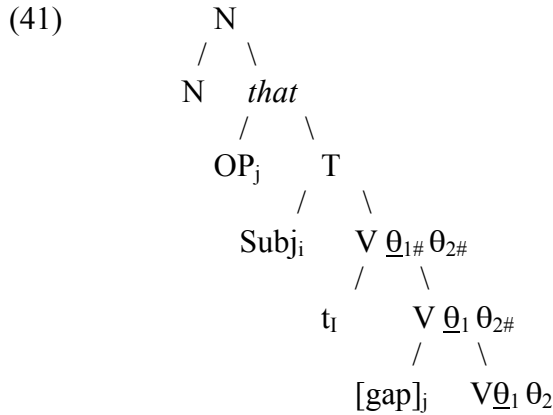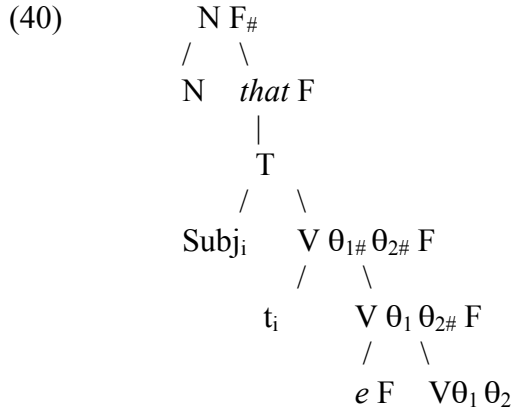
approach to clause structure.

I instead adopt Doherty's 1993 proposal, according to which the A-bar chain that turns the relative clause into a predicate is formed between the N constituent within the D constituent that the relative clause modifies and a base-generated empty category within the relative clause. To implement this idea, a proposal of Neeleman and van de Koot 2002 can be used according to which traces of movement are functions that need to be satisfied by an appropriate antecedent. The function introduced by the gap will be copied upwards until it is satisfied. It is important here that the relation between the gap in the relative clause and the modified noun is not a movement/chain relation; if it was a movement/chain relation, the chain condition in (1) would imply that the head noun could not be selected by the relative clause. This means that the function introduced by the gap in a *that*-less relative must differ from the function that is introduced by the gap in a *wh*-question, for example. This difference would correspond to the difference between the base-generated empty category in relatives and the *wh*-trace in Doherty's proposal.[30] In a *that*-less relative, the gap would then introduce a function F that is copied upwards until the node containing the function dominates the noun modified by the relative clause, or, in other words, until the head that carries the function is the sister of the modified noun. This structure is illustrated here:

(39)
$$
\begin{array}{c}
\text{N } F_{\#} \\
/ \quad \backslash \\
\text{N} \qquad \text{T F} \\
\quad / \quad \backslash \\
\text{Subj}_i \qquad \text{V } \theta_{1\#}\,\theta_{2\#}\, \text{F} \\
\qquad / \quad \backslash \\
\text{t}_i \qquad \text{V } \theta_1\, \theta_{2\#}\, \text{F} \\
\qquad\quad / \quad \backslash \\
e\,\text{F} \quad \text{V}\theta_1\,\theta_2
\end{array}
$$
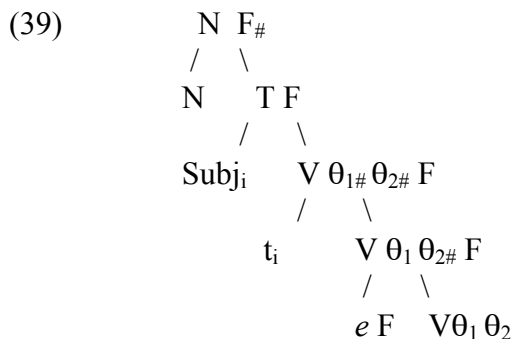
This analysis of *that*-less relative clauses entails the following pattern for English relative clauses. There is one possible structure for *that*-less relatives, namely that in (39). *Wh*-relatives can be analysed as in the standard analysis, with movement of the *wh* operator turning the relative clause into a predicate. However, there are two possible analyses of

that- relative clauses. The first one would be along the lines of that just proposed for *that-*less relative clauses, and the second would involve the standard null-operator movement. The two possibilities are illustrated here. (40) shows a relative clause with *that* without movement of a null operator; (41) shows a relative clause with *that* with operator-movement:

(40)
```
            N F_#
           /    \
          N    that F
                 |
                 T
               /   \
          Subj_i    V θ_{1#} θ_{2#} F
                    /    \
                 t_i      V θ_1 θ_{2#} F
                         /    \
                      e F    Vθ_1 θ_2
```

(41)
```
             N
           /   \
          N    that
               /   \
            OP_j     T
                   /   \
              Subj_i    V θ_{1#} θ_{2#}
                       /    \
                    t_I      V θ_1 θ_{2#}
                            /    \
                       [gap]_j    Vθ_1 θ_2
```

The second possibility, the one involving a null operator, becomes available in *that-*relative clauses because these involve the extra head position that contains *that*. This means that the null-OP can sit in the specifier position of *that*, and no adjunction/head chain structure needs to be created.

The choice between these two analyses does not affect the arguments about complementizer optionality here, so I will only provide a brief overview over some of the arguments that may affect the choice. The main difference between the two analyses is that on the operator-less view, *that* relatives are very similar to *that-*less relatives, while on the

---

[30] A similar distinction is also necessary in other frameworks. While movement of a null operator turns a

operator view they are more similar to *wh*-relatives. The arguments that I know of are inconclusive because in some ways *that*-relatives pattern with *that*-less relatives, while in others they pattern with *wh*-relatives.

There are range of data that suggest that the operator-movement analysis is more appropriate for *that*-relatives because they behave in a similar way to *wh*-relatives. Stacking is the term used to describe a structure in which more than one relative modifies the same head. Both *wh*- and *that*-relatives allow stacking, while *that*-less relatives do not. (The following examples are all from Doherty 1993:59.)

(42) a.   the man who Mary met who John likes

b.   the book that Bill bought that Max wrote

c.   the book which Bill bought that Mary disliked

(43) a.   * the man Mary met John likes

b.   * the book Bill bought Max wrote

The same pattern is found with regard to extraposition. *That*-less relatives cannot usually be extraposed, while both *wh*- and *that*-relatives can be:

(44) a.   A moose (that) Bill shot at appeared.

b.   A moose appeared that Bill shot at.

c.   * A moose appeared Bill shot at.

(45) a.   The man (who) Bill knew arrived yesterday.

b.   The man arrived yesterday who Bill knew.

c.   * The man arrived yesterday Bill knew.

Thus, these data suggest that *that* relatives should receive a parallel analysis to *wh* relatives.

However, there are also some data that go the other way. As is well known, restrictive relatives differ from non-restrictive relatives in that they can only be introduced by *wh*-phrases, i.e. there are no non-restrictive *that*- or *that*-less relatives (cf. McCawley 1988:418, Kayne 1994:163 note 69):

---

clause into a predicate, movement of a *wh*-operator does not need to.

(46) a.   The food, which had taken three hours to prepare, was a great success at the party.

    b.   * The food, that had taken three hours to prepare, was a great success at the party.

    c.   * The food, had taken three hours to prepare, was a great success at the party.

(47) a.   Frank, who they just fired, is really a nice guy.

    b.   * Frank, that they just fired, is really a nice guy.

    c.   * Frank, they just fired, is really a nice guy.

If the ability to function as a non-restrictive relatives is due to internal properties of the relative clause, this contrast seems to support an analysis of *that*-relatives along the lines of *that*-less relatives.

    Further data that illustrate differences between *that*- and *wh*-relatives have been observed by Lee 2001, who notes that only *that*- and *that*-less relatives allow existential *there* as their subject, while *wh*-relatives don't (Lee 2001: (18-20)):

(48) a.   The few salespeople [that there were] did nothing to help us.

    b.   The few salespeople [there were] did nothing to help us.

    c.   * The few salespeople [who there were] did nothing to help us.

Similarly, predicate nominals may appear as heads of *that*-relatives, but not of *wh*-relatives (Lee 2001: (43,44)):[31]

(49) a.   (He probably voted for Nader, being) the idealist that/*who he is.

    b.   He's no longer the athlete that/*who he used to be.

    Finally, Larson 1983: chapter 2 argues that data from adverbial relatives motivate distinct structures for *that*- and *that*-less relatives on the one hand and *wh*-relatives on the other. Since these data are, as Larson notes, "somewhat cumbersome to describe in prose" (1983:45), and the point that sometimes *that*-relatives do not pattern with *wh*-relatives should be clear by now, I refer to his work for discussion.

---

[31] The slight marginality of the *that*-less version in these examples may to the general reluctance to *that* omission in clauses embedded below N.

The approach to the optionality of complementizers from the perspective of the chain condition in (1) offers a natural explanation for the parallel behaviour of complement clauses and relative clauses. Complement clauses are implicated by the chain condition because they are selected, while relative clauses are implicated because they select. This seems to be a more explanatory way to capture the facts than simply to subsume both relative and complement clauses under the label of embedded clauses (Grimshaw 1997), or to propose that a relative clause "is not an adjunct, but counts as a complement – perhaps a second object of D (NP being the first object)" (Pesetsky 1998:356).

4.3 Summary

In this section, it was shown how different processes that occur inside different types of subordinate clause affect the distribution of complementizers. It was shown that where a pre-subject constituent, such as a sentence adverb or a topicalised phrase, occurs in English, the complementizer is obligatory. The same holds of verb movement. Where verb movement occurs inside a subordinate clause, the complementizer cannot be omitted. These patterns were derived from the prediction that a structure whose head is not in the deepest position within its chain cannot be selected and cannot select. The analysis is based on the condition on the selection of chains in (1) and the assumptions that clause structure does not contain abstract functional heads and that new additional positions are in principle freely generable. The next section will be concerned with two further constructions that bear on this generalisation, namely so-called "embedded V2 clauses" and apparent V2 relatives in German.

5 Embedded V2 clauses in German?

As discussed earlier, the theory implies that a clause headed by a moved verb cannot be selected. In section 2, the consequences of this prediction for complement clauses in languages with and without so-called V-to-I verb movement as well as the limited cases of embedded inversion in English were discussed. This ban on selection of a clause headed by a moved verb also makes an interesting prediction about a further type of examples, namely so-called EMBEDDED V2- (EV2-) clauses in a language like German.

It is well known that the propositional argument of certain verbs in German can be

realised as a V-final clause introduced by a complementizer or by a V2 clause without a complementizer:

(50) a.　Klaus sagte [dass Werner Herzog seine Schuhe gegessen hat]

　　　　Klaus said that Werner Herzog his shoes eaten has

　　　　'Klaus said that Werner Herzog ate his shoes.'

　　b.　Klaus sagte [$_{EV2}$ Werner Herzog hat seine Schuhe gegessen]

On a conventional analysis that treats V2 as movement of the verb into the C-position, there is no obvious reason why the structures in (50a) and (50b) should be treated differently. On such a view, the bracketed clauses would be of category CP in both cases. The examples in (50) would have structures like the following:

(51) a.　[Klaus [$_{VP}$ sagte [$_{CP}$ dass Werner Herzog seine Schuhe gegessen hat]]]

　　b.　[Klaus [$_{VP}$ sagte [$_{CP}$ Werner Herzog [$_C$ hat$_i$] seine Schuhe gegessen t$_i$]]]

Any systematic differences that may be found between the two constructions would require an explanation that is independent of the structural configurations involved.

　　　The present proposal makes a different claim about pairs like (50) since it implies the prediction in (9), repeated here:

(9)　A clause headed by a moved verb cannot be selected.

Since on the present view the V2 clause in (50b) is headed by a moved verb, (9) entails that it cannot be selected, and hence that it cannot occupy the same structural position as the V-final clause in (50a). A simplified structure of (50a) is given here:

(52)　　　　　　　　V $\underline{\theta}_1$ $\theta_{2\#}$
　　　　　　　　　／　　　　　＼
　　　　　　　dass　　　　V $\underline{\theta}_1$ θ_2
　　　　　　　△
　　　　　　　… V

Here, a head that didn't move, namely the complementizer *dass*, heads the embedded clause, and one of the matrix verb's θ-roles is satisfied by the embedded clause. This is the

normal case of selection.

In contrast, the bracketed V2 clause in (50b) is headed by a moved verb and has something like the following structure:

(53)   V $\underline{\theta}_{1\#}\,\theta_{2\#}$
            /   \
        Subj   V $\underline{\theta}_{1\#}\,\theta_{2\#}$
                   △

                  …
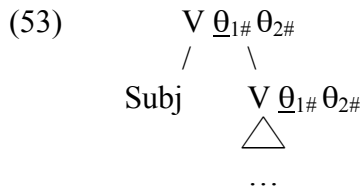
The head of this structure is the moved verb V $\underline{\theta}_{1\#}\,\theta_{2\#}$. On the assumption that all θ-roles must be satisfied, the prediction in (9) then entails that this structure cannot occur in a selected position because this structure cannot enter selectional relations. This means that (50b) cannot have the following structure, which would parallel that of (50a) given in (52):

(54)   * $V_{matrix}\,\underline{\theta}_1\,\theta_{2\#}$
                 /   \
         V $\underline{\theta}_{\#}\,\theta_{2\#}$ $V_{matrix}\,\underline{\theta}_1\,\theta_2$
          /   \
       Subj   V $\underline{\theta}_{\#}\,\theta_{2\#}$
                 △

                …
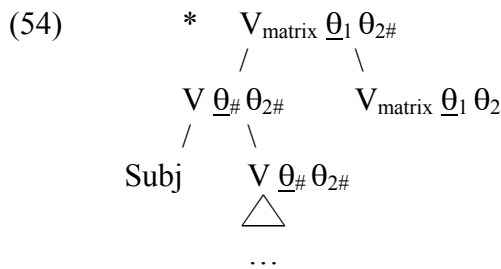
Instead the bracketed V2 clause in (50b) could be analysed as an independent clause. This independent clause could either be adjoined to the matrix clause, or the matrix clause could be analysed as a parenthetical. The examples in (50) would then have something like the structures in (50):

(55) a.   [Klaus [$_{VP}$ sagte [$_{CP}$ dass Werner Herzog seine Schuhe gegessen hat]]   = (51a)

   b.   [Klaus sagte] [Werner Herzog hat seine Schuhe gegessen]

Note that in a structure like (55b), there is no overt recipient for the object θ-role of *sagte*. Since the second clause in (55b) clearly is interpreted as the object of *sagte*, the theory implies that there must be some device that allows a linking of the apparently embedded V2 clause to the object θ-role of the apparent matrix verb.

Since in (55) the V-final example has a different structure from the EV2 example, it is expected that the two constructions differ in other ways than just the position of the verb.

In particular, on the standard analysis, the EV2 clause sits in a complement position, which suggests that extraction from it should be possible. In contrast, given a parenthetical analysis like (55b), extraction from the EV2 clause should be impossible.

One argument that has been used to support the claim that EV2 clauses can occupy the structural complement position is the apparent parallelism of EV2 clauses to verb-final clauses from which a phrase has been extracted: (This discussion follows Reis 1995.)

(56) a. Wieviel sagte sie dir dass er dafür bezahlen würde?
          How much said she you that he for this pay would
     b. Wieviel sagte sie dir würde er dafür bezahlen?
          How much said she you would he for this pay
          'How much did she say to you that he would pay for this?'

The example in (56a) can be analysed in terms of wh movement out of the embedded V-final clause into the highest specifier of the matrix clause:

(57)   Wieviel$_i$ sagte sie dir [t$_i$ dass er dafür e$_i$ bezahlen würde]?   Haider 1991: (10)
         How much said she you that he for this pay would

Haider 1991 (among others) suggests a parallel analysis of (56b), where the wh phrase apparently moves from an embedded V2 clause into the matrix clause.

(58)   Wieviel$_i$ sagte sie dir [e$_i$ würde [er dafür e$_i$ bezahlen]]?
         How much said she you would he for this pay
         'How much did she say to you that he would pay for this?'   Haider 1991: (10)

If such an analysis is correct, it follows that the V2 clause must be the complement of the matrix verb *sagte*, because otherwise such movement would be impossible.

While such an analysis may be appealing, it is clear that given the assumptions made here that led to the prediction in (9) above, such an analysis cannot be maintained. Since the EV2 clause cannot be a complement of *sagte* an analysis of the following type is necessary, where the apparent matrix clause is analysed as an independent clause:

(59)   [Wieviel$_i$ [sagte sie dir] würde er dafür e$_i$ bezahlen]?

         How much said she you would he for this pay

         'How much did she say to you that he would pay for this?'

The examples in (60) further illustrate the difference between the two types of analyses:

(60) a.   Wo glaubst du wohnt sie seit 1985?   Reis 1995:(1)

           Where believe you lives she since 1985

           'Where do you believe she has lived since 1985?'

    b.   In Bonn meint Franz wohnt sie seit 1985.

          In Bonn thinks Franz lives she since 1985

          'Franz thinks it is in Bonn that she has lives since 1985.'

The two types of analyses of these examples are contrasted here. (61) is an analysis in terms of extraction from an embedded clause; (62) treats the EV2 clause as an independent clause:

(61) a.   [Wo glaubst du [t' wohnt sie t seit 1985]]?   Reis 1995:(2)

    b.   [In Bonn meint Franz [t' wohnt sie t seit 1985]].

(62) a.   [Wo [glaubst du] wohnt sie t seit 1985]?   Reis 1995:(3)

    b.   [In Bonn [meint Franz] wohnt sie t seit 1985].

According to the first analysis above, a phrase moves out of V2 clause that sits in the complement position to a selecting verb. According to the second analysis, no such movement into the apparent matrix clause takes place; the two clauses are syntactically independent.

The following examples show that indeed the extraction analysis is not tenable, at least not for all cases of apparent extraction. This means that, for empirical reasons, any theory needs to assume that something like the parenthetical analysis, which is forced by the present model, is available for these structures. The example in (63) cannot be plausibly analysed in terms of extraction from a V2 clause complement, because, as the

examples in (64) show, the putative matrix verb *fragen* doesn't take V2 complements:

(63)   Wen   fragte Hans, wird der Chef entlassen?   Grewendorf 1988:84
        whom   asks Hans   will the boss fire
        'Who, Hans asks, is the boss going to fire?'

(64) a.   Hans fragte, ob der Chef ihn entlassen wird.
           Hans asked if the boss him fire would
           'Hans asked if the boss would fire him.'
   b.   * Hans   fragte, ihn wird   der Chef entlassen.
           Hans   asked him would the boss fire
   c.   * Hans   fragte, der Chef wird ihn entlassen.
           Hans   asked the boss would him fire

This means that an extraction analysis of (63) along the lines proposed by Haider for (56b) above, is not very convincing. Instead, an analysis like that illustrated in (65b) seems necessary:

(65) a.   * [Wen   fragte Hans   [t   wird der Chef entlassen t']]?   cf. (58)
               whom asks Hans   will the boss fire
   b.   [Wen   [fragte Hans]   wird der Chef entlassen t']?

      Further examples that clearly show that any theory will have to assume something like a parenthetical analysis are the following (from Reis 1995:46):

(66) a.   Wo(/In Bonn) wohnt sie **meint er** (/sagt Peter) mit dem Kind seit 1985?
           where(/in Bonn) lives she thinks he (/says Peter) with the child since 1985
   b.   Wo(/In Bonn) wohnt sie mit dem Kind **meint er** (/sagt Peter)seit 1985?
           where(/in Bonn) lives she with the child thinks he (/says Peter)since 1985
   c.   Wo(/In Bonn) wohnt sie mit dem Kind seit 1985 **meint er** (/sagt Peter)   ?
           where(/in Bonn) lives she with the child since 1985thinks he (/says Peter)

Here, the putative matrix *meint er* is preceded by a non-constituent. Thus, an extraction analysis would only be possible at the cost of introducing movement of non-constituents, which is not obviously desirable. These examples then add further evidence that a parenthetical analysis is independently necessary. The following examples show that such structures can also be recursive:

(67) a. Was könne   man   meint   er behaupte Karl nicht länger dulden?
         What can-SUBJ one   thinks   he claims-SUBJ Karl no longer tolerate
         'What does he think Karl claims can no longer be tolerated?' (Reis 1995:(29))

     b. Was glaubst du schätzt er, wieviel das Auto kosten wird?
         What believe you estimates he how much the car cost will
         'How much do you think he estimates that the car will cost?' (Reis 1995:(30))

A further range of data that highlight the special properties of EV2 clauses has been observed in Den Besten 1983. First, a verb of saying can introduce an indirect discourse that contains several V2 clauses in sequence that can all be in the subjunctive, which marks indirect discourse in careful registers (cf. Den Besten 1983:109).

(68) *Er sagte, er wäre   nicht damit einverstanden. Der Karl wäre   ein netter*
     he said   he was-SUBJ not with-it agreed   the Karl was-SUBJ a nice
     Bursche, wenn er nicht zuviel getrunken hätte. Aber man wüsste   ja, dass das
     guy   when he not too-much drunk   had   but one knew-SUBJ PRT that that
     normalerweise nicht der Fall wäre. Warum hätte man ihn überhaupt eingeladen?
     usually   not the case was-SUBJ why   had-SUBJ one him at-all   invited
     Der wäre   ja sonst nicht interessiert an Bürgerinitiativen.
     He was-SUBJ PRT   not interested   in Citizens' Committees
     'He said (that) he agreed. Karl was a nice guy when he hadn't drunk too much. But everyone knew that usually that wasn't the case. Why had he been invited anyway? Usually he doesn't care about citizens' committees.'

This example is relevant because in many languages subjunctive mood is restricted to complements of certain verbs, which suggests that it is licensed under selection. Thus, at first sight, such examples seem to provide evidence against a parenthetical analysis of EV2

constructions. However a look at more data shows that such examples indeed cannot involve selection.

While it might be plausibly argued that a sequence of declarative V2 clauses could constitute some coordination structure that sits in the canonical complement position of a selecting V, such an analysis is not possible for the above text. The important fact is that among the subjunctive V2 clauses there is a question, and after the question, the text continues with a further declarative. As the following example shows, such a question is not a possible complement of *sagen*:

(69)   * Er sagte, warum hätte man ihn überhaupt eingeladen?

However, if the sequence of subjunctive V2 clauses in (68) contains a clause that can't be a complement of the matrix verb, it follows that the whole sequence also can't be selected by this verb.

This point becomes even clearer in the following examples (cf. Den Besten 1983:122f):

(70) a.   *Das Telefon klingelte. Es war seine Chefin. Sein Kollege wäre   krank*
          the phone   rang   it was his   boss   his colleague is-SUBJ sick
          und er möchte   doch bitte zum   Büro kommen.
          and he may-SUBJ PRT please to-the office come.
          'The phone rang. It was his boss. She said his colleague was sick, and asked him if he could come to the office.'

     b.   *Aber er wollte nicht   mitmachen. Es wäre   ja   unerhört   dass man*
          But he wanted not   cooperate.   It was-SUBJ PRT outrageous   that one
          nicht verstünde,   dass er sich   weigerte   mit solchen
          not understood-SUBJ   that he himself   refused-SUBJ   with such
          Faulenzern   zu arbeiten.
          bums   to work
          'But he didn't want to cooperate. (He said that) It was outrageous that they didn't understand that he refused to work with such bums.'

(71) a.   * Das Telefon klingelte. Es war seine Chefin. Dass sein Kollege krank wäre,

           und er doch bitte zum Büro kommen möchte.

       b.   * Aber er wollte nicht mitmachen, dass es ja unerhört wäre, dass…

Here, the subjunctive indirect discourse in V2 occurs without an introducing verb of saying altogether. (The examples in (71) show that in the same context a V-final clauses is not possible.) Since in these examples V2 indirect discourse occurs without being licensed by a selecting verb, it is clear that this strategy may also be used in a context where a verb of saying occurs. Thus, these data show that regardless of the particular analysis that these EV2 clauses receive, the theory must provide a way to account for their occurrence as independent clauses.

The above examples have provided clear evidence that not all apparent V2 complement clauses can be analysed as structural complements and that instead some examples must be analysed as independent clauses. To the extent that this argument is correct, the unanswered question of how the apparent matrix verb's internal θ-role could be satisfied is then no longer an argument in favour of the extraction analysis. Since any theory will have to include a device that can deal with uncontroversial cases of missing complements like those in (65b) and (67) where the EV2 clause cannot be a structural complement, this device will also be available for the cases that could arguably be analysed in terms of extraction.[32] Moreover it may be noted that a conventional approach to clause structure is left with two possible analyses for EV2 clauses, namely EV2 clauses as complement clauses or an analysis of the relevant constructions as independent clauses. While this doesn't seem terribly elegant, it also raises the complex question of how one (the linguist or the learner) should choose between the two possibilities. In contrast, there is only one analysis of EV2 clauses that is compatible with the assumptions made here. Since this analysis predicts that there should be certain differences between V-final complement clauses and EV2s, this can be seen as an advantage of this approach.[33]

---

[32] See Banfield 1982:41-52, Reis 1995 for relevant discussion.

[33] EV2 constructions in German are restricted to bridge verbs. It is roughly the same class of verbs that licenses omission of the complementizer in Italian subjunctives (cf. Giorgi and Pianesi 1997, Poletto 2001). As noted in section 2, these examples appear to be an exception to Vikner's generalisation, which links complementizer optionality to verb movement. However, the fact Italian complementizer optionality is restricted to a similar class of verbs as EV2 constructions in German suggests that possibly a similar analysis

Before concluding the discussion of apparently embedded V2 clauses in German, a further range of data should be mentioned, namely apparent V2 relatives. Such examples are illustrated here:

(72) a. Es war einmal ein Mann, der verstand allerlei Künste.   Schuetze-Coburn 1984:(4)

        There was once a man the knew all-kind-of tricks

        'There once was a man who knew all kinds of tricks.'

   a'. Es war einmal ein Mann, der allerlei Künste verstand.

   b. Das Blatt hat eine Seite, **die ist   ganz schwarz**.   Gärtner 2001:(1a)

        The sheet has a page   that is   all black

        'The sheet has a page that is all black.'

   b'. Das Blatt hat eine Seite, **die ganz schwarz ist**.   Gärtner 2001:(1b)

   c. Bei den Büchern sind welche, **die gehören mir nicht**.   Gärtner 2001:(2a)

        Among the books are some   that belong me not

        'Among the books, there are some that don't belong to me.'

   c'. Bei den Büchern sind welche, **die mir nicht gehören**.   Gärtner 2001:(2b)

At first sight, the examples in (72a,b,c) differ from those in (72a',b',c') only in that the bold-faced second part of the examples has V2 rather than V-final order. Since the V-final clauses are relatives, these data seem to suggest that German has V2 relative clauses alongside V-final ones. However, given the discussion in the previous section, where it was argued that relatives select the noun they modify, it is clear that the existence of V2 relatives would be incompatible with the present proposal. The problem is that V2 clauses are headed by a moved verb, and that the condition in (1) states that selection by a moved element is impossible. If relatives select their modifiee, there should then be no V2 relatives. Indeed Gärtner 2001 provides a range of arguments to show that V2 clauses like those in (72a,b,c) are not relative clauses at all, but that they instead require some kind of parenthetical analysis. (I will not repeat Gärtner's discussion here.)

    In this section, I have discussed a range of apparently embedded V2 clauses in German in light of the prediction that embedded clauses cannot be headed by a moved

---

carries over to the Italian case. If Italian subjunctives could also receive a type of parenthetical analysis, they are no longer problematic for Vikner's generalisation.

verb. I provided a range of empirical arguments that showed that the relevant V2 clauses in fact should not be analysed as structural complements, but instead as some kind of parenthetical construction.[34]

6 Clause typing and verb movement

Much of this chapter has dealt with the effects that verb movement can have on the structure and distribution of clauses. This section explores yet another angle of this issue. In particular, I want to show how the interaction between the condition on chains in (1) and verb movement can throw some light on two well-known problems, namely the root-embedded asymmetry in German and the complementary distribution of verb movement and a complementizer in that language. First, I discuss in more detail why it may be worth thinking about these two problems after so many years of research devoted to them.

