

## CHAPTER 1

# Syntax, lexical categories, and morphology

## 1.0 Introduction

This book is an introduction to the basic concepts of syntax and syntactic analysis. Syntax is a central component of human language. Language has often been characterized as a systematic correlation between certain types of gestures and meaning, as represented simplistically in Figure 1.1. For spoken language, the gestures are oral, and for signed language, they are manual.

GESTURES ←————→ MEANING

Figure 1.1. *Language as a correlation between gestures and meaning*

It is not the case that every possible meaning that can be expressed is correlated with a unique, unanalyzable gesture, be it oral or manual. Rather, each language has a stock of meaning-bearing elements and different ways of combining them to express different meanings, and these ways of combining them are themselves meaningful. The two English sentences *Chris gave the notebook to Dana* and *Dana gave the notebook to Chris* contain exactly the same meaning-bearing elements, i.e. words, but they have different meanings because the words are combined differently in them. These different combinations fall into the realm of syntax; the two sentences differ not in terms of the words in them but rather in terms of their syntax. Syntax can thus be given the following characterization, taken from Matthews (1982:1):

The term ‘syntax’ is from the Ancient Greek *sýntaxis*, a verbal noun which literally means ‘arrangement’ or ‘setting out together’. Traditionally, it refers to the branch of grammar dealing with the ways in which words, with or without appropriate inflections, are arranged to show connections of meaning within the sentence.

First and foremost, syntax deals with how **sentences** are constructed, and users of human languages employ a striking variety of possible arrangements of the elements in sentences. One of the most obvious yet important ways in which languages differ is the order of the main elements in a sentence. In English, for example, the **subject** comes before the **verb** and the **direct object** follows the verb. In Lakhota (a Siouan language of North America), on the other hand, the subject and direct object both precede the

*An introduction to syntax*

verb, while in Toba Batak (an Austronesian language of Indonesia; Schachter 1984b), they both follow the verb. This is illustrated in (1.1), in which *the teacher*, *waʔspekhiye ki* and *guru i* function as subjects, and *a book*, *wówapi wə* and *buku* function as direct objects.

- |       |                                   |            |
|-------|-----------------------------------|------------|
| (1.1) | a. The teacher is reading a book. | English    |
|       | b. Waʔspekhiye ki wówapi wə yawá. | Lakhota    |
|       | teacher the book a read           |            |
|       | c. Manjaha buku guru i.           | Toba Batak |
|       | read book teacher the             |            |

The Lakhota and Toba Batak sentences also mean ‘the teacher is reading the book’, and in the Lakhota example the subject comes first followed by the direct object, whereas in the Toba Batak example the subject comes last in the sentence, with the direct object following the verb and preceding the subject. The **basic word order** in Toba Batak is thus the opposite of that in Lakhota. There are also languages in which the order of words is normally irrelevant to the interpretation of which element is subject and which is object. This is the case in the following Russian sentences.

- |       |                              |         |
|-------|------------------------------|---------|
| (1.2) | a. Učitel’nica čitaet knigu. | Russian |
|       | teacher read book            |         |
|       | b. Knigu čitaet učitel’nica. |         |
|       | book read teacher            |         |
|       | c. Čitaet učitel’nica knigu. |         |
|       | read teacher book            |         |

Again, all three of these sentences mean ‘the teacher is reading the book’, and in these Russian examples the order of the words is not the key to their interpretation, as it is in the sentences from the other three languages. Rather, it is the form of the words that is crucial. The *-a* on the end of *učitel’nica* ‘teacher’ signals that it is the subject, and the *-u* on the end of *knigu* ‘book’ indicates that it is the direct object. If the word for ‘teacher’ were the direct object in a sentence, then it would end in *-u*, as in (1.3).

