Chapter 5
Surface Structure

Once a message, or a fragment of a message, has been prepared for expression, the process of formulating can be initiated. Successive message fragments will trigger the Formulator to access lemmas, to inspect the message for functions, arguments, and modifiers, to specify grammatical relations, and to map these onto inflectional and phrasal structure.

This first stage of the formulating process was called "grammatical encoding" in chapter 1, and was distinguished from a second, phonological encoding stage in which word forms are accessed and prosodic patterns are generated. The present chapter will characterize the type of representation that forms the hinge between these two stages. It will be called "surface structure" (which involves an allusion but not a full commitment to particular grammatical theories). A surface structure is, by definition here, the output of grammatical encoding, and the input to phonological encoding. We will, however, stay rather close to Bresnan's (1982) notion of surface structure.

In order to understand the processes of grammatical encoding, which are discussed in subsequent chapters, we must have a sufficiently explicit specification of their target structures. It is, on the one hand, necessary to consider the way in which a surface structure expresses semantic relations through grammatical functions. This semantics-to-function mapping depends on the internal structure of lemmas, which are the terminal elements of a surface structure. It is, on the other hand, necessary to specify the way in which these grammatical functions are realized in a surface structure's hierarchical organization of phrases and in its case marking. This organization is essential input for phonological encoding.

These theoretical notions have, in general, not evolved from empirical analyses of the speaking process. They mostly stem from linguistics and computer science. Still, they do provide a much-needed framework for a theory of the speaker, which is not independently available. The theory we

will assume in the present chapter is somewhat along the lines of Bresnan (1978, 1982), for three reasons. First, it is an explicit theory that allows us to formulate explicit procedures of surface-structure generation. Second, it is—as will be discussed—lexically based, and that makes it an attractive starting point for a theory of grammatical encoding. The following chapters will argue that lemmas are the driving force behind the speaker's construction of the surface structure. It is in the lemmas of the mental lexicon that conceptual information is linked to grammatical function. A lexically based grammar is a natural companion in developing this notion. Third, Bresnan's theory combines well with the psycholinguistic theory of grammatical encoding developed by Kempen and Hoenkamp (1987), which we will follow in chapter 7. Both are lexically based, and both operate with dual constituent/functional representations, as will be discussed. It should be kept in mind, however, that no grammatical theory can claim to be the correct one. We will not be committed to all details of Bresnan's theory; it will only be used as an attractive explicit framework for approaching the problems of grammatical encoding. The present chapter will be limited to the phrasal and functional aspects of surface structure. The internal grammatical structure of lemmas will be taken up in the next chapter, where the speaker's accessing of lemmas will be discussed.

The assumption that grammatical encoding is in the first instance independent of phonological encoding in no way implies that a speaker generates complete surface structures for clauses or sentences before accessing the inventory of word forms and computing phonetic patterns. The notion of incremental sentence production is also valid at this level; any terminal fragment of surface structure that becomes available will immediately trigger its phonological encoding, following Wundt's principle. Some of the word forms may already have become available with their lemmas; others are still to be retrieved. It is likely that failure to find a word's phonological form within some critical period of time may trigger grammatical revisions (see Levelt and Maassen 1981 and section 7.5). As a first approximation, however, grammatical encoding can be treated as an autonomous process that takes messages as input and produces surface structures, unspecified for phonological form, as output.

5.1   Syntactic Aspects

5.1.1   Surface Structures as Expressions of Grammatical Functions
On a theory like Bresnan's (1982), grammatical functions such as "subject", "direct object", and "indirect object" are primitives and are the basis
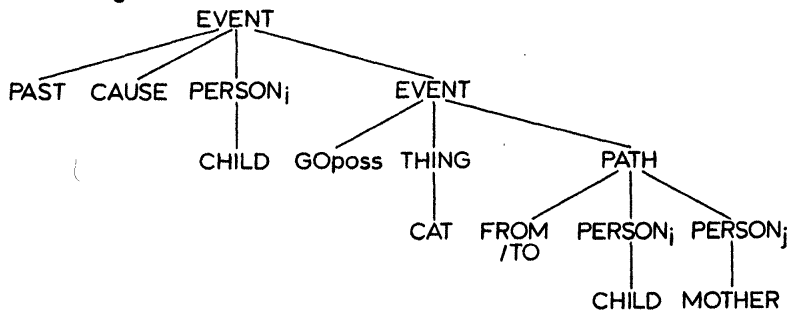
of syntactic structures in all languages. But these grammatical functions are realized differently in the surface structures of languages. They are, broadly speaking, realized either by case or by word order. Take, for example, the message in figure 5.1, and how it can be grammatically encoded in English and in Malayalam, a language spoken in southern India (Mohanan 1982). The message is an event proposition. There is no mood indicator, so it has the default mood DECLARATIVE. Its time indicator is represented as a proposition modifier; it is PAST. The entities CHILD, CAT, and MOTHER are all token referents, represented in the discourse model, but their accessibility index is ignored for the present purpose.

