

# 1 The Lexicon and Sentence Structure

---

## *Contents*

### Introduction and overview

- 1 The units of syntactic analysis
- 2 Words and phrases
- 3 Predicates and arguments
  - 3.1 Subcategorization
  - 3.2 Argument structure and thematic structure
    - 3.2.1 Argument structure in logic
    - 3.2.2 Argument structure in natural language
    - 3.2.3 Theta theory
- 4 The projection principle
- 5 The assignment of thematic roles
  - 5.1 Clausal arguments
  - 5.2 Expletives
    - 5.2.1 *It* and extraposition
    - 5.2.2 *There* and existential sentences
    - 5.2.3 Conclusion
  - 5.3 Main verbs and auxiliaries
- 6 The extended projection principle (EPP)
- 7 Thematic roles: further discussion
  - 7.1 The canonical realization of theta roles
  - 7.2 The subject theta role
- 8 Summary
- 9 Exercises



## Introduction and Overview

---

In the Introduction we saw that a grammar of a language is a coherent system of principles which determines the formation of the sentences of a language. The basic unit with which a grammar is concerned is the sentence. A grammar will specify what the components of the sentence are, how they interact, in which order they occur, etc. Partly, the principles formulated will be of a universal nature; partly, they will have to be parametrized to bring out language-specific properties of individual languages.

Grammars have nothing to say about units higher than the sentence, such as the paragraph, the discourse exchange, the text, etc. Such higher units will be the object of another type of enquiry.<sup>1</sup>

In this chapter we consider the relation between the structure of the sentence and the words that make up the sentence. We shall see that sentence structure is to a large extent determined by lexical information. As pointed out in the Introduction, it is assumed that the reader is familiar with the basic techniques and terminology of sentence parsing.

Chapter 1 is organized as follows: section 1 provides a brief discussion of the central concepts of sentence structure; section 2 focuses on the relation between lexical items and sentence structure; section 3 discusses the predicate-argument structure of sentences and introduces theta theory; section 4 sums up the link between lexical items and sentence structure and introduces the projection principle; section 5 explores the application of theta theory, concentrating on clausal arguments, expletive (non-argument) pronouns and auxiliary verbs; section 6 discusses the general constraint that sentences must have subjects; and in section 7 we consider the properties of the subject theta role.

## 1 The Units of Syntactic Analysis

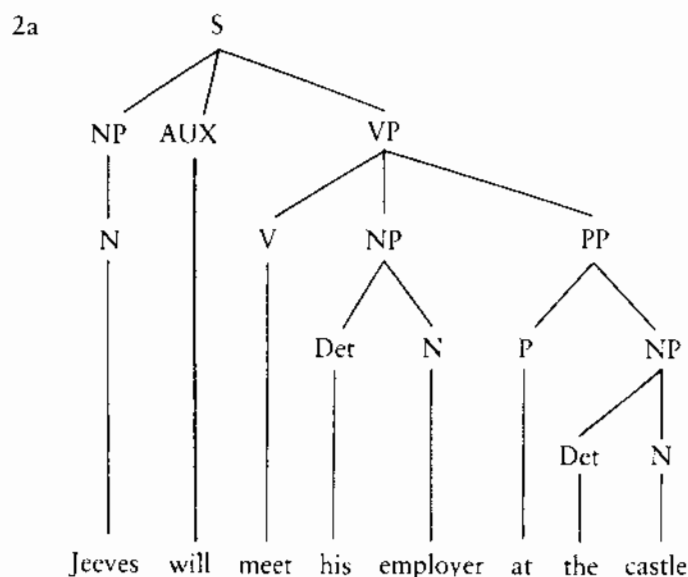
---

In this section we briefly recapitulate the basic notions of syntactic structure that will be the starting point for our discussion. Consider the following example:

<sup>1</sup> For an interesting approach to the study of sentences in discourse see Sperber and Wilson (1986) and Kempson (1988a, 1988b), who examines the link between Sperber and Wilson's theory of utterance interpretation and formal syntax.

1 Jeeves will meet his employer at the castle.

(1) is a grammatical English sentence. When we look for its component parts, the **constituents**, the units that perhaps come to mind first are the words of the sentence: sentence (1) contains eight words. But, as anyone familiar with traditional techniques of sentence parsing knows, words are not the **immediate constituents** of a sentence. Rather, they are the **ultimate constituents**. The words of the sentence are organized hierarchically into bigger units called **phrases**. In the framework of generative syntax the constituent structure of a sentence is represented in one of the following formats: by means of the **tree diagram** format as in (2a), by means of **phrase structure rules** or **rewrite rules** as in (2b), or by means of **labelled brackets** as in (2c).<sup>2</sup>



- 2b
- (i)  $S \longrightarrow NP - AUX - VP$
  - (ii)  $NP \longrightarrow (Det) - N$
  - (iii)  $VP \longrightarrow V - NP - PP$
  - (iv)  $PP \longrightarrow P - NP$
  - (v)  $N \longrightarrow Jeeves, employer, castle$
  - (vi)  $V \longrightarrow meet$

<sup>2</sup> For an introduction to parsing see Burton-Roberts (1986), Fromkin and Rodman (1988), Guéron and Haegeman (in preparation), and Wekker and Haegeman (1985).

- (vii) AUX → *will*
- (viii) P → *at*
- (ix) Det → *the, his*

2c [<sub>S</sub> [<sub>NP</sub> [<sub>N</sub> Jeeves]] [<sub>AUX</sub> will] [<sub>VP</sub> [<sub>V</sub> meet] [<sub>NP</sub> [<sub>Det</sub> his] [<sub>N</sub> employer]]] [<sub>PP</sub> [<sub>P</sub> at] [<sub>NP</sub> [<sub>Det</sub> the] [<sub>N</sub> castle]]]].

Representations such as those in (2) give us information concerning the structure of (1). They indicate, for instance, that the string *his employer* is a syntactic unit, a **constituent**. It is a **noun phrase** (NP), a constituent whose main element or **head** is the noun (N) *employer*. Analogously, the constituent *at the castle* is a **prepositional phrase** (PP); the head of this PP is the preposition *at*, which is followed by an NP, *the castle*. The constituent *meet his employer at the castle* is a **verb phrase** (VP), whose head is the verb *meet*, which is followed by the NP *his employer* and the PP *at the castle*.

The structural representations in (2) allow us also to describe syntactic operations that may affect sentence (1). Consider (3):

- 3a At the castle, Jeeves will meet his employer.
- 3b His employer, Jeeves will meet at the castle.
- 3c Meet his employer at the castle, Jeeves will (indeed).

The sentences in (3) are intuitively felt to be variations upon sentence (1); they are all paraphrases of (1). In order to capture the similarity between the sentences in (3) and that in (1) we assume that all these sentences have the same **underlying structure**, represented in (2). In each of the sentences in (3) one of the constituents identified in (2) has been moved to the beginning of the sentence, or **preposed**. Thus in (3a) the PP *at the castle* has been moved, in (3b) the NP *his employer* is moved, in (3c) the VP, *meet his employer at the castle*, is moved. The possibility for preposing elements of a sentence is structure-based: only constituents of the sentence such as NP and VP can be preposed; one cannot indiscriminately prepose any random string of words in the sentence:<sup>3</sup>

- 3d \*Employer at the, Jeeves will meet his castle.
- 3e \*Meet his, Jeeves will employer at the castle.

<sup>3</sup> For a formal discussion of operations such as preposing, see chapters 6 and 7.

Another operation that affects sentence constituents is the one that forms questions. If we form questions on the basis of (3) we see that again the constituent structure represented in (2) plays a crucial role.

We distinguish two types of questions: **yes-no questions** and **constituent questions**. The classification adopted is based on the type of answer expected. (4a) is a *yes-no* question: in normal circumstances we expect *yes* or *no* as an answer. The other questions in (4) are **constituent questions**: the answer to the question will be a constituent.

- 4a Will Jeeves meet his employer at the castle?
- 4b Who will Jeeves meet at the castle?
- 4c Where will Jeeves meet his employer?
- 4d What will Jeeves do?
- 4e Who will meet his employer at the castle?

Non-embedded *yes-no* questions are formed by moving the auxiliary (here, *will*) to the left of the subject. This process is usually referred to as **subject auxiliary inversion**, or SAI. In non-embedded constituent questions, SAI also applies and in addition a sentence-initial question word (such as *who*, *where*, *what*) substitutes for the constituent which is being questioned. In (4b), for instance, *who* substitutes for the subject NP.<sup>4</sup>

Operations such as preposing and question formation thus provide evidence for the role of phrase structure in syntax.

## 2 Words and Phrases

---

Although words are not the immediate constituents of the sentence, they play an important role as the ultimate building blocks of the sentence.

Words belong to different **syntactic categories**, such as nouns, verbs, etc., and the syntactic category to which a word belongs determines its **distribution**, that is, in what contexts it can occur. Normally, one cannot easily interchange words of one category for words of another. If you were to replace the verb *meet* by the semantically related noun *appointment* in (1) you would no longer obtain a grammatical sentence:

- 5 \*Jeeves will appointment his employer at the castle.

<sup>4</sup> Chapter 7 contains a detailed discussion of the formation of questions.

The grammar of English, and indeed of any language, will have to contain the categorial information associated with lexical items since this information plays a part in the formation of sentences.

We assume that the categorial information is also available to the native speakers of the language: they will agree that (5) is unacceptable and that the unacceptability is due to the inappropriate use of the N *appointment*. We postulate that speakers of a language are equipped with an internal 'dictionary', which we shall refer to as the mental lexicon, or *lexicon*, which contains all the information they have internalized concerning the words of their language. As seen above, this mental lexicon will have to contain, among other things, information on syntactic categories. We assume that each word of the language known by a speaker will be listed in his mental lexicon with its categorial specification. For instance, a native speaker of English will presumably have a lexicon containing the following information:

|    |                      |             |
|----|----------------------|-------------|
| 6a | <i>meet</i> :        | verb        |
| 6b | <i>employer</i> :    | noun        |
| 6c | <i>castle</i> :      | noun        |
| 6d | <i>at</i> :          | preposition |
| 6e | <i>the</i> :         | determiner  |
| 6f | <i>his</i> :         | determiner  |
| 6g | <i>appointment</i> : | noun        |

As we suggested in the Introduction, it would not make sense to claim that the native speaker's lexical knowledge, i.e. the mental lexicon, is innate. If lexical knowledge were completely innate, then human beings would have to be born equipped with the lexicons of all known or possible human languages. Rather, we assume that the lexicon of a language is learnt by each native speaker. The speaker learns the words of the language and what category they belong to. But this does not imply that he comes to this learning process totally unprepared. We assume that UG, our innate knowledge of language, contains, for example, the notion of syntactic category. When exposed to the words of a particular language, speakers will have some expectation as to which categories to discover. We shall not speculate further here as to the sort of knowledge this involves.