- |       |                                 |         |
|-------|---------------------------------|---------|
| (1.3) | a. Ženščina videla učitel’nicu. | Russian |
|       | woman saw teacher               |         |
|       | b. Učitel’nicu videla ženščina. |         |
|       | teacher saw woman               |         |
|       | ‘The woman saw the teacher.’    |         |

These changes in the form of the words to indicate their function in the sentence are what Matthews referred to as ‘inflections’, and the study of the formation of words and how they may change their form is called **morphology**. These examples illustrate the important relationship between syntax and morphology: something which may be expressed syntactically in some languages may be expressed morphologically in others. Which element is subject and which is object is signalled syntactically in the examples from English, Lakhota and Toba Batak, while it is expressed morphologically in the Russian examples. Syntax and morphology make up what is traditionally referred to as ‘**grammar**’; an alternative term for it is **morphosyntax**, which explicitly recognizes the important relationship between syntax and morphology. Even though

*Syntax, lexical categories, and morphology*

this book is focussed on syntax, morphology will nevertheless be an important part of the discussion.

Thus a more complex picture of the nature of language emerges than that given in Figure 1.1; it is summarized in Figure 1.2.



Figure 1.2. *Language as a correlation between gestures and meaning (revised)*

All of the examples looked at so far involve simple sentences, but one of the most important syntactic properties of language is that simple sentences can be combined in various ways to form complex sentences. In terms of Figure 1.2, one could say that syntax makes possible the formulation of expressions with complex meanings out of elements with simple meanings. One of the defining features of human language is its unlimited nature; that is, the number of meaningful expressions that can be produced by users of a human language is potentially infinite, and this expressive potential comes from the combination of the basic meaningful elements with syntactic principles. Much of the interest in language in psychology and cognitive science comes from what the study of the cognitive mechanisms underlying language use and acquisition can reveal about the human mind.

This book has three goals: first, to introduce the basic concepts of syntax; second, to elucidate the principles and tools of syntactic analysis, which make it possible for linguists to analyze the grammatical systems of human languages; and third, to give an overview of the typological range of phenomena found in human languages which syntacticians seek to describe. The content of this book is presupposed by more advanced courses in syntactic theory, and hence it is intended to prepare the reader for such courses. The perspective of the book is primarily **descriptive**, and theoretical issues will be raised only in chapter 6. To many people the term ‘grammar’ evokes bad memories of **prescriptive** rules learned in school, e.g. ‘don’t split infinitives!’ Since the early part of the twentieth century, linguistics has rejected the prescriptive tradition which underlies school grammars and focusses instead on describing what users of human language actually do, not on prescribing what they should do.

A central part of the description of what speakers do is characterizing the **grammatical** (or **well-formed**) sentences of a language and distinguishing them from **ungrammatical** or (**ill-formed**) sentences. Grammatical sentences are those that are in accord with the rules and principles of the syntax of a particular language, while ungrammatical sentences violate one or more syntactic rules or principles. For example, (1.1a) is a grammatical sentence of English, while *Teacher the book a reading is* would not be. Ungrammatical sentences are marked with an asterisk, hence *\*Teacher the book a reading is*. This sentence is ungrammatical because it violates some of the word order rules for English, that is (i) basic word order in English clauses is subject–verb–object, (ii) **articles** like *the* and *a* precede the **noun** they modify, and (iii) **auxiliary verbs** like *is* precede the **main verb**, in this case *reading*. It is important to note that these are English-specific syntactic rules; this word order is perfectly grammatical in Lakhota,

*An introduction to syntax*

as (1.1b) shows, and if the Lakhota words were arranged in the English order, e.g. \**Ki waŋspekhiye yawá wá wówapi* [the teacher reads a book], the result is thoroughly ungrammatical. Well-formed sentences are those that are in accord with the syntactic rules of the language; this does not entail that they always make sense semantically. For example, the sentence *the book is reading the teacher* is nonsensical in terms of its meaning, but it violates no syntactic rules or principles of English; indeed, it has exactly the same syntactic structure as (1.1a). Hence it is grammatical (well-formed), despite being semantically odd.

The organization of the book is as follows. In this chapter a number of distinctions that are relevant to the discussion in the remainder of the book are introduced. First, two aspects of syntactic structure are distinguished, one of which will be the main topic of chapters 2 and 3, and the other will be the main topic of chapter 4. Second, the traditional notion of parts of speech are reviewed, as these categories will be important throughout the book. Finally, a brief introduction to some of the basic concepts of morphology and morphological analysis is presented, with emphasis on those notions that will be especially pertinent to the discussion in the succeeding chapters.