In English, grammatical functions are largely expressed by word order and phrase order; they are assigned to positions in a phrase structure. The subject of a sentence, for instance, is the noun phrase (NP) directly dominated by the sentence node (S). In the English surface structure of the figure, the NP for *the child* is the only NP that branches directly from S, and hence *the child* is the subject of the sentence. Similarly, the indirect-object function is carried by the first NP in the verb phrase (VP), i.e., *the mother*. The direct-object function is assigned to the second NP in the verb phrase, *the cat*. In other words, the grammatical functions are expressed in the configurations of surface structure. Hence, languages such as English are called *configurational* languages (see Webelhuth 1985 for an excellent discussion of this notion).
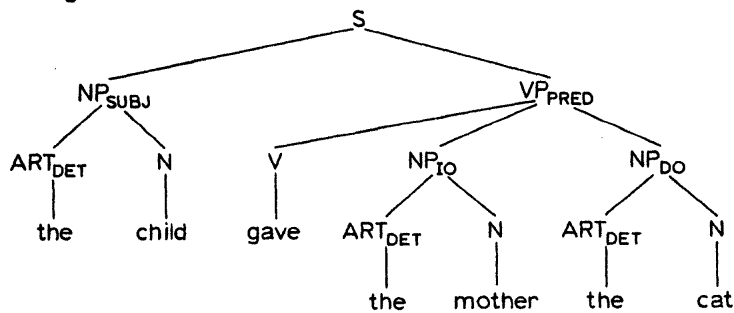
In Malayalam, however, grammatical functions are expressed through case-marking morphology (Mohanan 1982). The subject NP, for instance, is marked by assigning nominative case to the head noun (*kutti*). The subject NP is not the only NP that is directly dominated by S (as was the case in English); the two object NPs are also. Moreover, the order of the NPs is irrelevant for the assignment of grammatical function. The subject NP need not be in first position. The order of the two object NPs is also immaterial; they are grammatically distinguished by dative and accusative case markers. Compare this with the English example, where the order of the two object NPs does matter; *the child gave the mother the cat* and *the child gave the cat the mother* mean different things. Malayalam is called a *nonconfigurational* language because grammatical functions are not encoded in phrase-structural configurations. Nonconfigurational languages have great freedom of word order, and their surface phrase structures tend to be "flat" (i.e., without much hierarchy).

These examples show that the target structures of grammatical encoding are highly language-dependent; a language may be more or less configura-
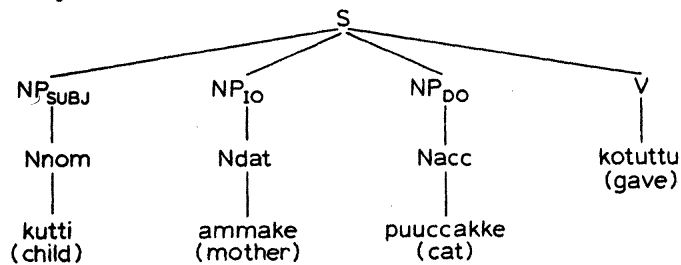
a message :

```
                        EVENT
        ┌──────┬───────┼──────────────┐
      PAST   CAUSE  PERSONᵢ          EVENT
                       │        ┌──────┬──────┐
                     CHILD   GOposs  THING    PATH
                                      │     ┌───┬──────────┬──────────┐
                                     CAT  FROM  PERSONᵢ   PERSONⱼ
                                          /TO      │          │
                                                 CHILD     MOTHER
```

b english surface structure :

```
                          S
            ┌─────────────┴──────────────┐
         NP_SUBJ                        VP_PRED
        ┌───┴───┐           ┌────────┬──────────┐
     ART_DET    N           V      NP_IO        NP_DO
        │       │           │     ┌──┴──┐      ┌──┴──┐
       the    child       gave ART_DET  N   ART_DET  N
                                 │      │      │      │
                                the  mother  the    cat
```

c malayalam surface structure :

```
                      S
      ┌──────────┬──────────┬──────────┐
   NP_SUBJ    NP_IO      NP_DO          V
      │          │          │           │
    Nnom       Ndat       Nacc       kotuttu
      │          │          │         (gave)
    kutti     ammake    puuccakke
   (child)   (mother)    (cat)
```

Figure 5.1
(a) A message. (b) Its expression in the surface structure of English, a configurational language. (c) Its expression in the surface structure of Malayalam, a nonconfigurational language.

tional, depending on how much grammatical function is encoded positionally and how much is encoded by morphological case marking. And many languages (such as German) encode certain functions both ways. The process of grammatical encoding, which generates these target structures, must be correspondingly language-dependent. It is a fascinating question how much, and along what dimensions, the psychological mechanisms of grammatical encoding can vary.

5.1.2 Surface Structures as Input to Phonological Encoding

A surface structure is not only the result of grammatical encoding; it is also the input for the subsequent phonological-encoding stage of the formulating process. It is not exceedingly clear how much phrase structure is needed for the accurate generation of segmental (lexical) and suprasegmental (prosodic) form information. Minimally, the following items are required:

• A string of pointers to form information in the lexicon for all content and function words in the sentence, in the correct grammatical order. Let us call these pointers *lexical pointers*. A lemma's lexical pointer indicates an address where the corresponding word-form information is stored. In figure 5.1 these lexical pointers are represented as words (*the, child*, etc). This is convenient as long as one is aware that these do not represent the word forms themselves, but only the addresses where they can be found.

• All diacritic features for each lexical pointer. The pointer to a form address may be indexed with various features that will affect the word form retrieved. We will call them *diacritic features*. The surface structure should indicate case, number, person, tense, aspect, definiteness, pitch accent, and whatever other features are to be morphologically or phonologically marked on the word form pointed to. Note that the inflections themselves are not yet specified; case, number, and so on are only abstract instructions for inflectional procedures to be run in the second phase. And the same holds for the pitch-accent feature. If a lemma is marked for pitch accent (for instance, because the corresponding concept is prominent), this will have various consequences for the subsequent phase of phonological encoding. It is generally recognized that semantically motivated pitch accent is a main determinant of sentence intonation and sentence rhythm (Bolinger 1972, 1983, 1986; Cutler and Isard 1980; Halliday 1967b; Ladd 1980; Schmerling 1976; Selkirk 1984a). We will return to the assignment of pitch accent in section 5.2.

