Nexus and Dependency

0. Introduction
In this chapter I develop and motivate a particular view of some of the basic relations among nodes in a syntactic tree, including predication (the syntactic and semantic relation of a subject to a predicate), subcategorization (the syntactic relation of a head to its complements), and selection (the thematic relation of a head to its arguments). The concepts developed in this chapter are put to use in Chapters 2 and 3 in specific analyses of some small clause constructions. This chapter is organized as follows: §1 develops the concepts of predication and nexus, §2 investigates notions of selection, developing a particular theory of c-selection and head-government, and in §3 I propose a formal sense of dependency which accounts for the failure of certain constituents to enter into anaphoric relations in tree structures.

1. Subject and Predicate
I have adopted the term NEXUS from the great Danish linguist Otto Jespersen (his use of the term is discussed in more detail in §1.1.4 below). Roughly, a nexus is a combination of a subject with a predicate; clauses and small clauses are nexuses. In this section I will develop the background assumptions about the nature of the nexus that I will use in the rest of this chapter as well as in Chapters 2 and 3. Section 1 is organized as follows: §1 consists of a more or less chronological discussion of some key works bearing on the notions of subject and predicate, the fundamental constituents of the nexus. §1.2 outlines the notions that will be used in this work, distilled from §1.

1.1. Previous work on predication
In this subsection I sketch some contributions to the notion of predication, as I understand them. This is by no means meant to be a complete history of ideas; I have tried only to give a summary of those works which have influenced my own assumptions in the proposals to come. The previous works are organized in nearly chronological order, beginning with Aristotle, whose work directly or indirectly influenced all of the others.

1.1.1. Two views on predication
1.1.1.1. Aristotle
In De Interpretatione, Aristotle notes that nouns and verbs represent concepts without themselves being true or false. He goes on to suggest that a noun is a symbol for something without reference to time, while a verb is a sign of what is said about something else; a verb also ‘has a time-reference,’ in the sense that it is only used in conjunction with tense. A proposition (apophansis) is the combination of a subject, a verb, and a tense; it has a truth-value, in the sense that it can be evaluated as true or false. There are two ways in which the three elements, a noun, a verb, and a tense, can be composed into a proposition: affirmation (kataphasis) and denial (apophasis). Multiple propositions can be combined into a single proposition by conjunction, but otherwise, there are no other ways to form a proposition. Thus, every proposition is the affirmation or denial of some property holding of some subject at some time (cf. Int. 17a 23-25).

As Horn 1989 points out, Aristotle’s notion of predication is syncategorematic; there are two ways in which to construct a proposition, and these two ways do not correspond to two different lexical items; denial may be expressed without the word not (as in Socrates never sleeps), and affirmation may be expressed with it (as in Socrates is not sick). Horn suggests that Montague’s 1970 syncategorematic introduction of tense and negation into sentence denotations may be seen as a formalization of Aristotle’s notion of predication. In more recent work, such as that of Pollock 1989, Laka 1990, and Zanuttini 1991, negation is consigned to a functional head position in the syntactic tree (Pollock’s Neg, Laka’s Σ); and in other work (e.g. Bowers 1993, discussed below), predication itself has been assigned a syntactic position. If the two ideas are conflated, and predication is conceived of as existing in two...
polarities, affirmation and denial, then we have something corresponding to a modern version of Aristotle’s predication. The development of this idea will be traced through the following sections.

1.1.1.2. Frege

Frege (1879, 1884, inter alia) developed a formal model in which it is possible to construct truth-conditional representations for sentences. Frege drew a fundamental distinction between OBJECT (Gegenstand) and CONCEPT (Begriff). Objects, typically the denotations of noun phrases, include real, imaginary, and abstract things, and are alike in that properties can be predicated of them. Concepts, on the other hand, are the denotations of predicates, and are by definition incomplete (ungenässigt or ergänzungsbedürfzig); no property can be meaningfully predicated of a concept (see especially Frege 1892b). Formally, a concept is an unsaturated function, which must combine with an object in order to be complete (its DENOTATION (Bedeutung) is then a truth-value, and it has as its SENSE (Sinn) a proposition; see Frege 1892a).

For Frege, a concept could be represented as follows: Φ(ξ); the capital Greek letter identifies the concept (say, the concept of being asleep), while the lower-case Greek consonant is a variable, holding a place for an object. Thus, if s is the symbol corresponding to the individual Socrates, then Φ(s) represents the proposition ‘Socrates is asleep,’ whose denotation would be True or False, depending on whether Socrates is asleep or not (cf. Frege 1894, 1903). The extension of the concept represented by Φ(ξ) (its Wertverlauf, often translated ‘course-of-values’) could be modelled as a set of ordered pairs, pairing individuals with truth-values; this set is an object, but the function itself, Frege stresses, could not.

4 For Laka 1990, Σ has two polarities, negation and affirmation; but for her, the positive value of Σ corresponds to emphatic assertion, as in She did so order coffee. On the view I am alluding to here, the positive mode of predication would simply be ordinary affirmation.

5 Thanks to Barbara Scholz for very helpful discussion.

6 Two-place predicates, or relations, were Φ(ξ, ζ). I am ignoring Frege’s more complex notational conventions.

7 Frege also allowed for the possibility that a proposition has no denotation, for example if some noun phrase in it is non-referring. Thus Socrates is asleep is neither true nor false if Socrates does not exist.

8 Frege maintained that every concept has an object as its extension, and that this object is a ‘proper’ object in being an admissible argument for any concept (that is, for any ‘first-level’ concept, a one-place predicate over objects, as opposed to such things as quantifiers, which are one-place predicates over concepts, hence ‘second-level’ concepts; see Frege 1894). This led to the paradox discovered by Russell that (continued next page)

A version of Frege’s formal truth-conditional representation for linguistic expressions has been profitably adopted in modern semantic theory, especially in the tradition that stems from Montague’s work (Montague 1970, 1974). In that tradition, there is no syncategorematic operation of ‘predication’: predicative expressions, such as verb phrases, are simply represented as functions. Predication is function application, and there is no basic distinction between affirmation and denial (negation is a one-place operation over the denotations of sentences, which has the effect of reversing the truth-value). This position, which Horn 1989 traces back to the Stoics and Cicero, has been very influential in subsequent work on predication (according to Horn 1989:463, “[t]he 2,300-year war between the term logic of Aristotle and the Peripatetics on the one hand and the propositional logic of the Stoics and post-Fregeans on the other seems to have ended in the complete rout of the former camp by the forces of the latter”), though since Horn, one might say Aristotle’s troops have rallied.

1.1.2. Jespersen

Jespersen’s work contrasts sharply with that of Aristotle and Frege. Jespersen, in his work on English grammar, dealt with an incredibly large amount of data, always with an eye to the historical development of a given construction or expression; the concepts he developed were not always formulated with mathematical rigor, but were intended to pick out very fine distinctions among different constructions. In fact, he criticizes unnamed logicians (Jespersen 1924:131 (PG))9 for analyzing all sentences as if they were copular constructions, charging that this confuses The man walks with The man is walking, which of course are different. This is not to say that Jespersen is blind to generalizations across constructions; on the contrary, he is extremely sensitive to them, and proposes many notions which not only permeate his analysis of English but are also intended as linguistic universals, such as his concepts of rank and nexus (see Jespersen 1924 (PG) inter alia for extensive discussion, and Jespersen 1937 (AS) for development and application of a formal notation of syntactic analysis which captures important generalizations; the concept of nexus is discussed immediately below).

4 Frege’s program of ‘logicism’; see e.g. Furth 1964, Bell 1979.

9 Jespersen’s works have appeared in many editions in different years, and are often cited by title rather than by year. For ease of reference I have included standard abbreviations in parentheses following the conventional citation. MEG is A Modern English Grammar on Historical Principles (vols. I-VII), PG is The Philosophy of Grammar, EEG is Essentials of English Grammar, and AS is Analytic Syntax.
1. SUBJECT AND PREDICATE

Thus Jespersen provides a data-sensitive approach to the notion of predication, and points out many subtleties which must be dealt with in any complete account.

Jespersen observed that subjects combine with predicates at a level below that of the clause; he coined the term NEXUS to refer in general to any combination of a subject and a predicate, not only as a sentence or clause, but also in contexts like those in (1). Examples (1a-e) are from Jespersen 1913:10 (MEG II), which predates the term nexus, but Jespersen comments that these examples contain “not two different objects, ... [but] one complex object, made up of two parts standing in the relation of a principal and its adjunct or predicate.” Examples (1f-g) are from Jespersen 1924:114-19 (PG). I have placed brackets around the nexus objects in each case.

(1)  a. They called [their boy Tom]
   b. We saw [John come]
   c. He made [John come]
   d. He made [his wife happy]
   e. He acknowledged [himself beaten]
   f. I like [boys to be quiet]
   g. He believes [me to be guilty]

It will turn out that the distinction between nexuses, as in (1), and propositions, which is what Aristotle and Frege were principally concerned with, is a crucial one (cf. Jespersen 1924:115 (PG)).

It will be convenient to follow recent usage and use Williams’ 1974 term SMALL CLAUSE for the bracketed sequences in (1a-e), to distinguish them from the (infinitival) clauses in (1f-g), which, like finite clauses, are also nexuses; the small clause will be defined in section §1.2.

Jespersen stresses (cf. Jespersen 1924:117-18 (PG), 1940:7ff (MEG V)) that the direct object of the verb in each of the examples in (1) is not the noun phrase immediately following it, but the entire bracketed sequence, consisting of a PRIMARY plus a SECONDARY (see especially his 1937 volume (AS) for a formalization of the syntactic structure). For Jespersen, a primary is an argument, typically a noun phrase, and a secondary is either an adjunct that modifies an argument, or a predicate predicated of one. The combination of a primary and an adjunct (as in the barking dog) is a junction, while the combination of a primary and a predicate is a nexus (as in the dog barks). It should be noted that Jespersen scrupulously avoids referring to the postverbal secondaries in (1) as predicates, as he reserves that term for tensed verb phrases, following grammatical tradition (see especially Jespersen 1937:120 (AS)); he occasionally made use of the term ‘adnex’ (e.g. 1924:97 (PG), 1933:95 (EEG)) or ‘predicative’ (e.g. 1937, Ch. 9 (AS), 1940:5 (MEG V)) for this element. However, in current work in syntax, the term predicate has acquired a general enough sense that it can be used for the embedded ‘secondaries’ in (1). Later I will use predicate in a more technical sense, and a terminological distinction will have to be introduced, but for now this can be safely ignored.

Jespersen points out that the predicate in a small clause can be of any category; it can be a verb, as in (1b-c), or an adjective or participle, as in (1d-e); or it can be a noun phrase, as in (2a-b), or a prepositional phrase, as in (2c-d) (1940:11-13 (MEG V)); in addition, he suggests that adverbs appear here, giving (2e) as an example, but presumably over could be analyzed as an adjective (cf. the participials ended, done, or finished), or following Emonds 1972 a preposition.

(2)  a. She made a pretty nurse, and his practised eye judged [her a good one].
   b. Your language proves [you still the child].
   c. Neighbors considered [the house in quarantine].
   d. That will mean [a heavy load off my mind].
   e. Savage then imagined [his task over].

It is crucial to distinguish nexus objects, or small clause complements, as in (1-2), from constructions with a direct object containing a postmodifier. The examples in (3) are ambiguous between a small clause (SC) reading and a reading where there is an DP direct object containing a postmodifier. The examples in (3) are from Chomsky 1955/1975:271.

(3)  a. I want the puppy in the window.
   b. They believe the revisions made by Stalinists.
   c. I can’t catch the dog climbing that tree.

On the SC reading of (3a), I express a desire that the puppy be in the window; this reading is easily paraphrased using an infinitive: I want the puppy to be in the window. On the DP reading, I express a desire to have a particular puppy, namely the one in the window. Similarly, the SC reading of (3b) is equivalent to They believe the revisions to have been made by Stalinists, and unlike the DP reading in no way entails that ‘they’ believe the revisions themselves. (3c), as Chomsky notes, has one reading where catch means ‘capture,’ and the dog climbing the tree refers to a dog, and another where catch means ‘discover’ and the dog climbing the tree refers to a kind of situation.
In addition, SC complements must be distinguished from sequences of a direct object and a depictive predicate, as in (4) ((4a-b) adapted from Jespersen 1940:29 (MEG V)).

In (4), the postverbal noun phrase is the direct object, and the depictive predicate characterizes a state enjoyed by that object at the time of the event described; thus, (4a) does not mean that ‘an ox whole’ is something that was roasted, rather that an ox was roasted while it was whole. I will assume that in these cases, the predicate is an adjunct (perhaps in a nexus with a PRO subject), as in Stowell 1981:263, Rothstein 1983, and McNally 1993. In principle, the adjunct can be predicated of either the subject or the object of the verb; (4d) in particular is plausible on both readings. This construction is discussed again briefly in §1.3 below.

Jespersen used the term INDEPENDENT NEXUS to refer to the familiar subject-predicate sentence; he contrasted this with DEPENDENT NEXUS, which was a nexus that was not independent. This included not only nexus objects, as in (1-2) above, but also nexus subjects as in (5a-b), nexuses as the complements of prepositions as in (5c-d), nexus adjuncts as in (5e-f), and even nexus predicates as in (5g-i). Again, I have added brackets for expository purposes.10

Jespersen’s term dependent nexus applies to a given token of nexus in a particular structure, rather than a kind of nexus; notice that the nexuses in (5) can be independent as exclamations, as in (6).

Jespersen used the term ‘nexus of deprecation’ to refer to certain independent small clauses like those in (6) (e.g. in Jespersen 1924, Ch. 9 (PG), 1937:82 (AS)); cf. also Akmajian’s 1984 ‘Mad magazine sentences’ (discussed recently in Zhang 1992). They require a certain context and intonation in order to be felicitous (hence the continuations in (6)). What they show is that there is no sense in which the nexuses in (5) are inherently ‘dependent’ as syntactic structures. In §3 of this chapter, I will develop a formal characterization of dependency which will have the result that some, but not all, of the complements in (1), and none of the small clauses in (5), will be called ‘dependent nexus’; as with the term ‘predicate,’ I am straying somewhat from Jespersen’s usage (though my nexus is essentially his).

Jespersen recognized a number of different categories of verbs taking nexus objects (cf. 1940:10-18 (MEG V)), giving numerous examples. Some of these classes will be discussed in Chapters 2 and 3. In addition, he noted that “any verb, whether transitive or intransitive, may be combined with a simple nexus to denote the result or consequence of the action or state implied in the main verb” (1940:23 (MEG V)); by this he meant what is now commonly called the RESULTATIVE construction.

---

10 (5a-b) is from his 1940:37 (MEG V), where he cites Curme 1931; (5c-d) is from his 1940:38-43 MEG V), from Kipling and Dickens; (5e) is from his 1940:50 (MEG V), from Galsworthy; (5f) is from his 1937:42 (AS); and (5g-i) are from his 1940:37 (MEG V), from Lie, Trollope, and Galsworthy. It should be noted that although I am using Jespersen’s examples in this section, all of these are grammatical in modern-day English, and similar examples can freely be generated.
(for recent analyses, see Carrier & Randall 1992, Fernald 1992). Some of Jespersen’s examples are given in (7).

(7) a. I could beat thy head flat as a dead frog.
   b. The dog ... would bite us all mad.
   c. They had bled the natives white.
   d. Strong winds had blown some [trees] nearly horizontal.
   e. They were drinking themselves drunk with the Zoyland cider.
   f. He slept himself sober.

Resultatives are different from the other constructions in (1) in that the verb in the resultative construction does not typically select a small clause complement; but neither can examples like those in (7) be analyzed as involving depictive predicates; in some cases even intransitive verbs occur as resultatives (as in (7f)). I assume, then, that resultatives are derived structures (as in Dowty 1979, Hoekstra 1988, Carrier & Randall 1992, Fernald 1992). I will return to these constructions briefly in Chapter 3, §5.

1.1.3. The Government and Binding framework

In this subsection I discuss some relevant proposals in the Government and Binding framework (GB). It is this framework, in essence, that forms the analytic basis for this work. The purpose of this subsection is not to provide the reader with an introduction to GB theory; it would unfortunately take too much space to accomplish that, and I must refer the reader unacquainted with the framework to one of the many introductory texts available, for example Lasnik & Uriagereka 1988 or Aoun 1991. Instead, the purpose is to establish what brand of GB I am assuming, and to introduce definitions for various terms used later in the dissertation, concentrating on those matters most pertinent to the syntactic structure of the nexus.

1.1.3.1. Chomsky

Chomsky’s work on these and related topics incorporates a number of elements important to what follows. Here I draw mainly on Chomsky 1981, 1982, 1986a, and 1986b and related work by other authors. First, there is the notion that X-bar theoretic structures can be used to define relations among lexical items and phrases in a construction. Second, Chomsky adopts a formal version of Fillmore’s 1968 Case Theory, more commonly known in its later incarnation as THETA THEORY. Arguments are associated with θ-roles corresponding to the basic roles an argument may play in an event or situation (in a very broad sense): a one-place predicate like *runs* has a single θ-role to assign, one which identifies an argument as the thing running, while a transitive verb like *loves* has two, corresponding to the lover and the loved. θ-roles are generally assumed to be drawn from a short list including, usually, Agent, Theme, Goal, Experiencer, and perhaps a few more (see Grimshaw 1990 for recent discussion and references). Chomsky proposes a one-to-one mapping of arguments and θ-roles, stated formally as in (8) (the notion CHAIN is used to prevent a single argument from receiving more than one θ-role by moving from one argument position to another; a chain includes a moved element and its trace). (See also Chomsky 1981:335 for an earlier formulation.)

(8) The THETA CRITERION (preliminary; adapted from Chomsky 1986b:97)\(^{11}\)
   a. Each argument appears in a chain containing a unique θ-position
   b. Each θ-position is in a chain containing a unique argument

θ-POSITIONS are positions to which θ-roles are assigned, e.g. the complement position of a transitive verb, or the subject position of anything that assigns a θ-role to a subject (an ‘external’ θ-role, in the sense of Williams 1980). Thus (8) requires that each chain containing an argument gets one and only one θ-role.

Another basic GB assumption is that there is a condition on noun phrases, namely the CASE FILTER of Chomsky 1980 (building on work by Vergnauad), which requires that a noun phrase have (abstract) Case. This notion of Case (as opposed to Fillmore’s) is closer to the older one of morphological case; the subject of a tensed clause gets nominative Case, the object of a verb gets accusative Case, and the object of a preposition gets oblique Case. For Chomsky (and Vergnaud), the Case Filter is a surface condition on all overt noun phrases, stated as in (9).