The next three chapters present basic syntactic phenomena from two different analytic perspectives and introduce the concepts and analytic tools used in each. Many of the same grammatical phenomena will be analyzed from each perspective. In chapter 5 the basics of writing a grammar to describe syntactic phenomena will be presented; the formulation of rules to express the generalizations arising from syntactic analysis and the role of the lexicon in a grammar will be discussed. Different linguistic theories make different sets of assumptions about the nature of syntactic structure and accordingly employ different analytic principles and tools. In chapter 6 the basic ideas of four linguistic theories will be summarized, and their approaches to important grammatical phenomena, including the formation of information questions (e.g. *What did you see?*) and the **passive voice** (e.g. *The bread was eaten by the mouse*), will be compared and contrasted. These two phenomena are especially revealing for a comparison of theories, because the accounts given by the various theories highlight the conceptual and analytic differences among them.

### 1.1 Aspects of syntactic structure

In the syntactic structure of sentences, two distinct yet interrelated aspects must be distinguished. The first one has already been mentioned: the function of elements as subject and direct object in a sentence. ‘Subject’ and ‘direct object’ have traditionally been referred to as **grammatical relations**. Hence this kind of syntax will be referred to as ‘**relational structure**’. It includes more than just grammatical relations like subject and direct object; it also encompasses relationships like **modifier–modified**, e.g. *tall building* or *walk slowly* (*tall, slowly* = modifier, *building, walk* = modified) and **possessor–possessed**, e.g. *Pat’s car* (*Pat’s* = possessor, *car* = possessed). Relational structure will be the primary focus of chapters 2 and 3.

The second aspect concerns the organization of the units which constitute sentences. A sentence does not consist simply of a **string** of words; that is, in a sentence like *The teacher read a book in the library*, it is not the case that each word is equally related to the words adjacent to it in the string. There is no direct relationship between *read* and *a* or between *in* and *the*; *a* is related to *book*, which it modifies, just as *the* is related

*Syntax, lexical categories, and morphology*

to *library*, which it modifies. *A* is related to *read* only through *a book* being the direct object of *read*, and similarly, *the* is related to *in* only through *the library* being the object of the **preposition** *in*. The words are organized into units which are then organized into larger units. These units are called **constituents**, and the hierarchical organization of the units in a sentence is called its **constituent structure**. This term will be used to refer to this second aspect of syntactic structure. Consider the eight words in the sentence *The teacher read a book in the library*. What units are these words organized into? Intuitively, it seems clear that the article *the* or *a* goes with, or forms a unit with, the noun following it. Is there any kind of evidence beyond a native speaker's intuitions that this is the case? Determining the constituent structure of sentences is the major topic of chapter 4, but a brief preliminary look at the kind of evidence needed follows.

If the article forms a unit with the noun that follows it, we would expect that in an alternative form of the same sentence the two would have to be found together and could not be split up. Thus in the passive version of this sentence, *A book was read by the teacher in the library*, the unit *a book* serves as subject, and the unit *the teacher* is the object of the preposition *by*. The constituent composed of a noun and an article is called a **noun phrase** [NP]; as will be shown later, NPs can be very complex. The preposition *in* and the NP following it also form a constituent in this sentence (*in the library*); it is called a **prepositional phrase** [PP]. The fact that the PP is a constituent can be seen by looking at another alternative form, *In the library the teacher read a book*. Finally, the verb plus the NP following it form a unit as well, as shown by a sentence like *I expected to find someone reading the book, and reading the book was a teacher*. The constituent composed of a verb plus following NP is called a **verb phrase** [VP]. As with NPs, VPs can be quite complex. In each of these alternative forms, a combination of words from the original sentence which one might intuitively put together in a single unit also occurs together as a unit, and this can be taken as evidence that they are in fact constituents. Using square brackets to group the words in constituents together, the constituent structure of *The teacher read a book in the library* may be represented as in (1.4). ('S' stands for 'sentence'.)

(1.4) [S [NP The [N teacher]] [VP [V read] [NP a [N book]] [PP [P in] [NP the [N library]]] PP] VP] S]

Note the nesting of constituents within constituents in this sentence, e.g. the NP *the library* is a constituent of the PP *in the library* which is a constituent of the VP *read a book in the library*. In chapter 4 constituent structure will be explored in detail.