- Phrasal information. This is, in the first place, important for the correct assignment of prosodic structure—particularly the allocation of sentence rhythm. Two utterances may involve the same string of words but different surface phrase structures. Compare, for instance, sentences 1 and 2, where major phrases are indicated by parentheses:

(1) (The widow) (discussed (the trouble)) (with her son)

(2) (The widow) (discussed (the trouble with her son))

A noncasual speaker pronounces such sentences differently (Levelt, Zwanenburg, and Ouweneel 1970; Wales and Toner 1980). There appears a slight rhythmic/melodic caesura after *trouble* in sentence 1, but not in sentence 2. Phonological encoding acknowledges such phrasal properties of surface structure. A sentence's rhythm reflects to some extent the grouping of words in surface structure. This phrasal information is certainly not the only determinant of rhythm or of intonation (pitch accent, for instance, is more important), but it does play a role, as we will see in chapters 8 and 10.

Further, phrasal information can be an important determinant of a word's pronunciation. A word's position in a phrase, for instance, can determine whether it will be phonologically reduced. A phrase-final preposition (as in *who were you thinking of?*) will not be as easily weakened as a non-phrase-final preposition (*a blade of* [əv] *grass*). These and many more examples can be found in Selkirk 1972. Like Selkirk, Kenstowicz and Kisseberth (1979) and Kaisse (1985) analyzed various other phrase-structural properties that may affect the final phonetic form of words in a sentence. These issues will be taken up in chapter 10.

So much can be said that, depending on the language, *some* phrase-structural information will be needed for the phonological encoding of a sentence. It is, however, unlikely that *all* such information is relevant for a speaker of a configurational language. The full phrase structure, with all its phrases categorized according to type (NP, VP, etc.), represents an upper limit. It is, however, still justifiable to consider linguistically fully specified surface structures as the hinging representations between the two stages of formulating, as long as one takes a procedural view. The generation of the minimally required information (listed above) will require procedural steps involving all the grammatical categories and phrasal relations represented in the "full" surface structure. They are essential *intermediary* results or outputs of grammatical encoding, as will become apparent in chapter 7.

5.1.3   Some Properties of Surface Phrase Structure

At least the following properties are essential ingredients of surface structure as far as a theory of the speaker is concerned:

(i) Surface structures are phrase structures; i.e., they can be represented as hierarchical tree structures without crossing branches (such as in figures 5.1b and 5.1c). Phrases and subphrases (also called *constituents*) are represented in their correct left-to-right order (Levelt 1974, volume 2). The terminal nodes of surface structures are of two sorts: lemmas with pointers to form addresses in the lexicon (these pointers are indexed with diacritic features), and empty elements, which have no lexical pointer but which do carry a grammatical function. An example of the latter appears in the sentence *the story is hard to explain (e) to Peter*, where (e) marks the position of the empty element. Here the empty element is the direct object of *explain*; it is, however, lexically expressed as the subject (*the story*) of the main clause.

(ii) Surface structures represent categorial information for all nodes. This involves, in the first place, four *major lexical categories*—noun (N), verb (V), adjective (A), and preposition (P)—that can have the grammatical function of head-of-phrase (see property iii below). In figure 5.1b, the lexical item *child* in the English surface structure is of category N and functions as head-of-phrase in the noun phrase (NP) *the child*. Similarly, *gave* is of category V and functions as head of the verb phrase (VP). There are also *minor lexical categories*, such as article (Art; e.g., *the*) and conjunction (Conj; e.g., *but*), which cannot be heads-of-phrase.

The phrases of which the major lexical categories N, V, A, and P can be heads are, respectively, noun phrase (NP), adjective phrase (AP), verb phrase (VP), and prepositional phrase (PP). These are called *phrasal categories* or *major constituents*. Examples in figure 5.1b are the NPs *the child*, *the mother*, and *the cat* and the VP *gave the mother the cat*. In turn, these phrasal categories can figure as heads of still more complex phrases, usually indicated by NP′, AP′, VP′, and PP′. In the sentence *the child gave the mother the cat on request*, the phrase *gave the mother the cat on request* is a VP′, with the VP *gave the mother the cat* as head. Finally, there is the phrasal category S, which may have as head a VP (or a VP′), as in English, or a V, as in Malayalam and in many configurational languages, probably including German and Dutch. When a language has a VP as head of S, the head-of-phrase function is sometimes called "predicate" (PRED), and the sentence is said to have a predicate as head-phrase.

The left-to-right order, mentioned under property i, is restricted by the syntactic category of the phrase. In chapter 7, we will consider phrasal categories as *procedures* or syntactic specialists that impose order on their parts.

It is not the case that every surface structure has S as a root. Almost any phrase can be an expressible surface structure. When the root is not S, the resulting utterance is said to be *elliptic*. So, *A church* as answer to the question *What did you see?* is elliptic. Its surface structure is of the category NP. We saw in chapter 3 that the underlying message is also elliptic; i.e., it is not a full proposition, but just an entity.

(iii) Surface structures represent functional information for all phrases except the root S. This functional information is of two kinds.

First, the surface structure indicates the *head-of-phrase function* for each phrase. The main head-of-phrase functions were specified under property ii. They assign a unique head to each constituent. What does it mean to be head-of-phrase? The notion was developed (see, especially, Jackendoff 1977) to account for the fact that one element in the phrase imposes some of its features on the phrase as a whole. For instance, if the head noun of a noun phrase is plural, then the whole noun phrase is plural; i.e., if *houses* is plural, then *the red houses* is also plural. This does not hold for the non-head elements. The noun phrase *the kings' crown* is singular in spite of the plurality of *kings*, because the head noun, *crown*, is singular. Similar relations between head and phrase exist for such features as gender, person, definiteness, and case. The head-of-phrase relation is also important in the generation of sentences. In chapter 7 we will see that the head-of-phrase can call the procedure for building the whole phrase around it.