In German verb movement takes only place in structures that are not introduced by a complementizer.[35] This pattern is illustrated here:

(73) a.   Der Affe   hat$_i$ die Kokosnuss geklaut t$_i$

the monkey has   the coconut   stole

'The monkey stole the coconut.'

   b.   Ich weiss [dass der Affe die Kokosnuss geklaut hat]

I know   that the monkey the coconut stole has

'I know that the monkey stole the coconut.'

The standard analysis of this pattern is generally attributed to Den Besten 1983. In updated terminology, the idea is that the moved verb in (73a) and the complementizer in (73b) occupy the same structural position, usually labelled C:

(74) a.   [$_{CP}$ Der Affe [$_C$ hat$_i$]   die Kokosnuss geklaut t$_i$]

   b.   [$_{CP}$ [$_C$ dass] der Affe die Kokosnuss geklaut hat]

---

[34] For more discussion of EV2 see e.g. Cinque 1989, De Haan 2001, Den Besten 1983, Frank 2000, Gärtner 2001, 2002, Heycock 2001, Iatridou and Kroch 1992, Öhl 2002, McCloskey 1992, Penner and Bader 1991, Reis 1995.

Given the assumption that movement to a position is only possible if this position is empty, the assumption that a verb can only move to an existing position, and the assumption that there is no other position available for the verb to move to, the presence of the complementizer *dass* will always block verb movement. On these assumptions, the complementarity of verb movement and complementizer is captured in a very simple and elegant way.

However, the view of clause structure has changed a lot in the last 25 years, and given certain current standard assumptions about clause structure and head movement, it is difficult to see how the elegance and simplicity of Den Besten's original proposal could be maintained. First, Den Besten's account crucially relies on the assumption that movement into a position that is not empty is impossible. Of course it is now a standard assumption that head movement is adjunction of a head into a position that already contains some material (either an affix, some features, or a separate word in the case of incorporation). This means that for a Den Besten-type account to work, it is at least necessary to assume that movement into the highest head position in a clause, which is usually taken to be the functional head C, differs in some fundamental way from other types of head movement. This problem gets more complicated if the results from earlier sections are to be taken into account which supported the claim that clause structure is not universal and that in different constructions heads of different categories can be the highest head in a clause. While in the present proposal such a difference follows from the way verb movement can create structure and its interaction with the chain condition, it was argued earlier that this correlation doesn't carry over to head adjunction theories in an obvious way.

Second, it is often assumed that the structural COMP position of Den Besten's proposal corresponds to a range of separate functional projections, as in Rizzi's 1997 multi-layered CP. Even if it is accepted that certain types of head movement are blocked if the targeted position is not empty, it would now be necessary to explain why the presence of a complementizer in a high position of the CP-layer should block movement of the verb into a lower C-position.[36]

---

[35] There are two complementizer-like elements that can introduce V2 clauses. *Denn* 'because' can only introduce V2 clauses. *Weil* 'because' in Standard German introduces V-final clauses only, but in spoken German it may also introduce V2 clauses.

[36] In order to maintain such an analysis, it could of course be argued that the features of the complementizer must be realised in a position lower than that which the moved verb occupies. In this case, movement of the

Third, at least since the proposal presented in Chomsky 1995a that clause structure is built up through the insertion of lexical material and that structural positions exist because they contain lexical items, the view that clause structure is made up of positional slots is very suspicious. In particular, it is not clear how there could be an empty position. If all structural positions correspond to lexical items, an empty structural position would also have to correspond to a lexical item, and such an item would have to have an empty lexical entry, and it is not clear how such an empty lexical item could be learned.

These arguments show that Den Besten's 1983 approach to the complementarity of verb movement and the complementizer seems difficult to maintain and loses much of its original appeal. Thus, it seems reasonable to look for alternative accounts of this pattern. Before getting to that, I introduce the second problem, which is closely related to the complementarity issue.

The second problem is the root/embedded asymmetry of verb movement. Verb movement is obligatory in German main clauses, but the verb usually remains in final position in embedded clauses. While the problem for complementarity has traditionally been taken to be the question 'what blocks verb movement in embedded clauses?', the problem here is 'what triggers verb movement in main clauses?'.

It is an idea of traditional grammar that verb movement marks main clauses. Leaving aside questions of technical implementation, this idea is still widely accepted in some form. In generative grammar, it has been interpreted in various way as the marking of main clauses for a particular syntactic or pragmatic property (see e.g. Weerman 1989, Wechsler 1991, Koeneman 1995). Thus, there are a range of different views on which property exactly is marked by verb movement, but there is broad agreement that this verb movement in Germanic does distinguish main clauses from embedded clauses.[37] In this section, I will not attempt to further define this "main clause" property, but I will discuss in more detail the theory of clause typing that is presumably responsible for the marking of a clause for a particular property.

---

verb across the complementizer would be blocked by the head movement constraint. However, on such an analysis, the complementarity of verb movement and the complementizer would be an accidental result of their lexical feature specifications, not an effect of structural principles. It is not clear in which sense such a proposal would be in the spirit of Den Besten's original proposal.

[37] The question of how languages in which such root/embedded asymmetries of verb movement are harder to find, like Yiddish and Icelandic, fit into the picture, remains a subject of future research.

In theories that assume that clause structure contains a range of abstract functional heads and that treat verb movement as head adjunction to a functional head, the projection(s) of the functional category C is (or are) responsible for the representation of Force or clause type information. That is, a clause is marked within its C-domain as a question, a declarative, etc. In such a framework, Cheng 1991 formulates the Clause Typing Hypothesis:

(75) Every clause needs to be typed. In case of typing a *wh*-question, either a *wh*-particle in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used thereby typing a clause through Spec-head agreement. (Cheng 1991:30)

Some particular clause type may be encoded as a V-feature on C and trigger movement of the verb to check this feature. These devices are of course not available in the present framework. I assume instead that clause typing involves the marking of the root node of a clause for some particular property at both PF and LF.[38]

Since clause typing requirements of languages obviously vary, they cannot be specified by UG. Hence the assumption that clause typing involves properties of both interface representations is plausible from the perspective of learnability. Clause type must be marked syntactically because e.g. a question receives a different interpretation from a declarative. Since clause type is not a universal property, it also must be marked phonologically, otherwise the syntactic marking could not be recovered, not by the learner and not by a hearer in a conversation.

Chinese has a special particle for the clause-typing of questions, and similar clause-typing particles can be found in languages like Arabic, Navajo, and Japanese (cf. Cheng 1991). In English, questions are marked through verb movement, while in colloquial French, there is no strictly syntactic marking of question, and instead intonation is used to distinguish questions from corresponding declaratives. The following examples from Chinese, English and French illustrate these patterns:[39]

---

[38] See Chomsky and Lasnik 1977, Cheng 1991, Bhatt and Yoon 1991, Öhl 2002, Rizzi 1997, Sportiche 1998 for general discussion.

[39] The Chinese data are from www.chinese-outpost.com/grammar/prtsmdls/as. Some diacritics are missing.

(76) a. Ni shi mei guo ren.

         You are (an) American.

     b. Ni shi mei guo ren ma?

         Are you (an) American?

     c. Jin tian shi xing qi wu.

         Today is Friday.

     d. Jin tian shi xing qi wu ma?

         Is today Friday?

(77) a. John should go home.

     b. Should$_i$ John t$_i$ go home?

(78) a. Jean est parti.

         Jean is   left

         'Jean has left.'

     b. Jean est parti?

         Jean is left

         'Has Jean left?'

English yes/no questions appear to be clause-typed by verb movement alone.

(79) a. John should cook pasta.

     b. Should John cook pasta?

However, it has long been argued that this is not quite accurate. Thus, it has been argued that in yes/no questions, verb movement is triggered by an abstract Q-morpheme (Katz and Postal 1964, Baker 1970), and the most common view now is that there is an empty question operator in the specifier of the moved verb.[40] The example in (79b) could then be represented as follows:

---

[40] Roberts and Roussou 2002 argue against an operator analysis. Instead, they propose that the relevant Q-features sit on an abstract functional head.

(80) 'Should John cook pasta?'

```
            should
           /      \
       OP        should
                /      \
           Johnᵢ       cook θ_# θ_2#
                      /       \
                  tᵢ          cooked θ_1 θ_2#
                             /       \
                         pasta       cooked θ_1 θ_2
```
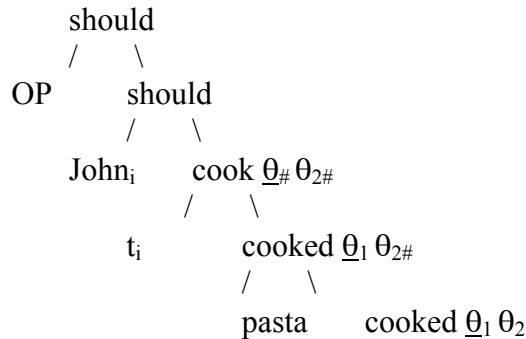
The clause-typing of yes/no questions in the English and Chinese examples in (76) and (80) illustrates an important aspect of clause-typing. While English uses verb movement (plus an empty operator), Chinese has a special particle for clause-typing yes/no questions. This in fact seems to be a general pattern, which suggests that for every positively set clause-typing parameter there are two ways to satisfy it; either through special particles or through verb movement (Cheng 1991). This suggests that we can assume that language variation with regard to clause-typing is determined by parameters of the following general form:

(81)   +/- Overtly mark the highest head of a clause for X.

If we take X to be yes/no questionhood, this parameter distinguishes between languages like English and Chinese on the one hand, and French on the other. The difference between English and Chinese need not be stated in terms of a parameter since it seems plausible to assume that wherever a language has a special clause-typing particle, it will employ it, and where a languages doesn't have a particle, it may resort to movement.

While the role of dedicated clause-typing particles is pretty clear, it is not so obvious how verb movement can get involved here. Some insight into this question can be gained from the typology of multiple wh questions. Ackema and Neeleman 1998 show that to describe the variation attested at least the following parameters are necessary:

(82) +/- All wh-words must take scope, i.e. move to the front of the clause.

(83) +/- Wh-questions must be clause-typed, i.e. the root node in a wh-question must be marked at both PF and LF.

These parameters entail the following logically possible language types:

(84)   a.   +*wh* movement, +verb-movement

        b.   +*wh* movement, no verb-movement

        c.   no *wh* movement, no verb-movement

        d.   *no *wh* movement, + verb-movement

Examples of the language types in (84a,b,c) are English, Czech, and Chinese respectively: (The examples are taken from Ackema and Neeleman 1998.)

(85) a.   Who did John ask for what?

       Who did you give which book?

    b.   Kdo koho   vid''el   Czech

       who whom   saw

       'Who saw whom?'

       Kdo co   rychle   komu   dal

       Who what   quickly to whom gave

       'Who gave what quickly to whom?'

    c.   Ni xihuan shei   Chinese

       you like who

       'Who do you like?'

       Wo xiang-zhidao Lisi   mai-le   sheme

       I   wonder   Lisi   bought what

       'I wonder what Lisi bought.'

However, the fourth type with a negative setting of the parameter that forces *wh* words to

move and a positive setting of the parameter that requires wh-questions to be clause-typed is missing. The pattern that these settings would yield is overt verb movement without wh movement. This gap can be understood if verb movement in wh questions takes place only in order to create a root node that can be marked under spec-head agreement with the wh-feature in its specifier.

Thus, the above paradigm suggests that verb movement only has a secondary role in clause-typing. Since a verb doesn't contribute any clause-typing features, the role of verb movement in clause-typing seems to be limited to creating the necessary specifier positions for clause-typing elements like *wh* phrases. This then seem to cast doubt on the idea introduced above that verb movement could mark main clauses, because there are no dedicated clause-typing elements.

In this light, consider the following data:

(86) a.   Mi oedd John yn chwerthin.   **Welsh**
      MI was   John in laugh
      'John was laughing.' (Jones and Thomas 1977)
    b.   Mae Mair yn gwybod [yr   oedd John yn chwerthin]
      is   Mair in know   that was John in laugh
      'Mair knows that John was laughing.'

(87) a.   (Ni)   chiciodd y bachgen mo'r bêl.
      NEG kicked   the boy   the ball
      'The boy didn't kick the ball.'
    b.   Gwn   na   chiciodd y bachgen mo'r bêl.
      I-know NEG   kicked   the boy   the ball
      'I know that the boy didn't kick the ball.'

In Welsh, main and embedded declarative clauses are marked by different particles.[41] Similarly, negatives in Welsh are marked differently in main and embedded clauses. Thus, Welsh main and embedded clauses have distinct clause typing particles. A similar pattern

---

[41] These particles can delete, but even if they do they trigger a morphophonological alteration on following consonants in certain environments; see chapter 3.

appears to hold in Hidatsa. Sadock and Zwicky 1984 report that Hidatsa has five particles that mark different types of declaratives and that do not occur in dependent clauses. They further note that while there is "no simple particle expressing merely declarative sentence type, any declarative sentence must have one of these five particles" (1985:166). A similar situation seems to hold in Korean (Bhatt and Yoon 1991). In contrast to Welsh and Hidatsa, English main clauses are not marked in special way, and English embedded clauses are only optionally marked through a complementizer:

(88)   a.   John was laughing.

         b.   We know (that) John was laughing.

The contrast between these languages suggests the following clause-typing parameter:

(89)   +/- Overtly mark the highest head of a main clause as syntactically independent.

A comparison of the paradigms for questions and for root clauses, in addition to Cheng's generalisation that clause-typing may involve special particles or syntactic movement, suggests that there should be a second type of language with a positive setting of (89). This language should lack clause-typing particles for main clauses, and instead use syntactic movement to mark main clauses. A language of the right type is obviously German. The root/embedded asymmetry of verb movement is illustrated in the following examples repeated from (73) above:

(73) a.   Der Affe   hat$_i$ die Kokosnuss geklaut t$_i$
           the monkey has   the coconut   stole
           'The monkey stole the coconut.'
     b.   Ich weiss [dass der Affe die Kokosnuss geklaut hat]
           I know   that the monkey the coconut stole has
           'I know that the monkey stole the coconut.'

The highest head in the embedded clause in (73b) is the complementizer *dass*, which can

be taken to be a clause-typing particle that marks subordination. The highest head in the main clause in (73a) is the moved verb *hat*. Thus, German has just the properties that a language with a positive setting of (89) should have if it lacks the appropriate clause-typing particles.

However, there is an important difference to the example of clause-typing by verb movement that is employed in wh questions in English. Crucially, this was shown to be only possible where the verb ends up in a specifier-head relation with a wh phrase. In contrast, there is no such functional element in German main clause declaratives. In the absence of an element that could provide a particular clause-typing feature, it is then not obvious in what way verb movement could contribute to the clause-typing of declarative main clauses. The answer to this question has to do with the prediction in (9), repeated here:

(9) A clause headed by a moved verb cannot be selected.

In the first two sections of this chapter, I argued that this prediction affects the distribution of complementizers within selected clauses. However, this prediction also implicates the distribution of clauses themselves. Thus, it implies that if a clause is headed by a moved verb, it must be an independent, main clause (since it cannot be selected or select). In a language that doesn't have main clause clause-typing particles of the Welsh type, a positive setting of (89) could be satisfied by verb movement in main clauses. Such movement would make a clause unselectable and by implication this would mark a clause as a main clause. This is what happens in German. By assumption, the language has a positive setting of the clause-typing parameter in (89). Thus, main clauses must be overtly marked in some way. Since German doesn't have appropriate clause-typing particles, this parameter setting triggers verb movement in main clauses. This then explains the root/embedded asymmetry of verb movement in German.

7 Conclusion

This chapter dealt with the effects of the restriction on selection to elements in chain-tail positions and its interaction with the view of phrase structure and verb movement developed in chapter 1. Section 2 showed how this model can provide a natural account for

Vikner's generalisation, i.e. the observation that only languages that lack obligatory verb movement may have complementizer optionality in complement clauses. Section 3 extended this analysis to topicalisation and adjunction to the highest position in an embedded clause. Section 4 showed how complementizer optionality in relative clauses follows the same pattern as complementizer optionality and that this pattern can be explained along the same lines. Section 5 dealt with apparently embedded verb-second clauses in German. It was demonstrated that, as the theory predicted, these clauses should not be analysed on a par with clauses introduced by a complementizer. Finally, section 6 expanded on the correlation between verb movement and selection, and addressed the complementarity of verb movement and complementizers in German from a perspective of clause typing.

CHAPTER THREE

Verb movement and derived heads

1 Introduction

In chapter 1, it was argued that verb movement always involves the pronunciation of the verb in the position of a head α, and that two verb movement structures can be distinguished. I used the term s-movement for structures in which α is a copy of the verb, and p-movement for structures in which α is a distinct category, either a separate lexical item or a distinct derived head. This chapter discusses in more detail some implications of the difference between these configurations. It proposes a new analysis of the obligatoriness of an initial specifier in front of a moved verb that is characteristic of the verb-second (V2) pattern and deals with the contrast between verb-initial (V1) and V2 languages with special attention to the trigger of verb movement in Breton and sentential particles in the (Brythonic) Celtic languages. In this section, I review some of the assumptions about clause structure made in chapter 1 and discuss why the above issues are particularly relevant with this background.

Most work inspired by Chomsky 1995c assumes a fundamental asymmetry between the operations of move and merge. Given merger of α and β, either of the two can in principle be the head of the constituent containing the two. In contrast, given movement that adjoins α, originally contained in β, to β, only (the head of) β can be the head. As argued in chapter 1, this asymmetry does not follow from any principle of phrase structure, and must hence be stipulated in some form elsewhere.[1] Given this stipulated asymmetry and the assumption that a head can only occur in a head position, the existence of head movement implies that certain head positions exist that can host a moved element.[2] If there were no head position to which a head X could move, movement of X would only be possible if the movement creates a new position. However, this would amount to saying that the moved head is the head of a structure, which, by assumption, is impossible.

Historically, similar considerations (possibly going to back Emonds' 1976 insight that

---

[1] In Chomsky's work this is done with the formulation of triggers for movement; for Brody 1995, 1998 this is part of the Generalised Projection Principle.
[2] Since phrases do not adjoin to heads, the formulation of phrasal movement doesn't require such an assumption. However, in theories that limit adjunction to one per projection (cf. Kayne 1994), phrasal movement may also require the assumption of extra heads.

116

transformations ought to be structure-preserving) gave rise to the introduction, and later the proliferation, of abstract functional categories whose main purpose is to act as landing sites for movement. At the same time, functional categories came to be used in the explanation of semantic as well as morphological properties of languages; and evidence from the three domains (syntax, semantics, and morphology) has been used to motivate the assumption of particular categories.[3] One result of this development is the lack of a principled theory of functional categories, which leaves open many questions about clause structure, e.g. about the (non-)universality of functional categories or the (in-)flexibility of clause structure within a single language.

Some of these questions can receive an answer if the stipulation is given up that a moved element cannot be the head of the constituent that contains it. On this view, verb movement does not have to be movement to some existing position, but verb movement can also involve the creation of a new position. If verb movement can give rise to new positions, it is no longer the case that a special type of category is necessary whose (main) function it is to make verb movement possible, and consequently this theoretical tool can be given up.

If verb movement is further assumed to vary across languages, it follows that clause structure is not universal, and that the clause structure of a particular languages must be learnable. This in turn means that a syntactic category can only be present in a language where there is evidence for this position.[4] Consequently, there will be one general constraint on potential syntactic positions. They must be learnable.

A consequence of the assumption that there is no fixed skeleton of abstract functional heads is that a particular empirical phenomenon cannot be related to a particular abstract functional category. On the standard view, beginning with den Besten 1983, V2 has been related to the category C(OMP). This has two consequences. First, if V2 is tied to the category C, the occurrence of V2 patterns in different languages would provide evidence that the clause structure of all of these languages contains the category C. To the extent that the existence of a particular category in a particular language is accepted as evidence for this category as part of a universal clause structure, such data would strengthen the

---

[3] See Hoekstra 1995, Webelhuth 1995, Chametzky 2000: chapter 2 for overviews of the development and chapter 1 for more discussion.
[4] See Iatridou 1990, Thrainsson 1996, Koeneman and Neeleman 2001, and chapter 1 for discussion.

evidence for the existence of this category, and hence also support the claim that there is a universal category C. This contrasts with the view that there are no such abstract functional categories, and that consequently there is no dedicated V2 category. If there is no functional category that is dedicated to the V2 pattern, the existence of V2 patterns in different languages cannot contribute to the discussion about the universality of clause structure (understood as a collection of functional projections), one way or another.

The assumption that verb movement can involve the creation of a new position also plays an important role for the analysis of the specifier requirement of V2. In particular, it will be argued that universally whenever verb movement involves a copy of the verb (s-movement), an additional specifier must be created, and that therefore s-movement always coincides with specifier-initial patterns. In contrast, where the verb is pronounced in the position of a distinct head (p-movement), no obligatory specifier is predicted. Indeed it will turn out that the principled account of the V2 pattern leaves only one possible analysis of certain V1 orders, and that this analysis derives the typological generalisation that V1 languages typically have preverbal particles.

The next section introduces the major puzzle of the V2 pattern, the presence of an obligatory specifier that can be of any category in front of a moved verb, and points out why this pattern has been so hard to tackle. In section 3, I introduce a new proposal that can deal with this specifier requirement in a straightforward way, and that makes the strong prediction that whenever verb movement is s-movement an additional specifier is necessary. This section also discusses briefly some patterns related to this prediction, namely the lack of V2 patterns despite verb movement in languages like French, and the existence of V1 and V3 orders in German. There will also be a brief discussion of the acquisition of V2 in Swedish. Given the principled relation of s-movement and an obligatory specifier, the big question is how the existence of languages with canonical V1 order can be explained. In section 4, I propose that canonical V1 orders are only possible, if a language has preverbal particles in whose position the verb can be pronounced, i.e. where p-movement is possible. In section 5, I take a closer look at the particle systems of, and verb movement in, Breton. In section 6, I discuss the implications of these proposals for the diachronic analysis of V2 in Welsh. In section 7, the trigger of verb movement in Welsh and Breton will be investigated, and the relevance of the findings for the theory of

clause structure discussed. Section 8 concludes this chapter.

2 Some properties of V2 structures

The verb-second phenomenon is a word order pattern found most famously in the Germanic languages. It involves a verb moved in front of the subject that is obligatorily preceded by a single constituent that can be of any category (hence also the subject). In (1), this pattern is illustrated with a fronted subject, a prepositional phrase, adverbs, a VP, and a subordinate clause:[5]

(1) a. [$_{DP}$ Die Kuh]   hat auf der Wiese   gestern gierig   Klee   gefressen
         the cow   has on the meadows yesterday greedily   clover   eaten
         'The cow greedily ate clover on the meadows yesterday.'

   b. [$_{PP}$ Auf der Wiese] hat die Kuh gierig Klee gefressen

   c. [$_{AdvP}$ Gierig]   hat die Kuh auf der Wiese Klee gefressen

   d. [$_{AdvP}$ Gestern]   hat die Kuh auf der Wiese Klee gefressen

   e. [$_{VP}$ Klee gefressen]   hat die Kuh

   f. [$_{CP}$ Dass die Kuh Klee   gefressen hat]   glaube ich nicht
         that the cow clover   eaten   has   believe I   not
         'I don't believe that the cow ate the clover.'

Moreover, in German verb movement, and hence the V2 pattern, is typically in complementary distribution with the presence of a complementizer:[6]

(2) a. Jeder   weiss   [$_{CP}$ [$_C$ dass Michael   oft   Milch   trinkt]]
         Everybody   knows   that Michael   often   milk   drinks

   b. [$_{CP}$ Michael$_k$ [$_C$ trinkt$_i$] [t$_k$ oft Milch t$_i$]]
         Michael   drinks   often   milk

---

[5] The most systematic overview of Germanic V2 is Weerman 1989; see also Haider and Prinzhorn 1986, Vikner 1995, Zwart 1997. Other languages with dominant V2 patterns are Kashmiri (Bhatt 1999) and Breton (see section 5 for references and discussion).
[6] This complementarity is discussed in chapter 2, section 6.

In the standard analysis of V2 (den Besten 1983 and much later work), the verb in matrix clauses moves to the position that the complementizer occupies in embedded clauses, C in (2). The initial phrase occupies the specifier of this position, Spec-CP. On the assumption that only one element can occur in the C position, the complementarity of verb movement and complementizer in a language like German then follows from the competition of the two heads for the same structural position.[7]

If V2 clauses had the structure in (2), it would not be difficult to formulate a trigger for verb movement in terms of properties of C. For example, within checking theory (Chomsky 1995b), it could be argued that the verb is attracted by a strong V-feature on C.[8]

However, the trigger of the specifier requirement of V2 is much harder to formulate within checking theory. As the examples in (1) above show, the initial phrase in a V2 clause can be of virtually any category. Unlike in the case of verb movement, it is therefore not possible to say that the C-head has a categorial feature X that needs to be checked and that triggers movement of a phrase to its specifier. Crucially, there is no single categorial feature that DPs, PPs, AdvPs, VPs, and CPs share. Instead, it is often assumed that the specifier requirement is triggered by a pragmatic feature, e.g. a [topic] feature, on C that needs to be checked by the initial phrase (e.g. Holmberg and Platzack 1995, Zwart 1997).

(3) a. Underlying structure:

[$_{CP}$ [$_{C}$ [+TOPIC, +V]] [Michael oft [Milch, +TOPIC] trinkt]]

b. Verb movement triggered by a +V-feature on C:

[$_{CP}$ [$_{C}$ trinkt$_i$ [+TOPIC, +V̶]] [Michael oft [Milch, +TOPIC] t$_i$]]

c. XP movement triggered by a +TOPIC-feature on C

[$_{CP}$ [Milch, +̶T̶O̶P̶I̶C̶]$_k$ [$_{C}$ trinkt$_i$ [+̶T̶O̶P̶I̶C̶, +V̶]] [Michael oft t$_k$ t$_i$]]

Such an analysis seems inadequate however, because it wrongly predicts that the initial phrase in a V2 clause is always interpreted in the same way. In (4a), the initial constituent

---

[7] This assumption is of course incompatible with the view that functional heads correspond to actual morphemes and that verb movement may be e.g. adjunction of the verb to a Tense morpheme. Thus, it is not clear how Den Besten's account of the complementarity of verb movement and complementizer can be maintained. For relevant discussion, see Zwart 1997, among others.

[8] This proposal was rejected above; see chapter 1 for general problems with checking theory and chapter 2 for an alternative trigger of verb movement. Further complications arise if not all V2 clauses are taken to be CPs; see below.

is a focused object, as indicated by the focus particle *nur*. In (4b), the initial object is not the focus but rather a topic. Here the postverbal subject is focused. In (4c), the initial constituent *alle Kühe* contains a quantifier, which means that it cannot be a topic, and it need not be focused ((4c) is a possible answer to the question *What happened?*).[9]

(4) a.   Nur das gr*ü*ne Gras   hat die Kuh gefressen.
           only the green gras   has the cow eaten.
           'The cow ate only the green gras.'
     b.   Die Disteln hat   nicht einmal die verr*ü*ckte Kuh gefressen.
           the thistles has   not even   the mad   cow   eaten
           'Even the mad cow didn't eat the thorny thistles.'
     c.   Alle   K*ü*he   haben   Gras   gefressen.
           All   cows   have   grass   eaten

Although it would still not be clear what feature could be responsible for the movement of a non-topic, non-focused constituent to the initial position, it would of course be possible to assume a different feature on C depending on the interpretation of the initial phrase. However, this would mean that the V2 pattern is an accidental result of the occurrence of different features on C and an additional condition that stipulates that C must contain exactly one such feature, not two, three or zero. More recent proposals according to which functional categories may or may not carry an EPP-feature, which if present forces movement of some phrase to the specifier of the category, (e.g. Chomsky 1999, Holmberg 2000), run into the same problem. An EPP-feature would have to be stipulated just in the cases where movement to the specifier position takes place. This does not appear to be very explanatory.