Lexical information plays a role in sentence structure because the syntactic category of a word determines its distribution. Let us take as an example sentence (1) and consider its syntactic representation (2a). In the tree diagram (2a) the word *appointment* will not be inserted in a position dominated by the node V because only verbs can be inserted under a node V, the same observation would apply to the other words in the sentence. Looking at the

tree diagram from top to bottom we can say that the terminal category labels such as N, V, etc., restrict which lexical elements can be inserted.

Looking at the tree from bottom to top, we see that the words that are inserted at the bottom of the tree determine the structure of the sentence. The inserted words will determine the syntactic category of the **head** of the phrase and hence they will ultimately determine the category of a phrase, the **phrasal category**. For instance, in our example (2a) the inserted N *employer* will be the head of a phrase of the type NP and not of a VP. Chapter 2 provides a more detailed discussion of the principles that regulate sentence structure.

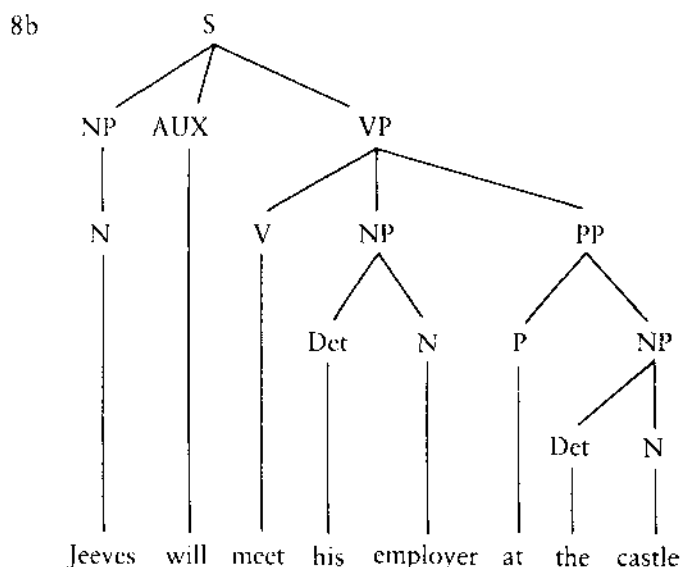
Clearly, the mere matching of lexical and phrasal categories is not sufficient to produce a good sentence. For instance, the random insertion of nouns in the slots provided for them in (2) produces odd results in (7b) and (7c):

- 7a Jeeves will meet his employer at the castle.
- 7b ?Jeeves will meet his castle at the meeting.
- 7c ?Jeeves will meet his castle at the employer.

The question arises whether (7b) and (7c) are ungrammatical: is their oddness due to a violation of a grammatical principle? A native speaker might say that (7b) is bizarre because the verb *meet* is followed by the string *his castle*. The oddness is due to the fact that the concept 'meet' usually involves an interaction between two animate participants, while 'his castle' refers to an inanimate entity which does not normally qualify to take part in an action of the type 'meet'. But if we were to endow the concept 'castle' with animacy the oddness would be removed. In a fairy tale where castles take a walk (7b) would become acceptable. What is wrong with (7b) is not a grammatical issue; its strangeness relates to our general knowledge of the world. Issues of language use which hinge on the interaction of the grammar with extra-linguistic information such as that just described must not be integrated in a grammatical description. Grammars do not contain principles about our beliefs about the world around us. (7b) may therefore be seen as grammatical but as bizarre in view of our encyclopaedic knowledge of castles as inanimate objects.

Let us return to sentence (1), repeated here as (8a), and its tree diagram representation (2a) repeated here as (8b):

- 8a Jeeves will meet his employer at the castle.



It is clear that some of the phrasal constituents of this sentence are more essential to the sentence than others. The PP *at the castle*, which specifies the place of the event, can be left out without any harm, but the NP *his employer* cannot.

8c Jeeves will meet his employer.

8d \*Jeeves will meet at the castle.

In the next section we shall try to account for the obligatory nature of certain constituents in a sentence.

### 3 Predicates and Arguments

---

In this section we explain which constituents of a sentence are minimally required, and why. We first provide an informal discussion and then develop a formal approach known as theta theory.

### 3.1 Subcategorization

Consider the following sentences:

- 9a Maigret will [<sub>VP</sub> imitate [<sub>NP</sub> Poirot] [<sub>PP</sub> with enthusiasm]].  
 9b Bertie will [<sub>VP</sub> abandon [<sub>NP</sub> the race] [<sub>PP</sub> after the first lap]].  
 9c Miss Marple will [<sub>VP</sub> reconstruct [<sub>NP</sub> the crime] [<sub>PP</sub> in the kitchen]].

The labelled bracketing in (9a), (9b) and (9c) shows that these sentences are structurally similar to sentence (8a) with its representation in (8b). In (8a) and in each of the sentences in (9) the VP contains a V, the head of the VP; an NP, the direct object; and a PP. In each of these the PP is optional (as illustrated in (8c) and in (10)): it provides information as to the manner, time or place of the event expressed in the sentence:

- 10a Maigret will imitate Poirot.  
 10b Bertie will abandon the race.  
 10c Miss Marple will reconstruct the crime.

In the traditional literature on parsing, optional phrasal constituents such as the PPs in (8a) and (9) are called *adjuncts*.<sup>5</sup> While the PPs in the examples above are optional, we see that the VP-internal NPs are not:

- 11a \*Maigret will imitate.  
 11b \*Bertie will abandon.  
 11c \*Miss Marple will reconstruct.  
 11d \*Jeeves will meet.

This does not mean, however, that each English sentence contains just one VP-internal NP, as (12) shows:

- 12a Hercule is dithering.  
 12b Wooster gave Jeeves the money.

Inserting an NP into the VP of (12a) renders the example ungrammatical:

<sup>5</sup> In chapters 7, 8 and 9 we shall turn to a technical definition of the notion adjunct, as it is used in the Government and Binding literature.

13a Hercule is dithering \*the crime/\*Agatha.

On the other hand, in (12b) the verb *give* must be followed by two NPs, or alternatively by an NP and a PP:

13b \*Wooster gave Jeeves.

13c Wooster gave [<sub>NP</sub> the money] [<sub>PP</sub> to Jeeves].

In traditional grammar the requirement that there should be or not be one or more NPs inside the VP is seen as a property of the verb involved. At least three classes of verbs are traditionally distinguished: transitive, ditransitive and intransitive verbs.<sup>6</sup> If a VP has a transitive verb as its head, one NP (the direct object) is required: the verb takes an NP complement. If a VP has a ditransitive verb as its head, two NPs or an NP and a PP (the direct object and the indirect object) are required. If a VP contains an intransitive verb as its head then no NP-complement is allowed. Whether a verb belongs to the group of transitive, ditransitive or intransitive verbs is treated as an idiosyncratic property of the verb.

Native speakers of English would agree on the judgements given for the sentences in (12) and (13). This means that they too must have internal knowledge of the principles that decide on the type of VP in which a verb may appear; the subclassification of verbs must constitute part of their lexical knowledge. Let us therefore integrate the information on verb classes in the mental lexicon which we have posited as part of the internal knowledge of the native speaker.

One way of encoding the information on the complement structure of the verb is by associating it directly with the lexical entry of the verb in question. This would lead to the following (partial) lexicon:

- 14a *meet*: verb; transitive
- 14b *imitate*: verb; transitive
- 14c *reconstruct*: verb; transitive
- 14d *abandon*: verb; transitive
- 14e *dither*: verb; intransitive
- 14f *give*: verb; ditransitive

<sup>6</sup> For discussion of the classification of verbs in the traditional literature, see Aarts and Aarts (1982), Burton-Roberts (1986), Huddleston (1984) and Quirk, et al. (1985).

The child acquiring English will have to learn not only the syntactic category of the words of his language, but also the subcategory the verbs belong to. Exposure to English sentences containing these verbs will offer positive evidence of this information: the verb will occur in the appropriate syntactic environment. The child exposed to a sentence like (15a) will thus be able to conclude that *sleep* is intransitive and will assign it the property 'intransitive' in its lexical entry (15b):

15a Mummy is sleeping.

15b *sleep*: verb; intransitive.

In the Chomskian tradition the notions transitive, intransitive, etc., are encoded in distributional frames. Verbs are classified according to the type of VP in which the verb typically occurs. For example, the verb *meet* requires an NP complement; its VP will contain an NP. This requirement can be represented as follows:

16 *meet*: V, [—NP]

(16) shows in which syntactic frame the verb *meet* can and must be inserted: *meet* is inserted in front of an NP. The verbs are characterized on the basis of the frames in which they occur. *Dither*, for instance, does not take any complement; *give* takes either two NPs (12b) or an NP and a PP (13c):

17a *dither*: V, [—]

17b *give*: V, [—NP, NP] or V, [—NP, PP]

The frames that identify subcategories of verbs are called **subcategorization frames**. We say that *meet* **subcategorizes for** or **selects** an NP.

### 3.2 Argument Structure and Thematic Structure

All we have done so far is classify verbs according to whether they require any VP-internal NP. We have not really attempted to explain anything. On the contrary, we have implied that the subcategorization frame of a verb, i.e. whether it is transitive or intransitive, etc., is an unexplained **primitive** property of the grammar, i.e. a property which does not follow from anything else. However, this seems intuitively wrong. Whether a verb is transitive

or not is not a matter of mere chance; it follows from the type of action or state expressed by the verb, from its meaning.