(9) The Case Filter: *NP\(_{[-\text{Case}]}\)

Aoun 1985 proposed a Visibility Condition which links Case-assignment to θ-assignment. For Aoun, a chain is VISIBLE if and only if it contains a Case-marked position (cf. Aoun 1985:76 and Chomsky 1986b:94). The

\(^{11}\) I have left out mention of ‘visibility’, to be discussed immediately below. Later, on p. 135, Chomsky gives the following restatement of the Theta Criterion:

[i] A CHAIN has at most one θ-position; a θ-position is visible in its maximal CHAIN

A CHAIN is a generalization over A-chains, A-bar-chains, and expletive-argument chains.
Visibility Condition has the advantage over the Case Filter that it does not apply to predicative noun phrases, on the reasonable assumption that they do not receive θ-roles (however, something like the Case Filter might still be needed to rule out adjunct noun phrases, depending on certain assumptions about adjuncts; cf. Larson 1985. Also, there is the problem that expletives appear to need Case, as argued by Lasnik 1992). Chomsky incorporates the Visibility Condition into the Theta Criterion, as stated in (10).

(10) The Theta Criterion (from Chomsky 1986b:97)
   a. Each argument α appears in a chain containing a unique visible θ-position P
   b. Each θ-position P is visible in a chain containing a unique argument α

Another basic tenet of GB theory that is relevant here is the Projection Principle, stated as in (11).

(11) The Projection Principle (adapted from Chomsky 1981:38)
   a. If YP is in X', then X0 selects YP.
   b. If X' selects YP, then X' selects YP at L0.
   c. If X' selects YP at L0, then X0 selects YP at all L.

In the definition, selection may be equated with θ-assignment (this is discussed in §2). L stands for syntactic Level of representation, ranging over D-structure, S-structure, and Logical Form (LF). Thus (11a) states that the only phrasal elements which may appear in X' are those which are selected by the head. (11b) states that if a head selects a category, then it selects it at some level; this requires that a selected argument be realized at one of the three levels (barring syntactic processes such as passivization which ‘internalize’ or ‘supress’ arguments). (11c) is the most important part; it states that selectional relations are realized uniformly across all levels. The idea behind this is that thematic relations should not be changed at any point during a derivation; D-structure is a “pure” representation of thematic structure, where the arguments of a head are assigned θ-roles in particular structural positions; complement θ-roles are assigned within the maximal projection of a head (hence ‘internal arguments’), and a subject θ-role is assigned to the subject position of the clause (hence the ‘external’ argument, external to VP). These arguments may be moved around at S-structure, but they will always remain ‘linked’ in a sense (by coindexed traces) with their D-structure positions, and at LF they will be interpreted as arguments of the heads which assigned them their θ-roles. Recall that the Theta Criterion was defined over chains, so that no argument could pick up an extra θ-role in the course of a derivation; the Projection Principle can be seen as a way of extending the Theta Criterion to D-structure and LF, since it will not be possible to insert or delete arguments over the course of a derivation (cf. Chomsky 1982:9).

Chomsky notes that the Theta Criterion and the Projection Principle are insufficient to account for the distribution of expletive subjects, which presumably receive no θ-role. For example, when a zero-place predicate like snows appears in a sentence, an expletive is necessarily inserted in subject position. Chomsky 1981 assumes that the obligatoriness of subjects in clauses follows from the obligatory expansion of S into [NP Infl VP]. In Chomsky 1982 it is proposed that the requirement that every clause have a subject be added to the Projection Principle, the combined requirement being called the Extended Projection Principle (EPP). However, it is not entirely clear that these two concepts are connected. The Projection Principle requires that certain relations (namely, thematic relations) be isomorphic across strata (to use Ladusaw’s 1985 term), while the requirement that every clause have a subject is clearly a condition on clauses at S-structure; expletives have no status at D-structure or LF. While the Projection Principle seems conceptually well-motivated, the requirement that every clause have a subject is not an obvious consequence of anything at all. Why should it not be possible to simply say, *Snows, or *Appears that he left? (Note that in those languages in which this is possible, Chomsky assumes that there is in fact a null pronoun). Other questions arise as well: does this requirement hold of infinitival clauses and small clauses, or just tensed clauses? These questions will be explored below.

At this point I think it would be appropriate to briefly mention the Minimalist program laid out in Chomsky 1993. The Minimalist program is the direct descendant of the GB framework, and preserves the basic GB outlook and many GB terms. The general thrust of that work has influenced the analyses adopted in this dissertation, especially in that I have avoided, wherever possible, relying on D-structure and S-structure conditions of the sort that were so familiar up until that paper began to be circulated.

In GB, as mentioned, there are three levels of syntactic representation: D-structure, which is generated by the base rules (X-bar theory, the Projection Principle) plus lexical insertion; S-structure, which is derived from D-structure by applying various transformations, and LF, which is in turn derived from S-structure. Since LF and S-structure are derived by similar mechanisms (Move alpha), they are different strata, rather than different levels, in the sense of Ladusaw 1985. Phonological
rules derive Phonetic Form from S-structure, and rules of interpretation derive a logical representation from LF. Various constraints, such as the Case Filter and the Extended Projection Principle, hold over individual strata of representation, or perhaps over all of them (the Projection Principle, the ECP (Empty Category Principle) in some formulations).

In the Minimalist program, there is no D-structure or S-structure. Instead, a mechanism (Generalized Transformation, or GT) builds trees from lexical items and other trees. At some point, a tree is converted into a Phonetic Form (Spell Out), after which point there is no further access to the lexicon. There is still LF, but there are no other strata on the same level as LF. PF and LF are different levels, subject to different constraints. There is no D-structure, since GT builds up trees from the lexicon, and there is no S-structure, since the point in the derivation at which Spell Out applies is not considered a separate stratum.

I make an effort in what follows to point out where GB and the Minimalist framework require significantly different treatments of the same data, wherever I am aware that that is the case.

1.1.3.2. Stowell

Stowell 1981, 1983, operating in the general GB framework as outlined in §1.1.3.1, proposed a direct mapping from syntactic structure to predicate structure: a predicate is always X' and subjects always occupy SpecXP. The clause, on this view, is a projection of Infl, the locus of tense and agreement features (this idea is also considered at several points in Chomsky 1981). Infl, like other heads, projects a two bar-level phrase, an IP; the subject of the sentence is located in the specifier or Spec position (the sister of X' and daughter of XP), which I will refer to as SpecIP, and Infl (or I') takes a VP complement, as diagrammed in (12).

(12) 

\[ \text{SpecIP} \quad \text{I'} \quad \text{IP} \quad \text{VP} \]

For Stowell, the small clauses discussed by Jespersen have a structure exactly parallel to that of the clause; except that there is no inflectional element; the small clause is a projection of a lexical head, and the subject is in its specifier, as in (13).\(^{12}\)

\[ \text{V'} \quad \text{PP} \quad \text{considered} \quad \text{DP} \quad \text{P'} \quad \text{the house} \quad \text{in quarantine} \]

Following fairly standard GB assumptions about Case-marking, government, and the distribution of empty categories, various facts about small clause structures fall out from Stowell’s proposal. The arguments of the small clause head (the house and quarantine in (13)) receive their \(\theta\)-roles from that head; as expected, the verb places no selectional restrictions on those arguments (but cf. Rizzi 1992 for an opposing view, with respect to perception verbs). The small clause subject is governed by the higher verb; therefore, the null element PRO is not allowed there (*John considers lucky), on the assumption that PRO must be ungoverned (cf. Chomsky 1981). Since the higher verb governs the small clause subject, it can assign Case to it; therefore, when the verb is passivized, the small clause subject may (in fact must) be promoted to main clause subject position, as in (14a-b) I have bracketed the small clause and marked the D-structural position of the small clause subject with a blank space.

(14) a. That house is considered \[\_ in quarantine\].
   b. The negotiators were made \[\_ uncomfortable\].

Stowell 1981, following previous work (see his fn. 27. page 294 for references) argues that various ‘raising’ structures may be subsumed under this analysis. The standard analysis of sentences like that in (15a) (at least since Kiparsky & Kiparsky 1970) holds that the main subject originates as the embedded subject and ‘raises’ to the higher position, as indicated by the structure in (15b).

(15) a. Max seems to be uncomfortable
   b. Max seems [IP \_ to be uncomfortable]

\(^{12}\) I am using DP for noun phrases in general, following Abney 1987; complement to in in (13) might be predicative, in which case it should be NP, following Stowell (continued next page)

1991b, Longobardi 1993 Grimshaw 1991 has proposed that such expressions as by car, at school might be cases of P selecting NP rather than DP.
Stowell suggests that examples like that in (16a), where *seem* appears followed by an adjective, also involve raising, as indicated in the example, and that some copular constructions involve raising as well, as in (16b) (cf. also Ross 1969a for a raising analysis of the verb *be*).

(16) a. Max seems *[AP ... uncomfortable].

   b. Jane is *[PP ... in her office].

In other words, the deep complement to the copula is a small clause, and the subject of that small clause raises to the main clause subject position for reasons of Case. Interestingly, this is very much like an analysis by L. L. Hammerich criticized in Jespersen 1937:132-5.13 Hammerich suggests that the sentence *Das Pferd ist krank*, 'the horse is sick,' be analyzed with an existential one-place predicate (*ist*) as its argument the small clause *das Pferd krank*; the logical motivation is that the sentence is an assertion that the situation of the horse's being sick exists. Jespersen objects that the analysis feels "unnatural." He points out that intuitively, the existential sense of the verb *be* (as in *God is*) is an extension of the copular use, rather than the other way around. He also points out that Hammerich’s analysis makes the wrong predictions regarding agreement (Hammerich would presumably have "The horses is sick from [the horses sick] is"). But of course this last complaint does not hold if the small clause is the deep complement to the copula. Jespersen's second objection, that the copula does not seem to have existential force, is especially valid in cases with individual-level predicates (discussed in Chapter 2, §4): *Otto is perceptive* does not mean 'a situation of Otto's perceptiveness exists.' As for the first objection, it may be the case that our tolerance for 'unnatural' explanations has increased since 1937. At any rate, Jespersen’s objections are not directed at the idea behind the raising analysis, only at the extension of a raising analysis to the copula. He himself proposed raising-type analyses for such predicates as *seem*, *sure*, and *happen* (in Jespersen 1913, §11.9 (MEG II), 1937 Ch. 16 (AS)).

I will adopt the raising analysis for some instances of *be*; this will be discussed further in §1.1.3.3 and §1.1.4 below.

Stowell's work suggests that the small clause is a basic structure in a number of ways. First, it is basic in the sense that it consists of the maximal projection of a lexical head, along with the satellites of that head, without any additional structure. Second, it is basic in the sense that many instances of predication involve an underlying small clause, including not only the obvious small clause complements but also copular constructions and (other) raising constructions. Another feature of Stowell’s work is that it suggests a direct mapping from X-bar structural representations at LF onto logical representations, in that subjects always occupy a Spec position and X' is always a predicate. Beyond this, however, no special status is accorded to subjects; a subject, in GB, is like an object except that it occupies a subject position (a Spec position; in later work such as Larson 1988, Chomsky 1993, and Bowers 1993, this distinction is blurred as objects begin to appear in Spec positions). Further work by Stowell on small clause structures will be brought up in later sections, as it becomes relevant; in particular, Stowell 1991a is discussed at length in Chapter 2.

1.1.3.3. The Internal Subject Hypothesis

Since Stowell postulated a subject position inside the VP for small clauses with verb heads, as in (17a), he paved the way for what has become known as the Internal Subject Hypothesis (ISH), where it is assumed that all subjects originate within VP, and later raise to SpecIP (Kitagawa 1986, Fukui 1986, Kuroda 1988, Fukui & Speas 1988, Koopman & Sportiche 1988).

(17) a. Jane saw *[VP the students riot].

   b. *[IP Jane [VP ... Max]]

   c. *[IP Max [VP ... danced]]

This allows a conceptually simpler system of θ-role assignment, as all θ-roles are assigned within the maximal projection of their θ-assigner; the ‘external’ θ-role is assigned to the specifier position of that head, rather than to SpecIP.

In fact, an even stronger position is sometimes assumed, following Stowell’s arguments that there is a subject position available in every lexical XP; on this view, all arguments originate within the projection of a lexical head. This view might be called the Lexical Internal Subject Hypothesis (LISH). The LISH is consistent with the analysis of copular constructions as raising constructions discussed in the previous subsection. In examples like those in (18) below ((18b) is repeated from (16b)), the main clause subject in each case receives its θ-role from the lexical head of the small clause.

---

1. SUBJECT AND PREDICATE

(18) a. Max is [AP _ uncomfortable].
   b. Jane is [PP _ in her office]

The strongest version of the LISH would be one in which all (non-predicative) noun phrases originate as arguments of lexical heads. Even possessors, then, would originate within NP. If the DP hypothesis (Abney 1987) is adopted, in which noun phrases with possessors are projections of a possessive head of the category D, then possessors might raise from within NP to SpecDP in languages like English, paralleling the movement in the clause from SpecVP to SpecIP.

(19) [DP Tom’s [NP _ friends]]

However, it is not clear that possessive noun phrases should in general be seen as receiving a θ-role from the noun. As Williams 1981a:88 notes, the possessor may hold virtually any relation to the head noun; for example, Geach’s ignorant lumpers might be the ignorant lumpers that Geach follows, the ignorant lumpers that follow Geach, or even the people Geach called ignorant lumpers (as the expression is in fact used in Horn 1989:508).

A weaker version of the LISH might still be maintained, if every non-expletive argument noun phrase receives a θ-role from some lexical head, within the projection of that head. However, this begs the question of what exactly an argument noun phrase is. Below I will show some more serious problems for the LISH, recently discussed in Heycock 1991.

1.1.4. Syntactic theories of Predication

In this section I discuss works by Williams, Rothstein, and Heycock which focus on predication and accord a special status to the subject. In the views discussed in §1.1.3 above, subjects have a special structural location (Spec) but are otherwise generally treated no differently from other arguments, and predication is a by-product of θ-assignment; the relation between a predicate and a subject is not fundamentally different from that between a verb and its direct object. In GB, every clause has a subject as a result of the stipulation that SpecIP is obligatorily projected (the EPP), which only accounts for the fact that small clauses have subjects if they too are assumed to be projections of Infl.

For the authors discussed here in §1.1.4, there is a difference between subjects and complements; for Williams, predication is a special case of θ-assignment, the assignment of an ‘external’ θ-role, subject to different conditions than the assignment of an ‘internal’ θ-role; for Rothstein and Heycock predication is a syntactic primitive, independent of θ-assignment, and the requirement that every clause have a subject is built into the conception of predication; it is the nature of a predicate to have a syntactic subject.

1.1.4.1. Williams

Williams 1980 (also 1982, 1983, 1987a, 1987b, 1994, inter alia) proposes a theory of predication couched in structural terms, but one different from Stowell’s. Specifically, he proposes that a predicate is always a maximal projection, and must be coindexed with a subject under mutual c-command. Taking the distinction between internal and external θ-assignment as basic, he argues that the external θ-role is assigned under coindexing with the predicate (in Williams 1994 it is proposed that this follows from principles of headedness; the external argument is the ‘head’ of the argument structure of the verb, and its index percolates to VP). The indexing requirement holds of all predicates in the broad sense; tensed VPs as well as selected small clause predicates, resultative predicates, and depictive predicates must be coindexed with a subject.

He also proposes (Williams 1987b:434) that predication is itself a species of θ-assignment; this suggestion will be taken up again in §1.1.4.3 below. Williams allows greater-than-binary branching; thus, even under mutual c-command, a subject need not form a constituent with its predicate in order to be coindexed with it. For example, he argues that the subject-predicate sequences that I referred to as small clauses in (1) above are not constituents at all, but are complements to the verb, sisters under V, as in (20a) (cf. Williams 1983); depictive constructions necessarily have the same structure, as in (20b).

---

14 As Barker 1991 points out, there are cases in which the interpretation of possessives is constrained (Barker’s LEXICAL POSSESSIVES); for example the table’s leg is possible, but *the leg’s table is not. What is important here is simply that there are some possessive constructions in which the relation of the head noun to the possessor noun phrase is not one of predicative head to argument. (Barker’s EXTRINSIC POSSESIVES).

15 Williams adopts Reinhart’s 1976 definition of c-command: A c-commands B iff every branching node which dominates A dominates B (the inverse of Klima’s 1964 ‘in construction with’).
The first half of the condition is intended as a universal condition; it is very similar equivalent to Williams’ 1980:206 C-Command Condition on Predication (which we might call the CCCP, that abbreviation being currently non-referring), but it defines a predicate as a non-θ-marked XP.

Rothstein 1983

Rothstein 1983 argues that predication is a syntactic primitive. She adopts Williams’ convention of indicating predication by coindexation, but does not link predication directly to θ-assignment. Instead, she proposes a pair of conditions, one syntactic and one semantic, on all syntactic structures. Like Williams, she adopts a version of X-bar theory (cf. §2.1 below) in which maximal projections (XPs) represent phrases and have a privileged status; syntactically, all XPs are either arguments or predicates. Arguments, which are always DPs or CPs (Rothstein’s NP and S’), are subject to the Theta Criterion of Chomsky 1981: each argument must receive a θ-role from a lexical head. Predicate XPs (which can be of any category, including CP and DP), are subject to the rule of PREDICATE-LINKING (RPL), which Rothstein states as in (21).

(21) **Rule of Predicate Linking** (for English)  (Rothstein 1983:11)

a. Every non-θ-marked XP must be linked at S-structure to an argument which it immediately c-commands and which immediately c-commands it.

b. Linking is from right to left (i.e., a subject precedes its predicate).

This has the effect of dividing XPs into two classes: arguments and predicates. This means that adjuncts are a kind of predicate (an intuition expressed many times throughout the history of grammatical investigation; see Jespersen 1933:114 ff. (EEG)). In fact, the picture turns out to be so simple: expletives are neither θ-marked nor predicative, and it is not clear to what extent adjuncts and predicates can be unified.

Another difference between the RPL and the CCCP is that Rothstein’s definition of c-command uses maximal projections rather than branching nodes (following Aoun & Sportiche 1982; a more recent term is M-COMMAND, M for ‘maximal projection,’ in Chomsky 1986b). This allows more flexibility in tree structures; for example, the depictive predicate in (20b) in the previous section can be adjoined to V rather than being its daughter, allowing a structural distinction between the two constructions that was unavailable to Williams (cf. her pg. 16).