At the beginning of this section it was noted that the two aspects of syntactic structure, relational structure and constituent structure, are 'distinct yet interrelated', and it is possible now to see how this is the case. For example, a VP was described as being composed of a verb and the following NP, but it could alternatively be characterized as involving the verb and its direct object. Similarly, a PP is composed of a preposition and its object. NPs, on the other hand, involve modifiers, and accordingly the relation between *the* and *teacher* could be described as one of modifier–modified. Thus, these two aspects of syntactic structure are always present in a sentence, and when one or the other is emphasized, the sentence is being described from one of the two perspectives. It will be seen later that different grammatical phenomena seem to be more easily analyzed from one perspective rather than the other.

*An introduction to syntax*

## 1.2 Lexical categories

In the discussion of the constituents of sentences, reference has been made to nouns and noun phrases, verbs and verb phrases, and prepositions and prepositional phrases. Nouns, verbs and prepositions are traditionally referred to as ‘parts of speech’ or ‘word classes’; in contemporary linguistics they are termed **lexical categories**. The most important lexical categories are noun, verb, **adjective**, **adverb** and **adposition**, which subsumes prepositions and **postpositions**. In **traditional grammar**, lexical categories are given **notional definitions**, i.e. they are characterized in terms of their semantic content. For example, *noun* is defined as ‘the name of a person, place or thing’, *verb* is defined as an ‘action word’, and *adjective* is defined as ‘a word expressing a property or attribute’. In modern linguistics, however, they are defined morpho-syntactically in terms of their grammatical properties.

Nouns may be classified in a number of ways. There is a fundamental contrast between nouns that refer uniquely to particular entities or individuals and those that do not; the best example of the first kind of noun is a proper name, e.g. *Sam*, *Elizabeth*, *Paris* or *London*, and nouns of this type are referred to as **proper nouns**. Nouns which do not refer to unique individuals or entities are called **common nouns**, e.g. *dog*, *table*, *fish*, *car*, *pencil*, *water*. One of the important differences between proper and common nouns in a language like English is that common nouns normally take an article, while proper nouns do not, e.g. *The boy left* versus *\*The Sam left* (cf. *\*Boy left* versus *Sam left*). Common nouns may be divided into **mass nouns** and **count nouns**. Count nouns, as the name implies, denote countable entities, e.g. *seven chairs*, *six pencils*, *three dogs*, *many cars*. Mass nouns, on the other hand, are not readily countable in their primary senses, e.g. *\*two waters*, *\*four butters*, *\*six snows*. In order to make them countable, it is necessary to add what is sometimes called a ‘measure word’, which delimits a specific amount of the substance, e.g. *two glasses/bottles/drops of water*, *four pats/sticks of butter*, *six shovelfuls of snow*. Measure words can be used with count nouns only when they are plural, e.g. *\*six boxes of pencil* versus *six boxes of pencils*, *\*two cups of peanut* versus *three jars of peanuts*. **Pronouns** are closely related to nouns, as they both function as NPs. Pronouns are traditionally characterized as ‘substitutes’ for nouns or as ‘standing for’ nouns, e.g. *John went to the store, and he bought some milk*, in which *he* substitutes or stands for *John* in the second clause. This, however, is true only of **third-person** pronouns like *he*, *she*, *it*, or *they*; it is not true of **first-person** pronouns like *I* or **second-person** pronouns like *you*. First- and second-person pronouns refer to or index the speaker and addressee in a speech event and do not replace or stand for a noun.

Verbs can likewise be categorized along a number of dimensions. One very important dimension which will be discussed in detail in chapters 2 and 3 is whether a verb takes just a subject (an **intransitive** verb), or a subject and a direct object (a **transitive** verb), or a subject, direct object and **indirect object** (a **ditransitive** verb). This will be referred to as the ‘valence’ of the verb. Another dimension concerns the kind of situation it represents. Some verbs represent static situations which do not involve anyone actually doing anything, e.g. *know* as in *Chris knows the answer*, or *see* as in *Pat sees Dana over by the bookcase*. Some symbolize actions, e.g. *run* as in *Kim ran around the track*, or *sing* as in *Leslie sang a beautiful aria*. Others refer to a change of state, e.g. *freeze* as in *The water froze* (the change in the state of the water is from liquid to solid), or *dry* as in *The clothes dried quickly* (the change in the state of the

*Syntax, lexical categories, and morphology*

clothes is from wet to dry). Some represent complex situations involving an action plus a change of state, e.g. *break* as in *Larry broke the window with a rock* (Larry does something with a rock [action] which causes the window to break [change of state]). This classification of verbs is quite complex and is more appropriately in the domain of semantics rather than syntax. However, some syntactically relevant aspects of the meaning of verbs will be investigated in chapter 2.