Since for each phrase it is uniquely determined what the head-of-phrase is, it is usually not explicitly marked in our graphic representations of surface structure. In figure 5.1b, this would have amounted to adding the subscript "Head" to V, the two N's, and VP. The head carries the grammatical function of the phrase as a whole. So when a noun phrase is subject of the sentence, thus carrying a particular case (nominative or ergative), then the head noun will carry the same case. Still, a particular grammatical function, such as subject, may be distributed over different phrases. This is especially apparent in nonconfigurational languages. Their free word order allows for discontinuous expressions. If a subject of a sentence is an adjective-noun pair, the two elements may be far apart in surface structure, and even be interwoven with the adjective and noun of the object, roughly

as in *Elephant mouse big small kill*. In Warlpiri, an Australian noncon-figurational language studied by Hale (1981), the two parts of the subject will both get one type of case morphology (ergative case), and the two object parts will both be in another type of case (called "absolutive"). A discontinuous expression doesn't have a unique head, but each of its continuous phrasal parts has one, and all of them are assigned the same grammatical function by means of case morphology, thus marking their functional togetherness.

Second, phrases are labeled with respect to their *grammatical functions*, if any. Examples in figure 5.1 are subject (SUBJ), predicate (PRED), two kinds of object (IO and DO for indirect and direct object, respectively), and determiner (DET). Apart from these grammatical functions, there can be obliques (OBL), such as the NP following *by* in passive sentences (*the child* in *the cat was given by the child*), and verbal complements (V-COMP), such as *steal the money* in the sentence *Grabber tried to steal the money*. These grammatical functions originate from specifications in the lemmas of verbs, nouns, adjectives, and prepositions; this will be taken up in the next chapter.

Heads-of-phrase are said to *subcategorize* other elements in the phrase. The verb *give* in figures 5.1b and 5.1c, for instance, can take three NPs, which express the conceptual arguments of giving: the one who gives, the object given, and the one who receives. These arguments are encoded as subject, direct object, and indirect object, respectively. Of these, only the last two appear in the verb phrase of which *give* is the head; they are the ones subcategorized by the verb. "Subject" is called an *external argument* of the verb.

Phrases can also have grammatical functions that do not express a conceptual argument of the head but rather express some *modification*. This is mostly the case for adjectives (A) or adjective phrases (AP), which modify an N or NP. In *the little child*, *little* has a modifying function; it is not an argument of *child*. But other phrases can also carry modifying functions. In *the child gave the mother the cat on request*, the prepositional phrase *on request* is such a modifying adjunct. Unlike the direct-object and indirect-object phrases (*the cat* and *the mother*), which *are* subcategorized by *give*, the oblique prepositional phrase *on request* is attached to the VP'-node in the surface structure, not to the VP of which *give* is the head. Modifier phrases generally attach one level up in the phrase struc-ture. Both subcategorized and modifier (or adjunct) phrases are called *complements*.

Finally, certain elements in the surface structure are *specifiers*. The NPs *the child*, *the mother*, and *the cat* all contain the article (*the*) which has the specifier function of determiner (DET). A determiner can, among other things, express definiteness or indefiniteness. There are also specifiers for number (NUM), as in *two children*, and for degree (DEG) (e.g., *very* in the AP *very large*), and there are other types of specifier. The category is rather heterogeneous and will not be elaborated here.

In summary: From the viewpoint of grammatical function, each constituent consists of at most four types of elements: the head, the complements that are subcategorized by the head and which express its conceptual arguments (if any), complements which express modifications (if any), and specifiers (if any). As a notational convention, these functions will be labeled as subscripts to the category name of the phrase, except for the function of head. Moreover, it is often convenient to leave out most or all of the function labels as long as no ambiguity arises.

(iv) The lowest-level or terminal nodes in a surface structure are (if not empty elements) lemmas with their lexical pointers indexed for diacritic features (as discussed in the first two items at the beginning of subsection 5.1.2).

This suffices as a specification of the syntactic aspects of surface structure. More extensive treatments can be found in Bresnan 1982, Gazdar et al. 1985, Jackendoff 1977, and Levelt 1974. The above notions will all return in chapter 7 when we deal with the speaker's generation of surface structure.

5.2   Prosodic Aspects

A surface structure has no prosody, but it does contain the information required in subsequent phases for the generation of prosodic patterns that will do justice to the speaker's intentions. In particular, the surface structure must contain specifications of *mood* and *focus*. Let us consider these in turn.

5.2.1   Mood and Modality

The preverbal message includes a marker for the intended mood of the utterance: declarative, interrogative, or imperative (see subsection 3.5.1). The intended mood has, of course, syntactic consequences in sentence generation. The messages DECL(PAST(LEAVE(JOHN))) and ?(PAST (LEAVE(JOHN))) lead to the generation of different surface structures,

namely those for the sentences *John left* and *Did John leave?*, respectively. The mood marker in a message will, like the temporal markers, lead to a feature of the tensed verb in surface structure. This, in turn, selects for a particular syntactic structure.

But there is an additional consequence of the mood marker: It will co-determine the melody of the sentence. There may, in fact, be a neutral or default way of intonating a declarative sentence in a language, and similarly for an imperative or an interrogative sentence. These default intonations are then largely realized in the sentence-final *boundary tone*—a falling tone for declarative and imperative moods, a rising tone for interrogative mood. If this is so, and the surface structure is the hinging representation between grammatical and phonological encoding, the surface structure should, for each of the three moods, contain some cue by which the Phonological Encoder can know what tone to select. Still, there is reason to be quite careful with such generalizations. It was mentioned in subsection 3.5.1 that a speaker's presuppositions and attitudes can, in complex ways, interact with what one would optimistically take to be "standard" tones for the major sentence moods. Whether there are such default tones is, honestly, an open issue. It seems, therefore, premature to specify how they are indicated in surface structure.