It is also not clear how such an analysis could be extended to embedded clauses. If the feature that triggers the specifier requirement is on C, it is expected that a focused or topicalised constituent should precede the complementizer. However, this is not the case. Topics and foci follow the complementizer in embedded clauses. The following examples show that a focused or topicalised element cannot occur before the complementizer in

---

[9] Examples like these were pointed out to me by Ad Neeleman.

embedded clauses:

(5) a. Sie sahen, dass die dornigen Disteln nicht einmal die verr*ü*ckte Kuh frass.

   they saw that   the thorny thistles   not even the mad cow ate

   'They saw that now even the mad cow ate the thorny thistles.'

    b. * Sie sahen, die dornigen Disteln dass nicht einmal die verr*ü*ckte Kuh frass.

(6) a. Sie sahen, dass nicht einmal die verr*ü*ckte Kuh die dornigen Disteln frass.

   they saw that   not even the mad cow the thorny thistles   ate

   'They saw that now even the mad cow ate the thorny thistles.'

    b. * Sie sahen, nicht einmal die verr*ü*ckte Kuh dass die dornigen Disteln frass.

    c. …dass sogar das grüne Gras die verrückte Kuh nicht gegessen hat.

        that even   the green gras the mad cow not eaten has

This issue seems to disappear if a complex CP-layer is assumed (e.g. Rizzi 1997, Roberts 2000). In that case, the complementizer could occupy a high position in the CP-layer, while topics and foci occupy lower positions. However, this proposal would have to be amended with further assumptions to account for the complementarity of verb movement and overt complementizers. Since, by assumption, the complementizer occupies a high position within CP, it is not clear what blocks movement of the verb into a lower position with CP. It is of course possible to propose that the overt complementizer and verb movement play a similar role in the checking of features in the CP-layer, and that this is the reason for their complementary distribution. However, this would amount to giving up Den Besten's structural approach to the complementarity, and be very similar to the functional explanation proposed in chapter 2, section 6. Moreover, on such an approach to clause structure, it is no longer clear why typically only a single constituent can occur in front of the moved verb in V2 languages, because in a complex CP-layer, several specifier positions would be available.[10]

It has been suggested that subject-initial V2 clauses are IPs, not CPs like other V2 clauses (Travis 1984, Zwart 1997). This appears to account for the tendency that initial

---

[10] Since the different categories within the CP-layer are associated with different features, the recent version of Relativised Minimality in Rizzi 2001, which is formulated in terms of features, does not obviously block movement of a topic across a focus.

objects typically receive a special interpretation while initial subjects typically do not. However, without additional assumptions such an analysis still cannot account for the data above. It also does not account for the fact that initial adverbs, like initial subjects but unlike initial objects, do not typically receive a special interpretation (Haider 1984). Thus, while in object-initial V2 clauses the object is typically (although not necessarily) focused, neither the initial subject nor the initial adverb in the following examples (repeated from (1)) need be focused:

(7) a. Die Kuh hat auf der Wiese   gierig   Klee   gefressen
        the cow has on the meadows   greedily clover   eaten
        'The cow greedily ate clover on the meadows.'

    b. Gestern hat   die Kuh auf der Wiese   Klee gefressen
        yesterday has   the cow on   the meadows   clover eaten

This approach also encounters a similar problem to the split-CP approach just discussed. If subject-initial clauses are IPs, it is not clear what blocks verb movement to I in embedded clauses introduced by a complementizer. Thus, the order complementizer-subject-verb-object is wrongly predicted to be grammatical in Dutch and German.

A further problem for Zwart's 1997 proposal is that his account of certain agreement patterns forces him to assume that all subject-initial clauses are IPs, even those with a focused subject. This means that focusing of subjects involves different structures from focusing of other constituents, which of course seems problematic for a uniform trigger of verb movement. (I ignore details irrelevant for this argument, such as the exact structure of IP and VP on Zwarts' account and whether there are any intermediate traces.)

(8)   Subject-initial V2 with and without focus
        a. $[_{\text{IP}}$ Michael$_k$ $[_\text{I}$ trinkt$_i$] $[t_k$ oft Milch t$_i$]]
            Michael drinks   often   milk
        b. $[_{\text{IP}}$ [MICHAEL, +FOCUS]$_k$ $[_\text{I}$ trinkt$_i$] $[t_k$ oft Milch t$_i$]]
            Michael drinks   often   milk

(9)   Focussed initial object

[$_{CP}$ [Milch, +FOCUS]$_k$ [$_C$ trinkt$_i$] [$_{IP}$ Michael oft t$_k$ t$_i$]]

If all subject-initial and non-subject-initial V2 clauses have essentially the same structure, the fact that subjects and adverbs need not be focused can be the result of their being the closest constituents available for movement to the initial specifier. In contrast, an object always has to move across the subject to get to that position.[11] (Further problems of asymmetric approaches to V2 are discussed by Schwartz and Vikner 1996, Williams 1997.)

While a feature checking approach to the specifier requirement in V2 is empirically inadequate, it is also conceptually problematic. As the examples in (1) show, the relevant feature cannot be a categorial feature. The examples in (4) show that the initial phrase does not receive a uniform interpretation, which means that there is no independent motivation for a pragmatic feature on C. Instead, the feature on C is assumed whenever a structure has V2 order. This suggests that an analysis of the specifier requirement in V2 in terms of feature checking is circular (see also Haider 1993:69f).[12]

3 The V2 proposal

3.1 S-movement and obligatory specifiers

In this section, I would like to suggest that in the framework proposed in chapter 1 nothing needs to be added to derive V2 patterns because in s-movement structures this pattern turns out to be the default.[13]

As argued in chapter 1, the assumption that phrase structure is a representation of structural hierarchies, but not of linear order, makes it possible to express phrase structure

---

[11] A alternative account for this pattern is suggested by Neeleman and Weerman 1999:178. They argue that an object can only appear in initial position in a V2 clause if it moves there. In contrast, they argue, both subjects and adverbs can in principle be base-generated in initial position. The focus pattern can then be related to whether or not the initial phrase is moved or not. It is not clear how this proposal could be made compatible with Zwart's approach, or the one developed below.

[12] As discussed by Fukui and Takano 1998:32f, a similar problem arises for LCA-based analyses of V-final languages (cf. Kayne 1994). There, a functional head needs to be assumed that forces the complement of V to move, regardless of its category. To the extent that the proposed analysis of the specifier requirement of V2 in the next section is successful, it is possible that the LCA problem could be addressed in a similar way. (Although of course, the proposal is independent of the validity of the LCA.)

[13] Chapter 1 contains of fuller statement of the theoretical assumptions underlying the analysis developed here.

as a set of objects, called TREELETS, that contain a category and specify the dominance relation of this categories with other categories. The dominance information for the category of each treelet is given as an unordered set, the DOMINANCE SET, that contains all the dominated categories. Thus, the following represents a treelet of category L whose dominance set is {X,Y}.

(10)   <L, {X,Y}>

The Inclusiveness Condition of Chomsky 1995a and Neeleman and van de Koot 2002 requires that the properties of a node X in a tree diagram must be recoverable. Thus, the category of a node X can correspond to (the copy of) a lexical item inserted in the structure, or it must be recoverable from the nodes that X immediately dominates. I assume that the same restriction holds of the categorial information contained in a treelet. That is categorial information must be recoverable either from dominated categories or from the lexicon.

Two treelets are identical if they belong to the same category and contain the same dominance set. Conversely, two treelets are distinct if they belong to different categories, contain different dominance sets, or both. The following examples illustrate the range of possibilities:

(11) $\alpha = $ <L, {X,Y}>
        $\beta = $ <K, {X,Y}>
        $\gamma = $ <L, {X}>
        $\delta = $ <L, {X,Y,Z}>

α is a treelet of category L with the dominance set {X,Y}. The treelet β is distinct from α, because it is of a different category. The treelet γ is distinct from α because its dominance set does not contain all of the categories that are contained in the dominance set of α, namely Y is missing. The treelet δ is distinct from α because its dominance set contains a category, namely Z, that is not contained in the dominance set of α.

In contrast to β, γ, and δ, the treelet ϵ below is identical to α.

(12) $\varepsilon = \langle L, \{X,Y,Y\}\rangle$

This is so because ε is of the same category as α and because its dominance set has the same members as the dominance set of α. This last point illustrates a principle of set theory according to which two unordered sets are identical if they contain the same members, regardless of how often they contain a particular member.[14] Since this principle will play an important role in the following discussion, it is convenient to give it a name here:

(13) PRINCIPLE OF SET IDENTITY: Two unordered sets are identical if they contain the same members, regardless of how often those members are represented in each set, i.e. $\{e\} = \{e,e\}$.

Dominance, which holds between categories, is the basic relation of phrase structure. I follow McCawley 1982, Higginbotham 1983, Partee et al 1990, Speas 1990:18-24, Chametzky 2000 and others, and define dominance as a reflexive, transitive and antisymmetric relation. So, dominance can be described by (at least) the following axioms (cf. Higginbotham 1983:(2), Speas 1990:20f):

(14) a. X dominates X (dominance is reflexive)
    b. if X dominates Y and Y dominates Z, then X dominates Z (dominance is transitive)
    c. if X dominates Y and Y dominates X, then X = Y (dominance is antisymmetric)
    d. if X dominates Z and Y dominates Z, then X dominates Y or Y dominates X (or both, if X=Y=Z)

Given these assumptions about treelets and dominance, the conventional tree structure in (15a) only has the status of a convenient representation of the set in (15b):

(15) a.   A
        / \
        B C

    b. $\{\langle A, \{A,B,C\}\rangle, \langle B, \{B\}\rangle, \langle C, \{C\}\rangle\}$

---

[14] See e.g. Wall 1972, Partee et al 1990 on set theory.

The first pair, <A, {A,B,C}>, in the set in (15b) states that there is a node of category A and that A dominates the categories A, B and C. The second pair states that there is a node of category B and that B dominates the category B. The third pair states that there is a node of category C and that C dominates the category C.

As discussed in chapter 1, this model of phrase structure cannot express certain differences that can be expressed in terms of conventional phrase structure trees. In particular, the trees in (15a) and that in (16) are distinct, but it is not possible to express the difference between the two here.

(16)   A
       |
      A
     / \
    B C
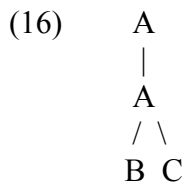
At first sight, (17a) looks like a phrase structure set that could be uniquely represented by the tree in (16). However, given the principle of set identity in (13), it is clear that this set can be expressed in a simpler way. Thus, the dominance set of the first treelet in (17a) contains the category A twice. This means that this set is identical to the set in (17b) where the dominance set of the first treelet contains A only once. However, the first two treelets in (17b) are identical; both are <A, {A,B,C}>. This means that this set is identical to the set in (17c), which contains <A, {A,B,C}> only once. But this set is identical to (15b), the set that is represented by the tree in (15a).

(17)   a.   {<A, {A,A,B,C}>, <A, {A,B,C}>, <B, {B,C}>, <C, {C}>}   cf. (16)

        b.   {<A, {A,B,C}>, <A, {A,B,C}>, <B, {B,C}>, <C, {C}>}   = (16a)

        c.   {<A, {A,B,C}>, <B, {B,C}>, <C, {C}>}   = (16b) = (15b)

So, it turns out that, on the view that dominance is reflexive, the distinct tree structures in (15a) and (16) represent the same phrase structure set, and are therefore identical in present model.

This view of phrase structure restricts the availability of verb movement in an interesting way. As discussed earlier, there is nothing wrong with structures headed by a moved element or, more generally, structures whose head is not in the lowest position in its

chain. This means that in principle it is possible that a structure can be headed by a copy of the verb. Conventional X-bar tree showing the relevant structures are given here:

(18)   a.   VP   b.   VP

```
(18)   a.        VP                          b.        VP
                /  \                                  /  \
               Y   V'                                V    VP
                  /  \                                    /  \
                 V   Z                                   Y   V'
                                                            /  \
                                                           V   Z
```

In a theory without categorial projection[15] but with phrase structure trees, the corresponding structures would be the following:

```
(19)   a.    V                          b.        V
            /  \                                   |
           Y   Z                                   V
                                                  /  \
                                                 Y   Z
```

These trees are of course structurally identical to those in (15a) and (16). Thus, on the present view of phrase structure, these trees, which are meant to correspond to structures with and without verb movement respectively, represent identical phrase structure sets and are therefore also identical. The relevant sets are given here: (Cf. the discussion of (17) above.)

(20)   a.   {<V, {V,Z,Y}>, < Z, {Z,Y}>, <Y, {Y}>}   cf. (19a)

        b.   {<V, {V,V,Z,Y}>, <V, {V,Z,Y}>, <Z, {Z,Y}>, <Y, {Y}>}   cf. (19b)

        c.   {<V, {V,Z,Y}>, <V, {V,Z,Y}>, <Z, {Z,Y}>, <Y, {Y}>}   =(20b) =(20a)

It is then clear that verb movement that involves a copy of the moved verb (i.e. s-movement) cannot be represented by a tree like (19b) since such a tree collapses to the tree without verb movement in (19a). The situation is different if the moved verb has its own specifier, a structure that would correspond to the following tree:

(21)
```
        V
       / \
      X   V
         / \
        Y   Z
```
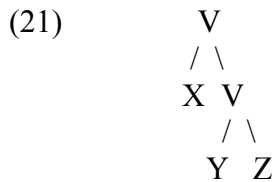
The presence of the specifier X of the moved V means that in the set that corresponds to this tree the dominance set of the moved V will contain one category that the dominance set of the lower V does not contain. Furthermore, the phrase structure set of course will contain a treelet of category X that corresponds to the specifier X. Thus, the set below that corresponds to the tree in (21) can be contrasted with the set in (20c) that is represented by the trees in (18).

(22)   {<V, {V,Z,X,Y}>, <Z, {Z}>, <V, {V,X,Y}>, <X, {X}>, <Y, {Y}>}   cf. (21)

The presence of the additional specifier means that this set is distinct from that representing a structure without verb movement, and hence that (21) is an appropriate structure to represent verb movement that involves the extension of clause structure through the addition of a copy of the verb.

Since verb movement that involves a copy of the verb, i.e. s-movement, is only possible if the moved verb has its own specifier the theory implies the following theorem:

(23)   S-movement structures must contain an additional specifier.

This theorem captures just that correlation of verb movement and obligatory specifier which has been the major problem for the analysis of V2 phenomena.[16] Since the specifier requirement stated here is triggered by the basic principles of phrase structure, this requirement will be blind to the category or interpretation of the additional specifier. Thus, in all s-movement structure where there are no independent restrictions on the content of the specifier position, a V2-like pattern is expected. In such contexts, V2 is then the default rather than the exception. This is the simplest possible explanation for the arbitrariness of

---

[15] See chapter 1: section 4 for discussion of the rejection of categorial projection.
[16] Earlier approaches to V2 in this spirit are Koeneman 1995, Bury 2000a, 2002a.

the initial constituent in V2 clauses illustrated in section 2.[17]

A somewhat simplified structure of a V2 clause is the following, where @ indicates the pronounced copy of the verb:

(24) a. $Klee_i$ $frassen_j$ Schafe t_i t_j.
         clover   ate sheep
         'Sheep ate clover.'

b.   {<frassen, {frassen, Klee, Schafe, ...}>, < Klee, {Klee}>, <V, {V, Schafe, ...}>, <Schafe, {Schafe}>, ...}

c.   frassen @
            /   \
       Klee   frassen
                  /   \
            Schafe   ...

While the theorem in (23) offers a straightforward explanation of the existence of V2 patterns, it also implicates a range of other verb movement phenomena that do not obviously conform to it. Most notably among them are languages with canonical verb-initial order which involve movement of the verb across the subject. Before dealing with V1 languages in some detail, the next section briefly addresses some other constructions.

3.2 Further verb movement constructions

3.2.1 The acquisition of V2 in Swedish. The theorem in (23) has implications for language acquisition. In particular, it implies a prediction about the acquisition of the V2 pattern that does not follow automatically in a theory that does not relate verb movement and the specifier requirement. A theory of the latter type makes no predictions about which of the two requirements involved in V2 patterns should be acquired first, the movement of the verb or the specifier requirement. On the other hand, in the present theory, the specifier requirement is triggered by verb movement, while the occurrence of an initial (pre-subject)

---

[17] In a structure where type-1 verb movement occurs and there are additional constraints on the content of the derived specifier position, it is possible that no V2 pattern occurs. Thus, in French type-1 verb movement is obligatory, but for independent reasons, presumably case theory, the subject must enter a specifier-head relation with the moved verb. Thus, rather than V2, subject-verb order surfaces.

XP does not (necessarily) trigger verb movement. Thus, the present theory does not rule out *XP – YP - verb* orders, but it predicts that the relevant verb movement structures without a specifier cannot be acquired. This means that, while the necessity of verb movement must be learned, the specifier requirement doesn't have to be; it will follow directly once verb movement is learned. In the acquisition of V2, V3 orders involving fronting of a constituent should then be possible during earlier stages, while V1 structures, in which the verb precedes the subject, should never occur (outside the contexts discussed for (adult) German below).

Swedish is a typical V2 language. A structure involving two preverbal phrases is ungrammatical:

(25) a. Idag köpte hon en ny bil.   **Swedish**
         today bought she a new car
         'She bought a new car today'. (Wechsler 1991:(1a,c,d))
     b. * Idag hon köpte en ny bil.
     c. * En ny bil hon köpte idag

However, in a study of the acquisition of V2 in Swedish, Santelmann (1999) shows that Swedish children produce structures that involve topicalisation without verb movement, which are ungrammatical in the adult language, as shown in (25b,c) above vs. (26), alongside V2 structures involving a moved verb preceded by a single phrase (27). Crucially, structures involving movement of the verb across the subject with no other element preceding the verb, are not found in Santelmann's study (* V_i [$_{VP}$ *Subj* t_i …).[18]

(26) a. Nu   han kör.   **Early Swedish**
         now he   drives
         'Now he drives.' (Santelmann 1999:(18e))
     b. Var   han bor?
         where he   lives
         'Where he lives?' (Santelmann 1999:(17a))

---

[18] There are some cases of apparent verb-initial orders that, as Santelmann (1999:286ff) notes, can be analysed on a par with verb-initial structures in adult Swedish, i.e. as involving an empty constituent in the derived specifier.

(27)   Nu   ska   jag gå.   **Early Swedish**

         now shall I   go

         'Now I'll go.' (Santelmann 1999:(14b))

Thus, the acquisition of V2 provides support for the present theory, and against theories that do not impose a restriction on the relative order in the acquisition of the two specifier requirement and the necessity of verb movement.

3.2.2 German V1 orders. Given the theorem in (23), the occurrence of verb-initial orders in a V2 language needs to be discussed. There are different types of V1 order in German.

The following examples illustrate the phenomenon known as TOPIC DROP. In these examples, one of the verb's arguments is missing and the constituent in the initial specifier is not pronounced. This constituent could either be occupied by a null pronoun or a full phrase that undergoes deletion at PF.

(28) a.   (Das   ist schön.)   Gebe   ich gerne   zu.

           that is nice   admit   I   gladly   PRT

           '(That's nice.) I'll happily admit that.' (*SZ* 23.09.2000)

     b.   (Ich habe Spaß gehabt, Geld   verdient und gedacht, das geht

           I have   fun had money   made   and thought that   continues

           einfach so   weiter.) Ging   es natürlich   nicht.

           simply   like-that on   went   it of-course   not

           '(I was having fun, made money, and thought that it would just go on

           like that.) Of course it didn't.' (*SZ* 28.09.2000)

These examples then are regular V2 clauses in which the specifier happens to be occupied by a constituent without phonological content. Since in either case the semantic content of this constituent is recoverable from the context the fact that the constituent is unpronounced is not problematic.

The examples in (29) illustrate a V1 pattern that only occurs at the beginning of narratives or jokes. In examples of this type, all arguments are overtly realized. Nevertheless, it is possible to subsume them under the topic-drop analysis. This requires

the assumption of an empty counterpart of the expletive *es* or the quasi expletive *da* that occupies the specifier, as in (29b).

(29) a. Kommt ein   Mann   in die Stadt
          came   a   man   to the city
          'A man came to the city.'
    b. Es/Da kommt ein Mann in die Stadt.

Finally, V1 order is possible in German in clauses that receive a marked, i.e. non-declarative, interpretation. This order is possible in conditionals, where V1 clauses alternate with V-final clauses introduced by *wenn* 'if', and obligatory in *yes/no* questions and imperatives.

(30) a. Trink   jeden Tag   ein Glas Milch!
          drink   every day   a glass   milk
    b. Trinkst du Milch?
          drink   you milk
    c. Verhindert   die Opposition aber   den Wahlakt, hat sie   noch längst
          prevents   the opposition though the election   has it   yet long
          nicht   gewonnen.
          not   won
          'Even if the opposition prevents the election, it does not mean that it
          will have succeeded.' (*FR* 05.10.2000)

An obvious analysis that goes back to the abstract Q- and I-morphemes of Katz and Postal 1964 and Baker 1970 is to assume that an empty operator occupies the specifier position of the moved verb. It is worth pointing out that the acquisition of such semantically contentful, phonologically null operators does not raise a particular problem. The model of phrase structure assumed here implies that universally in a verb movement structure of this type, the highest specifier must be filled. Consequently, a V1 structure involving s-movement must be analysed as containing an empty element in the derived specifier. Since these V1 structures all crucially involve a particular interpretation, the learner will be able

to recover the content of the empty element.[19]

3.2.3 German V3 orders. The theory of phrase structure only rules out verb movement structures with a verb in initial position. There is no restriction on the number of constituents that can occur in preverbal position. Indeed, this flexibility is necessary as witnessed by the availability of V3 orders in numerous languages:

(31) a.   [Marc] Ribot you ought to know through all his scary guitar
            work for Tom Waits. (*Time Out* 1554, 2000)

   b.   Aina   sinä olet myöhässä.   **Finnish**
            always you are late
            'You are always late.' (Holmberg et al 1993)

   c.   Aux enfants   je pardonne tout.   **French**
            'To   children I   forgive   everything.' (Cadiot 1992: (1a))

However, the existence of examples like (31) raises the question of why such V3 orders are not usually found in V2 languages.

Indeed, the fronting of more than a single constituent is ruled out by independent conditions on movement. In particular, A-bar movement across an A-bar specifier would violate Relativized Minimality (Rizzi 1990, 2001). This implies that the fronted verb can only be preceded by a single moved constituent. Further preverbal constituents are allowed if they are base-generated in their surface position.[20] This correctly predicts the contrast between the ungrammatical V3 clauses that involve multiple fronting in (32), and the grammatical ones in (33), where a coreferential pronoun indicates that the initial constituent cannot have moved out of the clause but must be base-genereated.[21]

---

[19] Roberts and Roussou 2002 argue against a null operator account of some V1 orders, and propose instead that the same content sits on an empty functional head.

[20] An exception that is problematic for any theory are marginal examples like (i), which may be acceptable in a particular context (the acceptability of (i) was confirmed to me by Corinne Iten, Tom Leu and Peter Oehl):
(i)   Dem   FRANK den   SPINAT sollst du geben, nicht dem   Thomas die Tomaten!
       the-DAT Frank   the-ACC spinach   should you give not   the- DAT Thomas the tomatoes
       'You should give the SPINACH to FRANK, not the tomatoes to Thomas.'

[21] The ungrammaticality of (i) is expected if the adverb is generated lower than the adjoined position (see Bury 2001):
(i)   *Gestern Frank hat   Kuchen   gegessen.
        yesterday Frank   has   cake   eaten

(32) a. *Frank$_i$ gestern hat t$_i$ den Kuchen gegessen

   Frank yesterday has the cake eaten

   b. *Gegessen$_i$ Frank hat gestern den Kuchen t$_i$

(33) a. (Er sagte) [der Hans]$_i$ der$_i$ habe   schon   wieder Hunger.

   he   said   the Hans   the have-SUBJ already again hunger

   '(He said that) Hans is hungry again already.' (Altmann 1981:149)

   b. (Christian meinte) [in der Stadt]$_i$ da$_i$ werde er es nicht mehr lange aushalten.

   Christian meant   in the city   there would he it no   more long bear

   '(Christian said that) in the city, he couldn't stand it much longer there.'

   c. [Eine Brigg]$_i$ was ist das$_i$?

    a   *Brigg*   what is that

   'A *Brigg*, what's that?' (Altmann 1981:148)

(34)   [Dass man über unterschiedliche Lösungswege diskutieren kann]$_i$

   that   one about different   solution-ways discuss   can

   Deutschlands Abiturienten   haben davon$_i$ nie gehört.

   Germany's highschool graduates have   of-it   never heard

   'Germany's highschool graduates never knew it was possible to discuss

   different ways to get to a solution.' (*Die Zeit* 48/2000)

In (33a,b), the first constituent is coreferential with a pronominal element in the derived specifier. In (33c,d) a coreferential pronoun sits in an argument position.[22]

   Similar restrictions on multiple topicalization appear to hold in non-V2 languages. In Italian or English, for example, all but one of the left-peripheral constituents must be base-generated in the surface position, usually licensed by a coreferential pronominal element (see e.g. the examples in Rizzi 1997:290f).

4 Verb-initial languages

The discussion in section 3.1 showed that s-movement always requires an additional derived specifier. This section investigates the conditions that hold of p-movement

structures. The following telescope tree illustrates a possible verb movement structure:

(35)   α
            / \
          (β) V

Verb movement occurs if V is pronounced in the position of α (Brody 2000). If α=V, i.e. if the verb is pronounced in the position of its copy, I talk about s-movement. In all other cases, i.e. if α≠V, I talk about p-movement.[23] The major conditions on verb movement are that both α and V must be within the same Extended Projection (cf. Grimshaw 1991, Frank 2002) and that α must immediately dominate V.

Configurations of s- and of p-movement differ in significant ways. First, only s-movement structures require an additional specifier. If an additional head position is required to make verb movement possible and no selectional requirements of the verb are satisfied in the new structure, the theorem in (23) requires that an additional specifier be created. Thus, for s-movement, β in (35) is obligatory. While p-movement is in principle independent of β, α can only be a *distinct* head derived from V, if there is a β that satisfies some selectional requirement of V. Otherwise α would be identical to V, and the structure would be a s-movement structure by definition. Thus, wherever verb movement targets the position of a derived head, this position must have a specifier, but where verb movement targets a distinct category, no specifier is necessary. Of course, even if α corresponds to a lexical category, it may have a specifier.

Second, if the structure is a p-movement structure, α will remain the head of the structure. Thus, pronunciation of V in the position of α (a PF effect) does not mean that the moved verb will be the highest head in the structure. Since only the phonological features of V are displaced in a p-movement structure, no head chain need or can be formed. The prediction is then that there should be no root-embedded asymmetries with regard to p-movement since the head α does not form a chain with the verb, and hence can be selected. If such movement is possible in a main clause, it should also be possible in an embedded

---

[22] See Rudin 1988 for an analysis of multiple *wh*-movement that is compatible with Relativized Minimality.
[23] For p-movement, there is one further distinction: α can be a derived head that is distinct from V (e.g. where β satisfies a selectional requirement of V) or it can correspond to a lexical item. I focus on the second case here, and comment on the first case where necessary.

clause.[24]

Third, in a structure where a head position in which the verb can be pronounced is available, the need for verb movement cannot trigger the creation of a copy of the verb to extend clause structure. In other words, s-movement is only possible, where p-movement is not, and vice versa.

Finally, as already mentioned, since clause structure is assumed to vary across languages, and hence the distribution of empty material is restricted to contexts where its content is recoverable (and learnable), the head α in the above structure is predicted to be phonetically realised where it corresponds to a lexical category. This means that p-movement should only be possible where the structure contains a visible head that dominates the verb. In contrast, s-movement is not possible if the verb is dominated by an extra head. Thus, in general, s-movement is expected to be only possible if there is no head preceding the verb, while p-movement is only possible if the verb is preceded by such a head.