A verb like *imitate* expresses an activity that involves two participants: the active participant, the person who imitates, and the passive participant, the person or thing that is imitated. This notion of 'participants in an activity' has been formalized on the basis of the approach commonly adopted in logic. In this section we first look briefly at the logical system of representation, then we apply it to language in terms of the general notion of argument structure and of the more refined notion of thematic structure.

### 3.2.1 ARGUMENT STRUCTURE IN LOGIC

Logicians have long been concerned with formulating representations for the semantic structure of sentences, or more correctly propositions.<sup>7</sup> In the notation of formal logic, (18a) is assigned the representation (18b):

18a Maigret imitates Poirot.

18b A (mp)

where A = 'imitate', m = 'Maigret' and p = 'Poirot'.

(18a) contains the NPs *Maigret* and *Poirot*, two referring expressions, i.e. expressions which serve to pick out an entity, a person, a thing, from those things we are talking about, the **universe of discourse**. It also contains a predicate *imitate*. The predicate does not refer to a person or thing but rather defines some relation between the referring expressions. In the logical notation in (18b) we see that the **predicate** 'imitate' takes two **arguments**, represented by m (for Maigret) and p (for Poirot). Predicates that require two arguments are **two-place** predicates. The transitive verbs of traditional syntax correspond approximately to the two-place predicates of logic. The arguments of a predicate are realized by noun phrases in our example: in (18a) the subject NP is one argument and the object NP is the second argument of the verb *imitate*.

Intransitive verbs correspond to **one-place** predicates: they take only one argument.

19a Maigret stumbled.

19b S (m)

where S = 'stumble' and m = 'Maigret'.

<sup>7</sup> For an introduction to logic written specifically for the linguist see Allwood, Dahl and Dahl (1977), McCawley (1981).

## 3.2.2 ARGUMENT STRUCTURE IN NATURAL LANGUAGE

Using the basic idea of formal logic outlined above, we can say that every predicate has its **argument structure**,<sup>8</sup> i.e. it is specified for the number of arguments it requires. The arguments are the participants minimally involved in the activity or state expressed by the predicate.

We could use a metaphor to summarize this: predicates are like the script of a play. In a script a number of roles are defined and will have to be assigned to actors. The arguments of a predicate are like the roles defined by the script of a play. For an adequate performance of the play, each role must be assigned to an actor. It will not do either to miss out on a part in the play or to have actors on the stage who have no part to play. Adjuncts might be compared to the parts in the script which are not central to the play.

We first discuss the argument structure of verbs and its relation to subcategorization frames. Then we also turn briefly to the argument structure of adjectives, nouns and prepositions.

The argument structure of the verb determines which elements of the sentence are obligatory. If a verb expresses an activity involving two arguments, there will have to be at least two constituents in the sentence to enable these arguments to be expressed. This conceptually defined argument structure can partly replace the classification of verbs in terms of either transitivity labels or subcategorization frames described above. If a speaker knows the meaning of the verb *meet*, in other words if he knows what activity is expressed, he will also know how many participants are involved and hence how many arguments the verb takes. 'Meet' involves two participants, and hence will be expected to take two arguments. If one argument is realized as the subject of the sentence (cf. section 6 for discussion), it follows that *meet* will select one VP-internal complement.

This does not mean that we can conclude that the verb *meet* necessarily subcategorizes for an NP. After all, the arguments might have been realized by categories other than NP.<sup>9</sup> The argument structure of the verb predicts the number of constituents needed but not necessarily their type. Let us assume for the moment that the type of constituent which realizes the argument must be lexically encoded. We can improve the lexical representation of verbs by specifying their argument structure, which is derived from their meaning, and the specification of the realization of the arguments. This notation replaces the labels transitive, intransitive and ditransitive, or the subcategorization

<sup>8</sup> For a more careful statement see section 5.3 where we discuss the difference between auxiliaries and main verbs.

<sup>9</sup> We return to this point in sections 5.1 and 7.1.

frames illustrated in (14) and (16)–(17) respectively.<sup>10</sup> We shall represent the arguments the verb takes by arabic numerals.

|     |                      |       |    |    |                  |
|-----|----------------------|-------|----|----|------------------|
| 20a | <i>meet</i> :        | verb; | 1  | 2  |                  |
|     |                      |       | NP | NP |                  |
| 20b | <i>imitate</i> :     | verb; | 1  | 2  |                  |
|     |                      |       | NP | NP |                  |
| 20c | <i>reconstruct</i> : | verb; | 1  | 2  |                  |
|     |                      |       | NP | NP |                  |
| 20d | <i>abandon</i> :     | verb; | 1  | 2  |                  |
|     |                      |       | NP | NP |                  |
| 20e | <i>give</i> :        | verb; | 1  | 2  | 3                |
|     |                      |       | NP | NP | NP               |
|     |                      |       | NP | NP | PP <sup>11</sup> |
| 20f | <i>dither</i> :      | verb; | 1  |    |                  |
|     |                      |       | NP |    |                  |
| 20g | <i>smile</i> :       | verb; | 1  |    |                  |
|     |                      |       | NP |    |                  |

Recall that in addition to the arguments of the verb, sentences may also contain adjuncts, constituents providing additional information, for instance with respect to manner, place, time, etc.

<sup>10</sup> There is an important distinction between subcategorization frames and argument structure. Subcategorization frames only specify the complements of the verb, i.e. the constituents that are obligatory inside the VP. The subject NP need not be mentioned in the subcategorization frame because all verbs have subjects, hence the property of having a subject does not create a subcategory of V, whereas the presence of objects does.

The argument structure lists all the arguments: it also includes the subject argument which is realized outside the VP. The thematic structure of the verb (see section 3.2.3) also lists *all* the arguments.

<sup>11</sup> *Give* allows for two types of realizations of its arguments:

- (i) I gave Bill the money.
- (ii) I gave the money to Bill.

The representation in (20e) in the text serves to indicate that the second argument of *give* is either realized as an NP (i) or as a PP (ii). As (ii) and (iii) show, a PP must follow the VP-internal NP. We turn to the relative order of VP-internal constituents in chapter 3.

- (iii) \*I gave to Bill the money.

In some cases it is less easy to determine the argument structure of predicates. Consider the following pairs of sentences:

- 21a Hercule bought Jane a detective story.  
 21b Hercule bought a detective story.

(21a) contains the verb *buy* with apparently three arguments. The argument *Jane* can be omitted, but as a result the meaning of the sentence changes subtly: in the unmarked context (21b) will be taken to mean that Hercule bought the detective story for himself. The action expressed in (21b) still implicitly involves someone for whom the book is bought. (21b) seems to contain an unexpressed or **implicit** argument. We shall encode the fact that some arguments may be left implicit by putting them in parentheses.

- 22 *buy*: verb;    1    (2)    3  
                   NP   NP   NP

So far we have only illustrated the argument structure of verbs. Other lexical categories too have an argument structure. Consider (23):

- 23a Poirot is restless.  
 23b Jeeves is envious of Bertie.  
 23c Jeeves envies Bertie.

In (23a) the predicate *restless*, an adjective, takes one argument. *Restless* is a one-place predicate. The adjective *envious* in (23b) takes two arguments analogously to the verb *envy* in (23c), which is semantically and morphologically closely related to the adjective. (23b) and (23c) are near-paraphrases. The respective arguments of the verb *envy* in (23c) are realized by the two NPs *Jeeves* and *Bertie*. The arguments of *envious* are realized by an NP and by a PP headed by *of*. At this point we merely note that the second argument of the adjective cannot be realized by a straight NP but that it requires the presence of the preposition *of*. The reason why this should be so is treated in chapter 3.

- 23d \*Jeeves is envious Bertie.

We cannot freely add new referring expressions to the sentences in (23):

- 24 \*Poirot is restless of the case.

Unlike verbal arguments, the arguments of adjectives can often be left implicit:

- 25a \*Poirot envies.  
25b Poirot is envious.

We shall again encode the argument structure of adjectives in the lexical information:

- 26a *envious*: adjective; 1 (2)  
NP PP  
26b *restless*: adjective; 1  
NP

The argument structure of lexical items is not always uniquely fixed. Take for instance the adjective *conscious* in the following examples:

- 27a Miss Marple is conscious of the problem.  
27b Sir Galahad is conscious.

We distinguish two argument structures for the adjective *conscious*. *Conscious* is either a two-place predicate (27a) or a one-place predicate (27b). It would not do to say that the second argument of *conscious* is left implicit in (27b) in the way that we argued that the second argument of *envious* in (25b) was implicit. In (25b) the adjective *envious* has the same meaning as in (23b), whereas there is a semantic difference between (27a) and (27b). In (27a) *be conscious* is near-synonymous with *know, be aware*. In (27b) it means 'not be in coma'. Depending on the meaning of the predicate we assume that a different argument structure is associated with it:

- 28a *conscious*<sub>1</sub>: adjective; 1 2  
NP PP  
28b *conscious*<sub>2</sub>: adjective; 1  
NP

*Conscious*<sub>1</sub> will be parallel to *know* or *aware*:

- 29a *know*: verb;           1    2  
                                  NP NP  
29b *aware*: adjective;   1    2  
                                  NP PP

Let us turn to nouns. Consider the following groups of examples:

- 30a Poirot will analyse the data.  
30b \*Poirot will analyse.  
30c \*There will analyse the data.  
  
31a Poirot's analysis of the data was superfluous.  
31b The analysis of the data was superfluous.  
31c The analysis was superfluous.

In (30) the verb *analyse* requires two arguments. The noun *analysis* is semantically and morphologically related to the verb *analyse* and on the basis of (31a) we assume it has the same argument structure.

- 32a *analyse*: verb;       1    2  
                              NP NP  
32b *analysis*: noun; (1) (2)  
                              NP PP

The two arguments of *analysis* are realized overtly in (31a); in (31b) the agent of the activity is left unexpressed and in (31c) both arguments are unexpressed. It is a typical property of nouns that both their arguments may be unrealized.