Rothstein intends that the second clause of the RPL be language-specific. But it is a necessary clause, since linking is not conceived of as a symmetric relation; although many predicates may be linked to a single argument, a single predicate cannot be linked to more than one argument. Rothstein is forced to make one exception to her rule of predicate linking (fn. 4 p. 22): pleonastics, which do not receive θ-roles, are nonetheless treated as arguments, in that they are coindexed with a predicate, rather than with an argument. However, a zero-place predicate like rains is not exempt from (21); it is not θ-marked and therefore must be linked to some argument; an expletive is inserted for this purpose.

Semantically, a non-argument is typically (i.e., apart from special cases like rains) an unsaturated one-place function in the Fregean sense. At LF, the argument that is syntactically linked with a predicate is interpreted as its subject. Modifiers, according to Rothstein, are basically predicate-like (for Montague, predicates and modifiers are functors), though they do not take arguments in the sense in which Rothstein uses that term (i.e. an argument must be θ-marked, and a VP is not θ-marked, but a VP modifier is a VP functor).

Rothstein introduces the terms PRIMARY PREDICATION and SECONDARY PREDICATION; her primary predication is that which forms a constituent, including adjuncture structures. Her secondary predication is the sort of non-constituent predication that Williams’ coindexing convention allows; thus the depictive adjunct is a secondary predicate (it is here she uses the term ‘small clause,’ which is close to Williams’
original use of the term, but I will be using it in its later sense). In other work (e.g. Holmberg 1986), the term secondary predicate has been used in different ways; I have used it to refer to an embedded predicate, without tense, to distinguish it from the main or clausal predicate, but allowing it to remain ambiguous between a predicative part of a small clause complement and a predicative adjunct.

For Rothstein, as for Williams, predication does not entail constituency of a subject and a predicate; and predication is further divorced from \( \theta \)-assignment and granted its own, primitive status.

### 1.1.4.3. Heycock 1991

Heycock 1991 follows Rothstein in arguing that predication is a primitive syntactic relation. Heycock also continues in the direction of separating \( \theta \)-assignment from predication, arguing that in some cases, noun phrases in subject positions are interpreted as arguments even though they receive no \( \theta \)-role. This requires a revision of Chomsky’s Theta Criterion. Recall that the Theta Criterion (stated in (10) in §1.1.3.1 above) required that each argument chain contain a \( \theta \)-position. One solution would be to follow Williams’ 1987b suggestion that predication is a form of \( \theta \)-assignment; but as Heycock points out, this would mean that expletives receive a \( \theta \)-role. If expletives received \( \theta \)-roles, then the concept of \( \theta \)-role would have to be completely rethought.

Instead, Heycock (1991:227, building on Browning 1987) proposes replacing the Theta Criterion with an Argument-licensing Principle which requires an argument chain to be licensed either by \( \theta \)-marking or predication. Anticipating my own analysis somewhat, I cast Heycock’s examples of non-\( \theta \)-marked subjects as challenges to the LISH (the requirement that every argument be \( \theta \)-marked by a lexical head), rather than as challenges to the Theta Criterion.

In one respect, Heycock’s conception of predication is more like the GB treatments and Jespersen’s nexus, namely in that a predicate always forms a constituent with its subject. The details, however, are different from Stowell’s: for Heycock, subjects are adjoined to their predicates.

The most important point in Heycock 1991, for my purposes, is the claim that there are referential subjects which do not receive \( \theta \)-roles. I will discuss each of her examples in turn. First, there is the tough-construction, discussed in detail in Browning 1987 (see that work for earlier references).

(22) a. It is tough for anyone to forget Janice.
   b. Janice is tough for anyone to forget.
1. SUBJECT AND PREDICATE

with the LISH. This structure, then, is not problematic. Now consider the structure in (25b). There, as Higgins and Heycock argue, the free relative is the predicate; yet there is no lexical head that could be understood as assigning a θ-role to the AP subject proud of himself. It is sometimes assumed that the verb be is in some cases a θ-assigner (e.g. in Higginbotham 1987), but as Heycock points out, specificational pseudoclefts are licit as small clause structures, without the copula ((26) is from Heycock 1991:213; as she notes, many examples are marginal).

(26) I consider novels what John writes best.

It seems, then, that we are left without a θ-assigner for the (deep) subject. Another case involves DP predicates (cf. Stowell 1991b for some discussion of NP and DP predicates in general).

(27) a. Mary is Max’s best friend.
   b. I consider Mary Max’s best friend.
   c. The man with the mask turned out to be Archie.
   d. We called him Archie.

The possessor indicates that the predicate is in fact a DP, following Abney’s 1987 analysis. If friend assigns a θ-role at all, we would expect it to be to the possessor, rather than to the nential subject (cf. Williams 1981a for discussion of the argument structure of nominals). The small clause structure in (27b) shows again that the copula is not the source of the subject’s θ-role. In (27c), a name is used as a predicate (Heycock cites Heggie’s 1988 USC dissertation for discussion), a possibility which is excluded from consider-type small clauses for some reason, although it is allowed with other verbs, if call as in (27d) can be assumed to take a small clause.


(28) Bunmeikoku-ga dansei-ga heikinzyumyoo-ga mizikai. (Jap)
    civ.countries-NOM male-NOM avg.life.span-NOM is.short
    ‘In civilized countries the average life span of men is short’

Presumably, the predicate mizikai ‘is short’ can assign a single θ-role; following the works cited, Heycock argues that the ‘extra’ nominative-marked noun phrases are subjects, but are not θ-marked; certainly they cannot be θ-marked under the ISH or the LISH, by any lexical head.

Finally, Heycock discusses constructions in (colloquial) English with like, including the examples in (29).

(29) a. It seems like it’s raining.
   b. It sounds like your car needs tuning.
   c. It looks like John failed the exam.

Here, it is non-referential; seem, sound, and look are raising verbs taking a complement headed by like with a single clausal argument. But just as with the tough construction, a referential subject can appear here. The examples in (30a-b) are from Rogers 1971, and (30c-d) are from Heycock (p. 269).

(30) a. IBM looks like someone’s been kiting it.
   b. Ermintrude looks like the cat got her tongue.
   c. That book sounds like its publication will cause quite a stir.
   d. That book sounds like everyone thinks it should be banned.

It should be noted that such examples have a colloquial flavor, even more so than the examples in (29). As Heycock notes, examples like (30d-c) suggest that the verb sound does not assign a θ-role to the subject in these constructions; (30c-d) are felicitous as comments on someone’s description of a book; the book doesn’t make any sound at all. Heycock also suggests that a coreferential pronoun in the embedded clause is not strictly necessary, offering the examples in (31a-d) (her pg. 271).

(31) a. That book sounds like everyone should own a copy.
   b. That restaurant sounds like new management would be an improvement.
   c. Those books sound like the covers should have been replaced long ago.
   d. That movie sounds like there should be a sequel.
   e. This place looks like the floor hasn’t been swept in weeks.
   f. This analysis looks like you’ve been reading too much Frege.

Again, these examples have a very colloquial ring to them (some similar examples border on the ungrammatical; cf. *John looks like it’s raining, *John looks like the exam was difficult), but I agree with Heycock that they are much better than the corresponding relative clause constructions (*the book that everyone should own a copy), suggesting that the
examples in (31) are not null operator constructions (see Heycock ibid for discussion and references to previous work). 17

We could add to Heycock’s litany of problems for the LISH the construction noted by Jespersen in which a nexus serves as a predicate (cf. 5f-h in §1.1.2 above, repeated here as (32a-c)).

(32) a. That’s another nightmare ended.
    b. That is one difficulty over.
    c. It’s a great weight off my mind. – It’s rather a weight on my wife’s, I’m afraid.
    d. Our biggest problem is pigs in the kitchen.

Here, too, the fact that such constructions are possible in small clauses means that the copula itself cannot be the θ-assigner for the subjects in these cases.

(33) a. I consider that another nightmare ended.
    b. I regard that as a great weight off my mind.
    c. I consider our biggest problem pigs in the kitchen.

The conclusion is that the LISH cannot be maintained; there are grammatical subjects which do not originate as the arguments of any lexical head. For Heycock, this motivates the status of predication as a primitive, not derivable from θ-assignment. In GB, it is sometimes taken as axiomatic that arguments have θ-roles; this view could be reconciled with Heycock’s if predication itself were sufficient to assign a θ-role; but as noted above, this creates problems with expletive subjects. In the next section, I discuss some recent proposals regarding predication that turn out to provide a way of dealing with this apparent impasse, as well as having some other appealing results.

To briefly summarize §1.1.4, the accounts discussed here all had in common that they treated predication as a syntactic device distinct from the taking of arguments. Various subject-object asymmetries support this general outlook. For instance, objects are often syntactically omissible, for example in elliptical constructions where a generic or indefinite sense for the object is contextually supplied (e.g. I wrote, I spoke, I painted, I lifted, etc.). A subject, on the other hand, cannot typically be omitted; in a causative-inchoative alternation, a missing causer is not contextually supplied but presumed nonexistent (cf. *The boat sank to collect the insurance). Also, a subject position cannot be left empty; an expletive must be inserted. Predication theory allows a more natural account for such facts than does GB, which treats subjects no differently from other arguments, except for their surface position in the syntactic tree; but of course, in GB, structural location is significant: the subject in Spec-c commands the other arguments and the head, and enters into an agreement relation with the head (Spec-head agreement). The accounts discussed in this section took different approaches to the interaction between phrase structure and predication. For Williams, predication occurs under sisterhood; for Rothstein, mutual m-command. For Heycock, the subject is adjoined to the predicate, which means that the subject and predicate form a constituent. In the next section, I explore a proposal that bridges the gap, in a way, between the GB accounts and the predication theory accounts.

1.1.5. A head in the small clause

1.1.5.1. Property Theory

In Chierchia 1985 and Chierchia & Turner 1988, a theory of properties is developed which reintroduces Frege’s distinction between properties and predicates, lost in the neo-Montagovian tradition of treating all predicates as functions. In Property Theory, a broad notion of entity-hood is adopted, so that not only are individuals entities, but so are properties and ‘information units’ (units of propositional content). Specifically, there are at least three basic sorts of type e: nominalized functions (sort nf), which correspond to properties, and urelements (sort u), which include individuals and information units (sort i).

As Chierchia 1985 notes, the inclusion of properties in the set of entities allows other properties to be predicated of them, which allows a simple analysis of such expressions as Swimming is fun and To learn French properly takes years (cf. also Aczel 1978, Bealer 1982).

Two operators are introduced, ‘∪’ and ‘∩’. ∪ is the predication operator, which takes a property and returns a predicate; it could be assigned the type <nf,<e,i>>. The nominalization operator, ∩, has the reverse effect: it takes a predicate and returns a property (and is thus type <e,i,nf>).

Given a Property Theory of this kind, we can assume that all cases of predication are mediated by ∪, reviving Aristotle’s notion that predication is an operation combining a subject with a predicate. Of

———17 One very odd type of example which is about as good as the ones in (31) has there as a main subject and there as an embedded subject:
[i] There looks like there’s a problem.

This seems to support Heycock’s contention that the predicate assigns no θ-role to the subject, but raises some other problems; for example, why must the embedded clause have a there subject? I will not discuss this sort of example further here.
course, Property Theory itself does not entail that there are no underived predicates, but this is what I will assume below.

In order to allow Heycock’s examples of predicates which do not assign θ-roles, we need only assume that each of those predicates can be construed as the sort of thing that the predication operator can apply to; for example, tough for anyone to forget must be construable as a property (sort pf), as must Max’s best friend or a great weight off my mind. This calls for an independently motivated definition of what it is to be a property, and an investigation of what sorts of things cannot be predicates (for example, Max intelligent is a good candidate. *That’s Max intelligent, *Our biggest problem is Max intelligent, etc.). This and related issues will be taken up in §3.

1.1.5.2. Bowers 1993

Bowers 1993 argues at length that there is a functional head in the small clause, which he calls Pr[edication] (I will use the abbreviation PRED, in part because it is less easily confused with P (for preposition) and Prt (for particle), standard abbreviations for elements which figure centrally in Chapter 3). In various previous works it had been assumed that there was a functional head in the small clause; in some early generative work, small clauses were derived from full clauses (Ss) by a rule of to be deletion (see, e.g. Borkin 1973, 1984). Later work continued to assume that small clauses were some sort of defective clause (e.g. Chung & McCloskey 1987), and when the notion became widespread that the clause is headed by a functional element, a null head was postulated for the small clause as well; thus, for Hornstein & Lightfoot 1987, the small clause is an IP with a null head; similarly for Taraldsen 1991b. Raposo & Uriagereka 1990, building on an idea of Kayne’s, proposed that adjectival small clauses have a null Agr[ement] head, allowing the adjectival predicate to agree with its subject. In Raposo & Uriagereka 1993 they develop this idea further, suggesting that there are two different types of Agr, and correspondingly two different types of small clause (this will be taken up again in Chapter 2).

Thus in this respect Bowers’ proposal is not new. However, his characterization of the nature of the functional head differs from the other proposals I have mentioned. Building on Chierchia 1985 and Chierchia & Turner 1988, Bowers suggests that (the denotation of) the XP ‘predicate’ of the small clause is not an unsaturated function, but a property. The functional head Pred has the effect of changing the type of this property, making it an unsaturated function. In other words, Pred is Chierchia’s ↓, relieved of its syncategorematic status and given a syntactic position in the tree.

1.1.6. Summary of §1.1

There are certain tensions among the accounts just discussed, but also certain harmonies. In general, it is observed that there is a basic relationship between a predicate and its subject which must be expressed in the grammar. The relation is nearly symmetrical; a predicate needs something to predicate over, and an argument needs some excuse for being in a sentence.

The main division is between those works in which predication is simply the filling in of a missing part of a predicate, as in Frege, Jespersen, Chomsky, Stowell, Williams, and Rothstein, and those works in which a predicate and a subject are combined by some external operation, as in Aristotle, Montague, Chierchia, and Bowers.

Another division is between works in which a nexus is necessarily a constituent, as in Jespersen, Chomsky, Stowell, and Bowers, and those in which it is not, as in Williams and Rothstein.

Finally, there is the tension regarding the proper licensing of arguments. The LISH provides a particularly strong constraint on argument licensing, one which turns out to be too strong, as Heycock shows. Instead, there must be some licensing of subjects (even referential ones) independent of the assignment of θ-roles by lexical heads. This can be accomplished using Chierchia’s predicator, which I will assume, following Bowers, is syntactically realized. My specific assumptions are outlined in the next subsection.

1.2. A distillation

Here I outline the assumptions I will be making in the rest of this dissertation. I address the following issues raised in the preceding sections: whether a predicate can be predicated of an argument with which it does not form a constituent; whether a nexus is necessarily headed, and if so, by what; and what is the status of the Theta Criterion in light of Heycock’s examples.

1.2.1. Constituency

One of the main objections to the analysis of small clauses as constituents was that they fail to undergo constituency tests, as shown in (34).

(34) a. *What we really considered was John an idiot.

b. *It was Bill as intelligent that we regarded.

Stowell’s (1981, 1983) account for these facts was based on Case theory (the subject of the displaced small clause cannot get Case), but there are problems for this assumption, as Pollard & Sag 1993 point out. However,
the theory of dependency that I develop in §3 below provides an account for the ungrammaticality of the examples in (34), as I detail in Chapter 2. This clears the way for us to adopt the analysis of small clauses in general as constituents, following Jespersen and Stowell.

Some further reasons for believing in the constituency of small clauses (particularly the small clause complements to verbs like consider) are worth mentioning here. First, it allows us to adopt a restrictive theory of predication in which subjects appear in the specifier of a predicate, as in Stowell 1981; in fact, it allows an even more restrictive conception of predication, since subjects will always appear in SpecPredP or SpecIP. In addition, it allows for isomorphism between the syntax and the semantics, since small clauses are interpreted as semantically coherent objects (cf. Chapter 2) (this isomorphism is sometimes assumed to be required by the Projection Principle). Finally, it is necessary to believe that at least some small clauses are constituents, because some small clauses, unlike those in (34), do appear in clefts and pseudoclefts, as observed by Pollard & Sag 1993.

(35) a. What we saw was Michelle drunk.
    b. It was Bill angry that we feared most.

Various additional arguments can be tendered, for example based on the distribution of expletives (the subject of the small clause may be an expletive), binding facts (the subject of the small clause behaves like a subject for binding purposes), extraction facts (extraction from the subject of the small clause is degraded, as it is with subjects). For discussion of these arguments, see Stowell 1981, 1983, 1991a; Heycock 1991; Aarts 1992. Another argument, based on coordination facts, is discussed in Chapter 2.

1.2.2. The head

Williams (1983) noted that the small clause predicate is a maximal projection. Thus, it can have a specifier, as in (36), and it can undergo wh-movement, as in (37) ((36a) and (37a) are from Williams 1983).

(36) a. John considers Bill Bob’s friend.
    b. The police consider him the city’s most dangerous criminal.
    c. The incident made him the city’s most popular criminal.

(37) a. What does John consider Bill?
    b. How do you want your eggs?
    c. How famous did the incident make the criminal?

This is a serious problem for Stowell’s analysis, in which the predicate of the small clause should be an X’, a projection almost universally acknowledged to be inert for movement (an exception is Rouveret & Vergnaud 1980). One solution is to postulate, as Koopman & Sportiche (1988) do, a third bar-level; the predicate is X”, and the small clause is X” (Koopman & Sportiche’s X^max). But this begs the question of what the formal properties of the various projections are; if a non-maximal projection can move, how can we tell, in the usual case, what has moved? And if the small clause subject is located under X”’, why is the possessor of a noun phrase located under X”? These questions do not arise if we adopt the analysis of the small clause as having a functional head. The moved constituent in (37) is the complement of the functional head, and like all complements it is a maximal projection. The functional head has a specifier, and like all functional specifiers it is a landing site for phrasal movement, and perhaps also a licit site for the base-generation of subjects.

One of Stowell’s (1981, 1983) arguments against a functional projection in the small clause was that a verb can select the category of the small clause (see Kitagawa 1988, Pollard & Sag 1993 for discussion). However, the analysis of non-local selection that I develop in §2.3 defuses this argument, and allows the verb to select the category of the predicate across the intervening functional head.

1.2.3. The nature of the head

For Bowers, the head of the small clause (Pred) is a kind of function which takes a property (a basic type, in Property Theory) and returns an unsaturated one-place predicate. I will adopt this position as well. It has the advantage that the head of the small clause has a crucial place in the interpretation of the construction; previous analyses have frequently posited a null head, on the basis of indirect evidence for its existence, but usually without being able to characterize the nature of that head except negatively (e.g. as a ‘defective Infl’).