On the assumption, introduced in chapter 1, section 7, that at PF two sisters must be linearly separated by their mother (which may be empty), this leaves just one possible derivation for true verb-initial (V1) structures, i.e. for V1 structures that do not contain a specifier filled with an empty element, and in which the subject precedes the object. Given this linearisation constraint, a structure like (36) cannot be linearised as VSO.[25]

(36)　　　　V2　　　　Possible orders of (36):　Subj –V2-[Obj - V1-δ]
　　　　　/　\　　　　　　　　　　　　　　Subj –V2-[δ- V1- Obj]
　　　　Subj　V1　　　　　　　　　　　　[Obj - V1-δ]-V2- Subj
　　　　　/　\　　　　　　　　　　　　　[δ- V1- Obj]-V2- Subj
　　　　Obj　　δ　　　　　　impossible:　*V2- Subj -[Obj - V1-δ]

If s-movement applies, the situation remains the same. S-movement requires an additional specifier. Consequently, the resulting structure (37) can only be linearised with either the specifier X of the higher V or the constituent headed by the lower V preceding its mother:

---

[24] The reverse does not necessarily hold, as will be discussed below.
[25] VOS is a possible linearisation if V is pronounced in δ or in V1.

(37) V   Possible orders of (37)   X-V-[$_V$ α-V-β]
      / \   X-V-[$_V$ β-V-α]:
     X   V   [$_V$ α-V-β]-V-X
        / \   [$_V$ β-V-α]-V-X
       α   β   impossible: *V-X-[$_V$ α-V-β]

If non-specifier initial orders can neither be base-generated nor derived through s-movement, p-movement is the only other option available. If V is dominated by an additional head F, as in (38), no additional specifier will be required and therefore V can be pronounced in the position of F, preceding both its daughters α and β, but following F.

(38)   F   Possible orders of (38) with movement of V to F:
           | F-V-α-β
           V   F-V-β-α
          / \
         α   β

In a theory that assumes a universal clause structure, F could of course be consistently empty. Consequently, it would be difficult to distinguish structures in which s-movement takes place and where an empty element occupies the derived specifier, and structures in which p-movement to an empty head position takes place. However, on the flexible view of clause structure adopted here, a category must be learnable and is therefore expected to have phonetic content. This means that in V1 orders derived by verb movement, the verb should be preceded by a preverbal particle. In other words, the theory implies the following typological prediction: [26]

(39) VSO languages have preverbal particles.[27]

This prediction seems to go the right way. A pattern like that predicted here has indeed been observed e.g. in the introduction of Carnie and Guilfoyle 2000:10, and in Carstairs-McCarthy 1999:158. Furthermore, Andrew Carnie (p.c.) informs me that, beyond the

---

[26] For (39) to make sense, it is of course necessary to disregard preverbal particles when counting the position of the verb. If the occurrence of preverbal particles meant that the verb is not in the first position, this statement would imply that there can be no verb-initial languages.

[27] This prediction carries over to VOS languages derived by p-movement. However, there are other possible derivations of this order, including base-generation and fronting of the verb phrase including the object, that do not entail the prediction in (39). Therefore evidence for the correct analysis is harder to find for VOS languaegs than for the VSO languages discussed here. See Chung 2002 for a survey.

Celtic languages, preverbal particles are found in Tagalog, Chamorro, Maori, Niuean, Arabic, K'ichee', Chinantek, Mixtex, and Turkana, among others. (Note that some of these languages are VOS languages.) The remainder of this chapter will be devoted to a discussion of preverbal particles in the Celtic languages.[28]

The following examples illustrate some aspects of the Welsh particle system.

(40) a.   Mi   oedd John yn chwerthin.
          PRT was John in laugh
          'John was laughing.' (Jones and Thomas 1977:358)

     b.   Mae Mair yn gwybod [yr   oedd John yn chwerthin]
          is   Mair in know   PRT was John in laugh
          'Mair know that John was laughing.' (Jones and Thomas 1977:358)

(41) a.   Ni   chiciodd yr bachgen y   pêl.
          NEG kicked   the boy   the ball
          'The boy didn't kick the ball.' (Louise Stanley, p.c.)

     b.   Gwyddwn na   chiciodd yr bachgen y   pêl.
          I-know   NEG kicked   the boy   the ball
          'I know that the boy didn't kick the ball.' (Louise Stanley, p.c.)

Main and embedded clauses are both VSO but dedicated particles distinguish the two clause types. *Mi* and *yr* in (40) introduce main and embedded clause declaratives, respectively. (41) shows that *ni* is used in main clause negatives, while *na(d)* is used in embedded negatives. Thus, the particles indicate whether a clause is a main clause or an embedded clause. This is a property typically associated with complementizers (Bresnan 1970), and it therefore seems reasonable to analyse these particles as complementizers (see Hendrick 1988, Willis 1998, among others).

The following examples illustrate a particularity of particles in Welsh (and in other Celtic languages). As is often the case, here, the preverbal particles are deleted. However, even if these particles delete, they trigger a morphophonological change, so-called

---

[28] The availability of V1 orders of the type found in German does not cloud the prediction. As noted earlier, in those examples the moved verb is "preceded" by an empty constituent in its specifier.

mutations, on following consonants. Thus, in (42a), initial /c/ changes to /g/ after the deleted affirmative complementizer, and in (42b) initial /c/ changes to /x/ after the deleted initial negative marker. These examples also illustrate that different particles trigger different mutations.[29]

(42) a. Gafodd hi anrheg ddoe.
        got she gift yesterday
        'She received a gift yesterday.' (Bury 2002a:226)
     b. Chafodd hi ddim anrheg.
        got she no gift
        'She didn 't receive a gift.'

Like Welsh, Irish has distinct particles for embedded and matrix clauses. However, the Irish particles are also inflected for tense.[30]

(43)

	embedded		matrix	
	negative	non-negative	negative	non-negative
past	nar	gur	nior	--
non-past	nach	go	ni	--

Note that there are no forms expressing the features [non-negative, non-past, matrix]. Pinker (1984) argues that paradigms form a crucial part of the entry of lexical items. If the Irish particles are stored in a paradigm, language acquisition offers an interesting account of the lack of a [non-negative, non-past, matrix] particle. The learner will crucially come across the following minimal pairs: (i) distinct negative particles for matrix and embedded clauses (*ni* vs. *nach*); (ii) distinct subordinating particles for negative and non-negative (*nach* vs. *go*). This means that the Irish particle paradigm (abstracting away form tense) will have dimensions for [matrix], [embedded], [negative], and [non-negative]. The learner can then conclude that the missing feature combination [matrix, non-negative] is realised by a null form of the particle.

In Breton, V1 order can be found in a variety of structures. If they are negated, both

---

[29] The deletion of preverbal particles in Celtic in general is discussed in Borsley and Roberts 1996:50, and for Breton in Denez 1973/1974:252f and Timm 1988:101 note 3, among others.
[30] The following table is adapted from a bigger list in Chung and McCloskey 1987.

main and embedded clauses are V1 and introduced by the particle *ne* as in (44). Relative clauses are also V1 and typically introduced by the particle *a*, although *e* and *ma* are also possible. This is illustrated in (45). Other embedded clauses are also V1. Complement clauses are introduced by the particle *e*, and clausal adjuncts can be introduced by the complementizers *ma* and *pa* among others.[31]

(44) a.   Ne gav ket e vota koad.

NEG find NEG his shoes wood

'He does not find his clogs.' (Stephens 1982:67)

   b.   Ne douin   ket [ne oa netra da damall d'am] distagadur

Neg I-swear Neg   Neg was nothing to blame to-my pronunciation

'I won't swear that there was nothing to complain of in my

pronunciation.' (Timm 1988:(29))

(45) a.   Hennezh a oa un den mat [a zifenne ar vrol]

that-one PRT was a man good PRT defended the country

'He was a good man who defended the country.' (Timm 1988:(45a))

   b.   an den [e komzan outañ]

the man PRT I-speak at-him

'The man to whom I am speaking.' (Timm 1988:(76))

   c.   Al labous [ma freuzhas Perig a neizh]

the bird detroyed Perig his nest

'The bird whose nest Perig destroyed.' (Stephens 1982:63)

(46) a.   Gouzout a   ra   Lenaig [e   lennas   Yann al lizher]

know   PRT does Lenaig PRT read-PAST Yann the letter

'Lenaig knows that Yann read the letter.' Stephens (1990:(3))

   b.   Mont a refomp er-maez [pa dorro an avel]

Go do-FUT-1P out when break-FUT the wind

'We shall go out when the wind drops.' (Stephens 1982:46)

---

[31] Unlike the particles *a* and *e*, "true complementizers" cannot be deleted (Timm 1988:84).

c.   Kemer un tamm bara [ma az peus naon]

Take a piece bread if have-2s hunger

'Take a piece of bread if you are hungry.' (Stephens 1982:46)

The examples in this section all illustrate the correlation between verb-initial order and preverbal particles. In the next section, a further aspect of Breton word order will be discussed, namely the fact that it typically display V2 order in main clauses.

5 Particles and verb movement in Breton[32]

5.1 Breton V2

The examples in the previous section showed that in a range of contexts, Breton clauses are V1, like those in other Celtic languages. However, it is well-known that V1 order is impossible in positive main clauses in Breton. Instead, a V2 pattern similar to that in Germanic is found. The following examples show that constituents of different categories can occur in preverbal position:

(47) a.   Annaig **a** gerc'ho bara evit koan.

Annaig PRT fetch(+FUT) bread for dinner

'Annaig will fetch bread for dinner.' Stephens (1982)

b.   Bara **a** gerc'ho Annaig evit koan.

bread PRT fetch(+FUT) Annaig for dinner

'Annaig will fetch bread for dinner.'

c.   Kerc'hat bara **a** raio Anna evit koan.

fetch bread PRT fetch(+FUT) Anna for dinner

'Annaig will fetch bread for dinner.'

(48) a.   Evit koan **e** kerc'ho Annaig bara.

for dinner PRT fetch(+FUT) Annaig bread

'Annaig will fetch bread for dinner.'

b.   Neuze   **e** teuas din da soñj va fezh-c'hoari.

Then PRT came-to-me to memory my play

'Then my play came to mind.' (Timm 1988:(7))

---

c. Sioul ha plijadurus   **e** tremenas an devez…

peaceful and pleasing PRT passed the day

'The day passed peacefully and pleasingly…' (Timm 1988:(8))

As in Germanic V2-languages, verb-initial orders are impossible in declaratives, and multiple preverbal constituents are usually ungrammatical:[33]

(49) a. * Prenn Soaz al levr.

buys Soaz the book

'Soaz buys the book.' (Jouitteau 2000: ch. 3)

b. *Al levr Soaz a brenn.

the book Soaz prt buys

'Soaz buys the book.' (Jouitteau 2000: ch. 4)

The similarity is not restricted to the impossibility of verb-initial order. It was shown in section 2 above that in German the obligatory preverbal constituent is not restricted to a particular discourse function. Discussing NP-initial clauses in Breton, Timm 1991:306 observes that initial NPs in Breton similarly can serve "a variety of Topicalization and Focalization functions". This is illustrated in the following texts from Timm 1991:(6) and (8).

(50) a. Surtoud amañ   **ez** eus   bezhin gwalc'he e gis-se…

above-all here   PRT be-3sg-presp-exist algae *gwalc'he* in form-that

'Here above all there is [found] the alga *gwalc'he* in that shape.'

b. An dra-se   **a** zo   mad d'ar   re … **a** rank mont pell

the thing-there PRT be-3SG-pres good to-the those PRT must go far

e gis paotred Plougerne

in form men Plougerne

'That's good for those who must go a long way such as the men from Plougerne.'

---

[33] Two exceptions are the verbs *emañ* 'to be' and *mont* 'to go', which can appear clause initially (Jouitteau 2000). Furthermore, Breton seems to allow V3 orders more freely than the Germanic languages (Schapansky 1996).

c. Ar re-se **a** rank kaoud bagou' brassoc'h heuse

those PRT must find boats bigger then

'They must find bigger boats, then.'

(51) a. Etre daou en em okupent

between two REFLX occupy-PL-IMF

'They divided themselves between two [occupations].'

b. Un tamm tiegezh   o doa oa   ket voian, da c'helloud gounid

a bit farm have-3PL-PAST be-SG-IMF NEG means to-be able   earn

awlc'h ewid bewañ gant an dra-se

enough for live with the thing-there

'They has a small farm, [otherwise] it was not possible to earn enough on the other [occupation of algae gathering].'

The Breton V2 pattern then poses the same problems for a conventional analysis in terms of feature checking as the German V2 pattern discussed at the beginning of this chapter. The V2 pattern involves a preverbal constituent that may be of any category, and this constituent can be interpreted in a range of different ways. This suggests that the pattern is not triggered by a particular syntactic or semantic/pragmatic feature. This means that the analysis of V2 as a consequence of the primitives of phrase structure should ideally be extended to Breton.

5.2 Breton particles and VSO order

Whether the earlier analysis of the V2 pattern can be extended to Breton depends on the analysis of a striking characteristic of the Breton V2 pattern. As the examples in (48) to (50) illustrate, the initial constituent in Breton V2 clauses is usually separated from the verb by one of the preverbal particles *a* and *e*, which also occurred in some of the V1 examples in the previous section. While they often delete, these particles are obligatory with finite verbs. This can be seen in the fact that the mutation of following consonants

that they induce remains even when they are deleted: [34]

(52)   a.   Soaz (a) brenn al levr.   Jouitteau 2000: ch. 4

          Soaz PRT buys the book

          'Soaz buys the book.'

       b.   * Soaz prenn al levr.

            Soaz buys the book

          'Soaz buys the book.'

The issue raised by these particles in the present context is the following. As argued in the previous section verb-initial orders can only be derived through p-movement, that is when there is a head position available in whose position the verb can be pronounced. The availability of (embedded) VSO orders in Breton in which the verb is preceded by one of the preverbal particles *a* and *e* then suggests that in these structures *a* and *e* are separate heads in clause structure and that verb movement in Breton embedded clauses is p-movement that targets the position of *a/e*. However, if the Breton V2 pattern is to be analysed in terms of s-movement, *a/e* in V2 clauses cannot be independent heads in clause structure. This is so because the presence of an independent head that dominates the verb would block s-movement, and instead p-movement would have to occur. This would make it impossible to derive the V2 pattern. I would like to suggest that the solution of this apparent paradox lies in the content of the particles *a/e*, and that indeed these particles are independent heads in VSO clauses, but not in V2 clauses.

This proposal relies on the idea that some lexical material, in this case the particles *a* and *e*, may enter a structure in some contexts as an independent syntactic head but in others only as part of a morphologically complex head. This is not a new idea. Borer 1998 discusses a range of such proposals from work on the morphology-syntax interface according to which morphological material may enter the syntactic component at different stages, and hence where the same morphological material may occupy different structural positions in different contexts. Depending on the point of insertion, this material may or may not have syntactic effects (see also Ackema and Neeleman 2002). For example, Borer

---

[34] *A* and *e* do not stand in opposition with forms of the verbs *kaout* 'have' and *emañ* 'be' (Timm 1988:82;

1998:182 argues that "while lexical compounds display idiosyncratic, drifted properties, compounds that are inserted later preserve argument structure (Japanese, Hebrew); allow their non-head members to be modified (Hebrew), and exhibit word-internal effects of post-syntactic phonology (Japanese), all absent from lexical compounds".

We should then take a closer look at the properties of *a/e*. The closest living relative to Breton is Welsh. It was argued in the previous section, based on the availability of different particles for root and embedded clauses, that the preverbal particles in Welsh should be analysed as complementizers. Just like Welsh, Breton has particles that cannot occur in a main clause, namely *ma* and *pa* in (46b,c) among others, and indeed in embedded contexts *a/e* are in complementary distribution with those heads. This suggests that *a/e* in embedded clauses should be analysed, like the unambiguous complementizers of Breton, along the lines of the preverbal particles in Welsh, namely as independent syntactic heads. In the following structure, Subj is the subject, V is the position of the verb before movement, α stands in for whatever other structure there may be (e.g. a complex verb phrase), and the root position is occupied by one of the heads that can introduce embedded clauses.[35]

(53)   a/e/pa/ma   (in embedded clauses)
                 |
                 V
                / \
            Subj   α
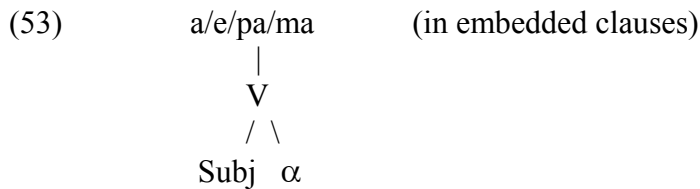
Given such a structure, the order particle-verb-subject can be analysed as the pronunciation of the verb in the position of the initial particle, i.e. as a p-movement structure.

There are independent considerations that suggest that an analysis of this type is the right one. In chapter 2, it was argued that the theory assumed here implies that a structure headed by a moved verb cannot enter into selectional relations and that therefore complement clauses and relative clauses cannot be headed by a moved verb. This condition is repeated here (see chapter 2 for discussion):

---

also Denez 1973/1974:257f).

(54) An embedded clause cannot be headed by a moved verb.

It is uncontroversial that verb movement across the subject is obligatory in Breton because in both embedded and in main clauses the verb precedes the subject[36] (see section 7 for a discussion of a possible trigger). Since embedded clauses cannot be headed by a moved verb, it follows that Breton embedded clauses cannot be s-movement structures, which would be structures in which the highest head is part of a complex chain. Now in an embedded clause introduced by *a/e*, the verb is not preceded by any other head (cf. (45), (46) above). This means that here p-movement is only possible if *a/e* are independent heads.

Given these considerations, it seems clear that *a/e* must be independent heads in embedded contexts. Since on such an analysis, *a/e* occupy the same structural position as Breton complementizers, it seems clear that they mark a clause as embedded clause.

Clearly in a main clause, *a/e* do not mark syntactic embedding. What then is the role of a/e in V2 main clauses? Where the function of these particles is discussed, there is wide agreement that their only role is to mark the function of the constituent preceding the verb, that is a type of agreement. Timm 1988:81 states that *a* indicates "that a preceding NP is a subject or [direct object]" and *e* is used "when anything else precedes the verb (including therefore all oblique cases of NPs)".[37] If all that a/e do in main clauses is to mark agreement, an analysis of these particles as independent heads would be suspicious, given Chomsky's 1995b discussion of agreement projections. On the assumption that clause structure only contains heads that receive an interpretation at LF (the Principle of Full Interpretation), material that only marks agreement cannot occur as an independent head. This suggests that in matrix clauses *a/e* are not independent heads.[38]

As noted above, lexical material may enter a syntactic structure in two ways. One possibility is that it may be inserted into a structure as an independent head. The alternative

---

[35] It is possible that V in (60) is a position created by type-1 verb movement. Thus, α may not only correspond to the traditional VP constituent, but also to a bigger structure, possibly corresponding to the traditional TP. Of course, such extra structure is only possible if the learner has evidence for it.

[36] Unless the subject is fronted in main clauses, of course.

[37] According to Stephens 1982, *a* occurs whenever an argumental NP, a non-finite verb form, or a VP precedes the inflected verb, while *e* occurs in all other cases.

[38] Schafer 1995 analyses *a* and *e* as heads of PolP, a syntactic projection containing information about the polarity of a clause. More recently, *a* and *e* have been analysed as overt realisations of light verbs in Jouitteau 2002.

is insertion as part of a complex head, generated outside syntax.[39] Since a/e cannot occur as independent heads in Breton main clauses, they must be part of a complex head. This conclusion converges with the view of Timm 1988:86, who proposes the following structure for the verbal complex:

(55)   V $\rightarrow$ [PRT + V]   PRT $\rightarrow$ a, e

The close relation of *a* and *e* to the verb is also reflected in the terminology of traditional Breton grammar. Jouitteau (2000: appendix) notes that the Breton term *rannig verb* translates as 'petit bout de verbes'. Thus, we can assume that while in embedded clauses a/e occur as independent heads, they are part of a complex head that also contains the verb in main clauses.

   If the particles and the verb form a complex head [$_V$ a /e + V], there is no additional syntactic head in the structure of Breton main clauses. Since verb movement across the subject is obligatory in Breton, main clauses must be s-movement structures. This explains the need for an additional specifier, and hence the V2 pattern. A Breton V2 clause would then have the following structure, where X indicates the material that occurs in the additional specifier:

(56)   [$_V$ a / e + V]
              /   \
         X   [$_V$ a / e + V]
              /   \
            Subj α

This analysis then can then account in a natural way for the V2 pattern of Breton main clauses, as well as for the ungrammaticality of unembedded particle-initial orders, like the following: [40]

---

[39] There are two possibilities for what 'outside syntax' could be. It could be either the word formation component, or it could involve late insertion in the sense of Distributed Morphology or earlier work (cf. Hudson 1976).

[40] Timm 1988:102f points out that there are some constructions that allow initial particles:

(i) A c'hellfe bezañ …   Similar examples were also brought to my attention by Melanie Jouitteau (p.c.). Maybe in those examples the particles are licensed as independent heads in discourse and through the special interpretation such structures receive. A different approach to these examples would be to say that they are in fact embedded but that the matrix clause has been deleted or is empty (cf. Ross 1970). However, the analysis of such constructions must await future research.

(57) a.　* a brenn Soaz al levr (Jouitteau 2000: ch.3)

　　　　PRT buys Soaz the book

　　　　'Soaz buys the book.'

　　b.　* e prennas Soaz al levr

　　　　PRT buys Soaz the book

　　　　'Soaz buys the book.'

Since s-movement structures always requires a filled specifier, the fact that the content of this specifier position is not limited to a particular syntactic category or to constituents that receive a particular semantic or pragmatic interpretation is expected.

Thus, the word order contrast between Breton main and embedded clauses results from the different structural positions that the particles a/e occupy in these clause types. Where a/e occur as independent heads, VSO order is possible; where they do not, obligatory verb movement triggers the V2 pattern.

This analysis can also account in a natural way for the word order of negative clauses in Breton. The following examples, repeated from section 4, illustrate that negative clauses are introduced by the negation marker *ne* and that they show VSO order whether they are embedded or not:

(58) a.　Ne gav ket e vota koad.

　　　　NEG find NEG his shoes wood

　　　　'He does not find his clogs.' (Stephens 1982:67)

　　b.　Ne douin　　ket [ne oa netra da damall d'am distagadur]

　　　　Neg I-swear Neg　　Neg was nothing to blame to-my pronunciation

　　　　'I won't swear that there was nothing to complain of in my

　　　　pronunciation.' (Timm 1988:(29))

Since the function of the particle *ne* is to mark the scope of negation, it is not expected to show any asymmetries between main and embedded clauses. Thus, if *ne* is analysed as an independent head, the absence of V2 effects in negative clauses in Breton is expected.

Before discussing what triggers verb movement in the first place in section 7, I turn now to the history of V2 and VSO in Welsh which provides further support for the approach to Breton particles proposed here.

6 The change from V2 to VSO in Welsh[41]

The restrictions the present theory imposes on verb movement configurations and the relation it implies on preverbal particles and verb-initial order entail the diachronic prediction in (59). Since V2 languages lack independent preverbal particles and VSO languages require them, a language that changes from V2 to VSO must develop such particles.

(59) If a language changes from V2 to VSO, it will develop preverbal particles that dominate the verb in clause structure.

This prediction can be tested against the development of Welsh, which changed from V2 to VSO within documented history (Willis 1998). The examples in (60) show that the Middle Welsh V2 pattern is very similar to the V2 pattern in Breton:

(60) a. Atref   y   doeth Arthur...
         homeward PRT came Arthur
         'Arthur came home.' (Willis 1998:52)

     b. Yn Harlech y   bydwch   seith mlyned ar ginyaw...
         in   Harlech PRT be-FUT-2P seven years   at dinner
         'In Harlech you will be at dinner for seven years...' (Willis 1998:51)

(61) a. Gwyssyaw   a   oruc Arthur milwyr yr ynys   honn...
         summon-VN PRT did   Arthur soldiers the island this
         'Arthur summoned the soldiers of this island...' (Willis 1998:52)

     b. A'r   ederyn a   doeth y'r   ynys   honn.
         And the bird   PRT came to the island this
         'And the bird came to this island.' (Willis 1998:51)

     c. Ef a   doeth makwyueit a   gueisson ieueinc y   diarchenu...
         it PRT came squires   and lads   young   to+3SM-GEN disrobe-VN
         'There came squires and young lads to disrobe him...' (Willis 1998:52)

Like Breton, Middle Welsh has particles that mark agreement with the initial phrase. Willis

---

[41] A version of this material appeared as Bury 2002a.

argues that *a* and *y(d)* are a form of "topic-agreement marking on the verb" (1998:62). *A* occurs after subjects, as in (60b) and (60c), direct object, nonfinite verbs, or VPs, as in (60a). *Y(d)* occurs mostly after adverbs, as in (61a) and (61b). Thus, these particles agree with the phrase occupying the specifier position of the fronted verb.[42]

It seems then reasonable to extend the Breton analysis (cf. (55)) to these examples, so that *a* and *y(d)* form a complex with the verb, which then creates a new position to be able to move and forces the creation of a specifier:

(62) a.   [$_V$ y+doeth] -->[$_{Agr}$ y] + [$_V$ doeth]

  b.   [$_{VP}$ [$_{AdvP}$ Atref] [$_{V'}$[$_V$ y+doeth]$_i$ [$_{VP}$ Arthur t$_i$]]]   (cf. (60a))

       homeward   PRT came   Arthur

As in Breton, negative sentences in Middle Welsh are VSO (Willis 1998:94). In (63) the fronted verb is only preceded by the negative particle *ny*.[43]

(63) a.   Ny   welei   ef y   twrwf   rac tywyllet y   nos.

       NEG saw-IMPF he the commotion for so-dark the night

       'He could not see the commotion because the night was so dark.'(Willis 1998:94)

  b.   [$_{NegP}$ [$_{Neg}$ Ny   [$_V$ welei]$_i$] [$_{VP}$ ef t$_i$ y twrwf rac tywyllet y nos]]

Another exception to the general V2 pattern in Middle Welsh are embedded clauses. Willis (1998) reports that in most embedded clauses the order is VSO. These two cases of V1 order are not surprising in this theory since in both cases the verb is preceded by an independent (because interpretable) particle. Embedded clauses in Middle Welsh then differ from those in Breton ones introduced by *a* or *e* to be discussed above in that they are introduced by real complementizers, as (64) illustrates. These complementizers and the Middle Welsh negative particle have interpretable features, just like the Modern Welsh

---

[42] A third agreement pattern is found after fronted adjectival or nominal predicates where the verb undergoes soft mutation; Willis 1998:52.

[43] See Willis 1998 for reasons to analyse *ny* as a head. This analysis may also explain why *ny* cannot co-occur with *a/yd*. Given a preference of move over merge, insertion of *ny* (if available) will always dominate the verb, and block the creation of an additional position. Consequently, the moved verb adjoins to *ny* and

particles. Thus, unlike *a* and *y(d)*, *chyt* can be an independent head. This means the verb is dominated by a higher head and p-movement is possible.

(64)   A   chyt   archo   ef yti   rodi   yr eil,   na   dyro...
       and though implore-SUBJ he to-you give-VN the second NEG give-IMPER
       'And though he implore you to give him the second, do not give (it)...'(Willis 1998:53)

We can now turn to the historical development.

Willis (1998) shows that in the sixteenth century, the system of preverbal particles underwent a number of changes. In particular, the particle *a*, which occurred with initial subjects and objects, was lost. Following the loss of *a*, initial subject pronouns became cliticised on the fronted verb. This change created an environment in which the expletive pronoun *fe* could be reanalysed as main clause complementizer. A sentence like (65a) would then be structurally ambiguous. The partial representations in (65b-c) are adapted from Willis 1998:179.