The situation is even less clear for the other modalities and attitudes a speaker can express in the prosody of his utterance. This issue is nicely treated by Cutler and Isard (1980), who show that proposals to the effect that a speaker has an "intonational lexicon" in which each contour has its own specific expressive meaning are untenable. One and the same contour can express very different speaker moods (such as surprise and indignation), and different contours can express the same emotional attitude of the speaker. These intonation "meanings" are highly context-dependent. We will return to them in chapter 8. Here it suffices to say that, eventually, intonational specifications for mood and modality may have to be introduced in surface-structure representations. But it remains to be seen how this should be done.

5.2.2 Prosodic Focus

The second message ingredient that will affect the prosody of an utterance is prominence. Subsection 4.5.2 mentioned three grounds for an entity to become prominent in the message: being newly introduced into the discourse model as a referent, contrasting in a focused role with a previously mentioned entity, or involving a new predication. These three grounds have in common that the entity has "news value" for the addressee. The present

subsection discusses how prominence becomes expressed in surface structure. Though one should not exclude the possibility that prominence is expressed differently in these three cases (Cutler and Isard 1980 presents evidence for a speaker's special treatment of contrastive prominence), we will ignore this for the present purposes.

Prominence, then, becomes expressed in surface structure through what is called *prosodic focus*. This is not to be confused with focus in the discourse record (the speaker's attentional "pointer"). There is some relation, of course; what is in the speaker's new focus will often be new information for the addressee and therefore be given prominence by the speaker. That prominence, in turn, results in prosodic focus. Where there can be no confusion in the discussion of prosodic focus, we will talk of just "focus."

Syntactic phrases that correspond to prominent parts of the message will be marked for focus; we will designate this by putting an f before the syntactic-category symbol in the surface structure. But when a syntactic phrase receives a focus marker, at least one of its constituent phrases will also be marked for focus, and so all the way down to the level of lemmas. Let us, by way of example, return to Seth and Marcia's ongoing conversation. At some point Marcia answers a question by Seth with the following utterance:
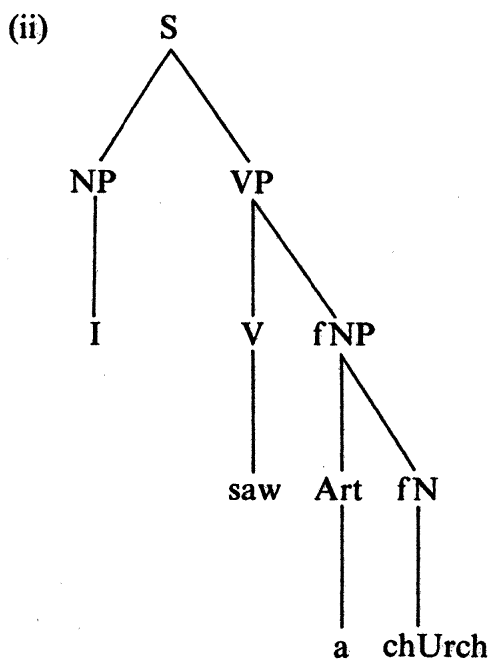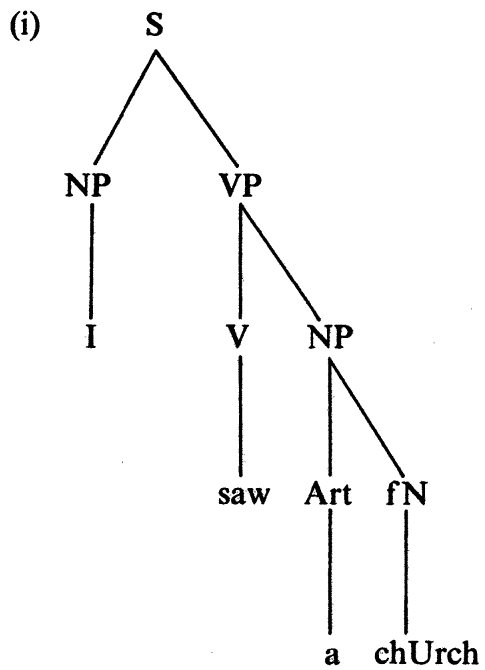
(3) I saw a chUrch.

What question was this an answer to? There are at least three different questions to which this could have been a natural-sounding reply:
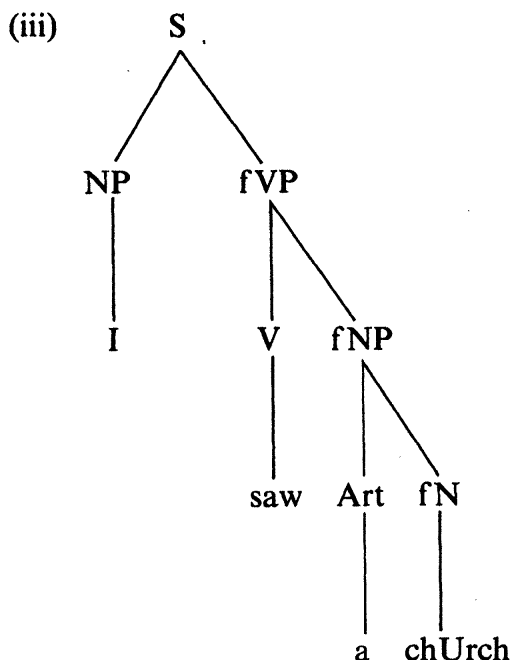
(4) Did you see a palace?