Actually, Pred could be seen as a kind of Infl, or perhaps, Infl could be seen as a kind of Pred. Infl is the locus of various features which are crucial to the interpretation of the sentence, including at least Tense; in some works, Infl also hosts Negation, Aspect, Modality, and/or subject Agreement features. If Predication is what combines the subject with its predicate, then presumably Infl performs this function as well (alternatively, these various features could be separated into distinct layers, as in Pollock 1989, Chomsky 1991, 1993). The location of Neg in Infl is particularly suggestive, if Pred is Aristotle’s affirmation or denial; this would mean that Neg is really just one polarity of Pred. This would
return us to something like the earlier view, noted above, that small
clauses are IPs, but for different reasons.

The picture turns out to be slightly more complicated, however.
Diesing 1992, building on earlier work by herself and work by Kratzer
1989, suggests that there are two types of Infl, to account for certain
contrasts between sentences with stage-level predicates and sentences
with individual-level predicates (in the sense of Carlson 1977); these
concepts will be discussed in more detail in Chapter 2, §3. Raposo &
Uriagereka 1993 suggest, for similar reasons, that there are two types
of small clause head. I will adopt a version of these proposals in Chapter 2.

1.2.4. Summary of §1

The interpretations of Aristotle and Frege of the logic of predication in
natural language are still at the heart of any current analysis. They both
focused on the logic of ‘judgment’ or assertion and denial, and Frege
showed how predicates could be seen as mappings from individuals to
truth-values. Jespersen, with an eye to the richness of the data, pointed
out that the same concept, predication, is central to many linguistic
constructions below the level of the clause, most notably in the small
clause.

GB theory provides a structural framework for stating syntactic and
semantic relations and constraints. Chomsky’s Theta Theory
formalizes the predicate-argument relation, a semantic one, in terms of θ-
roles, and requires that each argument be identified with a θ-assigner. θ-
assignment is subject to structural conditions, intertwining the syntax
with the semantics. The Predication Theory developed by Williams,
Rothstein, Heycock, and others separates predication from θ-assignment,
but does not make predication independent of structure; instead,
structural constraints are placed on predication which are determined
independently of the structural constraints on θ-assignment.

When predication is identified with a functional head, as in
Bowers, it is possible to maintain the Predication Theory distinction
between predication and θ-assignment, but to unify the structural
conditions on predication and θ-assignment, in the spirit of Chomsky and
Stowell, inter alios; for example, if we assume that Pred (or Infl) can in
some cases assign a θ-role (as argued by Diesing 1992), then we can
maintain the Theta Criterion in the face of Heycock’s data (though the
LISH will have to be abandoned, as Pred and Infl are not lexical heads).
Discussion of this move will be taken up in Chapter 2.

In addition, the predicator can be tied back to the earliest works on
predication; it can be used to capture the distinction Frege was concerned
with regarding predicates and properties (as argued by Chierchia &
2. Head and Complement

The material in this section is in some ways a further development of some ideas I presented in Svenonius 1993d. There, I explored a mechanism, which I identified with Chomsky’s 1965 Strict Subcategorization, which allows a head to determine certain kinds of features on its complements. In that paper I attempted to unify various sorts of head-complement relations under feature-checking of heads in derived head positions at LF, in the spirit of Chomsky’s 1993 analysis of the relation between T[ense] and V in English. Here I will take a weaker stance, distinguishing head-movement from another kind of head-head relation (head-chain formation) subject to very similar configurational requirements. I do not rule out the possibility that the two mechanisms can be unified, or even that the analysis in Svenonius 1993d is incorrect; but using head-chains instead of abstract head movement turns out to make the analysis here smoother in several respects; one problem in particular is avoided, a problem with coordination which I discuss in §2.1.3.

The section is divided into two main parts: §2.1 on selection, and §2.2 on the mechanism of head-chain formation.

2.1. Subcategorization and Selection

2.1.1. History

Chomsky 1965 proposed a system of STRICT SUBCATEGORIZATION, whereby a head specifies categorial features on its complements, and a distinct system of SELECTION, which was more semantic in nature and included the subject as well as the object of a verb. Grimshaw 1979, 1981 further motivated the distinction, showing that some verbs allow different syntactic categories with the same semantic function (e.g. ask takes a ‘question’ complement which can be DP or CP), while other verbs allow more than one kind of complement semantically but restrict it syntactically (e.g. care selects a proposition, exclamation, or question, but only subcategorizes for CP and not DP). Noting that proposition-selecting verbs almost always allow CP complements, but not always DPs, Grimshaw proposes that each semantic class has a CANONICAL STRUCTURAL REPRESENTATION (CSR), and that verbs generally allow the CSR of their selected class by default.

Pesetsky 1982, 1992 introduces the terms C-SELECTION (C for category) for Chomsky’s subcategorization and S-SELECTION (S for semantic) for selection. He proceeds to argue that c-selection can be eliminated, or at least severely reduced; he argues, for example, that whether a proposition-selecting verb appears with DP or not can be determined by whether it assigns Case or not, a property which Pesetsky argues is distinct from c-selection. For example, the failure of wonder and pretend to allow DP complements in (38) can be said to follow from their being specified as non-Case-assigners.18

(38) a. I wonder [what his purpose is/*his purpose].
   b. They pretended [that there was no problem/*the absence of a problem].

The question of whether c-selection for major syntactic categories exists is largely peripheral to my purposes; what is crucial is that c-selection for morphosyntactic feature specifications does exist. I think the issue is sometimes miscast as a question of whether there is ‘syntactic’ as opposed to ‘semantic’ selection; certainly, there are generally ‘semantic’ ways to characterize the sorts of complements that many verbs take. But the fundamental asymmetry between the kind of selectional restrictions that a head can place on its subject and the much greater degree of influence that a head has over its complement show that some sort of split between complements and subjects must be maintained. A particularly

18 Rothstein 1992 argues, contra Pesetsky, for c-selection, arguing for Burzio’s 1986 position that the property of accusative Case-assignment follows from properties of argument structure independent of subcategorization (viz. whether the verb has an external argument or not); she points out that verbs which for Pesetsky must be lexically specified as non-Case-assigners do assign Case in Resultative and Cognate Object constructions. Pesetsky 1993 is a brief reply to Rothstein’s paper, suggesting that the distinction between structural and inherent Case can be used to account for Rothstein’s cases of apparent mismatches between Case assignment and c-selection for DP. His arguments do not account for the fact that the selectional restrictions exhibited in examples like those in (38) persist even when the verbs are passivized, as seen in [i] (cf. also Rothstein 1992:123 for related remarks).

[i] a. What exactly his purpose was there was often wondered by the students.
   b. * His exact purpose was often wondered by the students.
   c. That there is no problem is often pretended by parents of drug addicts.
   d. * The absence of a problem is often pretended by parents of drug addicts.

This contrasts with the absence of non-derived verbs which require that their subjects be CPs (pace Sag & Pollard 1988; I discuss their putative counterexamples in Svenonius 1993b). What [i] shows is not so much that pretend and wonder c-select CP, but that Case-assignment is inadequate to distinguish them from DP-taking verbs like want, desire, feign, fake, and so on. It may be that, for example, pretend c-selects a category like ‘irrealis,’ which might be a possible value of some feature of C, but not a possible value for any feature of D, and that rules out DPs as complements to pretend.
clear case of this is the selection by a head for a particular preposition in its PP complement, for example as in (39) (cf. Williams 1981a:112).

(39) a. Max insists {on/*at/*to/*by/*with} his innocence.
    b. Rachel sticks {on/*at/to/by/with} her story.
    c. Geoff clings {*on/*at/to/*by/*with} false hopes.
    d. Sidney grasps {*on/at/to/*by/*with} straws.

Although verbs may appear with PP subjects (cf. §3 below), no verb requires a PP subject, nor does any verb specify a particular P in case it appears with a PP subject. Pesetsky 1992 calls the selection by a verb for a particular preposition L-SELECTION (L for lexical), which he suggests is the ‘residue’ of c-selection; he suggests that l-selection also allows a verb to select a particular complementizer (much as in Baltin 1989), which enables it in effect to determine finiteness and other features.

In the next subsections, I will develop an account for several different cases of a head determining features on its complement; the phenomenon of control over complements will be seen to be not entirely homogeneous, though certain principles unite the different cases. I will continue to use Pesetsky’s term c-selection to refer to a particular kind of influence a head has on its complement, as opposed to s-selection, which a head has on all of its arguments (s-selection can be thought of as influence a head has on its complement, as opposed to s-selection, which holds only of complements, was originally motivated by certain subject-object asymmetries. A verb always allows a DP subject, and if semantics permit, other categories may appear there as well. The kinds of restrictions a verb places on its subject are a subset of the kinds of restrictions a verb places on its complement; most notably, there are the interpretational restrictions associated with θ-roles, in that a verb may specify lexically that its subject will be interpreted as an Agent, an Experiencer, and so on. This is commonly represented (cf. Williams 1980, Stowell 1981) by associating with each verb a θ-GRID, a list of roles corresponding to the different arguments; alternatively, a system of lexical representation can be developed in which roles are determined by

relations to lexical primitives (e.g. in Jackendoff 1991, where an Agent is the first argument of the lexical primitive DO, a Theme the first argument of BE AT OR GO TO, and so on). There may be other sorts of s-selectional restrictions as well, for example the intransitive verb *feed may require its subject to be non-human (as in Chomsky 1965), or *disperse may require its subject to be plural or aggregate; alternatively, such matters could be separated from s-selection, perhaps regarded as part of the CONTEXT for felicitous use, as in Pollard & Sag 1993.

C-selection, on the other hand, allows a much wider range of feature specifications. For example, consider the complements to what might be broadly termed ‘causative’ verbs, for example make, get, and have in (40).

(40) a. Kate made the truck {start/*warm/*faster/*started/*using propane/*a creampuff}
    b. Ellen got the truck {*start/warm/*faster/started/using propane/*a creampuff}
    c. The instant the police arrived, Sam had the truck {start/*warm/*faster/started/using propane/*a creampuff}
    d. By the time the police arrived, Sam had the truck {*start/warm/*faster/started/using propane/*a creampuff}

I will assume that each of the verbs in (40) takes a small clause complement (I am deliberately using an inanimate DP as the subject of that small clause in order to rule out a possible control structure). For make, the predicate in that small clause can be a verb, or an individual-level adjective or noun phrase, but not a passive participle or a progressive participle. Get in (40b) allows a stage-level adjective, a passive participle, or a progressive participle, but not a noun phrase or a verb. Have in (40c) allows a verb when it has a dynamic interpretation, but not in (40d) where it has a stative interpretation. Many of these categories can appear as subjects, as I will discuss in detail in §3 (e.g. Using propane could damage the truck), but no verb exhibits the kinds of selectional patterns shown in (40) over its subject.

A similar point is made by the ‘aspectual’ verbs begin, start, continue, and keep in (41).
The semantic similarity of begin to start and continue to keep suggests that they should have similar s-selectional properties; but start and keep take small clauses while begin and continue do not; they take DPs (Francis began the game, Beth continued the job) and infinitivals (Francis began to play, Beth continued to work). In fact, start, when it takes a small clause, appears to require an ‘ING’ predicate. Once again, no verb exhibits such restrictions on its subject.

As another example, compare wish with want, where wish takes a finite CP complement, but want takes only a nonfinite CP. Similarly, understand and consider, but not realize, appear with an infinitival complement; for many speakers (myself included), consider does not occur in this sense with a finite CP.

There are no verbs that specify the finiteness of a CP subject. Irrealis modality is often necessary to license a for CP, and may prohibit a that CP, and similarly a factive context may prohibit a for CP which is inherently non-factive, but the verb does not exert the kind of arbitrary influence on its subject that creates the pattern of complements in (42). In (43a-c), different types of subject appear with a range of predicates consistent with propositional subjects; the modal would is used to adjust the modality in order to allow the for CP in (43a).

(43) a. For Pippi to defeat the pirates would {defy comment/bother the captain/suffice to impress Mr. Nelson}.

b. That Pippi defeated the pirates {defied comment/bothered the captain/sufficed to impress Mr. Nelson}.

c. Pippi’s defeating the pirates {defied comment/bothered the captain/sufficed to impress Mr. Nelson}.

d. Pippi’s victory {defied comment/bothered the captain/sufficed to impress Mr. Nelson}.

Of course, some verbs do not take propositional subjects, and therefore are illicit with the clausal subjects in (43a-c), as Vendler 1967 pointed out; for example occur, take place, begin, and so on; to say that a proposition occurred is as incoherent as saying that quadruplicity drinks or ideas sleep. This does not necessarily mean that propositions and events can be s-selected; but it means that having a duration can be s-selected, just as being conscious can (e.g. when a verb assigns an Experiencer θ-role). But nothing like the pattern seen in (42) occurs with subjects; finiteness, subjunctive, and so on are not s-selectable features, though they must be c-selectable.

In Chomsky 1965, such cases of c-selection were modelled by assigning to each verb a SUBCAT[eorization] frame, where the various restrictions that that verb places on its complements are specified (see Pollard & Sag 1993 for a particularly detailed treatment). The subcat frame, or list, can be seen as placing restrictions on sister nodes; a V’ headed by a verb with the subcat value <CP> must contain a CP as its only non-head daughter; or a verb with the subcat value <irrealis> must be sister to a node with the feature irrealis (or perhaps the value irrealis for the feature Modality); any other element will have the result that that V’ is ill-formed. In this conception, c-selection is local in a very strict sense. A slight variation on this is to assume that c-selection is a relation between a head and the head of its complement, as in Baltin 1989; allowing only head features (in the sense of Gazdar et al. 1985) to be specified on the subcat list has a nearly equivalent effect (as discussed in Svenonius 1993d).

This conception of c-selection makes a straightforward prediction: there should be heads which place c-selectional restrictions on a derived subject, such as the subject of an unaccusative or passive verb. I claimed above that a verb could c-select a for CP or a that CP, but that s-selection could not distinguish the two. If this is correct, then there might be
certain passive and unaccusative verbs which allow *that* CP subjects but disallow *for* CPs. This turns out to be the case, as seen in (44).

(44) a. That Pippi could defeat the pirates was [clear/improbable/likely/certain].
   b. * For Pippi to defeat the pirates {was/would be} [clear/improbable/likely/certain].
   c. That Pippi had defeated the pirates was [widely/often] [known/believed/rumored/asserted/claimed].
   d. * For Pippi to defeat the pirates {was/would be} [widely/often] [known/believed/rumored/asserted/claimed].

Following Bresnan 1972 and Cinque 1990b, I assume that the clausal argument to the adjectives in (44a) originates as a complement. This is clearly correct for the passive verbs in (44c) and the unaccusative verbs in (45).

(45) a. That Pippi had defeated the pirates came up in conversation.
   b. That Pippi defeated the pirates came to light this morning.
   c. That Pippi defeated the pirates will never appear in the history books.
   d. * For Pippi to defeat the pirates might come up in conversation.
   e. * For Pippi to defeat the pirates would have to come to light.
   f. * For Pippi to defeat the pirates would never appear in the history books.

The identification of unaccusative predicates in English is, of course, a tricky matter; but *come* and *appear* are fairly uncontroversial candidates. In contrast, every verb which is clearly not unaccusative and allows a CP subject fails to prohibit a *for* CP; this holds not only of the factive predicates in (46a-b) but also the psychological predicates in (46c-d) (cf. also (43) above).

(46) a. That Pippi can lift her horse {proves she’s strong/demonstrates her strength/clears matters up/settles our bet}.
   b. For Pippi to lift her horse would {prove she’s strong/demonstrate her strength/clear matters up/settle our bet}.
   c. That Mr. Nelson can read {astonished Anniken/dazzled Tommy/tittilated everyone}.
   d. For Mr. Nelson to be able to read would {astonish Anniken/dazzle Tommy/tittulate everyone}.

Thus it seems that there is a c-selectable feature that distinguishes *for* CPs from *that* CPs, but there is no such s-selectable feature. Furthermore, it seems that s-selectional restrictions are preserved under A-movement to subject position.

Another case of variation among complements that is not seen among subjects is the presence or absence of the complementizer. Certain heads allow the complementizer to be absent, as in (47a-c), while others require it to be present, as in (47d-f) (from Doherty 1993:55).

(47) a. I say (that) this is true.
   b. I’m sad (that) he’s not here.
   c. It’s a pity (that) he left.
   d. He chuckled *(that) you were mistaken.
   e. I’m flabbergasted *(that) he said that.
   f. It’s a tragedy *(that) they died like that.

This is sometimes characterized as a matter of c-selection (e.g. Bresnan 1972); the heads in (47a-c) can be seen as c-selecting IP, while other heads, such as those in (47d-f), c-select only CP (or possibly the CPs in (47d-f) are not c-selected at all; cf. Grimshaw 1990, Doherty 1993 for discussion). However, this is never preserved under A-movement to the subject position; the examples in (48) are quite impossible, though they would perfectly acceptable with the complementizer.

(48) a. * This is true has often been said.
   b. * He’s not here is sad.
   c. * He left is a pity.

This suggests, given the preceding discussion, that the possibility of complementizerless complements in (47a-c) is not a function of c-selection. In the account developed below, it will be seen that there is a close relation between c-selection and the kind of head-complement relation that allows complementizerless clauses, but the differences
between the two require some additional machinery beyond the device of c-selection.

Before moving on to the next subsection, I would like to make one more observation about c-selection. Recall that c-selection has always been cast as a local relation; either it holds between a head and its complement, or perhaps between a head and the head of its complement, but it reaches no further than that.

However, I argued in §1 that small clauses are headed by a functional head Pred, which takes a complement corresponding to the predicate of the small clause. If verbs can c-select the predicate of the small clause, as I suggested above (following Stowell 1981 inter alia), then the verb exerts c-selectional influence on something which is not its sister nor the head of its sister. One possibility is that there are several different types of Pred (so that a verb can c-select the Pred that takes an ING complement, etc.). This would introduce a multiplicity of kinds of Pred that is not independently motivated (recall that I suggested that there are at least two kinds of Pred; but they will be seen not to line up evenly with the categories of their complements). The other possibility is that the head of the complement moves into Pred, which brings it into a local enough configuration for c-selection. This requires that c-selection is not a D-structural relation, as in Chomsky 1981, but can be satisfied after D-structure, as in Larson 1988. In fact, I will assume below that c-selection is checked only at LF.

2.1.3. F-selection and head movement

Given that there is some mechanism of c-selection that allows a verb, an adjective, or a noun to specify various kinds of features on its complements, we might assume that the same mechanism accounts for the fact that C takes an IP complement, that Infl takes a VP complement, that D takes an NP complement, and so on; that is, the complementizer that will have a subcat value <IP,FIN> and the complementizer for will have a subcat value <IP,FEN>. Abney 1987 pointed out that the selectional variation for functional heads is more restricted than for lexical heads; in general, functional heads only appear with one category of complement; Abney coined the term F-SELECTION for this sort of selection.