(65) a.   Fe welodd Arthur farchog.
       FE saw   Arthur knight
       'Arthur saw a knight.' (Willis 1998:178)
    b. $[_{CP}[_C[_D \text{ fe}_i+\text{welodd}_j]] [_{F1P}[_{DP} \text{ t}_i] [_{F1'} \text{ t}_j [_{F2P} \text{Arthur } [_{F2'}...$
    c. $[_{CP}[_C \text{ fe}+\text{welodd}_i] [_{F1P} [_{F1'} \text{ t}_i [_{F2P} \text{Arthur } [_{F2'}...$

(65b) is the conservative structure. Willis assumes that the expletive construction was a lexical exception to the V2 constraint, and that *fe* was generated in SpecF1P and then adjoined to the verb in C. In the new structure, (65c), *fe* is inserted directly into C as a complementizer, and the verb adjoins to it. The two underlying structures derive identical surface structures with comparable interpretations.

The reanalysis of *fe* as main clause complementizer leads to "a sudden and quite dramatic increase" (Willis 1998:164) in use of this element towards the end of the sixteenth century. This increase in the use of *fe* coincides with the "rapid innovation during

---

cannot enter a specifier-head configuration with a possible topic. Thus, the agreement relation that *a/yd*

the [sixteenth; DB] century" of unmarked absolute verb-initial orders (Willis 1998:197). It is not clear what the exact figures are but Willis argues that by the seventeenth century "it is fair to characterise Welsh as a VSO language" (1998:205). Thus, the development of Welsh confirms the prediction of the present theory. VSO order becomes available at the same point as the language develops a VP-external head that can be targeted by verb movement, namely the main clause complementizer. Two examples of the new VSO order are given below:[44]

(66) a. Fe sugne   hi mer   nhw am arian.
        FE suck-COND she bone-marrow them for money
        '...she'd suck their bone-marrow for money.' (Willis 1998:230)

    b. Gorvüost   ar dy   elynion...
        overcame-2S on your enemies
        'You overcame your enemies...'(Willis 1998:196)

Consider now the analysis of the loss of V2 in Welsh proposed in Willis 1998. While Willis assumes the existence of a V2 parameter "which forces movement of the verb to C and movement of some phrasal constituent to SpecCP" (1998:57), he remains uncommitted as to the exact form of this parameter. The question at the heart of his discussion is what kind of evidence there was in Middle Welsh that triggered a positive setting of the V2 parameter, and he argues in some detail that there was a steady decrease of this evidence.

In Willis's view the most important change was the loss of the preverbal particles *a* and *y(r)* (or its older form *y(d)*). He assumes that these particles are generated in the C-position, which means that the phrase with which they agree must be in the specifier of C. The loss of *a* meant that subject-initial sentences could be assigned a structure in which the subject remains in the highest specifier position in the inflectional domain. Similarly, the loss of *y(r)*, "the single most important development precipitating to the breakdown of verb-second" (1998:188), meant that initial adverbs could be analysed either as occupying

usually mark cannot be established.
[44] There are two possible explanations for the alternation between *fe*-VSO and null-VSO clauses. Missing complementizers could be the result of PF deletion of *fe*, leaving the characteristic mutation. Alternatively, a true null-form could be learned through paradigm formation. See sections 3 and 4.

Spec-CP, or as being in a position adjoined to IP. Willis proposes that the learner will choose the latter analysis unless there is strong evidence that a more complex structure is necessary. The frequency of object-topicalization gradually decreased, and by the sixteenth century object-topicalization was only available under contrastive focus, which is essentially the Modern Welsh situation. Finally, the reanalysis of the expletive *fe* as main clause complementizer and the increase in use of the resulting C-VSO structure constituted a further exception to the V2 pattern. Ultimately the decrease of evidence for V2 brought about by these changes led to a resetting of the V2 parameter in the sixteenth century.

While compatible in many ways, Willis's proposal and the one developed above diverge in a number of respects. In Willis's account the loss of the agreement particles *a* and *y(r)* is of paramount importance to the loss of V2 because of its effect on the acquisition of the V2 rule. In the present proposal, this change is not directly related to the loss of V2. The loss of *a* is important in as much as its disappearance created an environment in which initial subject-pronouns could become clitics, which in turn led to the reanalysis of the expletive pronoun *ef* (later *fe*), but it has nothing to do with the loss of V2 itself. There are no theoretical reasons why the loss of the agreement particles should result in the loss of the V2 pattern, and there is no typological support for it either; as Willis notes, none of the Germanic languages have an agreement system comparable to that of Middle Welsh but yet they all are V2 languages (except, of course, for English).[45]

In Willis's account the reanalysis of the expletive *fe* as a main clause complementizer is only one among several changes that contributed to the decrease in the evidence for a positive setting of the V2 parameter. In the present proposal however, it is the one change that allowed Welsh to develop unmarked VSO orders. The loss of *a* was maybe a necessary but not a sufficient change, while the innovation of a main clause complementizer was both necessary and sufficient.

In this proposal, the chronology of the observed changes is predicted by the theory. (i) VSO orders could only be derived once there was a VP-external head in the functional domain of the language to which the verb could adjoin. (ii) The sharp increase in the use of VSO orders after the reanalysis of *fe* is expected because the innovation of the main clause

---

[45] Willis addresses this criticism but there is no room here for a discussion of his arguments.

complementizer means that there is no trigger for obligatory topicalization.[46] The fact that the reanalysis of *fe* and the development of unmarked VSO order coincide does not receive an immediate explanation in Willis's account. A further difference between the two proposals concerns the synchronic analysis of Middle Welsh. If Middle Welsh has a positively set V2 parameter, the VSO order found in negatives is unexpected. On the other hand, the above discussion showed that the present proposal predicts negatives to have VSO order.[47]

7 Verb movement and case

As noted in section 5, verb movement in Breton shows no root/embedded asymmetry and it can be realised either in a p-movement or an s-movement structure. We then need to find a trigger of verb movement in Breton that is compatible with these properties.

In a model of grammar where linguistic representations are evaluated against two sets of interface conditions, those of PF and those of LF, three types of conditions can be formulated, and used as triggers, without increasing the theoretical machinery: one type of triggers for each of the interfaces, and one further type can be formulated as conditions that hold of both interfaces. PF is directly accessible to the learner through positive evidence. This means that triggers formulated as conditions on PF are likely candidates for cross-linguistic variation. In contrast, it is usually assumed that learners have access to LF only by virtue of their knowledge of UG. This means that triggers formulated as conditions on LF can presumably not be parametrised. LF conditions can then only trigger universal operations. However, a condition that invokes concepts of LF as well as of PF may vary across languages. Thus, a particular range of syntactic (LF) properties may be selected as requiring overt (PF) realisation.

The clause typing requirement in German seems to be an example of the last type of trigger. It relies on properties of LF, such as the content of syntactic categories like complementizers and on notions like selection and chains, as well as on properties of PF, such as which of possible head positions are realised. Clause-typing is not an appropriate

---

[46] The decreasing frequency of topicalized objects was not matched by a corresponding increase in structures without topicalization.

[47] Indeed Willis speculates that the optionality of topicalization in negatives may be due to lexical differences between the negative and agreement markers. While this idea is promising, it is not clear how such an analysis will follow from a parametric theory.

trigger for verb movement in Breton because of the conditions formulated on LF. Thus, at the heart of the root/embedded asymmetry of verb movement in German that is triggered by the clause typing requirement in German is the formation of a head chain. In contrast, p-movement cannot be triggered by this type of requirement because it does not involve the formation of a head chain. If we assume that PF is blind to the category of a syntactic head and sees only the presence of a head position, the difference between s- and p-movement configurations disappears at this interface. Since both configurations allow for the pronunciation of the verb in the position of a higher head, PF seems to be the right place to look for a possible trigger of verb movement in Breton.

A condition that can function as a trigger of this type has been formulated by Neeleman and Weerman 1999: chapter 2 as part of their OV/VO parameter. The basic idea is that OV and VO languages differ not only in the direction of case licensing but also in the domain in which case licensing of noun phrases takes place (op. cit. p.26).[48] The direction of case licensing in OV languages is leftward and it is restricted to elements in a relation of mutual m-command. The direction of case licensing in VO languages is rightward and restricted to elements in the same prosodic phrase. Since case licensing holds at PF, the interface between syntax and phonology, both syntactic and phonological information must be available.

This proposal can be easily adapted for the purpose at hand as follows:

(67)   Case assignment (VO and VS)

        α may assign case to β iff

        a. α precedes β, and

        b. α and β are contained in the same prosodic domain $\{_\Phi \ \}$.

        c.   no possible case assigner γ intervenes between α and β.

Algorithms on the mapping between syntactic phrases and prosodic phrases have usually been formulated in terms of X-bar trees. Nevertheless, the formulation below does not cause too many problems.

---

[48] Neeleman and Weerman 1999:24 argue that a particular setting of directionality automatically implies a particular domain of case licensing, simplifying the task of the learner.

(68)　　Φ-formation (following Selkirk 1986)

　　　　Align the right edge of a prosodic phrase {} with the right edge of a

　　　　syntactic constituent [].

Consider the following structure, which is not specified for linear order and where the numbers represent positions corresponding to (derived heads of) the verb.

(69)　　　　　　3
　　　　　　　／ ＼
　　　　　Subj　2
　　　　　　　　／ ＼
　　　　　　　Obj　1
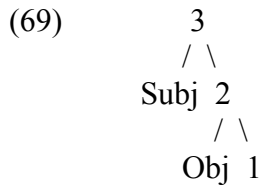
As argued in chapter 1, linearisation of such a structure preserves its constituency. This means that each part of a structure that is exhaustively dominated by a single node corresponds to a constituent. (69) can then be mapped onto the following linearised constituent structures:

(70)　　a.　[[Subj] 3 [[Obj] 2 [1]]]

　　　　b.　[[Subj] 3 [[1] 2 [Obj]]]

　　　　c.　[[[Obj] 2 [1]] 3 [Subj]]

　　　　d.　[[[1] 2 [Obj]] 3 [Subj]]

Given the algorithm in (68), the constituent structures in (70) can be mapped on to the following prosodic structures:

(71)　　a.　constituent str.　　[[Subj]　3 [[Obj]　2 [1]]]

　　　　　　prosodic str.　　　{ Subj}{3　Obj}{2　1}

　　　　b.　constituent str.　　[[Subj]　3 [[1]　2 [Obj]]]

　　　　　　prosodic str.　　　{ Subj}{3　1}{2　Obj}

c. constituent str. [[[Obj] 2 [1]] 3 [Subj]]

prosodic str. { Obj}{2 1} {3 Subj}

d. constituent str. [[[1] 2 [Obj]] 3 [Subj]]

prosodic str. { 1}{2 Obj}{3 Subj}

In theories assuming head adjunction, there is a distinction between empty head positions that correspond to traces and those corresponding to functional heads to which no verb has moved. In such a model, traces are the only empty heads able to assign case. This distinction can be preserved as a condition that a position corresponding to the verb or one of its derived heads is only a possible case assigner, if it has phonetic content or if it is preceded by a dominating head position that has phonetic content. This can be formulated without the disjunction as follows:

(72) A position can only be assigned case if it is preceded by a head position with phonetic content.

Of the above prosodic structures, only those in (71a,b,d) are then possible structures for a language in which the object needs to be assigned case by the verb. In (71c), the object is not preceded by an appropriate head position, none at all in fact, and therefore it cannot be assigned case.

The minimality condition on case assignment in (67c) above is necessary to account for examples where e.g. a verb selects a prepositional phrase, as in (73). In this structure, both the verb and the preposition are in the same prosodic phrase as the noun phrase. However, in such a structure, it appears always to be the case that the preposition assigns case to the noun phrase, not the verb.

(73) Constituent structure … [V [P [D …
      Prosodic structure …{V P D

Neeleman and Weerman 1999:197 propose that in certain VSO languages not only

objects need to be assigned case under the conditions in (67), but subjects as well.[49] However, the prosodic structures in (71) that are derived from the phrase structure in (69) do not provide an appropriate configuration in which the subject could be assigned case. On the assumption that a head position in prosodic structure must correspond to a head position in constituent structure, verb movement can then be triggered in order to provide an additional head position. This leads to the following phrase, constituent and prosodic structures: (The highest head position that would have be pronounced to assign case to the subject is typed boldface.)

(74)
```
            4
            |
            3
           / \
       Subj  2
             / \
          Obj  1
```
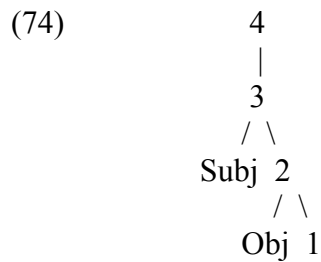
(75)   a.   constituent str.   [4 [[Subj] 3 [[Obj] 2 [1]]]]   VSO

          prosodic str.   {**4** Subj}{3 Obj}{2 1}   ok

       b.   constituent str.   [4 [[Subj] 3 [[1] 2 [Obj]]]]   VSO

          prosodic str.   {**4** Subj}{3 1}{2 Obj}   ok

       c.   constituent str.   [4 [[[Obj] 2 [1]] 3 [Subj]]]   VOS

          prosodic str.   {**4** Obj}{2 1}{3 Subj}   ?

       d.   constituent str.   [4 [[[1] 2 [Obj]] 3 [Subj]]]   VOS
          prosodic str.   {**4** 1}{2 Obj}{3 Subj}   ?

       e.   constituent str.   [[[Subj] 3 [[Obj] 2 [1]]] 4]   SVO

          prosodic str.   { Subj}{3 Obj}{2 1}{ 4}   *

---

[49] For alternative proposals see Adger 2000, Benmamoun 1998.

f.   constituent str.   [[[Subj] 3 [[1] 2 [Obj]]] 4]   SVO

prosodic str.   { Subj}{3   1}{2 Obj}{ 4}   *

g.   constituent str.   [[[[Obj] 2 [1]] 3 [Subj]] 4]   OVS, OSV

prosodic str.   {   Obj}{2 1}{**3** Subj}{4}   *

h.   constituent str.   [[[[1] 2 [Obj]] 3 [Subj]] 4]   OVS, OSV

prosodic str.   {   1}{2 Obj}{**3** Subj}{4}   *

The structures in (74a,b) are both possible structures in which the subject could be assigned case from the moved verb in the position of 4.

Neeleman and Weerman 1999 argue that this approach to case assignment can account for a pattern that holds of a range of VSO languages.[50] In finite clauses like (76a) the verb moves to assign case leftwards to the subject. A non-finite verb is not a possible case assigner, therefore does not move. Instead, the subject is assigned case by the preposition *i* in examples like (76b) (Borsley and Roberts 1996a):

(76) a.   Mae   S î on yn   yfed   llaeth.

is   S î on PROG   drinking milk

'S î on is drinking milk.'

b.   [Cyn i S î on yfed y llaeth] mae'n rhaid iddo fe fwydo'r ddraig.

before to S î on drinks the milk, is necessary to-him feed the dragon

'Before S î on drinks the milk, he has to feed the dragon.'

c.   *[Cyn yfed i S î on y llaeth] mae'n rhaid iddo fe fwydo'r ddraig.

before drink to S î on the milk, is necessary to-him feed the dragon

Neeleman and Weerman 1999 note in Welsh (and some other VSO languages) the fronted verb and the subject are always adjacent:

---

[50] The idea that verb movement in Welsh has to do with case assignment goes back to Sproat 1985, who observes that verb movement in Welsh contrasts with the presence of a potential governor, and suggest an analysis in terms of case theory.

(77) a. *Mi oedd   heddiw John yn chwerthin.

    PRT was   today   John   in laugh

   b. Mi oedd   John   yn chwerthin.

    PRT was   John   in laugh

    'John was laughing.' Jones and Thomas (1977)

On their approach to case assignment this pattern receives a straightforward explanation. In (78a) the subject is in the same prosodic domain as the finite verb, but in (78b) the adverbial closes off the prosodic domain containing the verb, and Case assignment to the subject is blocked:

(78) a. [Mi oedd [John] [yn chwerthin   cf. (77b)

    {Prt was   John} {in laugh}

   b. *[Mi oedd [heddiw]   [John][yn chwerthin   cf. (77a)

    { Prt was   today}{ John }{ in laugh}

    Breton displays a similar pattern with respect to verb movement and verb-subject adjacency as does Welsh. Stephens 1990 shows that if there is no verb movement, there must be an alternative case assigner for the subject. In (79a), the subject of the bracketed non-finite clause assigns case with the preposition *da*, while in (79b) it checks Case with the coordinative conjunction *ha*; the verb does not move in these examples:

(79) a. [Da Yann   da welout] e oant holl aze.

    {da Yann}   {da welout …

    to Yann   to see were-3PL all there

    'As far as Yann could see they were all there.'

   b. Ha Yann   lenn al lizher.

    {ha Yann} {lenn al lizher}

    PRT Yann to read the letter

As in Welsh, a fronted verb must in general be adjacent to the subject in Breton. The examples in (80) show that an adverbial intervening between the verb and the subject leads

to ungrammaticality.[51] The ungrammatical alternative of (80a) would receive the structure in (81), where the subject is not in a prosodic domain with the finite verb.

(80) a. Deus ar beure   e save   (*abred)   Yann (abred).
          in the morning   PRT got up   early   Yann early
          'Yann got up early in the morning.' (J. Stephens, p.c.)

      b. Deus ar beure   e lenne   (*abred) Yann   ar journal.
          in the morning   PRT read early   Yann   the paper
          'Yann read the paper in the morning.'

(81)   [VP [PP Deus [DP ar beure]]   e save   [AdvP abred] [DP Yann] …
          {Deus   ar beure}   {e save   abred}   {Yann …
               in   the morning   PRT got up   early   Yann

However, a refinement of the conditions for the formation of prosodic domains in Breton may be necessary. Thus, a number of prosodically light elements can intervene between the verb and the subject. These include negative markers, certain short adverbs, and inflected prepositions:[52]

(82) a. Ne lenn **ket** Lenaig al lizher.
          NEG read NEG Lenaig the letter
          'Lenaig didn't read the letter.' (Stephens 1990:(6))

      b. N' eus bet **biskoaz** ger   etrezomp
          Neg is   been never   word [between us]
          Il n'y a jamais eu le moindre mot (malentendu) entre nous.

(83) a. Ur wech an amzer e skrive e vamm dezhañ**.**
          Une fois   le temps PRT écrivait sa mère [à lui]
          De temps en temps, sa mère lui écrivait.

      b. Ur wech an amzer e skrive **dezhañ** e vamm.
          Une fois   le temps PRT écrivait [à lui]   sa mère
          De temps en temps, sa mère lui écrivait.

---

[51] Thanks to Janig Stephens (p. c.) for providing these examples.
[52] Thanks to Melanie Jouitteau for discussion and for pointing out the examples below to me.

c. Ne   skrive **ket dezhañ**   e gamaladez james

Neg   wrote not [to him] his friend   never

His friend never wrote to him.

(84) a. Breman e labour an dud   brav   eno.

now   PRT works the people good   there.

Maintenant, les gens y travaillent bien.

b. Breman e labour **brav** an dud   eno.

now   PRT works good the people there.

Maintenant, les gens y travaillent bien. .

As far as I understand these facts, the option to appear between the verb and the subject is restricted to very light elements.[53] Thus, if the adverb in (84b) is replaced by a polysyllabic adverb, or if it is modified, the resulting structure would be ungrammatical. This dependence on prosodic weight would then further support the proposal that Breton verb-subject adjacency is an effect of prosodic constraints.

In any case, such examples, and the prosody of Breton in general, require more detailed analyses, and this work remains the subject of future research.

8 Conclusion

This chapter dealt with the relation between V2 and V1 structures. Section 2 illustrated the basic problem for an account of the V2 pattern. Section 3 outlined an approach to clause structure in which V2 patterns are the expected for structures in which the moved verb is pronounced in the position of a syntactic copy of the verb, i.e. in s-movement structures. In section 4, it was argued that in V1 languages the moved is pronounced in the position of a (syntactically) independent head, i.e. that they involve p-movement structures, and this analysis was related to the generalisation that verb-initial languages, in particular VSO languages, have preverbal particles. Section 5 discussed V1 and V2 patterns in Breton and related the difference between the two to properties of the Breton particle system. Section 6 discussed the theory's prediction that the change from V2 to VSO is only possible where

---

[53] It may be relevant to note that Hendrick 1988:239 argues that *ket* is affixed to the verb.

a language develops a preverbal particle at the example of Welsh. Finally, section 7 considered the trigger of verb movement in Breton.

Free relatives and derived heads

1 Introduction

One of the main theses of the theory proposed in chapter 1 is the rejection of the assumption that moved material cannot head a structure. As noted there, this proposal makes available two types of structures that are unavailable on standard assumptions. In terms of conventional X-bar theory, these structures can be represented as in (1a,b). The structure in (1a) is created through movement of a head and subsequent projection of a new phrase by this head. The corresponding structures in a framework without categorial projection have been discussed in some detail in chapters 2 and 3. The structure in (1b) is created through movement of a phrase and subsequent projection of a new phrase by this phrase. Such a structure resembles a traditional adjunction structure where the constituent β is adjoined to the constituent XP, with the difference that in an adjunction structure XP would not have moved out of β. An example of a structure that has been analysed in this way are nouns modified by a relative clause. This analysis is illustrated in (1c), where the relative clause RC is adjoined to an NP. Usually, movement inside the relative clause of an operator that is coindexed with the modified noun, is assumed to turn the relative clause into a predicate.

(72) a.   XP   b.   XP   c.   NP
               / \   / \   / \
            X^0_i β XP_i β NP_i RC
               /\   /\   /\
            …t_i…   …t_i… OP_i … t_i…
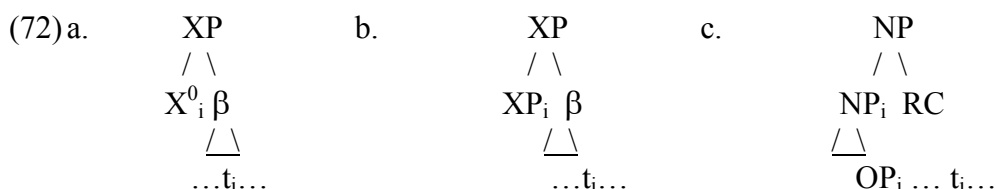
In this chapter, I exploit the similarity between structures like (1b) and (1c) and propose an analysis of free relatives as headed by a copy of the *wh* phrase in the specifier of the relative clause.[1]

The major problem for the analysis of free relatives is that a structure that looks like a relative clause does not have an antecedent. The denotation of *who you just met* in (2a) seems to be comparable to that of *the person who you just met* in (2b). However, while *who you just met* in (2b) has an antecedent, its counterpart in (2a) does not.

(73) a.   I don't like who you just met.

   b.   I don't like the person who you just met.

A great body of work has been devoted to free relatives, and good surveys with comprehensive bibliographies can be found in van Riemsdijk 2000 and Grosu 2002. Various analyses of free relatives have been proposed, two of which I will briefly discuss in this section. The first one builds on the superficial similarity of free relatives to embedded *wh* questions and assume that free relatives should be analysed as bare CPs, like embedded *wh* questions. This type of analysis is illustrated in (3) (cf. Jacobson 1995, Rooryck 1994).

(74) a.   I don't like [$_{CP}$ who you just met *t*]

   b.   I wonder [$_{CP}$ who you just met *t*]

However, as noted above, a free relative receives an interpretation similar to that of a noun phrase, not an interpretation similar to that of a clause. Therefore, if a bare-CP analysis is to be maintained, a semantic rule has to be formulated to ensure that the CP in (3a) is interpreted correctly. This rule would take a question and turn it into a referring expression. Although this analysis would be adequate in that free relatives indeed look like embedded *wh* questions on the surface, a rule that provides the correct interpretation would essentially be non-compositional. This drawback has already been observed by Schachter 1973:36, who argues that a similar proposal

> would, however, be incompatible with most current beliefs about the relation between underlying syntactic structure and meaning. One would be claiming that […] a single underlying structure could be mapped onto either of two semantically distinct surface constructions […] with nothing in the underlying structure to 'trigger' the [right interpretation].

The problem of non-compositionality may be circumvented by enriching semantic theory in other ways, but as it stands the simplest syntactic analysis of free relatives increases the

---

[1] An early incarnation of this idea can be found in Bury 1998.

complexity of the interpretational component.

Moreover it should be noted that such an analysis loses much of its appeal given a range of differences between free relatives and embedded questions (cf. especially Baker 1995: chapter 7). For example, while embedded questions can only occur with appropriate matrix verbs, free relatives can occur in basically all the positions that could be occupied by corresponding noun phrases. This contrast is illustrated in the following examples, where a free relative can be substituted by a noun phrase and a superficially identical embedded question can be substituted by a question introduced by *whether*, while the reverse is not possible (examples from van Riemsdijk 2000):

(75) a.   You should return what you have finished reading to the library.

   b.   You should return *War and Peace* to the library.

   c.   * You should return whether you have finished your book to the library.

(76) a.   Could you tell me what you have finished reading?

   b.   * Could you tell me *War and Peace*?

   c.   Could you tell me whether you have finished your book?

A second type of analysis, the COMP ANALYSIS, builds on the intuitive similarity of (2a) and (2b), and assumes that free relatives have a syntactic structure equivalent to DPs containing a regular relative clause. This requires the postulation of an empty head that functions as the antecedent of the relative clause, as in (6). The literature contains many proposals along these lines (cf. Groos & Van Riemsdijk 1981, Borsley 1984, and Grosu 1994, among others).

(77)   I don't like [$_{DP}$ *e* [$_{CP}$ who you just met *t*]]

Obviously, this analysis does not require the extension of the semantic component that is necessary if free relatives are bare CPs. It does lead to a complication of the syntax, however. The licensing of the null head in (6) forces modifications of the theory of empty categories. English is not a *pro*-drop language, and therefore some explanation is needed for the availability of an empty-headed DP in free relatives. Furthermore, free relatives are not like regular relatives in all respects. While in English regular relatives the CP-level

does not have to contain any phonetically realised material (or may be absent altogether, cf. chapter 2), free relatives always must be introduced by a *wh*-phrase (cf. 7). This would follow naturally from the bare-CP analysis, because *wh* questions cannot be introduced by null operators either. The null-head analysis, on the other hand, does not provide an immediate explanation.

(78) a.   I don't like [DP the person [CP 0 you just met *t*]]

b.   *I don't like [DP *e* [CP 0 you just met *t*]]

Thus, the simplest analysis from a semantic point of view leads to a complication of the syntactic component, while the simplest analysis from a syntactic point of view leads to complications of the semantic component. The analysis I propose in this chapter aims to combine the syntactic simplicity of the bare-CP analysis with the semantic simplicity of the null-head analysis.

The next section introduces the analysis, which involves two copies of the *wh* operator, one in the position of the head of the construction and one in the highest specifier of the relative clause. Section 3 deals with categorial matching effects. Section 4 addresses the question of what properties allow a *wh* word to introduce free relatives. Section 5 deals with case matching effects. Section 6 is concerned with the question of which of the two copies of the operator is pronounced. Section 7 is the conclusion.

2 The proposal

The reasoning behind the analysis of free relatives proposed below is the same as that behind the replacement of an analysis of verb movement in terms of adjunction to uninterpretable empty functional heads by an analysis in terms of s-movement structures, that is structures that contain an additional head position occupied by a copy of the verb.[2] In Chomsky 1995c and related work, the "target projects" condition stipulates that every movement must target a position that exists before the movement takes place (see chapter 1 for more detailed discussion). A consequence of this view is that some verb movements must be analysed in terms of adjunction to otherwise unmotivated heads. Thus, the "target

---

[2] The discussion here begins in terms of conventional X-bar structures, but returns to the framework developed in chapter 1 before the end of this section.

projects" condition forces a complication of the lexicon, which must on this view be allowed to contain elements that have no phonological and no semantic content. In contrast, the rejection of the "target projects" condition entails that no such radically empty heads are necessary. A moved verb can simply occur as the head of the verb movement structure.