(5) What did you see?

(6) What did you do?

In example 4 there is contrastive prominence; the concept CHURCH carries the news value by contrast to PALACE. Hence the noun *church* is given focus in surface structure. In example 5 it is the *theme* argument of SEE that is prominent in the message. This is expressed as a focus on the whole corresponding NP in surface structure: *the church*. In example 6, the whole ACTION was at issue (i.e., to go and see a church). Since the ACTION is mapped onto the surface VP, the whole VP *saw the church* is assigned focus. These three states of affairs can be represented as follows in terms of surface structure (trimmed of functional and other details which are irrelevant for the present discussion):

(i)
```
              S
          ╱       ╲
        NP         VP
        │        ╱    ╲
        │       V      NP
        │       │    ╱    ╲
        I      saw  Art    fN
               │     │      │
                     a   chUrch
```

(ii)
```
              S
          ╱       ╲
        NP         VP
        │        ╱    ╲
        │       V      fNP
        │       │    ╱     ╲
        I      saw  Art     fN
               │     │       │
                     a    chUrch
```

(iii)

```
                    S
                   / \
                  /   \
                 /     \
               NP      fVP
                |       |\
                |       | \
                |       |  \
                I       V  fNP
                        |   |\
                        |   | \
                        |   |  \
                       saw Art  fN
                            |    |
                            |    |
                            |    |
                            a  chUrch
```

In producing a sentence, a speaker will somehow have to realize the focus information as a pattern of *pitch accents*. Pitch accent is a prosodic parameter assigned to certain lemmas in the surface tree. All three surface trees above have pitch accent assigned to the lemma for *church*. As a notational convention, a lemma that carries pitch accent in surface structure will be printed with the accented vowel capitalized. It cannot be emphasized enough, however, that pitch accent is an *abstract parameter* of a lemma; the surface structure does not contain word-form information. When the lemma carries pitch accent, this means that the lemma's lexical pointer has a diacritic feature for pitch accent.

The way in which speakers "compute" the pattern of pitch accents will be taken up in chapter 7. Here I will only present some regularities that seem to govern the patterns of focus and pitch accents in surface structures. The presentation will largely follow Selkirk 1984a. The major issue is how the prominence of some message fragment becomes represented in the surface structure a speaker generates, and how it is handed down to the level of lemmas. The main rule seems to be the following.

Phrasal-Focus Rule
The prominence of a concept in the message is expressed by assigning focus to the surface-structure phrase representing it. In a focused phrase, either the head or any of the complements that express conceptual arguments of the head should be focused.

This rule guarantees, as will be discussed, that eventually there will always

be some lemma that receives focus when a higher-level phrase is focused. What happens to the lemma is given in the next rule.

Pitch-Accent Rule
A focused lemma receives pitch accent.

There is a subsidiary rule that distinguishes between heads of phrases and conceptual arguments of the head:

Focus-Interpretation Rule
Focus can be assigned to an argument only if that argument expresses a prominent concept in the message, but it can be assigned to the head irrespective of whether it represents a prominent concept.

These rules are fairly abstract, so let us consider how they work in practice by considering some examples. Examples 4–6 above are a good start.

When Seth asked *Did you see a palace?*, Marcia had to introduce *church* in a contrasting role; her having seen something indefinite was presupposed in the discourse. So Marcia gave prominence to the concept CHURCH in her message, and to nothing else. By the Phrasal-Focus Rule, the noun representing the concept was assigned focus; since the noun is a lemma, the focus was expressed as pitch accent (chUrch in diagram i) through the pitch-accent rule. Nothing else was given focus or pitch accent.

In example 5, Seth asked *What did you see?* The question carries the presupposition that something had been seen, and Marcia introduced the *theme* of her seeing (some church) in the discourse model by making it prominent in the message, and she made nothing else prominent. The theme in the message was mapped onto the object NP in the surface structure, and the NP was assigned focus in accordance with the Phrasal-Focus Rule. The NP contains no arguments, only a head and a specifier, so the same Phrasal-Focus Rule requires focus to be assigned to the head noun. The head noun is a lemma, and by the Pitch-Accent Rule it is given pitch accent (chUrch).

Example 6 is somewhat more complicated. Seth asked *What did you do?*, presupposing that some ACTION on the part of Marcia had taken place. Marcia's task was now to introduce her ACTION into the discourse. Since it was still absent from the discourse model, Marcia gave the ACTION prominence in the message. The ACTION got represented as a VP in surface structure, which received focus by the Phrasal-Focus Rule. The Phrasal-Focus Rule now requires the head or one of its subcategorized complements to receive focus—either the V (*saw*), or its object NP (*a church*), or both. It does not specify which. Marcia opted for the NP (why

she did so will be discussed shortly). From here on, the story proceeds as in example 5, leading to pitch accent for the head noun (chUrch).

Why did the NP get focused? The Focus-Interpretation Rule allows focusing of NP only in cases where it expresses a concept that is prominent in the message. Apparently, Marcia had marked not only the ACTION as prominent in her message, but also the *theme* argument (CHURCH). The Phrasal-Focus Rule required the NP expressing the prominent theme to be focused anyhow; the NP was thus focused for independent reasons. But, this being the case, the condition of the Phrasal-Focus Rule with respect to the VP was already fulfilled; there was a focused argument, and nothing more was required. There is an additional reason why *church* will become more stressed than *saw*: It receives *nuclear stress*, which will be discussed at the end of this subsection.