Some examples of adjectives in English include *red, happy, tall, sick, interesting, beautiful*, and many others. Adjectives typically express properties of entities, e.g. *a red apple, a tall woman, a beautiful sunset*. Some properties are inherent attributes of an entity; for example, some apples are red because they are naturally so, whereas some barns are red because they have been painted red, not because they are inherently red. Hence color is an inherent property of apples but not of barns. Some languages signal this distinction overtly. In Spanish, for example, the adjective *feliz* means ‘happy’, and whether it is an inherent or permanent property of the person referred to is signaled by the verb it is used with, i.e. *Maria es feliz* ‘Maria is happy (a happy person)’ versus *Maria está feliz* ‘Maria is happy (now, at this moment but not necessarily always)’. Spanish has two verbs meaning ‘be’, *ser* and *estar*, and one of the differences between them is that *ser* plus adjective (*es* in this example) is used to signify inherent or permanent attributes, while *estar* plus adjective (*está* in this example) serves to indicate non-permanent, transitory attributes.

English adverbs typically, but not always, end in *-ly*, e.g. *quickly, happily, beautifully, rapidly* and *carefully*. *Fast* and *friendly* are exceptions; *fast* is an adverb without *-ly* (it can also be an adjective), and *friendly*, despite the admonitions of road signs in Texas to ‘drive friendly’, is an adjective, e.g. *a friendly waiter*. Adverbs modify verbs, adjectives and even other adverbs, and they can be classified in terms of the nature of this modification; manner adverbs, for example, indicate the manner in which something is done, e.g. *The detective examined the crime scene carefully*, or *The ballerina danced beautifully*, while temporal adverbs, as the name implies, express when something happened, e.g. *Kim talked to Chris yesterday*, or *Dana will see Pat tomorrow*. *Yesterday* and *tomorrow* do not end in *-ly* and have the same form when functioning as an adverb that they have when functioning as a noun, e.g. *Yesterday was a nice day*, *Tomorrow will be very special*. The most common adverbial modifiers of adjectives and adverbs are words like *very, extremely, rather*, e.g. *a very tall tree, the extremely clever student, rather quickly*. This class of adverbs is referred to as **degree modifiers**.

Prepositions are adpositions that occur before their object, while postpositions occur after their object. English and Spanish have only prepositions, e.g. English *in, on, under, to*, Spanish *en, a, con*, whereas Japanese and Korean have only postpositions. German has both: *in dem Haus* ‘in the house’ (preposition *in*) versus *dem Haus gegenüber* ‘over across from the house’ (postposition *gegenüber*).

There are a number of minor categories. The category of **determiners** includes articles like *a* and *the*, and **demonstratives** like *this* and *that*. Determiners modify nouns in relation to their referential properties. Articles indicate roughly whether the speaker believes her interlocutor(s) can identify the referent of the NP or not; an **indefinite article** like *a(n)* signals that the speaker does not assume the interlocutor(s) can identify the referent of the NP, while a **definite article** like *the* indicates that the speaker does assume that the interlocutor(s) can identify it. Demonstratives, on the other hand, refer to entities in terms of their spatial proximity to the speaker; English