There are some possible exceptions to Abney’s claim that functional heads only appear with a certain category of complement. First, under Abney’s own analysis of gerunds, the possessor determiner must be able to select either NP or VP (to allow both his quietly singing the Marseilleise and his spatula), while most determiners select only NP (*every quietly singing the Marseilleise). Second, if we adopt Emonds’ 1985 proposal that P and C are the same category, then that functional meta-category consists of elements that select DP (the prepositions) and elements that select IP (the complementizers). A case in point would be during and while, which are arguably distinguished only by their subcat value; an element like before allows both types of complement. If it is accepted that prepositions are a functional category, then the possibility of small clause complements to with and without but not at and by also count as counterexamples. Finally, if we follow Abney in taking degree words like too and very to be functional heads (of the category Deg), and accept Rothstein’s 1992 arguments that too subcategorizes for a clausal complement (in too purple for us to eat), then it seems that Deg too must validate subcategorization frames, since e.g. very does not allow a clausal complement (*very purple for us to eat).

However, there are other cases of what Abney called f-selection for which it is probably unnecessary to posit a selectional restriction, given certain other assumptions. One such case is the selection of VP by Infl. Assume, following Emonds 1976, Chomsky 1981, and Pollock 1989, that V movement to Infl in languages like French is driven by the need for the verb to acquire or check inflectional features. There are two ways in which this could be wrought into a restriction on the distribution of Infl: one is to say that Infl needs to discharge its inflectional feature values, and must do so onto a verb. Another is to take the Minimalist route (Chomsky 1993) and say that Infl serves no other function than to check features on V, and so it will not be generated if there is no verb with features that must be checked, by principles of economy of representation. For concreteness let us take the first route, and say that in English the crucial feature hosted by Infl and necessarily discharged onto V is called Tense. We can say both that Infl must discharge Tense and that V must receive Tense.21 For the moment, we can follow Chomsky 1993 in assuming that the verb and Tense always combine under headmovement, which occurs at LF in English; I will return to this below. Furthermore, we can assume that even non-finite Infl has this abstract Tense (as in Stowell 1982) and imparts it to infinitival verbs; this will ensure that in the usual case, non-finite verbs do not occur without Infl (I will return to bare infinitival constructions in §3). This particular approach to the distribution of Infl and C will be further motivated in §3.

On this approach, a lower head (e.g. V) needs some features that it can only get from a higher head (in this case Infl). While this makes sense

---

21 It has been argued that Infl in some languages can appear with different categories (e.g. Aissen 1992 for Tzotzil, Doherty (to appear) for Irish), so in those languages either Infl can discharge its features onto other categories than V, or Infl does not need to discharge its features.
for f-selection, it would make less sense for c-selection. We saw that DPs and CPs can generally appear as subjects, which suggests that they do not need to undergo head-movement in order to acquire any particular features; yet their distribution as complements is restricted. However, given the notion of “obligatory discharge” introduced above, the cases that I have been calling c-selection could be subsumed under f-selection as described for the Infl-V relation; say that a verb that c-selects a finite CP, such as think, needs to discharge some features that it can only discharge onto a finite CP. This would have the same effect (some allowance would have to be made for optional complements). This has the positive effect that the overall number of grammatical mechanisms is reduced, since it is not necessary to posit a subcategorization relation in addition to the head-head relation.

Now I will briefly address the question of whether such a relation should be seen as being identified with abstract head movement or not. It has been suggested, e.g. in Baker 1988, Stowell 1991a (discussed further in Chapter 2), and den Dikken 1992 (discussed further in Chapter 3), that certain relations among non-incorporated heads can be seen as being due to abstract incorporation; Baker proposes that the Romance causative verb is co-superscripted with the head of its complement at S-structure, and that certain S-structural effects are licensed by this co-superscripting, which in turn is licensed by LF incorporation. The proposals of Stowell and den Dikken are similar in relevant respects. In Chomsky 1993, the Tense feature value on the finite verb in English is checked at LF when V moves into Infl (Chomsky’s T); thus, there, too, a surface effect (overt morphological tense inflection) is licensed by LF head movement. There is a problem for all of these analyses, one which may be very serious. A well-known constraint on coordinate structures is the COORDINATE STRUCTURE CONSTRAINT (CSC) of Ross 1967, which requires that an element not be moved out of only one conjunct. This accounts for the ungrammaticality of (49).

(49) a. * Has [John _ gone] and [Mary will leave]?
   b. * What how loudly did Mary [notice _ ] and [laugh _ ]?

When the same moved element binds gaps in both conjuncts, as in (49c-d), the sentences are grammatical (Ross 1967 called this ACROSS THE BOARD MOVEMENT, or ATB; cf. also Williams 1978).

(50) a. Will [John _ go] and [Mary _ stay]?
   b. What did Mary [say _ ] and [laugh about _ ]?

In fact, there are many counterexamples to the CSC; examples very much like that in (49b) are often considered acceptable. But one kind of
But of course for the main verbs in (54) to move into Infl, as Chomsky’s analysis would require, would be a violation of ATB.

One conceivable solution would be to resurrect the machinery of Conjunction Reduction, assumed in early generative work to derive sentences like those in (52) from coordinated sentences (e.g. One person quietly snuck into the cafeteria and one person quietly stole some popcorn). However, there are serious obstacles to such an analysis because the reduced construction does not have a reading that the coordinated sentences have (consider the problems with subjects like the same person, or different people, cf. Partee & Rooth 1983, Sag et al. 1985). In §2.2 I will show how sentences like (52) can be interpreted without postulating Conjunction Reduction and without allowing ATB violations at LF, but still capturing the similarity between languages like French, in which V moves to Infl, and languages like English, in which it does not.

2.2. Head Chains

Abstract relations between heads are often posited, for example in analyses of complex predicate formation (e.g. Rizzi 1978 for an account of Romance clause-union that does not involve movement; or van Riemsdijk 1978, Hornstein & Weinberg 1987 on pseudopassives, Christensen 1991 and references cited there on Scandinavian complex passive constructions, Williams 1980 on (subject-subject) raising as function composition). Such abstract relations are often described as ‘co-superscripting,’ following Rouveret & Vergnaud 1980; in this section I will develop, following work by den Besten and Sigurðsson, what might be thought of as a formal theory of co-superscripting.

Den Besten 1985 and Sigurðsson 1990 develop a formal theory of HEAD-CHAINS that allows certain featural specifications to be shared among heads in certain configurational relations. For example, consider the Icelandic Dative-Nominative structures in (55) (from Sigurðsson 1991:334).

(55) a. Okkur höfðu leðist strákarnir.
   ‘We had been bored by the boys’
   (Ice)

b. Okkur höfðu verðið sagðar sögurnar áður.
   ‘We had been told the stories before’

A Dative-Nominative structure is one in which the subject gets dative Case and the object nominative, in a finite clause; this happens in Icelandic with certain experiencer verbs, as in (55a), and certain passive verbs, as in (55b). The finite verbs in (55) agree in person and number with the nominative objects, rather than with the dative subjects. Following Sigurðsson, I will assume that a head-chain is formed from the finite Infl to the embedded verb, endowing it with the ability to assign nominative Case. The head-chain can be represented by co-superscripting the various heads, as in (56) (I use an embedded clause in (56), to avoid the complication of the CP level; I also ignore possible SpecVPs and the base-location of the dative DP okkur). Assuming that höfðu has moved from V to Infl, I identify its trace as a member in the head-chain.

(56)
features (notice that sagðar itself is third person feminine plural nominative, in agreement with súgurnar); alternatively we might assume that only the final element in the chain inherits the initial member’s Case-assigning ability; this will be taken up below.

This is a different phenomenon from the impersonal construction, where an expletive appears in subject position, as in (57)(from Sigurðsson 1991:350; the two sentences are synonymous).

(57) a. það mundu kannski einhverjir bátar hafa verið keyptr. (Ice) there would perhaps some boats have been bought
b. það mundu kannski hafa verið keyptr einhverjir bátar. there would perhaps have been bought some boats

‘There would perhaps have been some boats bought’

The DP einhverjir bátar ‘some boats’ is in the nominative case, as indicated. Following Safir 1982, an expletive-argument chain can be formed from the expletive to the embedded argument in such cases, transferring Case to the argument and agreement features to the expletive. Safir 1987 argues that this accounts for the definiteness effect in impersonal constructions (e.g. the DP einhverjir bátar in (57) could not be replaced with the definite báturinn ‘the boat’). But in (55) there is no expletive, and no definiteness effect.

Following Sigurðsson, head-chains are subject to conditions of government, relativized minimalism, and barriers; the primary observation to be captured is that head-chains are formed among heads in a head-complement configuration, and do not extend into specifiers or adjuncts. I will propose an explicit definition of head-chains below, slightly simpler than (but essentially equivalent to) that of Sigurðsson. Rather than using government, I will cast it in terms of LOCALITY, as defined in (58). The reason for this is that government, as conceived of in GB theory, is connected with matters of barrierhood and minimalism, and these notions play no role in the constructions I will be examining.

(58) Definitions of LOCALITY and STRICT LOCALITY

a. X is local to Y iff X is a head and some projection of Y is a sister to X.
b. X is strictly local to Y iff X and Y are sisters.

Two nodes are sisters if and only if there is no node dominating one of them but not the other. Strict locality is simply sisterhood (Barker & Pullum’s 1989:13 IDC-COMMAND), and is therefore symmetric, but locality is not symmetric. The restriction of X in (58a) to a head means that in practice, Y will always be some projection of a complement of X, or else it will be a head adjoined to Y. Subcategorization was cast above as a strictly local relation, i.e. a condition placed by a head on its sisters. Subcategorization for heads could be recast as a local relation, given the definition in (58).

For the purposes of discussion I adopt the simple restriction on head-chains in (59) as a partial definition (this will be refined later; see also Sigurðsson for a more detailed proposal).

(59) Condition on head-chains (preliminary):

An ordered n-tuple of heads <α₀, ..., αₙ> is a head-chain only if, for every i, 1 ≤ i ≤ n, αᵢ is local to αᵢ₊₁

In (56), the head-chain includes both auxiliaries, the trace of höfðu, and the passive participle. The terminal element in the chain (the terminus, in Sigurðsson’s terms) counts as the nominative Case assigner, and since it governs the DP, it may assign Case to it.22 A head-chain which enables a terminus to be a Case-assinger can be called a CASE-CHAIN; for Sigurðsson, a head-chain which allows the non-local control of a feature is called an F-chain, where F can be used generally or replaced with the name of the specific feature so controlled.

The formation of head-chains must be constrained. For example, we do not want to allow a noun which is the complement of an accusative-assigning verb to acquire the ability to assign accusative Case. In fact, it seems that only verbs and prepositions ever enter into Case-chains; we could restrict membership in an F-chain to potential F-assigners, meaning that only if X is a member of a category that assigns F can X enter into an F-chain; for example, we could say that the -N categories of Chomsky 1970 are potential Case assigners.

At this point it becomes crucial to distinguish between FEATURES and FEATURE VALUES. A FEATURE SPECIFICATION, in the sense of Gazdar et al. 1985 (especially Chapter 2), is an ordered pair of a feature and a feature-value, for example <Case, Acc>. This is usually represented [Case; Acc] or even just [Acc] (since only the feature Case can have the value Acc). If a syntactic element, say a noun, has a feature specification for the feature Case, this means that it has both the feature Case and a feature value appropriate for that feature. An element cannot have a feature value without having the feature of which it is a value (e.g. a noun cannot be Accusative and not have Case); however, an element

22 Actually, for Sigurðsson, the entire chain counts as the nominative Case assigner; for my purposes it is more useful to speak of the terminus as acquiring properties of the higher head.
can have a feature without having a value for that feature, in which case it does not have a feature specification. For example, we might suppose that it is a lexical property of nouns that they have the feature Case; but any value for that feature must be licensed by some Case-assigner. A predicate nominal, then, in a non-Case-assigned position, has the feature Case but no value for that feature.

Now, consider what it means for a verb or preposition to be a Case assigner. We could say that this means that that element has the feature specification [Case-assignment: +]. It will be important to distinguish non-Case-assigning verbs, which bear the feature ‘Case-assignment’ but have no value for it, from such things as nouns and adjectives, which do not bear the feature at all. Following Pesetsky 1982 we could say that whether a verb has a value for the feature ‘Case-assignment’ is stipulated for each verb; or following Burzio 1986 we could say that it follows from whether the verb assigns an external θ-role or not. Note that a Case-chain is really a Case-assignment-chain, since the feature value being passed down the chain is not a value for the feature Case, but a value for the feature Case-assignment. I will continue to call the chain a Case-chain, following standard usage.

Now we can define the notion ‘potential F-assigner’: a potential F-assigner is one that has the feature ‘F-assignment’ (and an actual F-assigner has a value for that feature). All members in a Case-chain must have the feature ‘Case-assignment,’ and the Case-chain has the effect of specifying the value for that feature in (at least) the bottommost member to match that of the topmost member.

But feature-chains are still more restricted. We do not see Case-assigning prepositions entering into Case-chains to assign nominative Case to their complements (i.e. Icelandic does not have Dative-Nominative structures where the nominative noun phrase is in a prepositional phrase). In fact, it seems that only non-Case-assigning verbs and prepositions enter into Case-chains (this will be discussed more in Chapter 3). Thus we want to say that any non-initial member of an F-chain must be a potential, but not an actual, F-assigner; or in the terms suggested above, the feature F must be unspecified on all non-initial members in the F-chain.

23 For languages which distinguish morphological cases, the feature Case-assignment might be a multi-valued feature. I used the value ‘+’ rather than, for example, ‘Acc,’ in order to avoid confusion with the values of the feature Case, which must be kept distinct.
Note that the condition on head-chains stated in (60) excludes chains into adjuncts. However, this may be something different, as Thráinsson shows examples of subjunctives in adjunct clauses controlled by a higher subjunctive Infl. In such cases, there is no overt morphological indication of Tense. Consider the contrast in (64).

(64) a. Tintin believes Captain Haddock was drinking.
   b. Tintin believes Captain Haddock to be drinking.

It is impossible to construe the embedded clause in (64b) as being evaluated at a time different from the utterance time. The present tense in the matrix clause necessarily binds the Tense of the lower clause. If the interpretation of the embedded clause requires a value for the Tense feature, then a Tense value must be supplied, so a Tense-chain is obligatorily formed in (64b) and the embedded Infl receives the same specification as the higher Infl. This will be expanded in Chapter 2 below.\(^{25}\)

Now recall the problem raised in §2.1.3 above with respect to LF head movement. There, it was pointed out that under Chomsky’s 1993 proposal, V moves to Infl at LF in English; the problem was that this results in ATB violations (in sentences like Westley sings and dances). Now, head-chains are clearly conceptually very similar to head movement; it is no accident that they are subject to the same configurational restrictions (this is especially clear in Sigurðsson’s formulation, less so in mine only because I used the notion of locality rather than government). If it is true that head-chains can control Tense features, then we can assume that V does not move to Infl in English at all. Instead, a Tense-chain is formed from Infl to V, which has the result that Tense feature values are specified on V. Since we allow feature

\(^{24}\)A similar relation may hold for the mood value: the clausal complement of a verb that heads a subjunctive clause may also be subjunctive (Thráinsson’s 1976, 1990 ‘domino effect’). However, this may be something different, as Thráinsson shows examples of subjunctives in adjunct clauses controlled by a higher subjunctive Infl. Note that the condition on head-chains stated in (60) excludes chains into adjuncts.

\(^{25}\)Note that using the perfect auxiliary have in the embedded clause does not actually make the tense of the embedded clause independent of the tense of the higher clause, e.g. in Tintin believes Captain Haddock to have been drinking) Although Captain Haddock’s drinking, in such cases, may have occurred prior to Tintin’s belief, the Captain’s state of having been drinking is simultaneous with Tintin’s belief.
specifications which are not morphologically overt, we can say that all non-participial verbs have a Tense feature but need a Tense feature specification. This they can only get from Infl. I will return to this in §3 below.

Formation of head-chains into coordinate structures will not result in ATB violations; since the feature values being specified are identical in both conjuncts, the head-chain is consistent with a version of the CSC modified to apply to structures in which no movement occurs; for example, F-chains from a head X may be formed with the heads Y and Z of a coordinate structure iff an F-chain from X to Y is licit and an F-chain from X to Z is licit.

2.3. Summary of §2

I have attempted to establish, in this section, that there is a relation between a head and its complement which goes well beyond the kind of semantic role determination usually associated with s-selection. This relation also goes beyond the simple determination by a head of the major category of its complement; a head can require specifications on its complement for a number of features, including at least such morphosyntactic features as inflectional form but also certain features not necessarily realized morphologically, for example ‘factive’ or ‘irrealis.’ I have used Pesetsky’s term c-selection to refer to this relation.

I have also proposed that the notion of head-chain (as formalized by Sigurðsson) be used to express certain other head-complement relations, including some analyzed elsewhere as involving abstract head-movement. The mechanisms of c-selection and head-chains as laid out in this chapter will be important in the analyses in §3 and in Chapters 2 and 3.

3. Dependency

Constructions analyzed as involving movement have been used since the earliest days of generative grammar as ‘constituency tests,’ based on the quite plausible assumption that only a syntactic constituent is able to undergo movement (equivalently in more recent non-derivational analyses, since the filler linked with a gap is always a constituent). Passing constituency tests must be seen as a sufficient but not a necessary condition on the constituent status of a string, since there are conditions on each constituency test, and on movement in general. In this section, I discuss some of those conditions, and develop a formal notion of DEPENDENCY which accounts for the failure of some constituents to undergo constituency tests.

The basic idea is that certain constituents are semantically ‘incomplete,’ and need something in order to be interpreted. This something will be supplied by a higher head through a head-chain. Developing some ideas of Chierchia’s (1985), semantically incomplete constituents do not correspond to basic types, and therefore cannot enter into anaphoric relations such as are imposed by constituency tests. Thus a dependent element is one that is semantically incomplete.

3.1. The failure of constituents to pass constituency tests

Some constituents pass some tests and fail others. For example, the bracketed PP is a constituent in (65); this is supported by the fact that it can be displaced in a Cleft, as in (65b). But a Pseudocleft will not displace a selected PP, as shown in (65c). On the other hand, a CP can generally be displaced in a Pseudocleft, whereas this is not possible for a Cleft, as shown in (66).