Prepositions too can be argued to have argument structure. The preposition *in*, for instance, will have two arguments; the preposition *between* will have three:

- 33a John is in London.  
33b *in*: preposition;       1    2  
                                  NP NP  
33c Florence is between Milan and Rome.  
33d *between*: preposition;   1    2    3  
                                  NP NP NP

## 3.2.3 THETA THEORY

Let us consider the argument structure of the verb *kill*.

34a *Maigret killed Poirot.*

34b *kill*: verb;    1    2  
                     NP   NP

In (34a), the two argument-NPs *Maigret* and *Poirot* are intuitively felt to stand in different semantic relationships with the verb. The argument-NP *Maigret* in the subject position refers to the entity that is the **AGENT** of the activity of killing. The argument NP *Poirot*, the direct object, expresses the **PATIENT** of the activity. We used the metaphor of the script of a play when discussing argument structure of predicates. A script of a play defines not only the number of parts to be assigned, hence the number of actors involved, but also what characters are involved, it specifies which roles these actors have to play. The more specific semantic relationships between the verb and its respective arguments may be compared with the identification of the characters in a play script. In the literature these relations between verbs and their arguments are referred to in terms of **thematic roles** or **theta roles** ( $\theta$ -roles) for short. We say that the verb *kill* takes two arguments to which it assigns a theta role: it assigns the role **AGENT** to the subject argument of the sentence, and the role **PATIENT** to the object argument. The verb **theta-marks** its arguments. Predicates in general have a **thematic structure**. The component of the grammar that regulates the assignment of thematic roles is called **theta theory**.

Although many linguists agree on the importance of thematic structure for certain syntactic processes, the theory of thematic roles is still very sketchy. For example, at the present stage of the theory there is no agreement about how many such specific thematic roles there are and what their labels are. Some types are quite generally distinguished. We discuss them informally here.

- 35a **AGENT/ACTOR**: the one who intentionally initiates the action expressed by the predicate.
- 35b **PATIENT**: the person or thing undergoing the action expressed by the predicate.
- 35c **THEME**: the person or thing moved by the action expressed by the predicate.
- 35d **EXPERIENCER**: the entity that experiences some (psychological) state expressed by the predicate.

- 35c BENEFACTIVE/BENEFICIARY: the entity that benefits from the action expressed by the predicate.
- 35f GOAL: the entity towards which the activity expressed by the predicate is directed.
- 35g SOURCE: the entity from which something is moved as a result of the activity expressed by the predicate.
- 35h LOCATION: the place in which the action or state expressed by the predicate is situated.

The inventory above is very tentative. Other authors amalgamate the roles PATIENT and THEME under the one role of THEME.

- 35i THEME<sub>2</sub>: the entity affected by the action or state expressed by the predicate.

We usually use the term THEME in this second interpretation.

The thematic roles are illustrated in (36):

- 36a Galahad gave the detective story to Jane.  
AGENT        THEME        BENEFACTIVE/GOAL
- 36b Constance rolled the ball towards Poirot.  
AGENT        THEME GOAL
- 36c The ball rolled towards the pigsty.  
THEME        GOAL
- 36d Madame Maigret had been cold all day.  
EXPERIENCER
- 36e Maigret likes love stories.  
EXPERIENCER THEME
- 36f Love stories please Maigret.  
THEME        EXPERIENCER
- 36g Poirot bought the book from Maigret.  
AGENT        THEME SOURCE
- 36h Maigret is in London.  
THEME        LOCATION

The identification of thematic roles is not always easy, as the reader can see for himself. However, intuitively the idea should be clear, and we shall be drawing on this rather intuitive approach to theta theory in subsequent discussion.

The information as to the semantic relationship between the predicate and its arguments is part of the lexical knowledge of the native speaker and should hence also be recorded in the lexicon. Rather than merely specifying the number of arguments of a predicate, one may envisage a representation which specifies the type of semantic roles of these arguments. In Government and Binding Theory this is represented by means of a **thematic grid**, or **theta grid**, which is part of the lexical entry of the predicate. *Kill* would be given the lexical representation in (37a):

37a *kill*: verb

| AGENT | PATIENT |
|-------|---------|
|       |         |

(37a) specifies that *kill* **assigns** two thematic roles (AGENT and PATIENT). We deduce that the verb is a two-place predicate, which requires two arguments to which these roles can be assigned. Some linguists propose that the syntactic category realizing the thematic role should also be specified in the theta grid of a predicate (cf. section 7.1 for discussion).

37b *kill*: verb

| AGENT<br>NP | PATIENT<br>NP |
|-------------|---------------|
|             |               |

Consider some examples:

38a Maigret killed the burglar.

38b \*Maigret killed.

38c \*Maigret killed the burglar the cellar.

We see that two arguments and no more than two are needed. In (38b) the absence of the second NP renders the sentence ungrammatical: the second

theta role cannot be assigned. In (38c), conversely, one extra NP is added to the sentence. This NP cannot be assigned a thematic role because *kill* only assigns two roles, which are already assigned to the subject NP and to the object NP respectively. In (38d) we have inserted the preposition *in*. The sentence is grammatical: the preposition *in* assigns the thematic role of LOCATION to the NP *the cellar*.

38d Maigret killed the burglar in the cellar.

One criterion for judging whether a sentence is grammatical is that the thematic roles associated with its predicate(s) must be assigned to arguments, these arguments must be structurally realized. Conversely, the referring NPs in the sentence must bear some semantic relation to a predicate. This semantic relation can be established via the assignment of thematic roles.

Each syntactic representation of a sentence is scanned for the predicate(s) it contains. Each predicate is tested with respect to its argument structure. Its arguments must be realized. More specifically the predicate is tested for its thematic roles: each role must be assigned to an argument.

Let us take as an example a sentence containing the predicate *kill*. *Kill* assigns the thematic roles of AGENT and PATIENT, hence it requires two arguments. When the theta roles can be assigned to arguments we say that they are saturated and we mark this by checking off the theta role in the thematic grid of the predicate. In order to identify the assignment of the respective thematic roles to the corresponding arguments, NPs are identified by means of an index, a subscript:

39a Maigret<sub>i</sub> killed the burglar<sub>j</sub>.

39b \*Maigret<sub>i</sub> killed.

39c \*Maigret<sub>i</sub> killed the burglar<sub>i</sub> the cellar<sub>k</sub>.

We shall not discuss the subscripting convention here. We hope that the intuitive idea is clear: an NP refers to an individual or an object and is identified by the referential index. Two NPs with the same index are said to be **coindexed**: they are interpreted as referring to the same entity.<sup>12</sup>

40a Maigret<sub>i</sub> said that he<sub>j</sub> was ill.

40b Maigret<sub>i</sub> hurt himself<sub>i</sub>.

<sup>12</sup> We return to coindexation in chapter 4 and in chapter 12.

In order to show how the theta roles of a predicate are assigned we enter the index of the argument to which the thematic role is assigned in the appropriate slot in the theta grid. For (39a) the saturation of the thematic roles can be represented as in (41):

41 *kill*: verb

| AGENT<br>NP | PATIENT<br>NP |
|-------------|---------------|
| i           | j             |

If we try to do the same for (39b) we see that one of the slots in the thematic grid will remain unfilled: one thematic role is not assigned. Conversely, in (39c) there is one referential index which cannot be entered on the grid, hence cannot be assigned a thematic role.

42a *kill*: verb

| AGENT<br>NP | PATIENT<br>NP |
|-------------|---------------|
| i           | ?             |

42b *kill*: verb

| AGENT<br>NP | PATIENT<br>NP |    |
|-------------|---------------|----|
| i           | j             | k? |

In (42a) corresponding to (39b) the thematic role of PATIENT is not assigned or not saturated. In (42b), corresponding to (39c), the argument-NP *the cellar* with the referential index *k* fails to be assigned a thematic role.

The requirement that each thematic role of a predicate must be assigned and that there must be no NPs that lack a thematic role is summed up in the **theta criterion**:

## 43 Theta criterion

43a Each argument is assigned one and only one theta role.

43b Each theta role is assigned to one and only one argument.

So far we have only discussed NP arguments. But other constituents may also be arguments: consider, for instance, (44a) and (44b):

44a The police announced the news.

44b The police announced that the pig had been stolen.

In (44a) *announce* is associated with two arguments, which will be assigned their thematic roles. The role AGENT is assigned to *the police*; THEME to *the news*. In (44b) the THEME role is assigned to a subordinate clause: *that the pig had been stolen*. Clauses too can thus be arguments of the predicate. We return to the issue in more detail in section 5.1.

Given the wide diversity in the labelling of thematic roles and their definitions it would be a difficult enterprise to fix the types of roles and their exact number. Even if we are unable to pin down the exact nature of the different roles involved we are usually quite clear as to how many arguments a predicate requires in a given reading. Hence, instead of specifying the exact type of thematic role for each predicate, we shall often merely list the number of arguments, identifying their roles by numbers rather than by role labels. Thus for the verb *kill* we shall use the following lexical representation, unless we need to refer explicitly to the thematic label.

45 *kill*: verb

|         |         |
|---------|---------|
| 1<br>NP | 2<br>NP |
|         |         |

The numerals 1 and 2 represent the thematic roles assigned by the verb whose labels need not concern us.

Research in this area suggests that it might not be necessary or desirable to refer to the thematic labels in the syntax,<sup>13</sup> and that indeed the representation

<sup>13</sup> For discussion, see Grimshaw (1979, 1981).

in (45) is the one we need. We do not go into that discussion here and we refer the reader to chapter 3, section 6.3 and to the literature.

## 4 The Projection Principle

---

Let us sum up what we have done so far. We have seen that the lexical items which are the ultimate constituents of a sentence play an important part in its syntactic representation. Section 2 shows that the lexical category of the head of a phrase determines the category of the phrase. Second, we have seen in section 3 that the thematic structure of a predicate, encoded in the theta grid, will determine the minimal components of the sentence. This idea that lexical information to a large extent determines syntactic structure is summed up in the projection principle:

### 46 Projection principle

Lexical information is syntactically represented.

The projection principle will play an important role throughout this book. For a discussion of the role of the lexicon in syntax see also Stowell and Wehrli (1992) and the references cited there.