A similar situation holds with regard to the Comp analysis of free relatives illustrated in (6). On standard assumptions, the only way free relatives can be analysed structurally on a par with nouns modified by a relative clause is if some empty element is assumed to occur in the position that corresponds to the modified noun in a regular relative clause construction. As already noted above, the occurrence of such an empty category is somewhat surprising in a language like English which does not usually freely allow for empty nouns in argument positions. Thus, the Comp analysis is forced to either assume a new type of empty category or to modify the licensing conditions of existing empty categories specifically to allow for the presence of an empty head in free relatives. While this is conceptually not very attractive, the discussion in this chapter indicates that the assumption of such a category makes it difficult to account for a range of empirical properties in an insightful way.

The rejection of the "target projects" condition makes available a different approach to the analysis of free relatives that can parallel the structure of a noun modified by a relative clause without increasing the inventory of empty categories or complicating the licensing conditions of existing empty categories. In particular, I propose that a free relative construction is headed by the *wh* operator and that a copy of the operator occurs in the highest specifier position in the relative clause. An operator movement-chain inside the relative clause turns the clause into a predicate, and the relative-clause external copy of the operator provides the antecedent for the clause.[3] The derivation of such a structure is illustrated by the X-bar representations in (8). First, a relative clause is formed by movement of *what* to the highest specifier inside the relative clause (cf. 8a). This relative clause needs an antecedent, which is provided by the copy of *what* that occurs outside the relative clause. The category of the resulting structure is that of the *wh* phrase, not that of

---

[3] Note that the relation between the copies of the operator in the head position and inside the relative clause is not a movement-chain, but that it is similar to the relation between the modified noun and the relative operator in a headed relative construction like (2a).

the relative clause.[4] The consequence is that the relative clause ends up as an adjunct to the external copy of the operator. The resulting structure in (8b) is similar to that of a headed relative construction like (2b). This structure of course can then become part of a more complex structure as in (8c).[5] (Multiple copies here and throughout this chapter are marked with angled brackets < >.)

(79) a. $[_{CP} [_{DP} <who>]$ you just met $<who>]$

b. $[_{DP} [_{DP} <who>] [_{CP}<who>$ you just met $<who>]]$

c. I don't like $[_{DP} [_{DP} <who>] [_{CP} <who>$ you just met $<who>]]$

This analysis avoids the problems raised by the analyses discussed earlier. In contrast to the Comp analysis, it does not rely on the postulation of a novel empty category (or a complication of the licensing conditions of existing empty categories). In contrast to the bare-CP analysis, it does not rely on a complication of the rules of semantic interpretation, and no ad hoc rule is required to assign a free relative to the correct semantic type. On this view, the major difference between a free relative and a headed relative construction is the target to which the relative clause is adjoined. While the target of adjunction in a free relative is the copy of an element that also occurs within the relative clause, the target of adjunction in a headed relative clause structure occurs only external to the relative clause.

To see how a structure like (8b) is interpreted, recall the proposal of Chomsky (1993:203f) that movement is a combination of a copy operation and one or more deletions. (I will explain below in what way these ideas are compatible with the model of phrase structure developed in the previous chapters.) Such deletions may take place at both LF and PF. The syntactic structure in (9a) results from copying of the *wh* phrase *which painter*. At PF the lower copy is deleted. In order to obtain a well-formed semantic representation, the operator part of the *wh* expression must delete in the lower copy. In addition, the restriction *painter* is deleted in either the higher or the lower copy.

---

[4] In most cases, a structure in which a moved specifier occurs as the head is ruled out on independent grounds. For example, an unaccusative object moved to subject position can not be the head of the structure because this would mean that the specifier-head configuration required for nominative case assignment is not available.

[5] The presence of copies of the *wh* phrase inside and outside the relative clause in this analysis is somewhat reminiscent of proposals by Schachter 1973:30, Kayne 1994:125, Donati 1997:150. However, in these

(80) a. syntax: [<which painter>] do you like [<which painter>] most

    b. PF: [<which painter>] do you like [<~~which painter~~>] most

    c. LF: [<which ~~painter~~>] do you like [<~~which~~ painter>] most

A similar idea can be used to account for the interpretation of free relatives. It is clear that even syntactically atomic *wh* expressions may have a more complex semantic structure. In particular, a *wh* phrase may contain an operator (op) as well a restriction (res). Thus, the word *who* is composed of a *wh* operator and the restriction [+human]. In these terms, the structure of a free relative like (8a) should be represented as in (10a). As in the *wh* question in (9), a number of deletions must take place here to get a well formed semantic representation. The copy in specifier of C must be interpreted as an operator, and therefore its operator part must be maintained, whereas the operator part of the other copies must delete. The copy that heads the construction cannot function as the antecedent of the operator in spec-CP if it lacks semantic content. Hence, the restriction of the *wh* phrase cannot be deleted in this copy. The restriction must then be deleted in the other copies. It follows that the lowest copy is just a variable:

(81) a. syntax: [$_{DP}$ [$_{DP}$ <op res>] [$_{CP}$ [$_{DP}$ <op res>] you just met [$_{DP}$ <op res>]]]

    b. LF: [$_{DP}$ [$_{DP}$ <~~op~~ res>] [$_{CP}$ [$_{DP}$ <op ~~res~~>] you just met [$_{DP}$ <~~op res~~>]]]

This is a good point to return to the theory of phrase structure proposed in the previous chapters, because this approach to the interpretation of free relatives converges in an interesting way with that theory.[6] As discussed in some detail in chapter 1, the copy theory of (specifier) movement is not compatible with the assumptions about phrase structure made in this dissertation. To see how much of the analysis in terms of the copy theory of movement in (10) carries over to the present framework, consider first the translation of (8b) into a simplified Telescope representation without categorial projection:

---

proposals the *wh* phrase moves to an existing position, a determiner or a complementizer. In contrast, here the CP-external position is the result of the copying of the *wh* phrase.

[6] To keep the presentation as clear as possible and to make it easier to see what is really important in a particular structure, I use some strongly simplified structures throughout this chapter. Thus, subject traces and adverbs are left out wherever they are not relevant to the argument, and often I use labelled brackets rather than phrase structure sets or Telescope tree diagrams.

(82)   who   cf. [$_{DP}$ <who> [$_{CP}$ <who> C [$_{TP}$ you T [$_{VP}$ met <who>]]]]
            |
            C
           / \
        who   T
              / \
          you   met θ$_{\#}$
                / \
             who   met θ
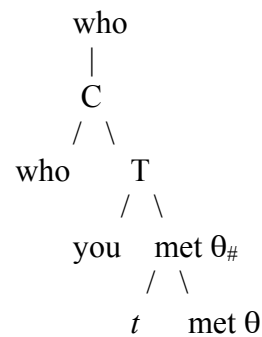
A convergence of this structure to the set notation that fully represents phrase structure highlights the familiar problems (see chapter 1 for detailed discussion of the theory). The two occurrences of *who* that occupy the specifiers of C and T, respectively, in the tree in (11) are both represented by a treelet <who, {who}> in the phrase structure set in (12a).[7] Of course, the principles of set theory entail that an unordered set S_1 containing two identical members M is identical to a set S_2 that differs from S_1 only in containing M only once, i.e. {M, M} = {M} (cf. the principle of set identity of chapter 1). First, this means that the multiple occurrences of *who* in the dominance sets of various treelets (given in bold face) in (12a) all collapse into a single occurrence of *who* for each dominance set. Second, this means that the double occurrence of the treelet <who,{who}> (underlined) in the phrase structure set in (12a) collapses into only a single occurrence. (12a) is then identical to (12b), which can be informally represented by the tree in (12b').

(83) a.   {<who,{**who, who**, C, T, you, met θ$_{\#}$, **who**, met θ}>, <C,{C, **who**, T, you, met θ$_{\#}$,
      who, met θ}>, <who,{who}>, <T,{T, you, met θ$_{\#}$, who, met θ}, {you,{you}>, <met
      θ$_{\#}$, {you, met θ$_{\#}$, who, met θ}>, <who,{who}>, <met θ,{met θ}>}

     b.   {<who,{who, C, T, you, met θ$_{\#}$, met θ}>, <C,{C, who, T, you, met θ$_{\#}$, met θ}>,
      <who,{who}>, <T,{T, you, met θ$_{\#}$, who, met θ}>, <you,{you}>, <met θ$_{\#}$, {you, met
      θ$_{\#}$, who, met θ}>, <met θ,{met θ}>}

b'.
```
            who
             |
             C
             |
             T
            / \
        you   met θ#
             / \
          who   met θ
```

The collapse of the two treelets <who, {who}> in (12a) to a single one in (12b) is the reason why specifier movement cannot be expressed in terms of multiple copies in the present framework. Thus, in order to be able to represent the chain linking the specifier of C to the specifier of met θ#, the lower copy of *who* in (11) must be represented by a trace (or some other, non-identical element; see chapter 1 for discussion). This is done in (13), which can be informally represented by the tree in (13').

(84) {<who,{who, C, T, you, met θ#, *t*, met θ}>, <C,{C, who, T, you, met θ#, *t*, met θ}>,

   <who,{who}>, <T,{T, you, met θ#, who, *t*, met θ>, <you,{you}>, <met θ#, {you, met

   θ#, who, *t*, met θ}>, < *t*,{ *t*}>, <met θ,{met θ}>}

(13')
```
            who                        cf. [DP <who> [CP <who> C [TP you T [VP met t]]]]
             |
             C
            / \
        who    T
              / \
          you   met θ#
              / \
            t    met θ
```

The impossibility of representing the lower position of the *who*-chain with an identical copy of *who* entails one of the conclusions about the interpretation of free relatives above. Thus, as discussed in the relation to (10), none of the categorial properties of the *wh* phrase are interpreted in the lowest member of the *wh* chain. The restriction [+human] is interpreted in the head position of the construction, while the operator part is interpreted in

---

[7] Recall that the first element in a treelet is the head category of the treelet and the second, the dominance set,

the specifier of C, which means that the lowest position can only be interpreted as a variable. This result is expected if the theory of phrase structure makes it impossible to represent the lowest position by a copy of the *wh* phrase.

It is striking that in contrast to the treelets corresponding to the two copies of *who* internal to the relative clause in (11), the treelet corresponding to the copy of *who* at the root of the structure does not collapse. This is so because the root occurrence of *who* dominates material not dominated by the lower occurrences of *who*, which means that the dominance set of the root *who* is distinct from the dominance sets of the lower *who*s. Since the dominance sets of the treelets are distinct, their categories can be identical without a collapse of the treelets. This approach to phrase structure is then compatible with the proposal above that a free relative is headed by a copy of the *wh* phrase that occupies the specifier of C inside the relative clause.

Note that the structure in (13) bears some similarity to the s-movement structures and adjunction structures proposed in earlier chapters (see e.g. chapter 1: section 5). In all of these structures, a derived head position is created with the result that the head category of a treelet dominates the identical head category of a distinct treelet. In the earlier structures, this configuration triggers the formation of a chain that contains the two categories in order to ensure compatibility with the sequence that restricts the order of functional categories within each clause (cf. Cinque 1999, Starke 2001:chapter 9). In (13), however, the two categories are not part of the same extended projection. On the assumption that extended projections delimit the scope of the functional sequence, no chain needs to be formed to make this structure compatible with the functional sequence. Thus, unlike the two identical categories in s-movement and adjunction structures, the identical categories that occur in the head position of the free relative structure and in the specifier of the relative clause do not form a chain. Instead, the relation is of the same type as the relation in a headed relative construction between a modified noun and the relative operator in the specifier of the relative clause.

The remainder of this chapter discusses some of the predictions that this analysis makes. The first prediction, discussed in the following section, is that the category of a free relative is identical to that of the *wh* phrase inside the relative clause because this phrase is

---

states which categories the head category dominates.

also the head of the free relative construction. The next prediction, discussed in section 4, concerns the content of *wh* phrases that can occur in free relatives. Since the restriction of a *wh* phrase is what allows it to appear as the antecedent of the relative clause, *wh* phrases that lack a restriction are not possible in free relatives. Since the *wh* head of a free relative is a copy of a *wh* phrase that is in a chain with the trace, the two *wh* phrases share the same grammatical features. Therefore, selectional requirements that refer to these features in the relative clause must be compatible with selectional requirements in the matrix clause. This is discussed in section 5.

3 Categorial matching

The analysis proposed in the previous section is similar in spirit to the HEAD ANALYSIS of Bresnan and Grimshaw 1978 (see also Larson 1987, Citko 2002). Bresnan and Grimshaw argue that in a free relative construction the *wh* phrase functions as the head, and that the relative clause is adjoined to its projection. The dependency between the *wh* phrase and the relative clause in encoded through a coindexed null pronominal. An updated version of Bresnan and Grimshaw's analysis would presumably assume null operator movement (cf. Roberts 1997):

(85)   I'll buy [$_D$ what [$_C$ OP you are selling *t*]]

The main difference between this analysis and the one proposed here is that, in (14), the specifier of C is occupied by an element distinct from the *wh* phrase in the head position. In contrast, in the present proposal (cf.(13)), both positions are occupied by an identical element, namely copies of the *wh* phrase:

(86)   I'll buy [$_D$ <what> [$_C$ <what> you are selling *t*]]

Both analyses provide a natural account for the matching effects observed by Bresnan and Grimshaw, in particular the observation that the category of a free relative is identical to that of the *wh* phrase. In any analysis that takes the *wh* phrase to be the head of the construction, this observation is a straightforward result.

The categorial matching effect is illustrated above and in Bresnan and Grimshaw's

examples below. The free relative construction is of category D if the *wh* phrase is of category D (cf. (15)); it is an adjective if the *wh* phrase is an adjective (cf. 16a), and it is an adverb the if the *wh* phrase is an adverbial (16b,c).

(87) a.   John will be [$_A$ [$_A$ <however tall>] [$_C$ <however tall> his father was *t*]]

b.   I'll word my letter [$_{Adv}$ [$_{Adv}$ <however>] [$_C$ <however> you word yours *t*]]

c.   I'll put my books [$_{Adv}$ [$_{Adv}$ <wherever>] [$_C$ <wherever> you put yours *t*]]

A similar account can be given for the case of 'missing prepositions' in certain free relatives. As discussed by Larson (1987), the lower preposition in (17) is realised optionally.

(88)   I'll talk to whoever you talk (to)

One way to form a free relative is to front the *wh* phrase, stranding the preposition. This results in a free relative of category D, which can in turn combine with a preposition, as in (18a). The second option is to pied-pipe the preposition along with *wh* phrase. This results in a free relative of the category P. Such a free relative can occur with a verb that selects a prepositional object, as in (18b).

(89) a.   I'll talk to [$_D$ [$_D$ <whoever>] [$_C$ <whoever> you talk to *t*]]

b.   I'll talk [$_P$ [$_P$ <to whoever>] [$_C$ <to whoever> you talk *t*]]

While these data follow straightforwardly under the analysis presented here, they require additional assumptions under the Comp analysis, where free relatives are assumed to be headed by an empty nominal category. Unlike free relatives, headed relatives do not display matching effects, that is the relative operator and its antecedent need not be of the same category. In (19a), for example, the antecedent is a noun, while the relative operator is prepositional. In contrast, a free relative with the same noun-preposition combination is not possible. In (19b), the relative clause introduced by *to whom* cannot combine with a nominal empty category.

(90) a.   [$_D$ the [$_N$ [$_N$ man] [$_C$ [$_P$ to whom] you talked yesterday *t*]]]

b.   * I talk to [$_D$ [$_D$ e] [$_C$ [$_P$ to whom] you talked yesterday *t*]]

Of course, the ungrammaticality of (19b) could be explained by assuming that the empty head inherits its categorial features from the relative operator (cf. Grosu 1996). But this assumption considerably weakens the predictions made by the null-head analysis. It would seem that the presence of a null head only has empirical effects if its content differs from its licenser, especially if the two are adjacent. In contrast, if the properties of the null head and the *wh* phrase are always the same, it will be very difficult to find independent support for the presence of the null head.

In the bare-CP analysis a similar problem arises, as it is impossible for the specifier of C to transfer its categorial features to the clausal C level. The only way to explain matching effects, therefore, would be to allow for downward selection. In (20), for example, one must allow the prepositional phrase in the specifier of C to satisfy the matrix verb's selectional requirements. Such downward selection is not attested elsewhere.

(91)   I'll talk [$_C$ [$_P$ to whoever] you talk *t*]]

In conclusion, categorial matching effects are most naturally explained if free relatives are headed by the *wh* phrase. There are two types of such an analysis. In the first one, proposed by Bresnan and Grimshaw 1978, the *wh* phrase occurs only outside the relative clause in the position that heads the free relative construction. In the alternative, proposed here, the *wh* phrase occurs both inside the relative clause and in the head position. There are two considerations that support the latter type of analysis. First, given the need for an operator chain within the relative clause, it is difficult to understand on a Bresnan and Grimshaw-type analysis why only null operators can occur occupy the specifier of C position in free relatives. Thus, it is not clear what rules out an example like the following, where *which* occurs inside the relative clause:

(92)   *I'll buy [$_D$ [$_D$ what] [$_C$ which you are selling *t*]]

The ungrammaticality of (21) follows from the present analysis, because the *wh* phrase that heads the free relative is a copy of the *wh* phrase in the specifier of C.[8]

---

[8] The fact that only one of the two copies of the *wh* phrase is pronounced should follow from general constraints on the spell-out of multiply occurring elements. Thus, it is often the case that languages exhibit a resistance against accidental repetition of morphemes, a phenomenon referred to as haplology (cf. Neeleman

A second problem for the Bresnan and Grimshaw analysis is the fact that the element that heads a free relative must be a *wh* phrase, and not for example a pronoun:

(93)   *I'll buy [D [D it] [C OP you are selling *t*]]

This also follows from the movement analysis, since a pronoun lacks the operator properties required for the formation of a chain internal to the relative clause.

4 The content of the relative operator

The second prediction that follows from the analysis developed in section 2 is that whatever operator heads a free relative must have a restriction. The following example illustrates the LF representation of the free relative in *I don't like who you just met*. The structure is well formed, because the restriction [+human] of the *wh* phrase means that the copy of *who* that heads the structure has semantic content. Consequently, this copy can function as the antecedent of the operator in specifier of C.

(94)   [D [D <~~op~~ res>] [C [D <op ~~res~~>] you just met *t*]]

What happens if the relative operator does not contain a restriction? This situation is illustrated in the following structure:

(95)   * [D [D <~~op~~>] [C [D <op>] you just met *t*]]

Here, the operator part of the *wh* phrase is interpreted in the specifier of C, as usual. However since there is no restriction, the clause-external copy of the *wh* phrase does not have any content. This implies that the structure contains a constituent [D ~~op~~] that does not have a semantic head, which in itself is probably sufficient to rule it out. Moreover this means but that the operator in the specifier of C does not have an antecedent. This suggests that only operators with a restriction can introduce free relatives. This section presents a range of data from both English and other languages that support this prediction.

Regular relative clauses in English can be introduced by a number of elements. A first

---

and van de Koot 2001). Whatever triggers haplology could presumably block the pronunciation of the two

option, illustrated in (25), is a null operator together with the complementizer *that*:

(96)   I don't like [the man [OP that you just met *t*]]

It is clear that null operators do not have a restriction. Relatives introduced by a null operator can take antecedents of any type. So, alongside examples like (25), where the antecedent is animate and human, there are examples like (26), where it is not.

(97)   I don't like [the book [OP that you just bought *t*]]

If the null operator does not contain a restriction, it is predicted that null operators cannot introduce free relatives in English. The following examples show that this prediction is correct. Regardless of whether the context is compatible with a [+human] or a [-human] constituent, a free relative headed by a null operator is not grammatical:

(98) a.   * I don't like [<OP> [<OP> that you just met *t*]]
     b.   * I don't like [<OP> [<OP> that you just bought *t*]]

A different picture is found with *wh* operators like *who* or *what*. *Who* or *what* clearly contain a restriction, since in questions *who* asks for a human and *what* for a non-human entity. Thus, the indicated discourses are deviant.

(99) a.   Speaker A: Who did you see?
            Speaker B: John / #a book
     b.   Speaker A: What did you see?
            Speaker B: #John / a book

The same restrictions can be observed in regular relative clauses. Although in Standard English *what* cannot be used as an operator in such contexts, *who* can only have a human antecedent:

(100)   a.   I don't like [the man [who you just saw *t*]]
         b.   *I don't like [the book [who you just saw *t*]]

adjacent operators in free relatives.

Given that *what* and *who* have restrictions, the analysis developed here correctly predicts that they can introduce free relatives:

(101)   a.   I don't like [<who> [<s̶w̶h̶o̶> you just saw *t*]]

   b.   I don't like [<what> [<s̶w̶h̶a̶t̶> you just saw *t*]]

These free relatives are interpreted in line with the properties of the *wh* operators that head them. Thus, the free relative in (30a) is assigned the LF representation in (31a), and can only refer to a human. The free relative in (30b) is assigned the LF representation in (31b), and can only refer to a non-human. Consequently, (30a) cannot mean 'I don't like the thing that you just saw', and (30b) cannot mean 'I don't like the person you just saw'.

(102)   a.   [$_D$ [$_D$ <s̶o̶p̶ human>] [[$_D$ <op h̶u̶m̶a̶n̶>] you just saw *t*]]

   b.   [$_D$ [$_D$ <s̶o̶p̶ nonhuman>] [[$_D$ <op n̶o̶n̶h̶u̶m̶a̶n̶>] you just saw *t*]]

Before discussing *which*, the third *wh* operator that can be used to form relative constructions in English, consider first operators of free relatives in other languages.

In Dutch, headed relative clauses are formed with the relative operators *die* and *dat*.[9] These operators have different grammatical genders. *Dat* is neuter and *die* is nonneuter. This grammatical gender dichotomy does not correspond to natural gender, however. Thus, nouns denoting things often combine with *die* in relative constructions, and nouns denoting persons often combine with *dat*, as illustrated by the examples in (32a,b). Examples comparable to (32b) are formed productively by the diminutive suffix *-je* (cf. 32c):

(103)   a.   Dat is de lamp die   we toen gekocht hebben.
             that is the lamp that-NONNEUT we then bought have
             'That is the lamp that we bought then.'

   b.   Zij is dat wijf   dat   me het leven zo   zuur gemaakt heeft.
         she is that woman that-NEUT me the   life so   sour made   has
         'She is the woman that spoiled my life.'

---

    c.   Hij is het mannetje dat   we toen ontmoet   have.

        he is the man-dim   that-NEUT   we then met   have

        'He is the little man we met that time.'

Given the lack of a semantic restriction in *die* and *dat*, the analysis predicts that they cannot be used to form free relatives. The examples in (33) show that this is indeed the case:

(104)   a. *Ik vind die   jij   ontmoet   hebt niet   aardig

          I find that-NONNEUT   you met   have not   nice

    b. *Ik vind dat jij   gekookt hebt   erg lekker

       I find that you cooked have very nice

Instead, free relatives are introduced by *wh* operators like *wie* 'who' and *wat* 'what'. This is illustrated in (34).

(105)   a.   Ik vind wie jij   ontmoet hebt niet aardig.

        I find   who you met   have not nice

        'I don't like who you met.'

    b.   Ik vind wat jij gekookt hebt erg lekker.

        I find what you cooked have very nice

        'I really like what you cooked.'

As expected, these *wh* operators differ from the relative operators *die* and *dat* in having a semantic restriction. Like their English counterparts *who* and *what*, *wie* and *wat* ask for a human and nonhuman entity respectively:

(106)   a. Speaker A: Wie heb je gezien?

                    who have you seen

        Speaker B: Jan / #een boek

                John / a book

b. Speaker A: Wat heb je gezien?

                 what have you seen

    Speaker B: #Jan / een boek

                 John / a book

Thus, the Dutch data further corroborate the prediction that operators in free relative constructions must contain a semantic restriction.

Similar arguments can be based on dialectal and Standard German. Regular relative clauses in Standard German are introduced by *der*, *die* or *das* (or paradigmatically related forms). As in Dutch, these operators express grammatical, but not natural gender:

(107)   a.   Sie ist der Weltstar den   ich gestern getroffen   habe.

             she is the-MASC star that-MASC I yesterday met   have

             'She is the star that I met yesterday.'

        b.   Er ist die   Mannsperson die   ich gestern getroffen habe.

             he is the-FEM man   that-FEM I   yesterday met   have

             'He is the guy that I met yesterday.'

        c.   Sie ist das   Weib   das   ich   gestern getroffen habe.

             she is the-neut woman   that-neut I   yesterday met   have

             'She is the woman that I met yesterday.'

It is therefore expected that, as in Dutch, free relatives cannot be introduced by the regular relative operators (cf. (37)). Instead, they are formed with *wh* operators like *wer* 'who' and *was* 'what' (cf. (38)), which are like their Dutch and English counterparts in having a semantic restriction.

(108)   a.   *Ich finde   den/die   du getroffen hast nicht nett

             I find   that-MASC/FEM   you met   have not nice

        b.   *Ich finde das du gekocht hast sehr lecker

             I find that-NEUT you cooked have very tasty

(109)   a.   Ich finde wen du getroffen hast nicht nett.

             I find who you met have not nice

             'I don't like who you just met.'

       b.   Ich finde was du gekocht hast sehr lecker.

             I find what you cooked have very nice

             'I find what you cooked very tasty.'

The same pattern is found in the Alemannic dialect spoken in the Breisgau region in southern Germany. In this language, headed relative clauses are consistently formed with the operator *wo*, irrespective of the properties of the head noun. This is illustrated in (37).

(110)   a.   Des isch dr   Kerli wo ich geschtert troffe   hab.

             this is the-MASC   guy that I   yesterday met   have

             'This is the guy I met yesterday.'

       b.   Des isch d'Frau   wo ich geschtert troffe hab.

             this is   the-FEM woman   that I yesterday   met have

             'This the woman I met yesterday.'

       c.   Sel isch 's   Wasser wo mã Schnapps drüs   macht.

             this is the-NEUT water that one Schnapps thereof makes

             'That's the water that is used for making schnapps.'

These examples clearly suggest that the relative operator *wo* does not have a semantic restriction. The prediction, then, is that *wo* cannot be used to form free relatives. This prediction is borne out by the data below.

(111)   a.   *Ich find wo dü troffe hesch nit nett

             I find that you met have not nice

       b.   *Ich find wo dü kocht hesch arg gued

             I find that you cooked have very nice

As in Standard German, free relatives in Alemannic are introduced by the *wh* operators *wer* 'who' and *was* 'what'. These do have a semantic restriction, and consequently they can be used to generate free relatives with well-formed LF

representations. Thus, the following examples are grammatical:

(112) a. Ich find wer dü troffe hesch nit nett.
            I find who you met have not nice
            'I don't like who you met.'
       b. Ich find was dü kocht hesch arg gued.
            I find what you cooked have very nice
            'I find what you cooked very nice.'

After this detour to Dutch and German, consider *which*, the remaining *wh* operator involved in the formation of English free relatives. Again, the question is whether *which* contains a restriction. In questions, this does not seem to be the case, as the data below indicate. (42a) asks for a person, and (42b) asks for a thing. Thus, *which* in questions is compatible with both human and nonhuman interpretations.

(113) a. Which is her favourite quarterback, John or Bill?
       b. Which is her favourite book, *The Unbearable Lightness of Being* or *The Trial*?

Further evidence for the underspecified nature of *which* comes from the fact that it can take a complement that provides it with a semantic restriction. This complement can essentially be any NP, irrespective of its semantic properties, as (43) illustrates.

(114)   Which woman/man/lamp did you see *t*?