26 Of course, some constituency tests are notoriously difficult to apply with any confidence, for example Coordination (cf. Gapping structures, as in Hankamer 1971) and Right Node Raising (cf. Abbot 1976, Swingle 1993). Such constructions as Clefts, Pseudoclefts, WH-questions, VP-Fronting, and Topicalization (including V2) are still generally taken to be reliable indications of constituency, however.
(65)  a. The quality of the pizza depends [on the mozzarella].
    b. It is [on the mozzarella] that the quality of the pizza depends.
    c. * [What/Where] the quality of the pizza depends is [on the mozzarella].

(66)  a. They think [that pizza is Italian].
    b. * It is [that pizza is Italian] that they think.
    c. What they think is [that pizza is Italian].

In such cases, the constituency of the bracketed string is not at issue; the applicability of a single uncontroversial constituency test is sufficient to establish that a string is dominated by a single node. I will not attempt to account for the restrictions on the Cleft and Pseudocleft constructions (for extensive discussion of the latter, see Higgins 1973). Instead, I will be concerned with constructions in which a string fails to undergo any constituency test, yet in which there are reasons to believe that that same string is a syntactic constituent.

One such example is VP in (Standard) English. The ‘stacked VP’ suggested by the bracketing in (67a) has been widely adopted (see Jespersen 1937/92 (AS), Ross 1969b, McCawley 1971, and Pullum & Wilson 1977 for arguments). However, the usual constituency tests do not seem to support it; for example, the VP headed by the perfect participle resists Pseudoclefting, Clefting, VP fronting, and Though-fronting, as shown in (67b-e).

(67)  a. Elvis has [left the building]
    b. * What Elvis has is [left the building]
    c. * It is [left the building] that Elvis has.
    d. * [Left the building], I suppose Elvis might have.
    e. * [Left the building] though Elvis has, we’ll stand here and yell anyway.

Now, there are two possible accounts for this fact. One is that the peculiar restrictions on the various constituency tests applied in (67) just happen to all be incompatible with the displacement of a VP headed by been. The other is that there is some basic principle at work which prevents (this sort of) VP from being displaced, for example VP might be ‘incomplete’ in the sense that X is incomplete. In the former case, we would expect to be able to find other languages in which VP displacement was possible, whereas in the latter case we would not.

In Norwegian and German, there is an auxiliary verb corresponding to (and cognate with) have, and its perfect participle VP passes constituency tests, as shown in (68).

(68)  a. Elvis har [forlat bygninga]. (Nor)
    b. Elvis hat [das Gebäude verlassen] (Ger)
    c. [Forlat bygninga] har Elvis allerede.
    d. [Das Gebäude verlassen] hat Elvis schon.

Since the past perfect structures are closely parallel (syntactically, morphologically, and semantically) in English, Icelandic, Norwegian, and German, the natural conclusion is that the ungrammaticality of (67b-e) is due to some language-particular constraint, or perhaps even several different language-particular constraints on the Cleft, Pseudocleft, Though-fronting and so on.

There are other cases, however, which call for a more principled account. IP cannot be displaced in any of the Germanic languages.

(69)  a. * Elvis has left the building, I doubt (that).
    b. * It was Elvis has left the building that I doubted (that).

(70)  a. * Elvis har forlat bygninga tror jeg ikke (at). (Nor)
    b. * Elvis das Gebäude verlassen hat glaube ich nicht (daß). (Ger)

In contrast, CP can be displaced in all of the Germanic languages.

(71)  a. At Elvis har forlat bygninga tror jeg ikke. (Nor)
    b. Daß Elvis das Gebäude verlassen hat glaube ich nicht. (Ger)

Since all of the Germanic languages allow IP complements (that is, they allow finite clausal complements with no determiner, which I assume are IP following Bresnan 1972, Webelhuth 1992, and Doherty 1993), and none of them allow those complements to be displaced, we need a grammar which rules out the displacement of IP in principle.

One proposal for the limited distribution of complementizerless clauses was put forth by Stowell 1981. Stowell adopted the position that subordinate clauses are always CPs, and suggested that if complementizerless subordinate clauses contain a null C, then that C should be subject to the ECP (the Empty Category Principle of Chomsky 1981:250, which requires that an empty category such as a trace be
properly governed; see Rizzi 1990 for more recent discussion). The distribution of complementizerless clauses will then be limited to governed positions such as the complements position; the sentences in (70) are correctly ruled out.

However, it is unclear why this null complementizer should be subject to the ECP; assuming that there is also a null complementizer in a controlled infinitival clause, that complementizer must escape this requirement. Furthermore, other null-headed constituents do not seem to be subject to the strict distributional limitations of complementizerless clauses; null determiners, for example, are widely assumed for names and bare plurals (Abney 1987, Stowell 1991b, Longobardi 1992), but determinerless DPs freely appear in ungoverned positions (at least in English; Longobardi 1992 argues that this is not the case in Italian). Doherty 1993:38-42 for a number of cases of ungoverned null complementizers.

Furthermore, as Doherty 1993 points out, non-complement clauses require complementizers even when they are in what are arguably properly governed positions. Doherty notes (p. 34) that Chomsky 1986a takes the relative acceptability of examples like that in (72a) to indicate that the subject of the embedded clause is L-marked, which entails that it does not form a barrier to government. This suggests, as Doherty observes, that the subject is properly governed, which in turn suggests that its head is properly governed; but the complementizer head of the embedded subject in (72b) (from Doherty 1993:35) cannot be omitted, as indicated.27

(72)  a. Who do you consider [friends of _] to be stupid?
    b. I consider [* (that) he is an idiot] to be obvious.

As a final objection to an ECP account for the distribution of complementizerless clauses, note that it is unclear why government should not be satisfied under reconstruction, at LF.28 For example, DPs

---

27 My hedging ("suggests") is due to the fact that there are so many different definitions of proper government. The definition given in Rizzi 1990, for example, states essentially that X properly head-governs Y if X is a lexical head, X c-commands Y; no barriers intervene, and Relativized Minimality is respected (cf. Rizzi 1990:25, 31-32). In (72), consider is a lexical head, consider c-commands that, no barriers intervene between consider and that, and Relativized Minimality is respected (i.e., there is no potential head-governor for that which c-commands that and does not c-command consider; cf. Rizzi 1990 for details). Thus the prediction of the ECP account seems to be that that should be deletable, contrary to fact.

28 Rizzi 1990:38 expressly proposes that head-government is not satisfied under reconstruction, but the data that motivates that stipulation receives an alternative (continued next page)
3. DEPENDENCY

(74) a. Pigs in the kitchen drives him crazy.
   b. Biff and Spike arrested spells trouble for the football team.
   c. Does he fear that Mandela free would pose a greater threat than Mandela behind bars?
   d. These two difficulties overcome makes the rest easy.

In (74a-b) and in Jespersen’s example, the singular agreement shows that the subjects are nexuses and not noun phrases with postmodifiers (as Safir notes). There are some other elements that appear in subject position at least in copular constructions, as shown in (75).

(75) a. PP\_loc [In the kitchen] is the best place to keep light bulbs.
   b. PP\_temp [After dinner] might be a good time to play Scrabble.
   c. PP\_man [With your toes] is a hell of a way to play the piano.
   d. AdvP [Gently] is the only way to handle a tarantula.
   e. VP\_FIN [Cry foul] is what he does best.
   f. AP [Proud of your dissertation] is a great way to feel.

I distinguish among locative, temporal, and manner PPs here simply in order to point out that subject PPs are not, as is often assumed, restricted to locative PPs. In the cases in (75), I think, it is difficult if not impossible to construct grammatical examples without the copula.

(76) a. * In the basement frightened the children.
   b. * After dinner drives him crazy.
   c. * With your toes shows you’re limber.
   d. * Slowly just angered him.
   e. * Win makes you arrogant.
   f. * Proud can get you fired.

This might be taken to suggest that the examples in (75) involve inversion of a predicate with the subject, and that the subject is in each case the postcopular DP; this sort of construction is modelled in (77a-c). However, as can be seen in (77d-f), the copular inversion structure is incompatible with raising verbs (including modals) and negation (cf. Higgins 1973).

(77) a. With the senator was a movie star.
   b. Playing saxophone is Geoff.
   c. In the kitchen is dessert.
   d. * With the senator wasn’t a movie star.
   e. * Playing saxophone turned out to be Geoff.
   f. * In the kitchen might be dessert.

This allows us to test the sentences in (75), and as shown in (78), they turn out to involve non-DP subjects.

(78) a. In the kitchen turns out to be the best place to keep light bulbs.
   b. After dinner might be a good time to play Scrabble.
   c. With your toes seems to be a hell of a way to play the piano.
   d. Gently isn’t the only way to handle a tarantula.
   e. Cry foul appears to be what he does best.
   f. Proud of your dissertation doesn’t seem like such a strange way to feel.

Thus, although these elements are restricted to equational or identification-type copular constructions, they really do behave as subjects in that environment. In this they can be contrasted with some other syntactic constituents which cannot occur in subject position at all, not even with the copula. Some are listed in (79) (the bracketed strings all appear as constituents in complement or adjunct positions).

(79) a. IP\_FIN * [IP can’t be a subject] is true.
   b. IP\_FIN * [Max to do the laundry] would be shocking.
   c. IP\_FIN * [To be intelligent] was seemed by the alien.
   d. NP * [Train] would be nice to arrive by.
   e. VP\_FIN * [Likes Mary] is funny.
   f. SC\_VP * [Mary leave] was seen by everyone.
   g. SC\_NP * [Sally a genius] was widely considered.
   h. Adv\_T * [Quite] would be big enough.

I have included two different infinitival IPs, one with an overt subject, and another contrived so as to have a trace in subject position; it seems that infinitival IP, like finite IP, never occurs as a subject, whether or not that IP contains a subject itself. The impossibility of NP subjects is contingent on the assumption, defended at length in Abney 1987, Stowell 1991b, and Longobardi 1992, that argument noun phrases are always DPs, and if no overt determiner appears there must be a null one; NPs may, however, be predicative, and appear as the complement to the
copula (e.g. Clinton is president; cf. *President is Clinton). Finite verb phrases do not actually occur as verbal complements, but they are constituents and occur as complements to Infl; they cannot be subjects, however. In (79f-g) I have included VP small clauses, which appear as the complements to causative and perception verbs, and DP small clauses; these will be discussed at length in Chapter 2 below, where I will suggest that the AP-PP vs. VP-DP distinction suggested by the above examples is not quite right. Finally, I include AdvP for ‘tertiary adverb,’ using tertiary to mean a modifier of a modifier (for Jespersen, VP adverbials were also teriaries). (79h) might be acceptable as a joke, but is not really grammatical.

Depending on one’s analysis, the direct object-PP sequence following some verbs is a constituent (e.g. Kayne 1984, Hellan 1988, Larson 1988), as is the indirect object-direct object sequence. However, these constituents, if they are such, can never be subjects.

(80) a. DO-PP * [A book to Mary] is rarely given.
b. IO-DO * [Mary a book] is even worse.

Now, it is striking that every constituent that can be a subject can also be displaced as in at least some constituency tests. This is demonstrated in (81-82), corresponding to (73) and (75).

(81) a. DP [Who] did you bark at?
b. CP_{FIN} What is strange is [that he didn’t recognize us.]
c. CP_{FIN} I’d like it very much [for pigs to be in my kitchen].
d. CP_{[PRO]} It was [to throw rocks] that John liked best.
e. VP_{ing} She still throws water balloons, but [throwing rocks], she stopped.
f. Acc What made me think of Alcatraz was [John tearing sheets].
g. SC_{PP} [Mary in Hawaii], we weren’t prepared for.
h. SC_{AP} It’s [workers angry about the pay] that the executives fear.

(82) a. PP_{loc} [In which room] do you keep the light bulbs?
b. PP_{temp} [After dinner] they said they might play Scrabble.
c. PP_{man} [With which toes] do you play the piano?
d. AdvP [How gently] must one handle a tarantula?
e. VP_{FIN} They said he would cry foul, and [cry foul] he did.
f. AP [How proud of yourself] are you?

Other examples can readily be constructed; in most cases, a variety of constituency tests can be successfully applied to the constituent in question. The same constituents and putative constituents that could not be subjects cannot be displaced, as shown in (83-84).

(83) a. IP_{FIN} * It is [snakes are mammals] that he doesn’t believe.
b. IP_{FIN} * [Max to do the laundry], I would really like.
c. IP_{FIN} * What the alien seemed was [to be intelligent].
d. NP * [How fast bicycle] would you like a?
e. VP_{FIN} * He said he sells turtles, and [sells turtles] he really.
f. SC_{VP} * What I saw was [Mary leave].
g. SC_{DP} * It was [Sally a genius] that they considered.
h. AdvV * How big it was was [quite].

(84) a. DO-PP * It was [a book to Mary] that we gave.
b. IO-DO * What we gave was [Mary a book].

The correlation is rather striking. Just the things which cannot have properties predicated of them cannot be displaced. A formal account for this is developed immediately below. I will refer to constituents which can be subjects and can be displaced as INDEPENDENT constituents, while those constituents which are syntactically constituents (at least those phrases in (83) certainly are) but which cannot be subjects and cannot be displaced and which are not interpreted as basic types as DEPENDENT constituents. Since IPs are always nexuses, and always dependent, non-root IPs are always dependent nexuses (as noted in §1, this is not exactly the same sense in which Jespersen used the term ‘dependent nexus’).

Notice that coordination must be excluded as a test for dependency, as there is nothing wrong with coordinated IPs, NPs, tensed...
VPs, and so on. This will follow from the account developed below. In addition, Right Node Raising and Gapping do not line up with the above tests;31 interestingly, both of these problematic constructions involve coordination (and prosody; cf. Swingle 1993). Null complement anaphora, however, does seem to be restricted to independent constituents, though I will not discuss it here.

3.3. The Functor Anaphora Constraint

The idea that I will make use of for accounting for the dependency of certain constituents is that an anaphoric element must be of a basic set theoretic type; this is essentially the idea behind Chierchia’s 1985 FUNCTOR ANAPHORA CONSTRAINT; I state two versions of it in (85) ((85a) is essentially Chierchia’s version, from his p. 429).

(85) FUNCTOR ANAPHORA CONSTRAINT

a. Weak: There are no variable-binding mechanisms for third-order functors

b. Strong: All variables are type e

(85) is a constraint only made possible by an enriched system of types like that proposed in Chierchia 1985 or Chierchia & Turner 1988 (cf. §1.5). The idea is that basic types like type e (entities, including urelements, information units, and nominalized functions) and second-order functors like type <e,e> (predicates)32 are the kinds of things that variable-binding mechanisms can operate on, while non-basic types like <e,t>, <e,t,t> (adverbial modifiers) and <e,t,t,t> (quantifier phrases) are not.33 In an apparent case of functor anaphora such as How does John play chess?, Chierchia suggests that the variable bound in the semantic representation is actually type e; the question corresponds roughly to What x is such that John plays chess in an x manner?34

31 That is, neither the ‘raised’ material in RNR constructions nor the conjuncts must correspond to independently motivated syntactic constituents (cf. Abbott 1976), and similarly neither the missing material nor the conjuncts in Gapping constructions must be syntactic constituents (cf. Hankamer 1971). Of course, there may be non-canonical constituents created by these processes.

32 Since information units are a sort (sort i) of type e, a function from individual to information units, type <e,i>, is also type <e,e>.

33 Chierchia & Turner 1988 assign quantified DPs to a basic sort Q. But assuming that QPs bind type e variables, I won’t need to make use of the sort Q.

34 Specifically, λp[∀x∃P[∀x∀ly’(P)(play chess’)(j)]], where ly’(P)(Q)(x)= x Qs in a P manner, P ranging over properties.

Given such translation rules, it is tempting to take the stronger position in (85b), namely that all anaphors be restricted to type e. Chierchia 1985:430 suggests that VP anaphora, as in Ezio hates Pavarotti and Nando does too, is a case of anaphora of a propositional function (type <e,t>); but notice that do-support retains the tense of the second conjunct sentence, suggesting that the null complement is not a tensed VP but an infinitival VP; in the system of Chierchia & Turner 1988 an infinitival VP translates as an nf, (nominalized function), a sort of e. It seems, then that (85b) will not rule out any cases of anaphora that we want to allow. The question then becomes, what constituents do not translate as basic types?

Empirically, we know which constituents are dependent (non-displaceable), from §3.2. At the risk of letting the tail wag the dog, I will assume that it is those constituents which translate as non-basic types. Consider, for example, IP. If the Functor Anaphora Constraint is to blame for its dependency, then IP must be a Functor. According to Chierchia and Chierchia & Turner, IP standardly translates as type i, which is basic; but noting (in fn. 14, p. 300) that IP never occurs as a subject, they suggest in passing that “we could say that information units (i.e. propositions) are not individuals but have individual correlates, just like functions. The complementizer might be, then, what maps information units onto their individual correlates.” This is exactly the sort of approach that I want to develop here.

Now, in order for the Functor Anaphora Constraint approach to be validated, we must have a well-motivated semantic analysis that assigns basic types to the independent constituents and not to the others. The fact that displaceable constituents appear as subjects is already a partial confirmation of this, in that it suggests that those elements are type e; but the question remains whether analyzing the non-displaceable strings as translating into non-basic types can be justified.

As a first approximation, let us assume that CPs denote propositions, and propositions are sort i (information unit), a sort of type e. Let us assume further that IPs do not translate directly as propositions; if a formula is represented as type t (crucially not a sort of type e), then the complementizer is a function from t to i, viz. <t,i>. Alternatively, we could directly represent the denotation of IP as a functor in the type theory, by positing a type “c” for the contribution of the complementizer; then IP translates as type <c,i>, a mapping from “c” to information units. But what about the fact that the verb believe can appear with a CP or IP complement, and the two are interpreted the same way? One possibility is to assume that believe is alternatively type <i,nf> (seeking an information unit (CP) to return a property) or type <<c,i>,nf> (seeking
something that’s seeking a ‘c’ in order to become an information unit). This latter type is explicitly ruled out by Chierchia & Turner’s first-order theory; it also has a real air of ad-hocity about it, but that might be an artifact of our using type-theoretic notation to represent dependency; equivalently, we could manipulate the rules of semantic interpretation slightly to allow IP-selecting verbs to ‘supply’ the missing ‘c’. This is the approach I will take below, using the notion of head-chains developed above.

The type-theoretic formulation allows a mechanical description of the dependency of IP, but what it lacks is an explanatory, well-motivated notion of what ‘c’ is and why it is necessary for propositional interpretation. This task is undertaken in §3.4 below.