## 5 The Assignment of Thematic Roles

---

In this section we look at the assignment of thematic roles in the syntax. We focus on three areas: section 5.1 discusses clausal arguments; section 5.2 discusses expletive pronouns; and section 5.3 considers the difference between lexical verbs, or main verbs, and auxiliaries.

### 5.1 *Clausal Arguments*

We have seen that the obligatory constituents of a sentence are determined by the semantic properties of the predicates (verbs, adjectives) and we have mainly discussed examples with NP-arguments. Sentences too may be arguments of a predicate.

- 47a Miss Marple has announced the news.  
 47b Miss Marple has announced that Poirot had left.

In (47a) the verb *announce* takes two arguments, realized by the NPs *Miss Marple* and *the news* respectively. In (47b) the arguments are realized by an NP and by the clause [*that Poirot had left*]. Consider also the following examples:

- 48a The robbery surprised all the inhabitants of Blandings.  
 48b [That the pig was stolen] surprised all the inhabitants of Blandings.  
  
 49a Jeeves' decision is very unfortunate.  
 49b [That Jeeves should be leaving] is very unfortunate.  
  
 50a Poirot asked three questions.  
 50b Poirot asked [whether anyone had seen the pig].  
  
 51a Maigret believes the story about the burglary.  
 51b Maigret believes [that the taxi driver is lying].  
  
 52a Constance is aware of the problem.  
 52b Constance is aware [that the pig is in danger].

The verb *surprise* takes two arguments. In (48a) both arguments are realized by NPs; in (48b) one argument is realized by a clause. Similarly, in (49a) the one argument of the adjective *unfortunate* is realized by an NP and in (49b) it is realized by a clause. In (50) and (51) we find further alternations between NPs and clauses as realizations of arguments. In (52) one of the arguments of the adjective *aware* is realized by an NP contained in a PP in (52a) and by a clause in (52b).

We conclude that the theta grid of predicates will not always specify a unique category to which a theta role can be assigned but will allow for a choice. We return to this point in section 7.1.

Let us consider clausal arguments a little more closely:

- 53a [That Galahad had left] is very surprising.  
 53b [For Galahad to have left] is very surprising.

- 54a Maigret<sub>i</sub> believes [this story]<sub>j</sub>.  
 54b Maigret<sub>i</sub> believes [that the taxi driver is innocent]<sub>j</sub>.  
 54c Maigret<sub>i</sub> believes [the taxi driver to be innocent]<sub>j</sub>.  
 54d Maigret<sub>i</sub> believes [the taxi driver innocent]<sub>j</sub>.

In (53) we see that the adjective *surprising* takes one argument, to which it assigns a thematic role. The argument is realized by a finite clause in (53a): the verb *had* is finite, it is inflected for the past tense and the clause is introduced by the complementizer *that*. In (53b) the argument of the main predicate is realized by a non-finite clause: *have* is in the infinitive and the sentence is introduced by the complementizer *for*. We return to the general principles of sentence structure in chapter 2.<sup>14</sup>

<sup>14</sup> Koster (1978b) argues that what looks like a clausal subject in (48b) and in (49b) is not in the subject position. Observe for instance that (48b) cannot be embedded as such.

- (ia) I wonder whether the robbery surprised all the inhabitants of Blandings  
 (ib) \*I wonder whether [that the pig was stolen] surprised all the inhabitants of Blandings.

(ib) can be made grammatical if we move the clausal subject to a final position and replace it by the pronoun *it*. As mentioned in the introduction, the pronoun *it* in examples such as (ic) seems to make no contribution to the semantics of the sentence. We return to this use of *it* in section 5.2.1.

- (ic) I wonder whether it surprised all the inhabitants of Blandings [that the pig was stolen].

Similarly, Stowell (1981) suggests that object clauses also do not occupy the same position as the object NP. Observe, for instance, that in English the object-NP *the situation* in (iia) can, and indeed must (cf. (ib)) precede the adverbial adjunct *very carefully*, while an object clause must follow it.

- (iia) He explained the situation very carefully.  
 (iib) \*He explained very carefully the situation.  
 (iic) \*He explained that he was not going to leave very carefully.  
 (iid) He explained very carefully that he was not going to leave.

There is further support for these observations from Dutch. In this language subordinate clauses have the SOV pattern: in (iia) the direct object NP *de waarheid* ('the truth') precedes the verb *zegt* ('tells'), the reverse order is not possible (iib):

- (iia) Ik verwacht dat Jan [de waarheid] zegt.  
 I expect that Jan the truth tells  
 'I expect Jan to tell the truth.'

In (54a) both arguments of *believe* are realized by NPs. In (54b) one of the arguments of *believe* is realized by a finite clause. As the bracketing indicates, the corresponding argument is realized by a non-finite clause in (54c).

The bracketing in (54c) is meant to show that we consider *the taxi driver* to form a constituent with *to be innocent*. The justification for this analysis is essentially one of analogy. If we compare the sentences (54b) and (54c) we see that they are very similar in meaning. In (54b) the verb takes two arguments: one argument which is realized by the subject NP, and one argument which is realized by a sentence. On the basis of examples like (54a) and (54b) we deduce that the lexical entry of *believe* has the following theta grid:

55 *believe*: verb

|         |           |
|---------|-----------|
| 1<br>NP | 2<br>NP/S |
|         |           |

In (54a) the arguments are saturated as in (56), where *j* is the index of an NP. In (54b), similarly, the saturation of the arguments can be represented as in (56), with *j* now seen as the index of a subordinate clause.

56 *believe*: verb

|          |           |
|----------|-----------|
| 1<br>NP  | 2<br>NP/S |
| <i>i</i> | <i>j</i>  |

- 
- (iiib) \*Ik verwacht dat Jan zegt [de waarheid].  
I expect that Jan says the truth.  
(iiic) \*Ik verwacht dat Jan [dat hij ziek is ] zegt.  
I expect that Jan [that he ill is] says  
(iiid) Ik verwacht dat Jan zegt [dat hij ziek is].  
I expect that Jan says that he ill is

For further discussion we refer the reader to the texts cited. However, these texts will not be accessible until we have covered the material in chapter 7.

Given the close similarity in meaning between (54b) and (54c), the minimal assumption is that the verb *believe* in (54c) is the same as that in (54b) and has the same theta grid. While in (54b) the second argument is associated with a finite clause, in (54c), the second argument is associated with a non-finite clause.<sup>15</sup> The theta roles in (54c) are saturated as in (56), with *j* standing for the non-finite clause.

(54d) is also very close in meaning to (54b) and (54c), so we postulate that the verb *believe* is unaltered and has the theta grid in (55). Given this assumption, we need to assign to (54d) a structure that allows the saturation of the argument roles 1 and 2. The bracketing in (54d) will do that adequately.

It is not immediately obvious how to label the structure [*the taxi driver innocent*]. In the traditional literature on parsing, the term 'verbless clause' is sometimes used. This term serves to indicate that we have a constituent which has a propositional meaning, i.e. the same sort of meaning as a full clausal structure has, but it lacks any verb forms. In (54d) the constituent [*the taxi driver innocent*] corresponds to the sentence [*the taxi driver to be innocent*] in (54c). In both sentences the NP *the taxi driver* is the subject of the predicate expressed by the AP *innocent*. In the Government-Binding literature, constituents such as [*the taxi driver innocent*] are called **small clauses**. We return to their structure throughout the book.

Non-finite clauses and small clauses are not normally<sup>16</sup> found as independent clauses: they can only be subordinate to some other main predicate. The italicized constituents in (57) are all small clauses:

<sup>15</sup> Note, however, that in this particular example, the non-finite clause cannot be introduced by the complementizer *for*. We shall return to this issue in chapters 2 and 3.

<sup>16</sup> Small clauses seem to be in frequent use in certain registers, such as informal notes or telegrams (i) or newspaper headlines (ii):

- (i) Mother in hospital.
- (ii) Hijackers under arrest.

Register-specific syntactic properties have not often been studied in the generative framework (see Haegeman, 1987, 1990; and Massam and Roberge, 1989).

We also find small clauses in colloquial expressions such as:

- (iii) What? Me angry?

For small clauses in acquisition see Radford (1990).

- 57a I consider *John a real idiot*.  
 57b The chief inspector wants *Maigret in his office*.  
 57c Emsworth got *Galahad in trouble*.

It is evident that the small clauses are of different types. In (57a) the small clause consists of an NP *John* and a second NP *a real idiot*. The first NP acts as a subject to the second one. In (57b) and (57c) the small clause is composed of an NP and a PP, where the NP is the subject with respect to the PP predicate.<sup>17</sup> That the italicized strings in (57) are constituents is supported by the fact that other material associated with the main verb of the sentence cannot occur internally to what we have called the small clause:

- 58 \*The chief inspector wants [Maigret [very much] in his office].

[\_\_\_\_\_]

In (58) the degree adjunct *very much*, which modifies the verb *want*, cannot intervene between the subject and the PP predicate of the small clause.

## 5.2 Expletives

Section 5.1 shows that not all arguments of a predicate are necessarily realized as NPs. In this section we shall see that the reverse also holds: some NPs in the subject position of the sentence are not assigned a thematic role, hence are not arguments of the predicate.

### 5.2.1 IT AND EXTRAPOSITION

The obligatory presence of certain constituents in a sentence can be accounted for in terms of the argument structure of the predicate of a sentence. Let us now extend our analysis to some further data:

- 59a The burglary surprised Jeeves.  
 59b That the pig had been stolen surprised Jeeves.  
 59c It surprised Jeeves that the pig had been stolen.

<sup>17</sup> For a discussion and further motivation of the analysis of small clauses, see Stowell (1983), Hornstein and Lightfoot (1987). However, these texts will only become accessible once chapter 8 has been covered.

From (59a) and (59b) we infer that *surprise* takes two arguments. Neither of these can be omitted:

- 60a \*The burglary surprised.  
 60b \*Surprised Jeeves.  
 60c \*That the pig had been stolen surprised.

*Surprise* will be associated with the thematic grid (61):

61 *surprise*: verb

|           |         |
|-----------|---------|
| 1<br>NP/S | 2<br>NP |
|           |         |

We cannot insert another NP in these sentences since this would not be assigned a theta role by *surprise*.