Note that such *which*+complement questions receive a different interpretation from parallel *what*+complement questions, as illustrated in the following discourses, (44) shows that *which man did you see?* asks for a particular man and not a type of man. (45) shows that *what man do you plan to marry?* does not ask for a particular man, but for a kind of man.

(115) Speaker A: Which woman do you plan to marry?
       Speaker B: Mary / #One with lots of money

(116)   Speaker A: What man do you plan to marry?

         Speaker B: #John / One with lots of money

This contrast may be due to the fact that *which* lacks a restriction while *what* does not. In the absence of a lexical restriction the complement of *which, woman* in (44), may be interpreted as the restriction of *which*. As argued above, there is clear evidence that *what* contains the restriction *nonhuman*. Since *what* already contains a restriction its complement cannot be interpreted as its restriction, and this may give rise to the 'kind of' interpretation of the question in (45).

   Since it seems reasonably clear that *which* lacks a restriction, a somewhat surprising pattern can be found in regular relatives introduced by *which*. Here, there seems to be a preference for *which* to combine with nouns referring to nonhuman entities. Thus, (46a), where a relative introduced by *which* modifies a human noun, is only marginally accepted. In contrast, (46b), where a *which*-relative clause modifies a nonhuman noun, is fine.

(117)   a.   %This is the quarterback which Mary likes best.
         b.   This is the book which Mary likes best.

How can the marginality of (46a) be explained? One possibility would be attribute it to the elsewhere principle, according to which more specific forms block less specific forms in contexts where both could be inserted. In relative constructions headed by a [human] nominal, both *who* and *which* can in principle be used to introduce the relative clause. Since *who* contains a restriction, it is more specific than *which*, and consequently insertion of *which* is blocked by the elsewhere principle. If this line of argumentation is adopted, it must be assumed that insertion of null operators is not in competition with insertion of *wh* operators, because otherwise (47a) would be blocked by (47b).

(118)   a.   The man OP that I saw *t*
      b.   The man who I saw *t*

This seems reasonable, as the elsewhere principle only concerns forms which share relevant features (such wh, comparative, tense, etc.). The defining characteristic of null operators is their lack of such features, and hence insertion of null operators is not in

competition with insertion of wh expressions. Obviously, there is more to be said about competition in relative clauses, but that will have to be done elsewhere.

With this background we can now consider the prediction the present theory makes with respect to free relatives introduced by *which*. Recall that the main claim made in this section is that only operators containing a restriction can introduce free relatives, because the restriction is necessary for semantic convergence. This implies that *which* on its own cannot head a free relative, but *which* accompanied by a syntactic restriction can. Indeed, the contrast in (48) bears out this prediction.[10]

(119)   a.   *I'll read whichever you suggest

       b.   I'll read whichever book you suggest.

(48a) is assigned the LF representation in (49a). This representation is ill-formed as the operator in spec-CP does not find an antecedent. The LF representation assigned to (48b), on the other hand, is well-formed. Here, a copy of the *wh* phrase *whichever book* occurs as the head of the free relative and the familiar deletions leave the restriction [book] in the head position as a possible antecedent for the operator in the specifier of C.

(120)   a.   *[$_D$ [$_P$ ~~whichever~~] [$_C$ [$_D$ <whichever>] you suggest [$_D$ t]]]

       b.   [$_D$ [$_D$ ~~whichever~~ book>] [$_C$ [$_D$ <whichever ~~book~~>] you suggest [$_D$ t]]] LF

Further confirmation for the requirement that a restriction must be present in free relatives comes from cases where *which* takes a pronominal complement. An example of this is given in (50). Here, *one* is interpreted as book on the basis of the preceding discourse.

(121)   Given the choice between these books, I'll read whichever one my lecturer advises.

---

[10] The example in (48a) should be interpreted as not embedded in a larger discourse; such embedding will be discussed below.

For many speakers, the pronominal element can remain silent. In (51) *which* seems to be used without a restriction. However, those speakers that accept free relatives of this type only accept them if they are placed in a context that provides a restriction for *which*. There is a sharp contrast between (51) and the earlier example in (48a), where no context is assumed.

(122)   Given the choice between these books, I'll read whichever *pro* my lecturer advises.

The conclusion then is that *which* in (51) can introduce a free relative because it is accompanied by a pronominal element that allows it to obtain a semantic restriction. Thus, the contrast between (50) and (51) supports the claim tested in this section.

5 Case matching

Section 3 dealt with the observation that the category of a free relative matches the category of the operator introducing it. So, if an operator of category D moves inside the relative clause, as in (52a), the free relative will be of category D as well. This implies that the selectional requirements of the embedded verb and the matrix verb must be compatible. It cannot be the case that one of the two selects a preposition and the other selects a determiner, as the contrast in (52b,c) shows.

(123)   a.   I don't like [$_D$ [$_D$ <who>] [$_C$ <who> you just met *t*]]

   b.   *I don't like [$_P$ [$_P$ <to whom>] [$_C$ <who> you just talked *t*]]

   c.   I don't like [$_D$ [$_D$ <who>] [$_C$ <who> you just talked to *t*]]

Given the analysis proposed in this chapter, it is expected that matching effects are not restricted to categorial matching. In particular, they should extend to case in languages in which case is realised morphologically. This can be illustrated using German examples. Consider the free relative in (53a). The moved operator must have accusative case features in order to satisfy the selectional requirements of *mitbringen* 'to bring along'. This implies that the entire free relative will have accusative case as well, because it is headed by a copy of the operator that sits in an accusative marked chain inside the relative clause. The prediction, then, is that the free relative can only appear in accusative positions. This

explains the contrast between (53b), where the matrix verb selects accusative, and (53c), where it selects dative.

(124) a.   [D [D <wen>] [C <wen>   du *t* mitgebracht hast]]

who-ACC who-ACC   you along-brought have

'who you brought along'

   b.   Ich hasse [D [D <wen>] [C <wen> du   *t* mitgebracht hast]]

I hate   who-ACC who-ACC you   along-brought have

'I hate who you brought along.'

   c.   *Ich half [D [D <wen>] [C <wen> du   *t* mitgebracht hast]]

I helped   who-ACC who-ACC you   along-brought have

'I helped who you brought along.'

The free relative in (54a) has nominative case because it is the copy of a nominative operator. Consequently, it can be inserted in (54c), but not (54b).

(125) a.   [D [D <wer>] [C <wer>   *t*   dich   mitgebracht hat]]

who-NOM   who-NOM   you-ACC along-brought has

'who brought you along'

   b.   *Ich hasse [D [D <wer>] [C <wer> *t*   dich   mitgebracht hat]]

 I hate   who-NOM who-NOM   you-ACC along-brought has

'I hate who brought you along.'

   c.   [D [D <wer>] [C <wer>   *t*   dich   mitgebracht hat]]   ist ein netter Kerl

Who-NOM   who-NOM   you-ACC along-brought has   is a   nice   guy

'Who brought you along is a nice guy.'

Like the categorial matching effects discussed in section 3, these case matching effects support an analysis in which the relative operator functions as the head of the free relative. Although these data are well known (see especially Groos and Van Riemsdijk 1981 and Pittner 1995), their implications are not sufficiently acknowledged. The important point is that in headed relative constructions, no matching effects are observed. In contrast to the examples of non-matching free relatives in (53c) and (54b), the examples of non-matching

headed relatives in (55) are grammatical. Here the noun and the relative operator need not agree in case.

(126)   a.   Ich half [$_D$ dem [$_N$ Mann] [$_C$ den   du *t* mitgebracht hast]]

          I helped the-DAT man   who-ACC you along-brought have

          'I helped the man who brought along.'

    b.   Ich hasse [$_D$ den [$_N$ Mann] [$_C$ der *t* dich   mitgebracht hat]]

          I hate   the-ACC man   who-NOM you-ACC along-brought has

          'I hate the man who brought you along.'

In view of these data, case matching effects are unexpected under a null-head analysis of free relatives. If a regular noun does not have to agree with the operator in the specifier of C, why would a null head have to in the largely parallel structures in (56)?

(127)   a.   *Ich half [$_D$ [$_D$ *e*] [$_C$ wen du mitgebracht hast]]

          Who-ACC who-ACC you who-ACC along-brought have

          'I helped who you brought along.'

    b.   *Ich hasse [$_D$ [$_D$ *e*] [$_C$ wer dich mitgebracht hat]]

          I hate   who-NOM you-ACC along-brought has

          'I hate who brought you along.'

Of course, additional assumptions can be made to capture case matching effects in terms of the Comp analysis. Thus, Groos and van Riemsdijk (1981) argue that the null head is in some sense invisible, so that it is the operator inside the relative clause, rather than the null head that satisfies the verb's selectional requirements. Grosu (1994) attributes case matching effects to the identification of *pro* by the operator inside the relative clause, a position adopted by Pittner (1995) and others. Clearly these proposals are descriptively adequate. However, they result in a situation in which it is hard to think of tests that might confirm or disprove the presence of the null head that they require. The null head does not appear to have any properties that exist independently of the context in which it occurs.

There are circumstances in which apparent case mismatches in free relatives seem to

be allowed. First, languages that lack morphological case also lack case matching effects. So, in both English and Alemannic, the counterpart of (54b) is grammatical:

(128)　　a.　I don't like [$_D$ [$_D$ <who>] [$_C$ <who> you brought along t]]

　　　　b.　Ich kann nit liede [$_D$ [$_D$ <wer>] [$_C$ <wer> dü mitbrocht hesch]]

　　　　　　I can　　not stand　who　　who　　you along-brought have

　　　　　　'I can't stand who you brought with you.'

This could be explained by saying that the features that encode case distinctions are acquired on the basis of morphological contrasts (cf. Pinker 1984). If there is no case morphology, the language lacks the relevant features, and therefore it will not show matching effects (cf. Neeleman and Weerman 1999). Perhaps this approach can be extended to mismatches that occur because of syncretism in the case paradigm (Groos and Van Riemsdijk 1981).

　　Second, even languages that have morphological case allow mismatches in certain constructions. However, as Pittner (1995) argues in great detail these mismatches follow a particular hierarchy, in that marked cases may overrule less marked ones, but not vice versa. Thus, accusative may overrule nominative, and dative may overrule both these cases. In the following example from Pittner, the dative assigned by *verpflichtet* overrules the accusative assigned by *lädt ein* in (58). As (53c) above showed, the reverse is not possible.

(129)　　　　Sie lädt ein [$_D$ [$_D$ <wem>] [$_C$ <wem> sie t zu Dank verpflichtet ist]]

　　　　　　she invites　who　　　　　　who she　to thank　obliged　is

　　　　　　'She invites those that she owes thanks.'

One interpretation of Pittner's case hierarchy would be that the more marked a case is the more features it contains. This means, for example, that if accusative is characterized as [F_1, F_2], dative should be characterized as [F_1, F_2, F_3]. Consequently, dative nouns can appear in accusative positions, as they have all the features required. Accusative nouns cannot appear in dative positions, however, as they lack the feature [F_3]. Usually, economy considerations prevent the merger of categories that have superfluous features, such as that

of a dative noun with a verb selecting accusative. But in contexts like (58) the [F₃] feature is not superfluous, as it is selected by the embedded predicate. From this perspective, the fact that certain languages with morphological case allow mismatches of this type is not too surprising (see Neeleman and Weerman 1999 for further discussion). Crucially, they do not allow mismatches of the type in (53c), and this is what underlies the argument for the analysis involving two identical copies of the relative operator in free relatives.

6 The position of the relative operator

The basic structure for a free relative like that in *I don't like who you just met* is repeated here:

(130) $[_D [_D$ <who>] $[_C [_D$ <who>] you just met $t]]$

Section 4 dealt with the question of how such a structure is interpreted at LF, and in particular with the role of the two copies of the *wh* operator. In this section, I consider the other interface, and in particular the question of which of the two copies of the *wh* operators is pronounced.

Since the two copies of the relative operator are adjacent, it seems difficult to find evidence concerning which one is pronounced. Nevertheless, there are some revealing facts from the extraposition of headed relative clauses that were first discussed by Groos and van Riemsdijk 1981. Headed relative clauses in Dutch and German need not always be adjacent to the head that they modify. The examples in (60a,b) show that they can be extraposed. However extraposition of the head noun along with the relative clause is not possible (cf. (60c)).[11]

(131)   a.   Der Hans hat [_D das Geld [_C das er gestohlen hat]] zurückgegeben

         The Hans has the money   that he   stolen   has   returned

         'Hans returned the money that he has stolen.'

    b.   Der Hans hat [_D das Geld *t*] zurückgegeben [_C das er gestohlen hat]

    c.   * Der Hans hat *t* zurückgegeben [_D das Geld [_C das er gestohlen hat]]

---

[11] Examples comparable to (62c) are grammatical in English: *He returned yesterday the money that he had stolen from the bank*. This difference between English and German is due to the fact that English allows heavy-NP shift rather freely, while German does not. But see below.

Since free relatives have been analysed here in terms of a structure that parallels the structure of headed relative constructions, there are two options. If the pronounced copy of the operator is the one in the head position, it is expected that the sequence *you-just-met* in (59) occurs in the extraposed position, leaving the operator *was* behind in object position. Alternatively, if the copy of the operator in the specifier of C position is pronounced, it is expected that the whole string *who-you-just-met* is extraposed.

If we consider another aspect of the extraposition of headed relatives, we get a clear prediction concerning which of the two possibilities should be realised. As noted in chapter 2, relative clauses can extrapose only when they contain (overt) material at the complementizer level.[12] Thus, (61) shows that a relative clause can occur with or without the complementizer *that* when it is adjacent to the noun it modifies, but when the relative clause is extraposed the complementizer is obligatory. Similarly, in (62) the relative operator *who* is optional where the relative is adjacent to the head noun, but obligatory when the relative clause is extraposed. (The examples are from Doherty 1993:59.)

(132)   a.   A moose (that) Bill shot at appeared.
         b.   A moose appeared that Bill shot at.
         c.   * A moose appeared Bill shot at.

(133)   a.   The man (who) Bill knew arrived yesterday.
         b.   The man arrived yesterday who Bill knew.
         c.   * The man arrived yesterday Bill knew.

If extraposition is only compatible with clauses that contain overt material at the complementizer level, there is a clear prediction for extraposition of free relatives. If the relative operator were pronounced in the head position, this would leave the extraposed clausal constituent without any overt material in its complementizer level. It is then expected that free relatives can only be extraposed if the copy of the operator inside the relative clause is pronounced.

---

[12] See Hawkins 2001 for some relevant discussion from a processing perspective.

This is indeed the pattern that is found. The extraposition pattern of related free relatives is illustrated in the following examples (Groos and van Riemsdijk 1981). In (63a) the free relative is in the canonical object position. (63b) shows that the complete free relative construction can be extraposed in a similar way to a relative clause in a headed relative construction. However, as (63c) shows, extraposition of the sequence *er gestohlen hat*, leaving behind the relative operator *was* is not possible.

(134)　a.　Der Hans hat [$_D$　was er　gestohlen hat] zurückgegeben

　　　　　The Hans has　　what he　stolen　　has returned

　　　　　'Hans returned what he has stolen.'

　　b.　Der Hans hat *t* zurückgegeben [$_D$ was er gestohlen hat]　　　　cf. (62c)

　　c.　* Der Hans hat [$_D$ was *t*] zurückgegeben [er gestohlen hat]　　cf. (62b)

As Groos and van Riemsdijk note this pattern is problematic for a proposal like that of Bresnan and Grimshaw 1978 in which the relative operator does not occur as a constituent of the relative clause.

(135)　　　　[$_D$ was [$_C$ OP er *t* gestohlen hat]]

Given such a structure, it might be argued that extraposition in (63a) is the result of heavy NP shift. While heavy NP shift in German is not as easily available as in English, the following examples show that it does occur:

(136)　a.　Der iranische Präsident Mohammed Chatami dagegen verurteilte

　　　　　the Iranian　president M. C.　　　　　in-contrast condemned

　　　　　scharf die Zerstörung historischer Buddha-Statuen als unislamisch.

　　　　　sharp the destruction historical B. statues　　as　un-islamic

　　　　　'In contrast, the Iranian president Mohammed Chatami sharply comdemned the destruction of the historic Buddha statues as un-islamic.' (*Berliner Z* 6 March 2001)

b. Die USA verurteilten scharf die   von den albanischen Extremisten

the USA condemend sharply the of the   Albanian   extremists

ausgehende   Gewalt.

emerging   violence (*Tagesspiegel* 25.03.2001)

'The USA condemned sharply the violence caused by Alabanian extremists.'

c. An der Haltestelle Lothringer   Strasse sind eben eingestiegen   in die

at the   stop   Lothringer   Street were   just   entered in the

4   vier Leute, zwei ältliche Frauen, ein bekümmerter einfacher

"4" four people two elderly women   a   worried   simple

Mann und ein Junge mit einer Mütze und Ohrenklappe.

Man and a boy   witha cap and ear muff (Alfred Döblin, *Berlin Alexanderplatz* p.42)

'At the stop *Lothringer Street*, four people entered the tram on Line 4: two elderly

women, a worried simple man, and a boy with a cap and ear muffs.'

d. Am 25. September verschied plötzlich   an einem Herzschlag mein

on-the 25th September passed-away suddenly   at a   heart-attack my

inniggeliebter Mann, unser lieber Vater, Sohn, Bruder, Schwager   und

dearly-loved man   our   dear father   son brother   brother-in-law   and

Onkel Max Rüst im   noch nicht   vollendeten Alter von 55 Jahren.

Uncle Max Rüst in-the not yet   completed   age of 55 years

'On September 25, passed away suddenly with a heart attack at less than 55 years

of age, Max Rüst, my beloved husband, our dear father, son, brother, brother-in-

law, and uncle.' (Alfred Döblin, *Berlin Alexanderplatz* p.42)

e. Wir kannten nicht sein unerhörtes Haupt, darin die Augenäpfel reiften.

We knew not his fantastic head, in-which the eye-apples ripened

'We never knew his fantastic head in which the eye-apples ripened.'

(Rainer Maria Rilke, 'Archaischer Torso Apollos')

A problem for such a proposal is that, as noted by Groos and van Riemsdijk 1981, heavy
NP shift, even though is not impossible, only seems to occur in stylistically marked

contexts. Thus the phrase *verurteilte*(*e*) *scharf* in (65a,b) is very typical of news reports. The example in (65b) occurs in the context of the obituary in (65c), and indeed the use of Heavy NP shift is fairly common in this genre. Finally, (65e) is from a poem. These restrictions are at odds with the lack of stylistic markedness in structures involving extraposed free relatives.

Moreover, it is not clear how an analysis like that in (64) could explain the ungrammaticlity of (63c). If the relative operator is in the position occupied by the head noun in a headed relative construction, there seems to be no obvious reason that would block extraposition of the relative clause that modifies it.

The extraposition facts discussed by Groos and van Riemsdijk 1981 then provide a strong argument against an analysis of free relatives along the lines of (64), but, as shown above, they are fully in line with the present proposal.[13]

7 Conclusion

This chapter dealt with the analysis of free relatives. The proposed analysis incorporates two features of earlier analyses. In particular, it preserves the structural similarity between free relatives and headed relative constructions of the Comp analysis (Groos and van Riemsdijk 1981), while at the same time it inherits the ability of the Head analysis (Bresnan and Grimshaw 1978) to account for matching effects in a straightforward way. Moreover, the major counter-arguments to these earlier analyses (matching effects and extraposition data, respectively) also find natural accounts. The claim that (copies of) the relative operator occurs both inside and outside the relative clause in a free relative construction also led to an explanation of the requirement on operators that introduce free relative that they contain a semantic restriction.

---

[13] If the suggestion in note 8 above that the fact that only one of the two copies of the operator is pronounced is due to a constraint against haplology is on the right track, it has implications for the analysis of extraposition. The constraint against haplology presumably holds at the PF interface. This means that it can only extend to the extraposition facts if extraposition occurs some time during spell-out after the haplology police has already left.

CHAPTER FIVE

Conclusion

In Chapter 1 of this dissertation, I developed a new model of phrase structure, whose major ingredients can be stated as in (1) to (3). In chapters 2, 3, and 4, I investigated the predictions of different aspects of the proposed theory in different empirical domains. In this chapter, I give a brief overview over how some of the analytical results follow from which of the assumptions, and talk a little bit about the implications of these proposals.

(1) Material that is not in the deepest position within its chain can head a structure (cf. Ackema et al 1993, Neeleman and Weerman 1999).
(2) The concept of categorial projection is rejected (Brody 1997a, 2000).
(3) Phrase structure does not express linear order, only hierarchical structure.

The assumption in (1) allows for movement that extends clause structure. Since e.g. verb movement varies across languages, this entails that clause structure is not universal. This in turn implies that the structure of a particular language must be learnable. This has the consequence that clause structure cannot contain any (lexical) material that remains consistently without overt content (except for the possible effect of universal semantic conditions). This conclusion directly rules out a whole range of otherwise plausible analyses. Thus, the derivation of Vikner's generalisation and most of the proposals in chapter 2 rely crucially on the idea that there is no empty counterpart to complementizers like *that* (except as the result of learnable, language-specific deletion rules such as the doubly-filled-Comp Filter). Similarly, the derivation of the generalisation that VSO languages must have (overt) preverbal particles would not be possible if empty heads were freely available. Beyond that it becomes possible to derive something like the generalisation proposed by Fukui and Takano 1998:54f that "a functional head H enters into feature checking only if H lacks phonetic content." In terms of present assumptions, such a conclusion follows automatically, namely that all derived heads, which are by definition heads without independent phonetic content, require their specifier position to be filled.

The assumption in (2) allows for the formulation of head chains that forms the basis of

the analysis of the interaction between verb movement and selection in chapter 2. (2) is also crucial in making it possible to generalise over structures involving verb movement and structures involving adjunction to their highest position in that chapter. Without such a generalisation, the identical effect of the two on the distribution of complementizers would be difficult to explain. (2) is also a central ingredient of the derivation of V2 effects in chapter 3.

The implications of (3) are somewhat broader (and possibly deeper) than those of (1) and (2), even though (3) is in some form assumed in most standard assumption of the three. (3) is what makes it possible to represent phrase structure in terms of sets of treelets. This makes it another vital ingredient of the analysis of V2 effects in chapter 3. As noted in various places, this way of looking at phrase structure also affects the possible analyses of specifier chains, an area that still requires more thorough exploration. The representation of phrase structures in terms of sets of treelets offers an interesting alternative to Grimshaw's 2001's way of looking at "economy of representation" effects, that is structures that are smaller than one might have reason to expect. This mode of representation entails that a whole range of feasible tree structures all collapse to representations of a single phrase structure set due to the principle of set identity of chapter 1 ($\{s\}=\{s,s\}$). Thus, nothing needs to be added to the theory to rule out structures that would involve vacuous projections in conventional frameworks. While this view of phrase structure could presumably be modified to allow only for antisymmetric structures, I think it is a virtue of the proposal that *a priori* structures are fully symmetric. This means that both head-final and head-initial languages can be analysed in a straightforward way and without the complications that arise for theories that take one of the two to be universally underlying (cf. Kayne 1994, Fukui and Takano 1998). Anti-mirror effects, like the overwhelming preference for leftward movement, can then be derived from independently motivated constraints, e.g. on processing (Ackema and Neeleman 2001, Davis and Alphonce 1992).

REFERENCES

ABELS, K. 2001. Move? Generals paper, University of Connecticut.

ABNEY, S. 1996. A grammar of projections. Ms., Univerity of Tübingen.

ACKEMA, P. 1999. *Issues in Morphosyntax*. Amsterdam: John Benjamins.

ACKEMA, P., A. NEELEMAN, and F. WEERMAN. 1993. Deriving functional projections. *Proceedings of NELS* 23:17-31.

ACKEMA, P., and A. NEELEMAN. 1998. Optimal questions. *Natural Language and Linguistic Theory* 16:443-90.

ACKEMA, P., and A. NEELEMAN. 2000. M-selection and phrasal affixation. *UCL Working Papers in Linguistics* 12:307-344. University College London.

ACKEMA, P., and A. NEELEMAN. 2001. Effects of short-term storage in processing rightward movement. To appear in *Storage and Computation in the Language Faculty*. Studies in Theoretical Psycholinguistics, Kluwer, Dordrecht.

ACKEMA, P., and A. NEELEMAN. 2002. Interfaces in word formation. Ms., University of Nijmegen and University College London.

ADGER, D. 2000. VSO clause structure and morphological feature checking. In R. D. Borsley, *The Nature and Function of Syntactic Categories*, 79-100. New York: Academic Press.

ADGER, D., S. PINTZUK, B. PLUNKETT, and G. TSOULAS, eds. 1999. *Specifiers: Minimalist Approaches*. Oxford: Oxford University Press.

ALEXIADOU, A., and G. FANSELOW. 2002. On the correlation between syntax and morphology: The case of V-to-I. Ms., University of Potsdam.

ALSINA, A. 1999. Where's the Mirror Principle? *The Linguistic Review* 16:1-42.

ALTMANN, H. 1981. *Formen der 'Herausstellung' im Deutschen: Rechtsversetzung, Linksversetzung, freies Thema und verwandte Konstruktionen*. Tübingen: Niemeyer.

ARNOLD. M. D. 1995. Case, periphrastic *do*, and the loss of verb movement in English. PhD dissertation, University of Maryland.

ARNOLD. M. D. 1997. Indirect grammatical pressure driving language change. Ms., University of Maryland.

BACH, E. 1964. *An Introduction to Transformational Grammars*. New York: Holt, Rinehart, and Winston.

BAKER, C.L. 1970. Notes on description of English questions: The role of an abstract question morpheme. *Foundations of Language* 6:197-219.

BAKER, C. L. 1995. *English Syntax*. Second edition. Cambridge, Mass.: MIT Press.

BAKER, M. 1988. *Incorporation*. Chicago: University of Chicago Press.

BAKOVIC, E., and E. KEER. 2001. Optionality and ineffability. In G. Legendre et al, *Optimality-theoretic syntax*, 97-112. Cambridge, Mass.: MIT Press.

BALTIN, M. 1989. Heads and projections. In M. Baltin and A. S. Kroch, eds., *Alternative Conceptions of Phrase Structure*, 1-16. Chicago: University of Chicago Press.

BALTIN, M. 1991. Head movement in Logical Form. *Linguistic Inquiry* 22:225-49.

BALTIN, M., and C. COLLINS, eds. 2001. *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.

BANFIELD, A. 1982. *Unspeakable Sentence: Narration and Representation in the Language of Fiction*. Boston: Routledge and Kegan Paul.

BELLETTI, A. 2001. Agreement projections. In Baltin and Collins 2001, 483-510.

BELLETTI, A., and L. RIZZI, eds. 1996. *Parameters and Functional Heads: Essays in*

Comparative Syntax. Oxford: Oxford University Press.

BENMAMOUN, E. 1998. Agreement asymmetries and the PF interface. Ms., University of Illinois, Urbana-Champaign.

BESTEN, H. DEN 1983. On the interation of root transformations and lexical deletive rules. In W. Abraham, ed., *On the Formal Syntax of the Westgermania*. 47-131. Amsterdam: John Benjamins.

BHATT, R., and J. YOON. 1992. On the composition of Comp and parameters of V2. *Proceedings of WCCFL* 10:41-52.

BOBALJIK, J.D. 2000. The rich agreement hypothesis in review. Ms., McGill.

BOBALJIK, J.D. 2002. A-chains at the PF-interface: Copies and 'covert' movement. *Natural Language and Linguistic Theory* 20:197-267.

BOECKX, C., and S. STJEPANOVIC. 2001. Head-ing for PF. *Linguistic Inquiry* 32:345-55.

BOLINGER, D. 1977. Another glance at main clause phenomena. *Language* 53:511-519.

BORER, H. 1998. Morphology and syntax. In A. Spencer and A. Zwicky, eds., *The Handbook of Morphology*, 151-90. Oxford: Blackwell.