### 3.4. Some specific cases of Dependency

A very straightforward case of dependency, as conceived here, is the dependency of a transitive verb: the semantic type of the translation of a transitive verb will be \(<e, \text{nf}>\), a non-basic type in the relevant sense. No element which translates as type \(<e, \text{nf}>\) should be able to enter into an anaphoric relation, given the Strong FAC proposed above. Consider a sample translation of a sentence containing a transitive verb, Hilda is reading the Principia; assume for the moment that IP translates as something of type t (the correct translation of IP will be the topic of §3.4.2).

(86) IP
    \[\text{is-reading}'(h, p), t\]
    \[\text{DP}\]
    \[I\]
    \[\lambda x[\text{is-reading}'(x, p)], <e, t>\]
    Hilda
    \[I^o\]
    VP
    reading(p), nf
    is
    V
    DP
    p, e
    reading
    the Principia

To the right of each branching node is a semantic translation provided by the mapping rules, followed by the semantic type of the constituent.35

---

35 To be fully explicit, I am assuming for the moment the following translations:
- Hilda translates into \(h\), a constant of type \(e\).
- the Principia translates into \(p\), a constant of type \(e\).
- reading translates into \(\lambda x[\text{reading}'(x)]\), a functor of type \(<e, \text{nf}>\).
- is translates into \(\lambda x\lambda y[\text{is-x}(y)]\), a functor of type \(<\text{nf}, <e, t>>\).

\(P(x)(y)\) is rewritten \(P(y,x)\).
Some questions remain; for example, does head-movement create an anaphoric relation (the relation of the head with its trace), and if so, how does the movement of a transitive verb (e.g. in Norwegian V to C movement) satisfy the FAC? Another question has to do with the interpretation of tensed VPs; given the suggestion in §3.2 (cf. also Chierchia 1985) that tensed VPs are dependent, they must not translate as sort nf. How, then, do they translate? I will attempt to provide an answer to this last question below, but first I will approach the question of the dependency of IP. As for the question regarding head-movement, I offer only some speculative remarks, also below.

3.4.1. The dependency of VP

As noted above, tensed VPs are dependent. This means, on the current approach, that they do not translate into entities. In the translation in (86) above, I represented the auxiliary verb as occupying the Infl position, and interpreted it as if it were Pred (cf. §1.1.5). But in the absence of an auxiliary, Infl is unoccupied. Thus a similar translation scheme for the sentence Rome fell runs as in (90).36

(90) \[ IP \rightarrow \text{fell}'(r), t \]
\[ DP \rightarrow \lambda x[\text{fell}'(x)], <e, t> \]
\[ \text{Rome} \rightarrow \lambda x'[\text{fell}'(x)], <e, t> \]
\[ \text{fell} \rightarrow \lambda x[\text{fell}'(x)], <e, t> \]

This gives us a VP that translates as an nf, a sort of type e; hence there is no explanation for the dependency of the VP (and the badness of *What Rome did was fell). But there is reason to believe that correct translation of the verb is not as given in (90). The translation in (90) completely ignores the morphological form of the verb, and translates fell into fell, just as the translation in (86) gave \( \lambda x[\text{reading}'(x)] \) as the translation of reading. Obviously, this glosses over the semantics of tense and aspect. Let us instead assume a semantics for the different inflectional forms, and encode this in the translation. Let the morphological form of a verb be represented as a feature specification on the verb, which we might abbreviate VFORM (following Gazdar et al. 1985), thus a progressive participle like reading is represented as reading[VFORM: Prog], and a past tense form like fell can be represented as fell[VFORM: Past]. VFORM values can be assigned semantic values as type-preserving functions over verb stems, so that the argument structure of the verb stem can determine the semantic type of the inflected verb. This is stated as a rule in (91).

(91) Verb translation rule

\[ \text{If C is a verb with stem B and VFORM value A, and the} \]
\[ \text{translation of B is } \beta, \text{type } \delta, \text{and the translation of A is } \alpha, \text{then} \]
\[ \text{the translation of C is } \alpha(\beta), \text{type } \delta. \]

The translation of fell[VFORM: Past] is now Past(fell'), or Past-fall'. Of course, this translation rule is largely devoid of content unless the semantic values for \( \alpha \) and \( \beta \) are specified. For an idea of how the values for \( \alpha \) might appear, I refer the reader to Dowty 1979. The values for \( \beta \) must be determined on a verb-by-verb basis.

Now consider the fact that two VFORM values in particular are closely tied to the Infl node, namely the Tense values Present and Past; in Chomsky 1991, Tense lowers from Inf1 to V, and in Chomsky 1993, V raises to Inf1 (at LF) to check its Tense feature. In fact, we can provide a semantic motivation for this close connection between Infl and the tensed verb. Unlike the participles, the full translation of a tensed verb makes reference to a point in time with respect to which the clause is evaluated. The meaning of the operators associated with the participles have to do with different aspectual organizations of the event referred to, but the meaning of the tense operator has to do with the relation between the time of utterance (Reichenbach’s 1947 S, for POINT OF SPEECH) and the time of what is being described (Reichenbach’s R, for POINT OF REFERENCE). Put simply, the meaning of the present tense is something like ‘there is some R equal to S at which...’ while the meaning of the past tense is something like ‘there is some R prior to S at which...’ A more formal representation of the tense values is given in (92), using restricted quantifiers; \( t \) is a variable over points of time, the subscripts correspond to Reichenbach’s terms, and ‘<’ indicates the relation ‘precedes’ (see Dowty 1979 for formulations similar to those in (92)).

(92) a. Pres: \( \exists t_0; t_0 = t \)
   b. Past: \( \exists t_0; t < t_0 \)

Now, the full translation of fell[VFORM: Past] is \( \exists t_0; t < t_0[\text{fall}'] \). This representation is clearly missing something: there is no variable bound by the quantifier (cf. Chomsky’s 1986b discussion of the absence of vacuous quantification in natural language). We could insert a time-variable in the

---

36 Specifically:
\( \text{Rome} \) translates into r, a constant of type e.
\( \text{fell} \) translates into fell', a constant of type nf.
\( P \) translates into \( \lambda x y [x(y)], \) a functor of type <nf, <e, t>>.
representation of the verb by caveat. Alternatively, we could allow this variable to be introduced by Pred. Setting that question aside for a moment, note that the value of t_r is not always directly determined by the time of the utterance of the sentence. Consider the sentences in (93).

(93) a. Juliet said Romeo was in her room.
    b. Juliet will say Romeo was in her room.

In the embedded clauses in (93), the past tense picks out a time not prior to utterance of the sentence as a whole, but prior to the time of Juliet’s reported (or predicted) utterance. That is to say, (93) is a true prediction on Monday if it comes to pass that Juliet says on Wednesday that Romeo spent Tuesday in her boudoir. This suggests that t_r may be specified for each IP; but it will be specified by factors external to the tense-inflected verb itself (i.e. the root context, or the embedding verb, or perhaps some other factors). In this sense, the interpretation of the Tense morphemes is connected to the higher structure in a way that the interpretation of the aspectual inflections is not. Compare the progressive participle in some different contexts in (94).

(94) a. Romeo is humming a tune.
    b. Romeo was humming a tune.
    c. Romeo has been humming a tune.
    d. Juliet will say that Romeo was humming a tune.

The progressive aspect has the effect that the humming event is treated as covering an interval (cf. Dowty 1979), and may or may not be completed (in Reichenbach’s terms, we might say that E spans an interval and is possibly unfinished, and R is located within E). This interpretation remains the same in all of the various contexts in (94), indicating that the interpretation of the progressive is not affected by feature specifications in the higher structure. The same thing can be demonstrated for the perfect participle.

Thus, there is semantic evidence that Tense is connected to the higher structure in a way that non-tensed inflectional values are not. Following this intuition, let us say that a tensed verb does not actually have a VFORM value in the sense relevant to the translation rules in (91), until the value for t_r is specified. This can be stated formally as in (95). (95) is not intended to be a special rule of translation, just a refinement of the special case of (91) instantiated by the values of α represented in (92).

(95) Verb translation rules for tensed verbs
   a. If C is a verb with stem B, VFORM value Pres, and t_r value ε and the translation of B is β, type δ, then the translation of C is ∃t_r: t_r ∈ [β(t_r)], type δ.
   b. If C is a verb with stem B, VFORM value Past, and t_r value ε and the translation of B is β, type δ, then the translation of C is ∃t_r: t_r ∈ [β(t_r)], type δ.

We can assume for the moment that t_r values are base-generated in Infl, and a tensed verb must be specified for those values during the course of the derivation. In this case, the base-generated structure for the sentence Romeo fell is not as in (90) above, but as in (96) below.

(96) Here, the verb is not specified for the feature t_r. The rules in (91) are unable to assign a semantic translation to it, because the value for the Past morpheme is incomplete; fell simply does not translate in this structure. However, there is a mechanism by which we can specify a value for the feature t_r on the in the structure in (96). Recall the conditions on head-chains from (60) in §2.2, restated here in (97).

(97) Conditions on head-chains:
   a. An ordered n-tuple of heads <α_1, ... α_n> is a head-chain only if, for every i, 1 ≤ i < n, α_i is local to α_{i+1}; α_1 is the INITIAL of the head-chain and α_n is the TERMINUS of the head-chain.
   b. A head-chain H is an F-chain, for any feature F, iff:
      i. all members of H bear the feature F.
      ii. the initial of H has a value specified for F, and
      iii. no non-initial member of H has a value specified for F.
   c. In an F-chain, the terminus inherits the initial’s specification for F.
A head-chain can then be formed between Infl and the verb, creating the tree in (98). The translation rules apply as desired.

\[
\begin{array}{c}
\text{IP} \\
\exists t; t < \alpha \{ \text{fall}('t, t') \}, t \\
\text{DP} \quad I' \\
\lambda x [ \exists t; t < \alpha \{ \text{fall}('t, t') \}], \langle e, t \rangle \\
\text{Rome} \\
\end{array}
\]

This idea could be refined in various ways, for example the \( t \) variable could be equated with Davidson’s 1969 event argument, and could be assumed to be part of the argument structure of the verb. The main idea which needs to be preserved for my account of dependency is that a tensed verb, but not a participle, must enter into a head-chain with the verb in order to have its Tense feature licensed or checked. If this licensing or checking does not occur, the verb fails to receive a proper translation.

Now consider in what sense this accounts for dependency of a tensed VP. The VP actually translates as something of type nf, so it may appear that we are back where we started. But crucially, it receives this translation only by virtue of a feature specification that it got via a head-chain. In a displaced context, the head-chain could not be formed, and the VP would remain uninterpretable.

Now compare a structure with an auxiliary verb, diagrammed in (99).

\[
\begin{array}{c}
\text{IP} \\
\text{Pres-Perf-fall}('r), t \\
\text{DP} \quad I' \\
\lambda x [ \text{Pres-Perf-fall}('x) ], \langle e, t \rangle \\
\text{Rome} \\
\end{array}
\]

Here I simply represent the VFORM values as unanalyzed prefixes (or functions) in the logical representation. I represent has as having raised from a verb node to Infl. Here, its feature specifications are unified with those of Infl, and it is specified as \([t; \alpha t] \). This specification allows the translation of the present tense morpheme, and the translation proceeds as shown. Alternatively, has might be interpreted in its base-position, \( t _ i \) in (99). This could be seen as a special case of a head-chain passing feature specifications down to a lower head, except in this case the head-chain would be passing all of the values for some feature like HEAD or CONTENT.\(^{37}\)

\[\text{What is more relevant to present concerns is the observation that the lower VP in (99) is independent. Its translation as an entity of sort nf is not licensed by any head-chain or movement operation, and any hindrance to its displacement must be due to factors other than dependency.}\]

3.4.2. The dependency of IP

As noted above, a CP is not dependent, and therefore must denote a kind of entity (specifically, an information unit). If a verb takes a CP as its complement, then that verb must denote a relation between individuals and information units. But many such verbs also admit IP complements, and I claimed above that the dependency of IP indicated that it does not translate as a kind of entity. How can IP be a complement to a belief verb? In this subsection I develop an account for this. The basic intuition is that an information unit consists of a proposition plus an anchor. Without the anchor, it is interpreted only as a formula, which is incoherent in an embedded context. In order to develop this account, I first adopt a general framework for discussing the context and content of propositions.

In Montague grammar, a simple declarative sentence can be translated into a formula, which can be seen as a mapping from

\(^{37}\) I.e., if the trace bears some feature like HEAD and all of the information necessary for interpretation is encoded in feature-specifications which constitute values for the feature HEAD.
circumstances to truth values. The truth value of the sentence in a given circumstance is its extension; the set of all the possible circumstances in which the sentence is true is its intension. It is common to model circumstances as possible worlds, where a world consists of a variety of individuals with various and in various relations to each other (I am ignoring time for the moment). The utterance of such a sentence is typically interpreted as expressing an assertion on the part of the speaker that the formula accurately characterizes the real world, in the context of that utterance. Even ignoring Gricean pragmatic issues, the belief, knowledge, and attitude of the speaker as well as the context of utterance are crucial in evaluating the truth of a simple declarative sentence. For example, the speaker’s association of referents to pronouns may affect truth conditions, and so may the immediate physical context of the utterance, for example when various deictic terms are used. In a broader sense, the context of the utterance also includes various tacit agreements between the speaker and the hearer, for example the familiarity of various individuals and the domain of quantification (cf. Karttunen 1973).

Following Stalnaker 1974 and 1978, we can speak of the COMMON GROUND of an utterance, the set of propositions that both the speaker and hearer take for granted. If propositions are the intensions of sentences, then the common ground is a set of sets of possible worlds. Each assertion made during a conversation is added to the common ground, if it is not challenged or retracted. Still following Stalnaker, the intersection of all the propositions in the common ground is the CONTEXT SET (a set of possible worlds, hence a proposition). This is the set of possible worlds that the speaker and hearer consider, for the purposes of the conversation, to be viable realities, i.e. it is assumed that the real world is in the context set. The context set must not be empty, so contradictory assertions cannot be accepted.

This model of discourse gives us a way to describe what happens when a participant in a conversation utters a simple declarative sentence. A sentence $P$ may be translated as a set $\phi$ of possible worlds, but the speaker is not simply naming that set. The speaker, in stating $P$, adds $\phi$ to the common ground, which amounts to a proposal that all worlds incompatible with $\phi$ be removed from the context set. This can be stated as in (100).

(100) If a speaker asserts $\phi$ and the common ground is $c$, $\phi$ is added to $c$.

(100) represents the basic case. More complicated circumstances will be addressed below. This model allows a simple treatment of presupposition (see Chierchia & McConnell-Ginet 1990, Chapter 6, for discussion); when a proposition $\phi$ presupposes a proposition $\psi$, this means that there is a condition on the felicitous utterance of $\phi$ that the context set be contained within $\psi$, i.e. $\psi$ must be in the common ground.

In addition, this model provides a useful tool for the evaluation of statements about different individuals’ beliefs. As participants in a discourse, we generally assume that everyone believes everything they say; anyone of whom this is not true is a liar. As long as we believe everything that a speaker says, that speaker’s utterances are added to the common ground. But if we differ, then we know that there is a difference between what we believe and what that person believes. We can model this, when it is relevant, by keeping track of another set of propositions alongside the common ground, namely the belief-set of each participant in the discourse. Thus (100) above can be modified as in (101) below.

(101) If a speaker $x$ asserts $\phi$ when the common ground is $c$ and $x$’s belief-set is $b$, $\phi$ is added to $c$ and to $b$.

Now it is also assumed, in a typical conversation, that listeners believe what the speakers say, if they do not object to it. This could be added as well to (101). But consider what happens if the speaker, $x$, says something to which a hearer, $y$, objects. $y$ might allow the contentious assertion to be added to the common ground anyway, for the purposes of the conversation. Or $y$ might challenge the validity of the assertion, for example by saying, “I don’t believe that.” Now, a more complicated thing happens. $y$ has made an assertion, which, assuming that all of the participants in the discourse believe it, is added to the common ground, in accordance with (101). But in addition, the previous proposition is removed from $y$’s belief-set, or its negation is added. All things being equal, a participant’s beliefs can be assumed to be a superset of the context set; but as soon as some claim is made about belief, a difference will be established between the context set and the individual’s belief.

Thus an individual’s beliefs can be modelled as a proposition, the intersection of all of the propositions in the belief-set.38 This requires that individuals not hold contradictory beliefs, which has of course been challenged, but I will nonetheless assume it to be true. Another issue that arises is whether individuals must believe all of the entailments of their beliefs. In general this must be the case, so I will assume that in general, when reasoning about the belief-set of an individual $x$, if that set contains $\phi$, and $\phi$ entails $\psi$, then $x$’s belief-set also contains $\psi$, provided that $x$ has

38 Note that this allows intensional contexts within beliefs. Thus, an individual who believes that Norway might be part of Sweden has as an object of belief a set of worlds, in some of which Norway is part of Sweden and in some of which Norway is not a part of Sweden.
a proper understanding of all of the component parts of the sentence translated as $\psi$. In the vast majority of cases, this seems to produce the desired results. If Mary believes that she has a sister, then Mary believes that she has a sibling, provided she understands the concept of sibling, because the proposition *Mary has a sister* entails (is a proper subset of) the proposition *Mary has a sibling*.

There are certain well-known problems for this view. The most often cited one is that of necessary truths, for example mathematical truths: if Mary has a few basic mathematical beliefs, then this view seems to entail that she has much more complex ones as well. Say, for example, that she knows something about negative numbers. She believes that $2 + (-2) = 0$. Can we infer that she believes that $2 - (-2) = 4$? Apparently not. But upon discovering that she rejects the second equation, we are inclined to patiently explain that she is wrong. If successful, we manage to show her that she holds inconsistent beliefs about $-2$; there are many propositions in her belief-set, but their intersection is empty.

Chierchia & McConnell-Ginet 1990 point out a slightly different sort of problem. In their example, Mary believes that everyone ate chicken, but she hasn’t considered the possibility that anyone ate tofu. She might not even know that tofu exists. It does not seem that we can fairly assert that she believes that nobody ate tofu; but nor can we say that she holds contradictory beliefs. The intersection of her belief-set contains no worlds in which anyone ate anything other than chicken, but it might not contain tofu at all. In other words, Mary has no beliefs at all regarding tofu. This is different from a situation in which John understands the concept of tofu, but doesn’t believe that it exists. In such a case, his belief also contains no worlds in which people ate tofu, but it does contain some worlds in which there is tofu (perhaps imagined worlds, or dream worlds).