- 62a \*The burglary surprised Jeeves *it*.  
 62b \*That the pig had been stolen surprised Jeeves *it*.

In (62a) or (62b) the NP *it* cannot be assigned a thematic role and thus the sentence violates the theta criterion (43). The theta criterion specifies that theta roles are assigned uniquely. Hence one could not, for instance, propose that in (62a) theta role 1 is assigned both to the subject NP *the burglary* and to the NP *it*.

Now let us look at (59c) repeated here as (63a):

- 63a It surprised Jeeves that the pig had been stolen.

(63a) is a paraphrase of (59b). We deduce that *surprise* in (63a) has the theta grid given in (61) with two theta roles to be associated with two arguments. How are these arguments realized? If we capitalize on the equivalence between (63a) and (59b) then the easiest thing would be to say that in both (59b) and (63a) one theta role, say 1, is assigned to the clause [*that the pig had been stolen*] and the other one, 2, to *Jeeves*. This hypothesis leaves us with the NP-constituent *it* in the subject position of (63a) unaccounted for. This NP is not optional:

63b \*Surprised Jeeves that the pig had been stolen.

On the other hand, *it* cannot be assigned a thematic role since *surprise* only assigns two thematic roles already saturated as described above.

One element in the discussion is that the choice of a filler for the subject position in (63a) is very limited: indeed no other NP (pronominal or not) can fill the position:

64a \**This* surprised Jeeves that the pig had been stolen.

64b \**He* surprised Jeeves that the pig had been stolen.

Moreover it is not possible to question the element *it* in (63a), nor can *it* receive focal stress:

64c \**What* surprised Jeeves that the pig had been stolen?

64d \**IT* surprised Jeeves that the pig had been stolen.

In fact, the pronoun *it* in (63a) contributes nothing to the meaning of the sentence, (63a) being a paraphrase of (59b). *It* is not a referring expression: it does not refer to an entity in the world, a person or an object; it cannot be questioned; it cannot receive focal stress.

On the basis of these observations we formulate the hypothesis that *it* plays no role in the semantic make-up of the sentence and that its presence is required in (63a) simply for some structural reason. The relevant explanation for the presence of *it* in the subject position in (63a) will be argued to be that English sentences must have an overt subject (see section 6 and chapter 2 for more discussion). The pronoun *it* in (63a) acts as a mere slot-filler, a dummy pronoun without semantic contribution to the sentence; it is a place-holder for the otherwise unfilled subject position. In the literature such a dummy pronoun is often called an **expletive** pronoun. The term **pleonastic *it*** is also used. Expletives are elements in NP positions which are not arguments and to which no theta role is assigned.<sup>18</sup> Note that, unlike adjuncts, expletives contribute nothing to the sentence meaning.

<sup>18</sup> The reader will recall from the Introduction that in *pro*-drop languages such as Italian the subject pronoun can be omitted (ia) and that in such languages there are no expletive pronouns (ib):

(ia) (Io) parlo Italiano.  
(I) speak (Isg) Italian

In an example like (63a) it is sometimes said that the sentential subject is **extraposed** and that it is **in construction with** an expletive. (65) contains some more examples of extraposition patterns. The extraposed clause and the expletive are italicized:

- 65a *It* worries Maigret *that Poirot should have left*.  
 65b *It* is unfortunate *that Poirot should have said that*.  
 65c *It* is out of the question *that Jeeves should be fired*.  
 65d I consider *it* odd *that Poirot should have left*.

The expletive *it* cannot just appear in any type of sentence. Consider for instance the following pair:

- 66a An announcement about the robbery worried Maigret.  
 66b \*It worried Maigret an announcement about the robbery.

(66b) shows that the expletive *it* cannot become the place-holder for an extraposed NP.<sup>19</sup>

### 5.2.2 THERE AND EXISTENTIAL SENTENCES

Now let us turn to another sentence pattern which poses problems for our theory outlined so far.

- 67a Three pigs are escaping.  
 67b There are three pigs escaping.

- 
- (ib) (\*Ciò) è chiaro che Louisa non partirà.  
 (\*it) is clear that Louisa not leave (fut, 3sg)  
 'It is clear that Louisa won't go.'

In *pro-drop* languages the subject pronoun is only present when it receives stress (for instance when contrasted or focalized):

- (ic) *Lei* parte e *lui* arriva.  
*she* leaves and *he* arrives

Since expletive pronouns cannot receive focal or contrastive stress it follows that they will not be used in *pro-drop* languages.

<sup>19</sup> A good survey of the literature on extraposition can be found in Williams (1980). For different views see also Bennis (1986), Grange and Haegeman (1989) and Postal and Pullum (1988).

In (67a) the verb *escape* has one argument, realized by the NP *three pigs*. In (67b) the sentence contains one more element: the pronominal *there*, which occupies the subject position. First note that *there* is not an adjunct of place. In (67b) *there*, unlike place adjuncts, cannot be questioned:

- 68a I saw Bill there last week.  
       Where did you see Bill last week?  
       There.  
 68b \*Where are three pigs escaping? There.

Also, unlike the place adjunct, *there* in declarative (67b) cannot be omitted freely:

- 69a I saw Bill last week.  
 69b \*Are three pigs escaping.

But *there* does not really contribute anything to the meaning of (67b), which has the same meaning as (67a). Again the data suggest that *there* is required for structural reasons: it fills up the subject position. Unlike the place adjunct *there*, *there* in (67b) cannot receive focal stress:

- 70a I saw him (right) *there*.  
 70b \**There* are three pigs escaping.

As was the case with the pronominal *it* discussed before, we call *there* an expletive. In contrast with *it*, *there* is associated with NP-subjects which have been moved to the right in the sentence, and it cannot be associated with clausal subjects:

- 71a \*There surprised Jeeves [that the pig had been stolen].

The construction with *there* has many intriguing properties. For instance the *there*-construction is only allowed if the moved subject NP is indefinite. There are also heavy restrictions on the type of verb that can occur in this construction. Transitive verbs, for instance, are disallowed.<sup>20</sup>

<sup>20</sup> For the discussion of the *there*-construction, see Belletti (1988), Milsark (1974, 1977) Moro (1989) and Stowell (1978). These texts will be accessible after chapter 6 has been covered.

71b \*There are the three pigs escaping.

71c \*There saw three children the pigs.

### 5.2.3 CONCLUSION

We have seen that there are two pronouns in English, *it* and *there*, that can be used without being assigned a thematic role. They are expletives filling the subject position for structural reasons. We turn to those structural reasons in section 6.

Expletives always turn up in the subject position of the sentence, i.e. the NP position for which the verb does not subcategorize. Indeed the theory outlined so far predicts that expletives will never turn up in subcategorized positions. Expletives are elements lacking a theta role. The positions a verb subcategorizes for are determined by the thematic structure of the verb. Whenever a verb requires a complement NP, this is because the verb has a theta role to assign to the NP. Inserting an expletive NP in an object position would miss the point, because the expletive element would not be able to receive the theta role. In (72) we find a pronoun *it* as the object of *believe*, but this pronoun is not an expletive: it is assigned a thematic role by the verb. In such examples *it* can substitute for other NPs:

72 Poirot believes *it*/this/this story/the announcement.

The prediction of the theory outlined is thus that expletives can only occur in NP positions that are not subcategorized for, i.e. the subject position of the sentence.<sup>21</sup>

## 5.3 Main Verbs and Auxiliaries

So far we have implied that all verbs assign thematic roles. However, it is well known that the class of verbs can be divided into two sets: (a) lexical verbs or main verbs like *eat*, *sleep*, *walk*, and (b) auxiliaries: *be*, *have*, *do*, and the modal auxiliaries *will*, *shall*, *can*, *may*, *must*, *ought*. All these elements are inflected for tense.<sup>22</sup>

<sup>21</sup> See Postal and Pullum (1988) for a different view.

<sup>22</sup> In the present tense, verbs and the auxiliaries *have* and *be* are also inflected for person and number. Modals are not inflected for person and number. For discussion cf. chapter 11. For the development of modals see Lightfoot (1979).

| 73 | Verb         | Present tense |        | Past tense |
|----|--------------|---------------|--------|------------|
| a  | <i>eat</i>   | eat           | eats   | ate        |
| b  | <i>sleep</i> | sleep         | sleeps | slept      |
| c  | <i>walk</i>  | walk          | walks  | walked     |
| d  | <i>be</i>    | am/are        | is     | was/were   |
| e  | <i>have</i>  | have          | has    | had        |
| f  | <i>can</i>   | can           | can    | could      |
| g  | <i>do</i>    | do            | does   | did        |

Auxiliary verbs have some special properties distinguishing them from lexical verbs. In (74) and (75) we have paired sentences containing a lexical verb in (a) and an auxiliary in (b). The reader can check that auxiliaries and main verbs behave differently in negative and interrogative patterns:

- 74a John eats chocolate.  
       \*John eatn't chocolate.  
       John doesn't eat chocolate.
- 74b John has eaten chocolate.  
       John hasn't eaten chocolate.  
       \*John doesn't have eaten chocolate.
- 75a John eats chocolate.  
       \*Eats John chocolate?  
       Does John eat chocolate?
- 75b John has eaten chocolate.  
       Has John eaten chocolate?  
       \*Does John have eaten chocolate?

The negation element *n't* follows the auxiliaries (cf. (74b)), whereas it must precede the lexical verb (cf. 74a)). In a *yes-no* question the auxiliary and the subject of the sentence are inverted (see chapter 2 for discussion). Lexical verbs do not invert with their subjects: in both negative sentences and in questions the auxiliary *do* is needed. We return to the positions of lexical verbs and of auxiliaries in chapter 7 and especially in chapter 11. Now let us consider the thematic structure of auxiliaries and main verbs.

- 76a Poirot accuses Maigret.  
 76b Poirot has accused Maigret.  
 76c Poirot is accusing Maigret.  
 76d Poirot does not accuse Maigret.