Would it ever be possible for the head verb, but none of its complements, to become focused when the ACTION is prominent in the message? Ladd (1980) and Selkirk (1984a) argue that this can happen, and provide various examples. In our discussion between Seth and Marcia, the following exchange might arise:
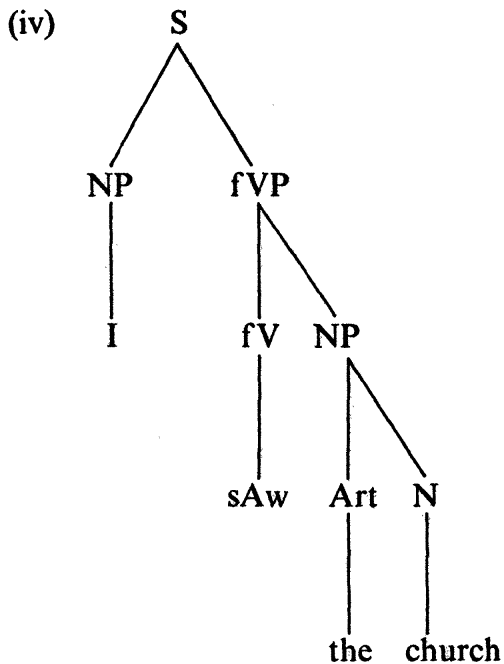
(7) Seth: Did you have a chance to see the Saint Peter?
    Marcia: I sAw the church.

Neither the seeing nor its conceptual argument (*the Saint Peter*) was newly introduced in the discourse by Marcia. She could figure that both were in Seth's focus of attention. This is apparent not only from the fact that *the church* is an anaphor for *the Saint Peter* but also from the fact that Marcia could even have answered *Yes, I did*, which would have involved a reduction of both function and argument. Still, sAw was given pitch accent. Marcia made the ACTION prominent in her message for the purpose of confirmation, which was the new information asked for by Seth. The Focus-Interpretation Rule now excluded the possibility of giving prosodic focus to the complement NP *the church*, since it was not prominent in the message. More generally, an anaphor—a substitute for an expression used earlier (as *church* is for *Saint Peter*)—will, as a rule, be deaccented. This tells the listener that the referent is at hand. The only case where an anaphor can become focused in surface structure is when the referent is marked for contrastive prominence in the message. An instance of this kind was discussed in subsection 4.5.2: *Tessie went after SImon, and then hE chased hEr*. But the head V (*saw*) in the present example could receive focus, and thus pitch accent, in spite of the fact that it represents a nonprominent part of the message. This follows from the Focus-Interpretation Rule. Selkirk

calls this "default accent." The resulting surface focus structure is shown in diagram iv.

(iv)
```
              S
             / \
           NP   fVP
            |    |\
            |    | \
            I   fV  NP
                |   |\
                |   | \
              sAw  Art N
                    |   |
                    |   |
                   the church
```
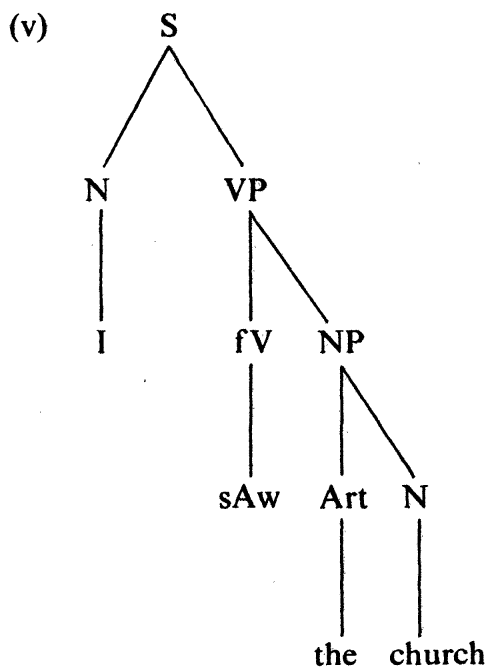
Notice that this differs essentially from the situation where *saw* receives contrastive accent. That would have happened in the following exchange:

(8) Seth: Did you PAINT the Saint Peter?
    Marcia: No. I only SAW the church.

The focus structure here is as in diagram v, with "narrow" focus on the verb only.

(v)
```
              S
             / \
            N   VP
            |    |\
            |    | \
            I   fV  NP
                |   |\
                |   | \
              sAw  Art N
                    |   |
                    |   |
                   the church
```

The possibility of default pitch accent on the head-of-phrase exists not only for verb phrases but also for the other three main phrasal categories. In example 9, there is default accent on the prepositional head of a PP. (The acceptability of this example depends to some degree on the dialect of English one speaks.)

(9) Shopkeeper: Did you manage to deliver the wine to Mr. PEters?
    Assistant: Yes sir, I got it tO him.

Here, the preposition is in no way contrastive, and the assistant can assume that Mr. Peters is in the shopkeeper's center of attention. Thus, neither the *direction* nor the *goal* is prominent in the assistant's message. What *is* made prominent, for reasons of confirmation, is the PATH, which is mapped onto the PP (*tO him*). Notice that the same PATH was also the prominent part of the shopkeeper's question, but there Mr. Peters was also prominent for an independent reason: He was being newly introduced into the discourse. Thus, it was possible for the shopkeeper to realize the PP focus by focusing its NP complement (*Mr. PEters*), according to the Phrasal-Focus Rule.

Before we leave the issue of focus assignment, a word should be said about the role of modifiers. The Phrasal-Focus Rule mentions the conceptual arguments of the head, but not its modifiers. Can a constituent be focused by focusing a modifier complement or a specifier? Is it indeed necessary that the complement represent a conceptual *argument?* Probably so. Focusing a modifier or a specifier tends to involve narrow focus only, as is clearly the case in a deviant case such as the following one.