One way to capture this distinction would be to posit a notion of accessibility (worlds in which people ate tofu are accessible to John, but not to Mary), but I will not attempt to do so. Instead I will simply assume that for the purpose of evaluating the truth-conditions of statements about belief, individual’s beliefs can be modelled as sets of propositions which have a non-null intersection. An individual whose belief-set is constructed to have a null intersection can rightly be accused of inconsistency.39

Adapting the notation of Farkas 1992b, let us represent a set of possible worlds as $w$; the context set is $w_R$ (R for ‘real,’ as $w_R$ is assumed for the purposes of the discourse to contain the “real world”). The set of propositions that we take an individual to believe is that individual’s belief-set, and the intersection of x’s belief-set is $w_b(x)$. Say that a sentence P translates into a proposition $\phi$. If uttered in the context $w_R$, then the complete information unit represented by P in that context can be modelled in the following way: $A(w_R, \phi)$, where A signifies ‘assert,’ and the proposition is added to the anchor (various other notations could be adopted, for example the context could be made into a superscript, or the context and the proposition could be represented as an ordered pair).

There are also other modes of anchoring, for example ‘presuppose,’ which could be represented as $\Pi(w_R, \phi)$ if $\phi$ is presupposed in $w_R$.40

The context of a proposition embedded under a belief-predicate is $w_b(x)$, where x is the subject of the belief-predicate. I will speak of an asserted proposition as being anchored to a particular context, and I will speak of the context as the anchor of the asserted proposition (this is somewhat different from Farkas’ use of the same terms; for Farkas, the anchor of $w_b(x)$ is x; for me, the anchor of $A(w_b(x), \phi)$ is $w_b(x)$). If a proposition is an assertion, then it is added to the common ground, from which $w_b$ is calculated; I will describe this more loosely by saying that a proposition is added to its anchor, i.e. $\phi$ is added to $w_R$. Thus, presented with an information unit of the form $A(\alpha, \beta)$, we can say that $\beta$ is added to $\alpha$ (more precisely, $\beta$ is added to the set from which $\alpha$ is calculated). Propositions can be mapped onto information units by rules of translation. For example, we can state one such rule as in (102).

(102) Root Rule:

If P is an root indicative declarative sentence uttered by x, and the intension of P is $\phi$, and the context set at the time of utterance is $w_R$, then P translates into $A(w_R, \phi) \land A(w_b(x), \phi)$.

The Root rule anchors the intension of the root IP to the context set ($w_R$, the intersection of all the previous assertions in the discourse, plus various entailments of those assertions, plus various other salient propositions that I can safely assume are believed by my audience) and

question.

39 Here I disagree with, for example, Katz 1986, who argues that people can knowingly believe contradictory things; he suggests that there are “…Christians who believe God’s creation to contain unnecessary and undeserved suffering and who see this as entailing that God cannot be both all good and all powerful but who still believe God to be both.” (Katz 1986:66). My suspicion is that the Christians in question would challenge the entailment, although I suppose that’s an empirical (continued next page)

40 The modes of anchoring A and $\Pi$ might be extended to definite and indefinite noun phrases; an indefinite noun phrase would be anchored with A, meaning that its referent should be added to the domain, while a definite noun phrase would be anchored with $\Pi$, meaning that there is a felicity condition on its use that its referent already be in the domain.
the belief-set of the speaker. IPs in other contexts will receive their anchoring in different ways. Consider the interpretation of a finite CP in a subject position, as in (103).

(103) a. That senators take bribes disturbs me.
   b. That Jason has made cookies makes sense.

Kiparsky & Kiparsky 1970 point out that that CPs in subject position are generally factive, that is, their truth is presupposed. As a first approximation of the meaning of the complementizer that, suppose that it has the function of anchoring the content of the IP to \( w_R \). The anchoring takes the form of a presupposition, rather than an assertion; in other words, the relevant proposition is not added to the context set, it is merely checked against the context set for compatibility. In a sense this makes that like the determiner the, which has the effect that the entity denoted by the DP is presupposed to be in the domain of discourse (cf. Bresnan's 1972 observation that the complementizer that is 'definite'). This can be represented as in (104).

(104) Complementizer rule:
If P is an IP, and the intension of P is \( \phi \), and the context set at the time of utterance is \( w_R \), and P is the complement of the complementizer that heading a non-complement CP Q, then Q translates into \( \Pi(w_R, \phi) \).

Of course, when that CPs appear as the complements to certain verbs, they are not factive. Just as with IPs, when that CPs are embedded under epistemic verbs, the verb determines not only the world the proposition is anchored to, but also the type of anchoring. The examples in (105) show finite CPs interpreted non-factually; their truth is not presupposed.\(^{41}\)

(105) a. It was reported by the UPI that most senators take bribes.
   b. Chris claims that Jason has made cookies.

Thus, if I state Jack believes Max is afraid of spiders, the whole proposition is anchored to the context set. The embedded proposition Max is afraid of spiders is anchored to Jack's belief-world, \( w_b \). There are also fictitious worlds, for example the world of a dream; Zack dreamt there were bees in his hair anchors the proposition there were bees in his hair to \( w_d(z) \), Zack's dream-world. Note that the anchoring is not the whole of the content of the verb; Jack doubts Max is afraid of spiders anchors Max is afraid of spiders to \( w_b(j) \), but not in the same way as believe does.\(^{42}\)

The function of that, then, is not simply to anchor the intension of the IP, \( \phi \), to \( w_R \); but to anchor \( \phi \) to some \( w \); a verb of propositional attitude can determine what \( w \) \( \phi \) is anchored to (and how), or \( w_R \) can be assumed by default if the CP is not embedded under an epistemic verb. A couple of verb-specific anchoring rules are given in (106).

(106) Anchoring: If P is an IP, and the intension of P is \( \phi \), and the context set at the time of utterance is \( w_R \), then:
   i. Belief: if P is the complement of the complementizer that heading a CP Q which is the complement of believe, and \( x \) is the subject of believe, then Q translates into A(\( w_b(x) \), \( \phi \)).
   ii. Dream: if P is the complement of the complementizer that heading a CP Q which is the complement of dream, and \( x \) is the subject of dream, then Q translates into A(\( w_d(x) \), \( \phi \)).

Some regular rules apply to the common ground, all belief-sets and all dream-sets: if a proposition is added to the common ground, then a new context set is calculated from the new common ground. Similarly, if a proposition is added to \( x \)’s belief-set, a new \( w_b(x) \) is calculated. Belief and dreaming are different in that individuals are assumed to have one \( w_b \), while they may have as many \( w_d \)’s as they have dreams. The extension of this pattern to other verbs is fairly obvious; there can be a claim-set, the set of propositions an individual claims (in a given setting; if I say that Frank told Sally one thing and me another, then I must posit two different claim-sets for Frank). As long as it is assumed that individuals are not lying, propositions added to their claim-sets will also be added to their belief-sets. All verbs of saying can be assumed to

\(^{41}\) Kiparsky & Kiparsky 1970:167 note that CPs in subject position tend to be interpreted factively, even when they are logically the complements of non-factive verbs; they give the pair in [i].

[i] a. That Smith had arrived was reported by the UPI.
   b. It was reported by the UPI that Smith had arrived.
I agree with their observation that the CP in [ia] is most naturally interpreted as factive, while the same CP in [ib] is not. However, the correlation does not seem to be absolute. I think that the sentences in [ii] are possible, without any presupposition of the truth of the CP.

[ii] a. That the earth is flat was widely believed until recently.
   b. That Clinton will run for re-election is likely.

\(^{42}\) More 'modes of anchoring' might be introduced to encode this variation; for example, if A(\( \alpha \), \( \beta \)) means 'add \( \beta \) to \( \alpha \)', then D(\( \alpha \), \( \beta \)) might mean 'remove \( \beta \) from \( \alpha \)' to capture the difference between, say, believe and disbelieve, or claim and deny.
contribute to claim-sets (deny would subtract its complement proposition from the claim-set, or contribute its negation); all verbs of belief would contribute to the belief-set (know includes a presupposition that the context set is contained in the complement proposition). Imagine is like **dream** in that multiple imagination-sets can readily be produced, there being no expectation that the different things an individual imagines be consistent. Thus, the representation of information units as pairs of propositions and anchors plus a mode of anchoring is completely general; the specific way in which a proposition is connected to its anchor is determined by the embedding verb or the Root Rule or the Complementizer Rule. Obviously, more rules are needed. Presupposed propositions can be added to the common ground by accommodation, if they are not considered controversial or surprising. Assertions which go unchallenged are typically added to the hearer’s belief-set as well. But the rules in (106) give a general idea of the utility of contexts in calculating truth-values. However, I will improve on these rules below, using more general mechanisms already available.

Now, I established in §3.3 that the dependency of IP could be characterized by saying that IP lacks something which I called “c”, the semantic contribution supplied by the complementizer. Let us say now that the semantic contribution of the complementizer is the anchoring itself. This can be formalized in the following way. Say that the property of being potentially modally anchored is represented by a feature φ, just as the property of being potentially assigned Case is represented by the feature Case; then all elements that can potentially be modally anchored must bear the feature φ. Values for the feature φ are pairs consisting of a mode of anchoring and an anchor. An IP with an intension φ and a φ specification <A, wR, φ> will be translated as the information unit A(wR, φ). Since one proposition might have more than one anchor, we can allow the value of φ to be a set of ordered pairs. This is encoded in the rule stated in (107).

(107) If an IP P has intension φ and a φ value consisting of pairs of the form <α, β>, then P is an information unit of the form α(β, φ) (for each α and β).

Now we can restate the Root Rule from (102) above as in (108) below.

(108) **Root Rule:**

If P is an root indicative declarative sentence uttered by x, and the context set at the time of utterance is wR, then specify the φ value of P as {<A, wR>, <A, wR(x)>}. The Root Rule specifies the anchoring for a root IP, and (107) applies to any IP with a specified φ value. Now we can also restate the Complementizer rule from (104) above as in (109) below.

(109) **Complementizer rule:**

that has φ value <Ω, wR>.

Now, if we assume that Infl is base-generated with a φ feature with an unspecified value, an IP embedded under the complementizer that will not be interpretable as an information unit, unless it forms a head-chain with the complementizer and receives the complementizer’s φ feature specification. Once Infl bears a φ feature, the IP is specified for φ (assuming that φ is a head feature in the sense of e.g. Gazdar et al. 1985) and the rule in (107) will assign it a licit translation as an information unit. This takes care of cases where the complementizer contributes a presuppositional anchoring.

Now consider the case where a CP is embedded under an epistemic verb. We can assume that each epistemic verb bears a φ feature specification. When CP is embedded under an epistemic verb, we want to copy the verb’s value for φ onto the complementizer, and then onto the embedded Infl. However, recall from §2.2 that all non-initial members of an F-chain must bear F but be unspecified for it. This means that the complementizer that, when embedded under an epistemic verb, must bear an unspecified φ; in other words, (109) must be constrained to act as a default rule, only applying when no other value is available.

The initial representation of a sentence like (110a), prior to head-chain formation, might be something like that in (110b), ignoring the φ feature on the main clause Infl.

(110a) a. Max believes that Helen read the Investigations.

b. Max believes [CP that <Ω, wR> [IP Helen [I that] read the Investigations]]

c. Max believes [CP that <Ω, wR> [IP Helen [I that] read the Investigations]]

In (110b), a head-chain has been formed which consists of the epistemic verb, the complementizer, and the embedded Infl; the complementizer and Infl inherit the φ feature value of the epistemic verb, and the IP is interpreted as a anchored proposition, i.e. an information unit. Consider a step-by-step translation of (110), given in (111) (ignoring tense in order to simplify the representations).
The translation proceeds from the bottom up. The components of the embedded IP yield the formula read(h, i), which is type t, but since the IP has the feature specification [ω: Π, w(i)m+1], the formula is mapped onto the information unit A(w6(y), read(h, i)), which is type i. The Root Rule specifies the ω value on the root Infl (not shown). When the root IP is composed, rule (107) gives the information unit given on the first two lines, and the proposition is added to the context set and to the speaker’s belief-set (the speaker is indicated with ‘s’). The role of that is really just to pass along the value for ω. The rules in (106) can now be completely replaced by some lexical specifications of the ω values for the various epistemic verbs.

The case of the bare IP complement to an epistemic verb now follows directly from what has already been said. Consider the representation in (112). There, believe is fully specified as it was in (110b) above. The complement of believe is an IP, specified exactly as the embedded IP in (110b) was specified. In (112c), the head-chain is shorter, including only the verb believe and the embedded Infl, but the effect is the same; the embedded IP will be translated as an information unit.

In (114b), that and the embedded Infl have their base-generated values, i.e., they have no value for ω. No head-chain can be formed between the displaced CP and the verb believe, so its values cannot be copied onto the complementizer or Infl. Instead, the Complementizer Rule applies, specifying that as [ω: Π, w(i)m+1]. A head-chain is formed between that and Infl, and Infl is also specified as [ω: Π, w(i)m+1]. This means that the preposed CP is interpreted as an information unit, but it is not anchored to
Jack’s belief-world. Instead, it is presupposed true. Now, the displaced information unit is coindexed with its trace, and will be interpreted as the direct object of believe, so the proposition Jack believes that Max is afraid of spiders will be added to the common ground and to the speaker’s belief-set. This might be enough to establish Max is afraid of spiders as belonging to Jack’s belief-set, perhaps by accommodation. But it does seem that (114) has a presuppositional sense that the corresponding non-displaced version does not. Now compare (115), where what is fronted is an IP.

(115) a. * Max is afraid of spiders, Jack believes.
   b. [IP Max [θ]_{af}] is afraid of spiders], Jack believes_{[θ]} [θ_{af}]

In (115), no basic interpretation can be assigned to the fronted IP, because it does not bear ω. It does not correspond to a basic type, so it cannot be coindexed with a trace. There is no default rule for non-root IP, and no head-chain can be formed. The sentence cannot be translated.

A for CP is slightly different, but the same principle holds. Infl needs a ω feature specification; the ω feature value provided by for does not identify the denotation of the IP as simply containing wR but introduces a sense of irrealis (cf. Jespersen 1924); for might provide EXTENSIONAL anchoring, in the sense of Farkas 1992b, meaning that the IP is anchored not to a single possible world but to a set of alternative possible worlds. Certain heads can relativize the irrealis sense of for, for example hate and like seem to impose some requirement on the situation denoted by the infinitival IP that it be something that has been experienced, and such verbs as manage definitely anchor a proposition to wR. Once again, all of these different anchorings can be seen as specifications for ω features, which are provided either by for or by a governing verb, under chain-government.

Now we have an alternative to saying, as I did in §2, that a verb like want c-selects a non-finite CP; we could say instead that it supplies ω feature specifications to its complement which are incompatible with tense, but which are compatible with a non-finite nexus, and are compatible with the complementizer for. These issues will be taken up again in Chapter 2; I will continue to refer to such relations as c-selection, though it can be seen that they are quite different from the original conception in Chomsky 1965 of strict subcategorization for syntactic categories.

3.5. Summary of §3

In many constructions, a feature specification on a given node is determined by some property of a neighboring node. For example, the
dative Case on the Icelandic DP in (116a) is determined by the preposition. Similarly, the participial form of the verb leave in (116b) is determined by the governing auxiliary. Or, as I discussed in §2, the finiteness of the CP complement to believe in (116c) is determined by that verb.

(116) a. Hestarnir eru í túninu. (Ice)
   the.horses are in the.field.DAT
   ‘The horses are in the field’
   b. Elvis is leaving the building.
   c. John believes that pigs can fly.

None of these relations induce dependency; the dative noun phrase, the progressive participle, and the finite CP can all be displaced.

(117) a. Túninu eru hestarnir í. (Ice)
   the.field.DAT are the.horses in
   ‘The field, the horses are in’
   b. Leaving the building is Elvis.
   c. What John believes is that pigs can fly.

All of these relations can be analyzed as a form of feature-specification via head-chains, or they could be modelled by specifying feature values on a subcat list: the auxiliary be might c-select an ING complement, or it might be that a progressive participle has a certain feature, call it Aspect, which must be specified, and be has a value for that feature which can be copied onto the participle through chain-government. The important thing is that whatever relation this is, it is not disrupted by displacement, be it A-movement, A-bar movement, or some more indirect relation. The specifications provided for features by head-chains are assumed to be transmitted through A-chains and A-bar chains, or perhaps are satisfied at LF under reconstruction.

I have suggested that in some cases, the feature specifications determined by head-chains are crucial to the interpretation of an element. This corresponds in certain cases to their being non-basic types semantically, in set-theoretic notation: an object of type <e, e> is not an entity but a function; if all variables are type e, then a variable can never be interpreted as corresponding to a function, and any incomplete constituents will have to remain in situ in order to be properly interpreted. This could be modelled as a constraint on reconstruction; I have cast it as a limitation on the inventory of variables, following the spirit of Chierchia 1985 and Chierchia & Turner 1988.
4. Conclusion

To a certain extent, this chapter represents an attempt to build up a framework within which certain empirical problems, introduced in Chapters 2 and 3, can be resolved. As a result, the connections among some of the parts of this chapter may not be apparent to the reader at this point. This brief overview might help to pull some strands together.

In §1, I discussed previous work on predication. I adopted the position that every nexus, in Jespersen’s sense, is built around a predicator of the sort proposed by Chierchia and Bowers. This predicator has the effect of turning a property into a predicate. Infl, it was suggested, is one such predicator, but there is also a tenseless species of predicator, Pred, which heads the small clause.

The general program of stating constraints on syntactic representations in structural terms was adopted from GB. Certain issues regarding the Theta Criterion of GB were left unsettled, particularly the question of whether certain subjects (in Heycock’s examples) get θ-roles, and if so, from where. This matter will be dealt with in Chapter 2 below.

In §2, some specific constraints were explored, namely those having to do with the relation of a head to its complement. I argued that one such relation be identified with c-selection. I furthermore argued for a particular conception of head-chains, a mechanism for copying or sharing features across different parts of a syntactic tree, subject to strict structural constraints. I proposed that head-chains, under certain conditions, could supply a lower node with feature specifications, sometimes even when the lower node would not otherwise bear that particular feature (the Composition chain).

Finally, in §3, I developed the notion of dependency, in order to account for the failure of certain constituents to be displaceable. I pointed out that constituents which are not displaceable are also not licit as subjects. I argued that this was crucially tied in with limitations on the rules of semantic interpretation. A constituent which does not translate with a basic type cannot have a property predicated of it, and so cannot be a subject; in addition, it cannot be coindexed with a trace or dummy element, and so cannot be displaced. The theory of head-chain developed in §2 was employed to allow a proper translation of those elements in complement positions; only in complement position is the necessary configuration for head-chain formation (locality) available.

All of these proposals will be brought to bear on the data in Chapters 2 and 3.