In (76a) the assignment of the thematic roles of *accuse* is straightforward: one thematic role will be assigned to the NP *Poirot* and the other one to *Maigret*. In addition to the lexical verb *accuse*, (76b) contains the perfective auxiliary *have*. The sentence is grammatical, which must mean that all thematic roles of the predicate(s) are assigned and that all referring NPs in the sentence have a thematic role assigned to them. Given that there are just as many NPs present in (76b) as in (76a), we are led to conclude that the auxiliary *have*, though morphologically like a verb in that it is inflected for tense, person and number, does not assign any thematic roles of its own. If *have* did assign any thematic roles then we would expect (76b) to contain one or more NPs in addition to those in (76a), which would be assigned the thematic roles of the auxiliary. The same argument can be applied to the auxiliaries *be* in (76c) and *do* in (76d). We conclude that auxiliaries do not assign thematic roles.

A related problem appears in connection with the copula *be* in (77) (cf. example (54)).

77a *Maigret*<sub>i</sub> believes [that the taxi driver is innocent]<sub>i</sub>

77b *Maigret*<sub>i</sub> believes [the taxi driver to be innocent]<sub>i</sub>

77c *Maigret*<sub>i</sub> believes [the taxi driver innocent]<sub>i</sub>

In (77) *believe* assigns one theta role to *Maigret* and it assigns the second one to the bracketed clausal constituents. We are concerned with the internal predicate argument structure of the clausal argument. In the discussion above we have argued that the finite complement clause in (77a), the non-finite one in (77b) and the small clause in (77c) all basically mean the same thing: in all of them the property *innocent* is ascribed to the referent of the NP *the taxi driver*. We have also seen that adjectives, like verbs, have an argument structure. Let us first turn to (77c). Inside the small clause *the taxi driver innocent* the NP *the taxi driver* must have been assigned a thematic role, by virtue of clause (43a) of the theta criterion. We deduce that the NP is assigned a thematic role by *innocent*. The adjective will have the thematic grid (78):

78 *innocent*: adjective

|   |
|---|
| 1 |
|   |

In the small clause in (77c) the theta role is assigned to the NP *the taxi driver*. With respect to (77a) and (77b) *innocent* must also be able to assign its thematic role and by analogy with (77c) we assume that it assigns it to the NP *the taxi driver* and hence we conclude that the copula *be*, like auxiliaries, does not assign any thematic roles. Interestingly, the copula *be* also has the other syntactic properties of the auxiliaries:

- 79a The taxi driver is innocent.  
 79b The taxi driver isn't innocent.  
 79c Is the taxi driver innocent?

The formal differences between main verbs on the one hand and auxiliaries and the copula *be* on the other are matched by a semantic property: neither auxiliaries nor the copula *be* assign thematic roles.<sup>23</sup>

## 6 The Extended Projection Principle (EPP)

---

Our discussion reveals that sentence constituents may be required for two reasons.

In the first place, the argument structure and the theta grid of the predicate determine the minimal composition of the sentence. Sentence structure is thus partly lexically determined. This property of syntactic representations is summed up in the projection principle ((46), section 4).

Second, expletive elements are required to fill the subject position in certain constructions (section 5.2). The structural requirement which necessitates the insertion of expletives is that the subject position of a sentence must be filled, i.e. sentences must have subjects.<sup>24</sup> This requirement is not one that is specific to individual lexical items, but it is a general grammatical property of all sentences. In this respect the structural requirement that sentences have sub-

<sup>23</sup> See Pollock (1989) for an explanation.

<sup>24</sup> In chapters 5–8 we shall see that the subject may be non-overt. The reader will recall that in *pro*-drop languages the subject pronoun can be omitted:

- (i) Parlo italiano.  
     speak (1sg) Italian

This might seem in contradiction with the subject requirement discussed in the text. For sentences lacking an overt subject we will propose that there is a non-overt pronoun in the subject position. We discuss this type of non-overt subject in *pro*-drop languages in chapter 8.

jects is an addition to the projection principle. Not only must lexical properties of words be projected in the syntax, but in addition, regardless of their argument structure, sentences must have subjects. The latter requirement has come to be known as the **extended projection principle (EPP)** (80). The phrase structure rules of our grammar (cf. (2b)) will specify that every sentence has a subject. (We return to a discussion of phrase structure in chapter 2.)

#### 80 Extended projection principle

$S \longrightarrow NP - AUX - VP$

Consider (81):

81a *Maigret* accused *Poirot*.

81b \*Accused *Poirot* *Maigret*.

In both (81a) and (81b) the two arguments of *accuse* are realized by the NPs *Maigret* and *Poirot*. The ungrammaticality of (81b) follows from the extended projection principle (80): the subject position is not filled. Insertion of the expletive *there* is not possible because *there* cannot be associated with definite NPs and also it cannot be used with transitive verbs. Similarly, *it* cannot be inserted since this expletive cannot be in construction with an NP.

82a \**There* accused *Maigret* *Poirot*

82b \**It* accused *Maigret* *Poirot*.

## 7 Thematic Roles: Further Discussion

### 7.1 The Canonical Realization of Theta Roles

Syntacticians do not agree about the question whether the lexical information associated with a lexical item must specify the categorial realization of the thematic roles, i.e. whether it must specify by which syntactic categories the thematic roles such as AGENT, THEME etc., are realized. It has been noted, for instance, that thematic roles such as AGENT are always realized by an NP. It is then said that the **canonical** realization of AGENT is NP, i.e. that this theta role is normally realized by NP. This means that the theta grid of a verb which assigns a theta role AGENT will not need to specify by which syntactic category the theta role is realized. In other words, this view proposes

that categorial selection or c-selection follows from semantic selection or s-selection. However, there is not always a perfect match between certain types of thematic roles and the corresponding syntactic realization. Compare, for instance, the examples in (83) (from Rothstein 1992):

- 83a I asked        what the time was.  
                       the time.  
 83b I inquired    what the time was.  
                       \*the time.

*Ask* and *inquire* are near-synonymous: both take a semantic complement of the type which we could informally describe as 'Question'. We could say that questions are canonically realized by interrogative clauses, and thus we expect that both *ask* and *inquire* take a clausal complement. However, the verb *ask* can also select an NP-complement (*the time* in (83a)), while *inquire* cannot. In (83a) the NP *the time* is interpreted as equivalent to a question ('what the time was'); NPs interpreted as questions are called **concealed questions**. Grimshaw (1981) uses data as those in (83) to propose that there are predictable relations between the semantic characteristics of the complement of a lexical item and the canonical realization, which she labels **canonical structural representations (CSR)**. Categorial information concerning the syntactic representation of an argument of a certain head will only be required in those cases where this argument is realized by a non-predictable category. For instance, categorial information is needed to specify that the complement of *ask* may be realized by an NP, though not to specify that it may be a clause. The latter information is derived from the semantic information that the complement of *ask* is a question. Grimshaw stresses that we cannot eliminate c-selection from lexical information because we need to specify the non-canonical realizations. We briefly return to this point in Chapter 3 (section 6.3).

In our discussion we often omit the specification of the syntactic category to which a thematic role is assigned and use representations as (83c). This convention is adopted for convenience' sake and implies no decision with respect to the relation between categorial selection and thematic roles.

83c *inquire*: verb

|   |   |
|---|---|
| 1 | 2 |
|   |   |

## 7.2 The Subject Theta Role

With respect to the assignment of thematic roles we have treated arguments in the subject position in the same way as arguments in the object position. For instance in (84) we would have said that both the NP *Maigret* and the NP *the taxi driver* are assigned a theta role by the verb *accuse*.

84 *Maigret accused the taxi driver.*

Two related observations are often advanced in the literature for treating subject arguments as different from object arguments. On the one hand, the choice of the object affects the thematic role of the subject while the choice of the subject argument does not affect the role of the object, and on the other hand, there exist 'object idioms' with the subject as a free argument while there are no subject idioms with a free object.<sup>25</sup>

In (85) we see how the choice of the object may determine the theta role of the subject:

85a *John broke a leg last week.*

85b *John broke a vase last week.*

In both (85a) and (85b) the verb *break* takes an NP complement. The choice of the complement determines the thematic role of the subject: while in (85b) *John* could be considered AGENT, in (85a) this is not the case: John is the one who undergoes the event. Consider also (86). While in (86a) the literal meaning of the verb *kill* is intended, the other three examples are idioms with free subjects. The idiomatic interpretation of the sentence depends on the combination of the verb *kill* and its object.

86a *kill an insect*

86b *kill a conversation*

86c *kill a bottle (i.e. empty it)*

86d *kill an audience (i.e. wow them)*

(examples from Bresnan, 1982: 350)

The theta role assigned to the subject is assigned **compositionally**: it is determined by the semantics of the verb and other VP constituents. Roughly,

<sup>25</sup> A discussion of the grammatical functions in the Government and Binding framework is found in Marantz (1981). For a discussion of some problems raised by the approach, see Bresnan (1982).

the verb assigns an object role first, the resulting verb–argument complex will assign a theta role to the subject. The subject argument is as if it were slotted in last.

If a predicate assigns a thematic role directly to some constituent we shall say that the predicate **theta-marks** the constituent **directly**. If the predicate theta-marks an argument compositionally we call this **indirect theta-marking**.

As mentioned above, there is no agreement as to whether the difference between the types of theta roles should be considered as syntactically relevant. However, most linguists agree that the thematic role assigned to the subject must somehow be set apart from the other thematic roles. One quite popular proposal due to Edwin Williams (1981)<sup>26</sup> is that the argument which must be realized in the subject position and hence will be theta-marked indirectly is singled out lexically. The lexical entry for the predicate signals explicitly which argument must be outside the VP. Given that this argument is projected onto an NP outside the VP, it is referred to as the **external argument** and conventionally the external argument is indicated in the thematic grid by underlining:

87 *accuse*: verb

|          |   |
|----------|---|
| <u>1</u> | 2 |
|          |   |

Theta roles assigned to internal arguments will be referred to as **internal theta roles**; that assigned (indirectly) to the external argument is often referred to as the **external theta role**. (Cf. chapter 6, sections 3 and 5 for further discussion.)

## 8 Summary

---

In this chapter we have considered the extent to which sentence structure is determined by lexical properties. As a basis for the formation of sentences we have adopted the projection principle:

<sup>26</sup> Williams (1981) shows the relevance of the distinction between external and internal theta role for the domain of morphology. See also exercise 6.