(10) Seth: What did you see today?
     Marcia: I saw an Old church.

The adjective *Old* cannot carry the focus of the NP (*an old church*); it can only be contrastive to another adjective (such as *new*) in previous discourse. Here is an example in which the temporal adverbial cannot carry the focus for the verb phrase as a whole:

(11) Seth: What did you do?
     Marcia: I visited the church todAY.

This is by no means a full account of how prominence in the message is grammatically encoded in surface structure. The essential point, however, is that focus and (in its wake) pitch accent are conditioned by the prominence structure of the message, i.e., of its functions, arguments, and modifiers. There are no independent syntactic reasons for the assignment of pitch accent; syntax has only a mediating role.

The assignment of focus should not be confused with the assignment of nuclear stress. It is a general rule of English that, in any major phrase, one word receives more stress than any of the others. This word is the rightmost non-anaphoric word of the rightmost constituent in the phrase. When a speaker opens a conversation by saying

(12) Peter's father had a terrible accident

father will normally receive more stress in the NP *Peter's father* than *Peter's*, because *father* is the rightmost constituent of the noun phrase. Similarly, *accident* will be most stressed in the NP *a terrible accident*. In the VP *had a terrible accident* the word *accident* will be most stressed, because it is the most stressed word in the VP's rightmost constituent, namely the NP. In the sentence as a whole, *accident* will be the most stressed word, because it is the most stressed word in the rightmost constituent of the sentence, the VP (see also subsection 10.2.2).

When Marcia answered Seth's question *What did you do today?* with *I saw a church*, the Phrasal-Focus Rule allowed for the focusing of both *saw* and *church*. If Marcia had focused both, *church* would still have become more accented than *saw*, owing to the assignment of nuclear stress. We will return to these issues in chapters 8 and 10.

The treatment of focus in this section was based largely on facts of English. It should be kept in mind that rules can be rather different for other languages. (See Bierwisch 1965 for a classical analysis of the rule system for intonation in German.)

Summary

This chapter discussed the representation mediating between grammatical encoding and phonological encoding. It was called "surface structure." The processes of grammatical encoding project the concepts and their relations in the preverbal message onto a phrase-structural organization of lemmas and grammatical relations.

The first section reviewed various syntactic aspects of this organization and their relevance for phonological encoding. The way in which surface structures represent grammatical relations ranges between two extremes in the languages of the world. Configurational languages express grammatical functions in the hierarchical and left-to-right organization of surface constituents. Nonconfigurational languages use case marking on words to express grammatical functions; word order is less relevant. Phrase structure and inflectional features are important for phonological encoding. The

lemmas in a surface structure point to addresses where the corresponding word-form information is stored. These pointers are indexed for the inflections the word stem should undergo. The way in which a word becomes articulated also depends on its position in the phrase, and on other properties of the phrase structure. The phrase structure is, in particular, relevant for the generation of sentence rhythm and melody.

The constituents of a surface structure are of different syntactic categories and fulfill different grammatical functions. The major phrasal categories are Sentence, Noun Phrase, Verb Phrase, Prepositional Phrase, and Adjective Phrase. Each of these phrases has a head. In English the head of a sentence is the VP; it fulfills the function of predicate. The other heads are Noun, Verb, Preposition, and Adjective; they may be subcategorized within the same phrase for complements, such as the objects of verbs or prepositions, and they may have modifiers and specifiers. Heads play an important role in the generation of sentences. A head can call the syntactic procedures that will build its characteristic phrase around it, assigning the correct grammatical functions to its complements within the phrase.

The second section reviewed the way in which the surface structure embodies instructions for the generation of sentence prosody. This is done by means of indicators for mood and modality, and by focus markers. How mood and modality are indicated in surface structure is not well known; we suppose that such indicators are recognized by the Phonological Encoder, which generates the appropriate pattern of intonation (in particular, the characteristic boundary tones).

Focus is, initially, assigned to each syntactic constituent that expresses a prominent concept in the preverbal message. These focus markers will "percolate down" to the lemma level. In the end, each focused constituent contains at least one focused lemma. The lexical pointer of that lemma will be indexed for pitch accent; the word will be pronounced in a prosodically prominent way. The prosody of an utterance, finally, depends on the phrasal organization of its surface structure. The assignment of nuclear stress was given as an example.

Chapter 6
Lexical Entries and Accessing Lemmas

A main thesis of this and the following chapters will be that formulation processes are lexically driven. This means that grammatical and phonological encoding are mediated by lexical entries. The preverbal message triggers lexical items into activity. The syntactic, morphological, and phonological properties of an activated lexical item trigger, in turn, the grammatical, morphological, and phonological encoding procedures underlying the generation of an utterance. The assumption that the lexicon is an essential mediator between conceptualization and grammatical and phonological encoding will be called the *lexical hypothesis*. The lexical hypothesis entails, in particular, that nothing in the speaker's message will *by itself* trigger a particular syntactic form, such as a passive or a dative construction. There must always be mediating lexical items, triggered by the message, which by their grammatical properties and their order of activation cause the Grammatical Encoder to generate a particular syntactic structure.

The crucial role of the mental lexicon in the generation of speech makes it necessary to consider in some detail the internal structure and organization of entries in the mental lexicon. This is done in the first two sections of this chapter. Section 6.1 deals with the structure of lexical entries and their mutual relations. Section 6.2 analyses in more detail the aspect of lexical entries that we called "lemmas" in chapter 1. After these more structural sections we will turn to issues of processing. Section 6.3 reviews some major theories of lemma access in speech. This theoretical section is followed by two more empirical ones. Section 6.4 addresses accessing failures, their taxonomy and their potential causes. The time course of accessing lexical items is the subject of section 6.5.