BORSLEY, R. D. 1984. Free relatives in Polish and English. In J. Fisiak, ed., *Contrastive Linguistics: Prospects and Problems*, 1-18. Berlin: Mouton.

BORSLEY, R. D. 1992. More on the difference between English restrictive and non-restrictive relative clauses. *Journal of Linguistics* 28:139-148.

BORSLEY, R. D. 1997. Relative clauses and the theory of phrase structure. *Linguistic Inquiry* 28:629-647.

BORSLEY, R. D., and A. KATHOL. 2000. Breton as a V2 language. *Linguistics* 38:665-710.

BORSLEY, R.D. and I. ROBERTS, eds. 1996. *The Syntax of the Celtic Languages*. Cambridge: Cambridge University Press.

BOŠKOVIĆ, Ž. 1994. D-Structure, θ-Criterion, and movement into θ-positions. *Linguistic Analysis* 24:247-86.

BOŠKOVIĆ, Ž. 1997. *The Syntax of Non-finite Complementation*. Cambridge, Mass.: MIT Press.

BOUCHARD, D. 1995. *The Semantics of Syntax*. Chicago: University of Chicago Press.

BRESNAN, J. 1970. On complementizers: Toward a syntactic theory of complement types. *Foundations of Language* 6:297-321.

BRESNAN, J. 1994. Locative inversion and the architecture of universal grammar. *Language* 70:72-131.

BRESNAN, J., and J. GRIMSHAW. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9:331-391.

BRODY, M. 1985. On the complementary distribution of empty categories. *Linguistic Inquiry* 16:505-546.

BRODY, M. 1987. Review of Chomsky 1986b. *Mind and Language* 2:165-177.

BRODY, M. 1993. θ-Theory and Arguments: *Linguistic Inquiry* 24:1-24.

BRODY, M. 1995. *Lexico-Logical Form: A Radically Minimalist Theory*. Cambridge, Mass.: MIT Press.

BRODY, M. 1997a. Mirror theory. Ms., University College London.

BRODY, M. 1997b. Perfect chains. In L. Haegeman, ed., *Elements of Grammar*. 139-67. Dordrecht: Kluwer.

BRODY, M. 1998. Projection and phrase structure. *Linguistic Inquiry* 29:367-398.

BRODY, M. 2000. Mirror theory: Syntactic representation in perfect syntax. *Linguistic*

Inquiry 31:29-56.

BURY, D. 1998. Is what you see what get? Matching effects and the derivation of free relatives. MA dissertation, University College London.

BURY, D. 2000a. Particles, V2 and the ungrammaticality of verb-initial structures. *UCL Working Papers in Linguistics* 12:367-386. University College London.

BURY, D. 2000b. Self-attachment and the status of functional projections. *Snippets* Issue 2. (http://www.ledonline.it/snippets/)

BURY, D. 2001. German V3 and the origin of adverbs. *Snippets* Issue 4. (http://www.ledonline.it/snippets/)

BURY, D. 2002a. A reinterpretation of the loss of verb-second in Welsh. In D.W. Lightfoot, ed., *Syntactic Effects of Morphological Change*, 215-231. Oxford: Oxford University Press.

BURY, D. 2002b. Phrase structure and inversion in verb-second and verb-initial languages. *Proceedings of 2002 LSK International Summer Conference*. Kyung Hee University, Seoul. (http://web.kyunghee.ac.kr/~lsk2002/)

CARNIE, A. 2000. On the definition of X^0 and XP. *Syntax* 3:59-106.

CARNIE, A., and E. GUILFOYLE., eds. 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.

CHAMETZKY, R. A. 1995. Dominance, precedence, and parametrization. *Lingua* 93:245-264.

CHAMETZKY, R. A. 2000. *Phrase Structure: From GB to Minimalism*. Oxford: Blackwell.

CHENG, L. 1991. On the typology of wh-questions. PhD dissertation, MIT. Published 1998 with Garland.

CHIERCHIA, G. and S. MCCONNEL-GINET. 2000. *Meaning and Grammar: An Introduction to Semantics*. Second edition. Cambridge, Mass.: MIT Press.

CHOMSKY, N. 1955/1975. *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.

CHOMSKY, N. 1957. *Syntactic Structures*. The Hague: Mouton.

CHOMSKY, N. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

CHOMSKY, N. 1965. *Aspects of the Theory of Syntax* . Cambridge, Mass.: MIT Press.

CHOMSKY, N. 1966. *Cartesian Linguistics*. New York: Harper and Row.

CHOMSKY, N. 1970. Remarks on Nominalization. In Jacobs and Rosenbaum 1970, 184-221.

CHOMSKY, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

CHOMSKY, N. 1986a. *Barriers*. Cambridge, Mass.: MIT Press.

CHOMSKY, N. 1986b. *Knowledge of Language*. Westport, Conn.: Praeger.

CHOMSKY, N. 1991. Some Notes on Economy of Derivation and Representation. In R. Freidin, ed., *Principles and Parameters in Comparative Grammar*, 417-54. Cambridge, Mass.: MIT Press. Also in Chomsky 1995b.

CHOMSKY, N. 1993. A minimalist program for linguistic theory. In K. Hale and S. J. Keyser, eds., *The View from Building 20*, 1-52. Cambridge, Mass.: MIT Press. Page numbers refer to the version in Chomsky 1995c.

CHOMSKY, N. 1995a. Bare phrase structure. In G. Webelhuth, ed., *Government and Binding Theory and the Minimalist Program*, 383-439. Oxford: Blackwell.

CHOMSKY, N. 1995b. Categories and transformations. In Chomsky 1995c, 219-394.

CHOMSKY, N. 1995c. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

CHOMSKY, N. 1999. Derivation by phase. Ms., MIT. Published in: 2001. M. Kenstowicz, ed., *Ken Hale: A life in language*. Cambridge, Mass.: MIT Press.

CHOMSKY, N. 2001. Beyond explanatory adequacy. Ms., MIT.

CHOMSKY, N. and H. LASNIK 1977. Filters and control. *Linguistic Inquiry* 8:425-504.

CHOMSKY, N. and H. LASNIK. 1993. Principles and Parameters Theory. In J. Jacobs, A von Stechow, W. Sternefeld & T. Vennemann, eds., *Syntax: An International Handbook of Contemporary Research*. Walter de Gruyter, Berlin. Page numbers refer to the version in Chomsky 1995c.

CHUNG, S. 2002. VOS languages. Ms. To be included in the Syntax Companion SynCom.

CHUNG, S. and J. MCCLOSKEY. 1987. Government, barriers, and small clauses in Modern Irish. *Linguistic Inquiry* 18:173-237.

CINQUE, G. 1981/1982. On the theory of relative clauses and markedness. *The Linguistic Review* 1:247-294.

CINQUE, G. 1989. On embedded verb second clauses and ergativity in German. In D. Jaspers et al, eds., *Sentential Complementation and the Lexicon*, 77-96. Dordrecht: Foris.

CINQUE, G. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford: Oxford University Press.

CITKO, B. 2002. (Anti)reconstruction effects in free relatives: A new argument against the Comp account. *Linguistic Inquiry* 33:507-511.

COCCHI, G., and C. POLETTO. 2001. Complementiser deletion in Florentine: The interaction between merge and move. Ms., Universita di Firenze and Universita di Urbino/CNR Padova.

COLLINS, C. 2003 Eliminating labels. In S. D. Epstein and T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell.

CORMACK, A. 1995 The semantics of case. *UCL Working Papers in Linguistics* 7:235-276.

CORMACK, A., and N. SMITH. 1997. Checking features and split signs. *UCL Working Papers in Linguistics* 9:223-252.

CORMACK, A., and N. SMITH. 2000. Fronting: the syntax and pragmatics of 'Focus' and 'Topic'. *UCL Working Papers in Linguistics* 12:387-416.

CORMACK, A., and SMITH, N. 2002. Compositionality, copy theory, and control. . *UCLWorking Papers in Linguistics* 14:355-375.

CULICOVER, P. 1991. Topicalization, inversion, and complementizers in English. Ms. University of Utrecht.

CULICOVER, P. and R. JACKENDOFF. 2001. Control is not movement. *Linguistic Inquiry* 32:493-512.

DAVIS, H. and C. ALPHONCE 1992. Parsing, linear asymmetry, and *wh*-movement. *Proceedings of NELS* 22:87-100.

DEN DIKKEN, M. 2002. Lexical Integrity, Checking and the Mirror: A checking approach to syntactic word formation. Ms., CUNY. To appear in *Journal of Comparative Germanic Linguistics*.

DE HAAN, G. 2001. More is going on upstairs than downstairs:; Embedded root phenomena in West Frisian. *Journal of Comparative Germanic Linguistics* 4:3-38.

DENEZ, P. 1973/1974. A structural approach to Breton grammar: The so-called relative pronoun in Breton. *Studia Celtica* 8-9:252-67.

DEPREZ, V. 1994. A minimal account of the *that*-t effect. In G. Cinque et al, eds., *Paths towards Universal Grammar: Studies in homor of Richard Kayne*. Georgetown University Press, Washington, D.C.

DI SCIULLO, A.-M., and E. WILLIAMS. 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.

DOHERTY, C. 1993. Clauses without *that*: The case for bare sentential complementation in English. PhD dissertation, UC Santa Cruz. Published 2000 with Garland.

DOHERTY, C. 1997. Clauses without complementizers: Finite IP-complementation in English. *The Linguistic Review* 14:197-220.

DOHERTY, C. 2000. Residual verb second in Early Irish: On the nature of Bergin's construction. *Diachronica* 17:5-38.

ELSNESS, J. 1984. That or zero? A look at the choice of object clause connective in a corpus of American English. *English Studies* 65:519-33.

EMONDS. J. 1976. *A Transformational Approach to Syntax*. New York: Academic Press.

EPSTEIN, S. D. et al. 1998. *A Derivational Approach to Syntactic Relation*s. Oxford: Oxford University Press.

ERNST. T. 2002. *The Syntax of Adjuncts*. Cambridge: Cambridge University Press.

FABB, N. 1990 The difference between English restrictive and non-restrictive relative clauses. *Journal of Linguistics* 26:57-77.

FRANK, N. 2000. Probleme lexikalischer Selektion und abhängige Verbzweitsätze. *Linguistische Berichte* 184:469-483.

FRANK, R. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, Mass.: MIT Press.

FUKUI, N. and H SAKAI. To appear. The visibility guideline for functional categories: Verb raising in Japanese and related issues. *Lingua*.

FUKUI, N. and M. SPEAS. Specifiers and projection. *MIT Working Papers in Linguistics* 8.

FUKUI, N. and Y. TAKANO. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 7:27-86.

GÄRTNER, H.-M. 1998. Review of Nunes 1995. *Glot International* 3:17-20.

GÄRTNER, H.-M. 2001. Are there V2 relatives in German? *Journal of Comparative Germanic Linguistics* 397-141.

GÄRTNER, H.-M. 2002. On the force of V2 declaratives. *Theoretical Linguistics* 28:33-42.

GAZDAR, G., and G. PULLUM. 1982. Subcategorization, constituent order and the notion 'head'. In M. Moortgat, H. v. d. Hulst, and T. Hoekstra, eds., *The Scope of Lexical Rules*, 107-23. Dordrecht: Foris.

GIORGI, A. and F. PIANESI. 1997. *Tense and Aspect*. Oxford: Oxford University Press.

GREEN, G. 1976. Main clause phenomena in subordinate clauses. *Language* 52:382-397.

GREEN, G. 1996. Distinguishing main and subordinate clause: The root of the problem. Ms. University of Illinois.

GRIMSHAW, J. 1979. The structure-preserving constraint: A review of Emonds 1976. *Linguistic Analysis* 5:313-343.

GRIMSHAW, J. 1986. A morphosyntactic explanation for the Mirror Principle. *Linguistic Inquiry* 17:745-49.

GRIMSHAW, J. 1991. Extended Projections. Ms., Brandeis University.

GRIMSHAW, J. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28:373-422.

GRIMSHAW, J. 2001. Economy of structure in OT. Ms., Rutgers University.

GROOS, A. and H. VAN RIEMSDIJK. 1981. Matching effects in free relatives: A parameter of

the core grammar. In A. Belletti, L. Brandi, and L. Rizzi, eds., *Theory of Markedness in Generative Grammar*, 171-216. Pisa: Scuola Normale Superiore.

GROSU, A. 1994. *Three Studies in Locality and Case*. London: Routledge.

GROSU, A. 2002. Strange relatives at the interface of two millenia. *Glot International* 6:145-167.

HAIDER, H. 1991. The Germanic verb-second puzzle. Review of Weerman 1989. Linguistics 29:703-717,

HAIDER, H. 1993. *Deutsche Syntax Generativ*. Tübingen: Gunter Narr.

HAIDER, H. 2000. Adverb placement: Convergence of structure and licensing. *Theoretical Linguistics* 26:95-134.

HAIDER, H. and M. PRINZHORN, eds. 1986. *Verb-Second Phenomena in Germanic*. Dordrecht: Foris.

HARRIS, Z. 1957. Co-occurrence and transformation in linguistic structure. *Language* 33:283-340.

HAWKINS, J. A. 2001. Why are categories adjacent? *Journal of Linguistics* 37:1-34.

HENDRICK, R. 1988. *Anaphora in Celtic and Universal Grammar*. Dordrecht: Kluwer.

HEYCOCK, C. 2001. Embedded root. Ms., University of Edinburgh. To be included in the Syntax Companion SynCom.

HIGGINBOTHAM, J. 1983. A note on phrase markers. *Revue quebecoise de linguistique* 13:147-166.

HOEKSTRA, T. 1995. The function of functional categories. *Glot International* Volume 1 Issue 2.

HOLMBERG, A. 2000. Scandinavian Stylistic Fronting: How any category can become an expletive. *Linguistic Inquiry* 31:445-83.

HOLMBERG, A. and C. PLATZACK. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford: Oxford University Press.

HOOPER, J. and S. THOMPSON. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4:465-497.

HORNSTEIN, N. 1999. Movement and control. *Linguistic Inquiry* 30:69-96.

HUCK, G. 1985. Exclusivity and discontinuity in phrase structure grammar. *WCCFL* 4:92-98.

HUDSON, R. 1976. Lexical insertion in a transformational grammar. *Foundations of Language* 12.

HUDSON, R. 1980. Constituency and depdency. *Linguistics* 18:179-198.

HUDSON, R. 1990. *English Word Grammar*. Oxford: Blackwell.

IATRIDOU, S. 1990. About Agr. *Linguistic Inquiry* 21:551-577.

IATRIDOU, S. and A. KROCH. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Linguistics* 50:1-24.

JACOBS, R. and P. ROSENBAUM, eds. 1970. *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn.

JACKENDOFF, R. 1977. *X'-Syntax*. Cambridge, Mass.: MIT Press.

JACKENDOFF, R. 1997. *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.

JACOBSON, P. 1995. On the quantificational force of free relatives. In E. Bach, E. Jelinek, A. Kratzer, and B. H. Partee, eds., *Quantification in Natural Languages*, 451-486. Dordrecht: Kluwer.

JONES, B. M. 2001. The negative word *dim* in informal Welsh. Ms., University of Wales, Aberystwyth.

JONES, M., and A. R. THOMAS. 1977. *The Welsh Language: Studies in its Syntax and Semantics*. Cardiff: University of Wales Press.

JOUITTEAU, M. 2000. DEA thesis. University of Nantes.

JOUITTEAU, M. 2002. Nominal Properties of vPs in Breton: Implications for the Typology of VSO langages. Handout for the Workshop on Verb-Initial Languages. University of Arizona, February 2003.

KATZ, J. and P. POSTAL. 1964. *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press.

KAYNE, R. 1982. Predicates and arguments, verbs and nouns. Paper presented at GLOW, Paris.

KAYNE, R. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

KAYNE, R. 1998. Overt vs. covert movement. *Syntax* 1:128-191.

KITAHARA, H. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Mass.: MIT Press.

KOENEMAN, O. 1995. Flexible phrase structure: On the role of verbal projections in Dutch, English and Yiddish. MA dissertation, Utrecht University.

KOENEMAN, O. 2000. The flexible nature of verb movement. PhD dissertation, Utrecht University.

KOENEMAN, O. and A. NEELEMAN. 2001. Predication, verb movement, and the distribution of expletives. *Lingua* 111:189-233.

KOOPMAN, H. and A. SZABOLCSI. 2000. Verbal complexes. Cambridge, Mass.: MIT Press.

KOSTER, J. 1987.*Domains and Dynasties*. Dordrecht: Kluwer.

KRACHT, M. 2001. Syntax in chains. *Linguistics and Philosophy* 24:467-529.

KRUG, M. 2000. *Emerging English Modals: A Corpus-Based Study of Grammaticalization*. Berlin: Mouton de Gruyter.

LANGENDOEN, T. 2002. Merge. Ms., University of Arizona.

LARSON, R. K. 1983. Restrictive modification: Relative clauses and adverbs. PhD dissertation, University of Wisconsin-Madison.

LARSON, R. K. 1988. On the double object construction. *Linguistic Inquiry* 19:335-91.

LASNIK, H. 1981. Restricting the theory of transformations: A case study. In N. Hornstein and D. Lightfoot, eds. *Explanation in Linguistic Theory*, 152-73. London: Longmans. Also in H. Lasnik 1990: *Essays on Restrictiveness and Learnability*. Dordrecht: Kluwer.

LASNIK, H. 2001. A note on the EPP. *Linguistic Inquiry* 32:356-362.

LEE, F. 2001. Relative clauses without *wh*-movement. *NELS* 31:321-31.

LYONS, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

MANZINI, M.R. and A. ROUSSOU. 2000. A minimalist theory of A-movement and control. *Lingua* 110:409-447.

MATUSHANSKY, O. 2002. The status of head-movement in linguistic theory. Ms., MIT/CNRS.

MCCAWLEY, J. 1968. Concerning the base component of a transformational grammar. *Foundations of Language* 4:243-269.

MCCAWLEY, J. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13:91-106.

MCCAWLEY, J. 1988. *The Syntactic Phenomena of English*. Chicago: Chicago University Press.

MCCLOSKEY, J. 1992. Adjunction, selection, and embedded verb second. Ms., University of California, Santa Cruz.

MCCLOSKEY, J. 1996a. On the scope of verb movement in Irish. *Natural Language and Linguistic Theory* 14:47-104.

MCCLOSKEY, J. 1996b. Subjects and subject positions in Irish. In Borsley and Roberts 1996, 241-283.

MEISEL, J., and M. PAM, eds. 1979. *Linear Order and Generative Theory*. Amsterdam: John Benjamins.

MÜLLER, G. 2002. Verb-second as *v*P first. Ms., IDS, Mannheim.

MUYSKEN, P. 1982. Parametrizing the notion head. *Journal of Linguistic Research* 2:57-75.

NASH, L., and A. ROUVERET. 1997. Proxy categories in phrase structure theory. *NELS* 26:287-304.

NEELEMAN, A. and F. WEERMAN. 1999. *Flexible Syntax: A Theory of Case and Arguments*. Dordrecht: Kluwer.

NEELEMAN, A., and H. VAN DE KOOT. 2001. Syntactic haplology. Ms. UCL. To be included in the Syntax Companion SynCom.

NEELEMAN, A., and H. VAN DE KOOT. 2002. The configurational matrix. *Linguistic Inquiry* 33:529-574.

NILSEN, Ø. 2000. V2 and Holmberg's generalization. Ms., OTS Utrecht.

NUNES, J. 1995. The copy theory of movement and linearization of chains in the minimalist program. PhD dissertation, University of Maryland, College Park.

OH, C.-K. 1977. A node does not dominate itself. *Linguistic Inquiry* 8:586.

ÖHL, P. G. 2002. Economical computation of structural descriptions in natural language: A minimally radicalist theory. PhD dissertation, University of Stuttgart.

OUHALLA, J. 2000. Clitics and functional categories in Berber. Paper presented at the Workshop on African and Asian Languages, SOAS, London.

PARTEE, B., A. TER MEULEN, and R. WALL. 1990. *Methods in Mathematical Linguistics*. Dordrecht: Kluwer.

PENNER, Z. and M. BADER. 1991. Main clause phenomena in embedded clauses: The licensing of embedded V2-clauses in Bernese Swiss German. *The Linguistic Review* 8:75-95.

PESETSKY, D. 1998. Some optimality principles of sentence pronunciation. In P. Barbosa et al, eds., *Is the Best Good Enough?*, 337-383. Cambridge, Mass.: MIT Press and MIT Working Papers in Linguistics.

PINKER, S. 1984. *Language Learnability and Language Development*. Cambridge, Mass.: Harvard University Press.

PITTNER, K. 1995. The case of German relatives. *The Linguistic Review* 12:197-231.

POLETTO, C. 2001. *The Higher Functional Field*. Oxford: Oxford University Press.

POLLOCK, J. Y. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365-424.

POSTAL, P. 1964. *Constituent Structure*. The Hague: Mouton.

REINHART, T. 1976. The syntactic domain of anaphora. PhD dissertation, MIT.

REIS, M. 1995. Extraction from V2 clauses in German? In U. Lutz and J. Pafel, eds. *On Extraction and Extraposition in German*, 45-88. Amsterdam: John Benjamins.

RIEMSDIJK, H. VAN. 1998. Categorial feature magnetism: The endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2:1–48.

RIEMSDIJK, H. VAN. 2000. Free relatives. Ms., Tilburg University. To be included in the Syntax Companion SynCom.

RIZZI, L. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

RIZZI, L. 1997. The fine structure of the left periphery. In L. Haegeman, ed. *Elements of Grammar: A Handbook in Generative Syntax*, 281-337. Dordrecht: Kluwer.

RIZZI, L. 2001. Relativized minimality effects. In Baltin and Collins, 89-110.

RIZZI, L., and I. ROBERTS. 1989. Complex inversion in French. *Probus* 1:1-39. Also in Belletti and Rizzi 1996. Page references are to the 1996 version.

ROBERTS, I. 2000. Principles and parameters in a VSO language: A case study in Welsh. Ms., University of Stuttgart.

ROBERTS, I. 2001. Head movement. In Baltin and Collins 2001, 113-47.

ROBERTS, I, and A. ROUSSOU. 1997. Interface interpretation. Paper presented at the LAGB meeting, Edinburgh.

ROBERTS, I, and A. ROUSSOU. 2002. The Extended Projection Principle as a condition on the tense depdency. In P. Svenonius, ed., *Subjects, Expletives, and the EPP*, 125-155. Oxford: Oxford University Press.

ROBERTS, T. 1997. Pashto free relatives and triply-filled Comp: Evidence for a headed analysis. *Lingua* 102:77-85.

ROBINSON, J. 1970. Dependency structure and transformational rules. *Language* 46:259-285.

ROCHEMONT, M. 1989. Topic islands and the subjacency parameter. *Canadian Journal of Linguistics* 34:145-170.

ROORYCK, J. 1994. Generalized transformations and the wh-cycle: Free relatives as bare CPs. *Groninger Arbeiten zur germanistischen Linguistik*:195-208.

ROUVERET, A. 1994. *Syntaxe du gallois: Principles généreaux et typologie*. CNRS Editions, Paris.

ROSS, J. R. 1970. On declarative sentences. In Jacobs and Rosenbaum 1970, 222-272.

ROUSSOU, A. 2000. On the left periphery: Modal particles and complementisers. *Journal of Greek Linguistics* 1:65-94.

RUDIN, C. 1988. On Multiple Questions and Multiple Wh-Fronting. *Natural Language and Linguistic Theory* 6:445-501.

SAITO, M., and N. FUKUI 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29:439-74.

SANTELMANN, L. 1999. The acquisition of verb movement and spec-head relationships in Child Swedish. In Adger et al 1999. 271-298.

SCHACHTER, P. 1973. Focus and relativization. *Language* 49:19-46.

SCHAFER, R. 1995. Negation and verb-second in Breton. *Natural Language and Linguistic Theory* 13:135-172.

SCHAPANSKY, N. 1996. Negation, referentiality and boundedness in Breton: A case study in markedness and asymmetry. PhD dissertation, Simon Fraser University.

SCHUETZE-COBURN, S. 1984. On the Borders of Subordination. *Proceedings of the BLS* 10:650–659.

SCHWARTZ, B. and S. VIKNER. 1996. The verb always leaves IP in V2. In Belletti and Rizzi 1996, 11-62.

SELKIRK, E. 1982. *The Syntax of Words*. Cambridge, Mass.: MIT Press.

SELKIRK, E. 1986. *Phonology and Syntax*. Cambridge, Mass.: MIT Press.

SPEAS, M. J. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.

SPEAS, M. J. 1995. Economy, agreement, and the representation of null arguments. Ms., University of Massachusetts.

SPENCER, A. 1991. *Morphological Theory*. Oxford: Blackwell.

SPORTICHE, D. 1998. Sketch of a reductionist approach to syntactic variation and dependencies. In D. Sportiche, *Atoms and partitions of clause structure*, 379-419. London: Routledge.

SPROAT, R. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3:173-216.

STARKE, M. 2001. *Move dissolves into merge: A theory of locality*. PhD dissertation, University of Geneva.

STEPHENS, J. 1982. Word order in Breton. PhD dissertation, SOAS, London.

STOWELL, T. 1981. Origins of phrase structure. PhD dissertation, MIT.

STUMP, G. 2001. *Inflectional Morphology*. Cambridge: Cambridge University Press.

THIERSCH, G. 1978. Topics in German syntax. PhD dissertation, MIT.

THOMPSON, S. A., and A. MULAC. 1991. The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics* 15:237-251.

THRAINSSON, H. 1996. On the (non-)universality of functional categories. In W. Abraham et al, eds., *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, 252-81. Amsterdam: John Benjamins.

TIMM, L. 1988. Relative clause formation in Breton. *Word* 39:79-107.

TIMM, L. 1991. The discourse pragmatics of NP-initial sentences in Breton. In J. Fife and E. Poppe, eds., *Studies in Brythonic Word Order*, 275-310. Amsterdam: John Benjamins.

TRAVIS, L. DE MENA. 1984. Parameters and effects of word order variation. PhD dissertation, MIT.

URIAGEREKA, J. 1998. *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, Mass.: MIT Press.

VIKNER, S. 1991. Relative *der* and other C^0 elements in Danish. *Lingua* 84:109-136.

VIKNER, S. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. Oxford: Oxford University Press.

VIKNER, S. 2001. V^0-to-I^0 movement and *do*-insertion in Optimality Theory. In G. Legendre et al, *Optimality-theoretic syntax*, 427-464. Cambridge, Mass.: MIT Press.

WALL, R. 1972. *Introduction to Mathematical Linguistics*. Englewood Cliffs, New Jersey: Prentice Hall.

WATANABE, A. 1993. Agr-based case theory and its interaction with the A-bar system. PhD dissertation, MIT.

WECHSLER, S. 1991. Verb-second and illocutionary force. In K. Leffel and D. Bouchard, eds., *Views on Phrase Structure*, 177-191. Dordrecht: Kluwer.

WEERMAN, F. 1988. *The V2 Conspiracy*. Fordrecht: Doris.

WEISLER, S. 1980. The syntax of *that*-less relatives. *Linguistic Inquiry* 11:624-631.

WILLIAMS, E. 1980. Predication. *Linguistic Inquiry* 11:203-38.

WILLIAMS, E. 1994. *Thematic Structure in Syntax*. Cambridge, Mass.: MIT Press.

WILLIAMS, E. 1997. Review of Zwart 1997. *Journal of Comparative Germanic Linguistics* 1:263-272

WILLIS, D. 1998. *Syntactic Change in Welsh: A Study of the Loss of Verb-Second*. Oxford: Clarendon/Oxford University Press.

WILTSCHKO, M. to appear. On the interpretability of tense on D and its consequences for case theory. *Lingua*.

ZWART, C. J. W. 1997. *Morphosyntax of Verb Movement*. Dordrecht: Kluwer.

ZWART, C. J.W. 2001. Syntactic and phonological verb movement. *Syntax* 4:34-62.