In chapter 6 we shall discuss an alternative approach to the status of the subject NP.

### 1 Projection principle

Lexical information is syntactically represented.

The type of lexical information with which we have been mainly concerned in this chapter is the thematic structure of the predicate, i.e. the number and types of arguments which the predicate takes. The thematic structure associated with lexical items must be saturated in the syntax, as stated in the theta criterion:

### 2 Theta criterion

2a Each argument is assigned one and only one theta role.

2b Each theta role is assigned to one and only one argument.

The theta roles of a predicate are represented in a grid-format. The assignment of thematic roles is registered by means of referential indices which are associated with thematic roles.

Independently of the argument structure of the main predicate, it is a general property of sentences that they must have subjects. This property is stated in the extended projection principle (EPP):

### 3 Extended projection principle (EPP)

$S \longrightarrow NP - AUX - VP$

In order to satisfy the EPP, so-called expletives may have to be inserted in the subject position of a sentence. Expletives are pronouns such as *it* and *there* in English which are not assigned a thematic role.

## 9 Exercises

---

### Exercise 1

We have seen that lexical verbs are specified for the number and types of theta roles they assign: *work* assigns one thematic role (AGENT), *destroy* assigns two (AGENT, THEME) and *give* assigns three (AGENT, THEME, BENEFICIARY). Provide five more examples for each type of verb.

*Exercise 2*

Discuss the argument structure of the verbs in the following sentences:

- 1 Poirot promised Maigret the job last week.
- 2 Emsworth is walking the dogs.
- 3 That Poirot had left disappointed the crowd immensely.
- 4 The huge pig frightened the spectators.
- 5 I have received the books this morning.

*Exercise 3*

The following examples illustrate how arguments of predicates can be realized in different ways. Discuss the syntactic realization of the arguments in the examples.

- 1a I prefer very much that the students should leave first.
- 1b I prefer very much for the students to leave first.
- 1c I prefer the students out of the way.
- 1d I prefer the students to leave first.
  
- 2a I want hot chocolate.
- 2b I want my coffee to be piping hot.
- 2c I want my coffee piping hot.
- 2d I want the students out of my office.

*Exercise 4*

Sentences 1–10 below are all grammatical. On the basis of the examples establish the theta grid for the main verb of each sentence:

- 1 Mary is eating an apple.
- 2 John is washing the dishes.
- 3 The baby is drinking a glass of whisky.
- 4 John has never met Mary.
- 5 The President is kissing his wife.
- 6 The professor is writing a book on syntax.

- 7 The new secretary pleases all the students.
- 8 This analysis leads us to an unexpected conclusion.
- 9 Poirot smokes cigars.
- 10 Louise is expecting a visitor.

It is not necessary to give the label for each role identified; the number of arguments is the most important property to establish.

Now consider the following examples. They are also all grammatical. What problems do they raise for your treatment of the examples 1–10?

- 11 The children are eating.
- 12 Mary is washing.
- 13 John drinks.
- 14 These two students have never met.
- 15 The professor and his wife were kissing.
- 16 My father writes.
- 17 The students are easy to please.
- 18 This analysis led to quite unexpected conclusions.
- 19 Poirot is smoking.
- 20 Louise is expecting.

### *Exercise 5*

From sentences 11–20 in exercise 4 we conclude that certain thematic roles can be implicit. In (a) below, for instance, *eat* has an understood object, which would correspond to the explicit object in (b):

- a The children are eating.
- b The children are eating lunch.

This is not a general property of transitive verbs, though:

- c The children are devouring their food.
- d \*The children are devouring.

Consider the following paired examples. Again in variant (a) there seems to be one more argument present than in variant (b). Try to characterize the semantic relation between the two sentences. You are

not asked to give a very technical discussion, but simply to provide a description of the difference and/or similarity between the examples:

- 1a Mary is cooking dinner.
- 1b Dinner is cooking.
- 2a Maigret opened the door.
- 2b The door opened.
- 3a Poirot does not grow artichokes.
- 3b The artichokes are not growing.
- 4a Maigret has arrested the criminal.
- 4b The criminal has been arrested.
- 5a Mary is eating too much cake.
- 5b Mary is overeating.
- 6a Poirot was smelling the envelope.
- 6b Your feet smell.
- 7a Maigret is washing his shirts.
- 7b These shirts wash well.
- 8a They are already closing down the new cinema.
- 8b The new cinema is already closing down.
- 9a Poirot is reading the announcement.
- 9b Poirot is reading.
- 10a The guard marched the prisoners round the square.
- 10b The prisoners marched round the square.

Provide a classification of the examples above according to the variation in their thematic structure.

### *Exercise 6*

It is generally accepted that morphological processes may affect thematic structure.<sup>27</sup> Consider the following examples:

- 1a I understand his position.
- 1b His position is understandable.
- 2a This shirt is too wide.
- 2b He has widened the shirt.
- 3a They arrest the criminal.

<sup>27</sup> Williams (1981) contains an important discussion about the interaction of morphology and thematic structure.

- 3b The criminal has been arrested.
- 4a Their activities are not regular.
- 4b They are regularizing the activities.
- 5a He read the book.
- 5b He reread the book.

Discuss the impact, if any, of the affixation of *-able*, *-en*, *-ed*, *-ize*, and *re-* on the thematic structure of the stems to which they attach. For each affix, provide five more examples of the affixation process and check whether your conclusion holds.

### Exercise 7

The following text belongs to the **register** of instructional writing: it is a recipe. Consider the thematic structure of the verbs in the text and try to identify which syntactic properties characterize this register:

Beat two eggs and leave for three minutes. Add milk and mix thoroughly. Cover with grated chocolate. Bake in a moderate oven for 20 minutes. Serve immediately.

The sentences below are other examples of the register of instructional writing. Do they pattern like those in the preceding text?

- 1 Cross now.
- 2 Shake well before using.
- 3 Open here.
- 4 Push.

As you can see it is typical of the register of instructional writing that complements of verbs can be left implicit. Consider the interpretation of the implicit objects in the preceding examples. Discuss what enables people using this register to leave the objects of verbs implicit and how the reader can interpret these sentences correctly.<sup>28</sup>

<sup>28</sup> For discussion see Haegeman (1987) and Massam and Roberge (1989). Both articles will be accessible after reading chapter 8.

*Exercise 8*

Discuss the assignment of thematic roles and the problems, if any, raised for the theta criterion by the following examples:

- 1 John, I really don't like him.
- 2 Which detective will Lord Emsworth invite?
- 3 Which book do you think Poirot will read first?
- 4 Which assistant do you think will reveal the secrets?
- 5 The new assistant appears to have revealed the secrets.
- 6 Which articles did Poirot file without reading?
- 7 *Italian*  
Ho visto Maria.  
have (1sg, pres) seen Maria  
'I have seen Maria.'
- 8 *Spanish*  
Lo vimos a Juan.  
him see (1pl, pres) to Juan  
'We see Juan.'
- 9 *French*  
Quel livre a-t-il acheté?  
which book has-he bought
- 10 Quel livre Poirot a-t-il acheté?  
which book Poirot has-he bought  
'Which book has Poirot bought?'

*Exercise 9*

Although it is not always possible to define the nature of the thematic roles assigned by the verbs, the role of AGENT is one that is fairly well understood. In (1) below, for instance, the subject NP *Poirot* is assigned the AGENT role and in its passive variant (2) AGENT is assigned to an NP inside a PP introduced by *by*:

- 1 Poirot bought the pigs.
- 2 The pigs were bought by Poirot.

There are certain adjuncts which seem to require the presence of an AGENT in the sentence:

- 3 Poirot bought the pigs deliberately.
- 4 Poirot bought the pigs to annoy his mother.
- 5 \*Poirot liked England deliberately.
- 6 \*Poirot liked England in order to annoy his mother.

Adjuncts such as *deliberately* and *in order to annoy his mother* cannot modify predicates such as *like*. This is because these adjuncts imply intentionality, a notion which is not easily compatible with involuntary activities or states such as 'liking'. Consider the following examples: what conclusions can you draw with respect to the thematic structure of the predicates in the sentences?<sup>29</sup>

- 1a The enemy sank the ship deliberately.
- 1b The ship was sunk by the enemy deliberately.
- 1c The ship was sunk deliberately.
- 1d The ship sank.
- 1e \*The ship sank deliberately.
  
- 2a We sold the books to raise money.
- 2b These books will be sold by the schools to raise money.
- 2c These books will be sold to raise money.
- 2d These books sell well.
- 2e \*These books sell well to raise money.

We discuss passive sentences in chapter 3 and especially in chapter 6.

### Exercise 10

In the text we suggest that there are only subject expletives and that there are no object expletives. Do you think that the examples in (1) and (2) provide counterevidence for this claim?

- 1 I consider it likely that Louisa will not leave.
- 2 I thought it stupid that she should have gone out in the rain.

<sup>29</sup> For discussion see Hale and Keyser (1986, 1987) and Roberts (1987). These texts presuppose the contents of chapter 6.

Now consider the following Italian sentence.

- 3a Ritengo probabile che Maria rimanga.  
 I consider likely that Maria stay (subj)  
 3b (\*Ciò) è strano che Maria sia venuta.  
 it is strange that Maria be (subj) come  
 'It is strange that Maria should have come.'

If we assume that in (1) and in (2) *it* is the subject expletive of a small clause (cf. section 5.1) then we deduce that such expletive subjects are unavailable in Italian small clauses. We might then be tempted to conclude that the absence of an expletive subject in (3a) is another effect of the *pro*-drop parameter discussed in the Introduction (section 4). Recall that Italian does not have expletive subjects in finite clauses (3b). On the basis of the following French example, discuss whether the unavailability of an expletive in (3a) should be made to follow directly from the *pro*-drop parameter:

- 4a Je trouve bizarre qu'elle soit là.  
 I find strange that she be there  
 'I find it strange that she is there.'  
 4b \*Est étrange qu'elle soit là.  
 is strange that she be there

We return to the structure of small clauses in chapter 2, section 3.5.