*An introduction to syntax*

*this* refers to an entity close to the speaker, while *that* refers to one farther away. (*Which book do you mean? This one here or that one over there? versus \*This one over there or that one here?\*) Many languages make a three-way distinction: close to the speaker (English *this*, Spanish *esta* [FEM]), away from the speaker but not far (English *that*, Spanish *esa* [FEM]), and farther away from the speaker (archaic English *yon*, Spanish *aquella* [FEM]). These distinctions are also expressed by locative demonstratives, e.g. English *here*, German *hier*, Spanish *aquí* versus English *there*, German *da*, Spanish *ahí* versus English *yonder*, German *dort*, Spanish *allí*. **Quantifiers**, as the label implies, express quantity-related concepts. English quantifiers include *every*, *each*, *all*, *many*, and *few*, as well as the numerals *one*, *two*, *three*, etc., e.g. *every boy*, *many books*, *the seven sisters*. **Classifiers** serve to classify the nouns they modify in terms of shape, material, function, social status and other properties. They are found in many East and Southeast Asian and Mayan languages, among others. They are similar in many respect to the measure words that occur with English mass nouns, but they occur with all nouns regardless of the count–mass distinctions, e.g. Cantonese *yāt būi séui* [one CL water] ‘one cup of water versus *yāt jā séui* ‘a jug of water’, versus *yāt jēun séui* ‘a bottle of water’ with a mass noun, *nī ga dihnlóuh* [this CL computer] ‘this computer’ (classified as machine) versus *nī bouh dihnlóuh* ‘this computer’ (classified as model) versus *nī go dihnlóuh* ‘this computer’ (classified as object) with a count noun (Matthews and Yip 1994). **Conjunctions**, like *and*, *but* and *or*, serve to link the elements in a conjoined expression. There are conjoined NPs, e.g. *a boy and his dog*, conjoined verbs, e.g. *Leslie danced and sang*, and conjoined adjectives, e.g. *Lisa is tall and slender*. All major lexical categories can be linked by conjunctions to form conjoined expressions; this will be discussed in more detail in chapters 3 and 4. **Complementizers** mark the dependent clause is a complex sentence, e.g. English *that* as in *Sally knows that Bill ate the last piece of pizza*. The final category is **particles**, which is a classification often given to elements which do not fall into any of the other categories. Many particles have primarily **discourse** functions, e.g. English *indeed*, German *doch*, Spanish *entonces*.*

There is an important opposition that divides lexical categories into two general classes, based on whether the membership of the class can readily be increased or not. Languages can usually increase their stock of nouns, for example, by borrowing nouns from other languages or creating new ones through **compounding** (e.g. *black* + *board* yields *blackboard*) or other morphological means (e.g. *rapid* + *-ly* = *rapidly*), but they do not normally create or borrow new adpositions, conjunctions or determiners. Lexical categories such as noun and verb whose membership can be enlarged are termed **open class** categories, whereas categories such as adposition, determiner or conjunction, which have small, fixed membership, are called **closed class** categories.

The definitions of lexical categories given so far are primarily the notional ones from traditional grammar. These definitions seem intuitively quite reasonable to speakers of Indo-European languages, and they seem to correlate nicely with the syntactic functions of the different parts of speech. Let us define three very general syntactic functions: **argument**, modifier and **predicate**. In a sentence like *the teacher read an interesting book*, *the teacher* and *an interesting book* are the arguments, *read* is the predicate, and *the*, *an* and *interesting* are modifiers. Similarly, in *Kim is tall*, *Kim* is the argument and *is tall* is the predicate. The term ‘argument’ here includes NPs and PPs functioning as subject, direct object or indirect object. The notions of predicate and



*Syntax, lexical categories, and morphology*

argument will be discussed in more detail in the next chapter, but for now one can say simply that in a sentence the predicate expresses the state of affairs that the referents of the arguments are involved in. (The terms ‘predicate’ and ‘argument’ are also used in semantics with a different meaning; they are being used here and elsewhere to refer to syntactic notions, unless otherwise noted.) It is usual to distinguish 1-place, 2-place and 3-place predicates, depending on how many participants there are in the state of affairs depicted by the predicate. Being sick is a state of affairs involving only one participant, hence *be sick* is a 1-place predicate which takes one argument, e.g. *Kim is sick*. In *the teacher destroyed the note*, there is an action of destroying involving a teacher and a note. Destroying involves a destroyer and something destroyed; hence *destroy* is a 2-place predicate and takes two arguments. Finally, giving involves a giver, something given and a recipient, and therefore *give* is a 3-place predicate and takes three arguments, e.g. *The teacher gave an interesting book to Kim*. Given these distinctions, it seems intuitively clear that nouns would be arguments, verbs would be predicates and adjectives would be modifiers, and this is in fact the case very often.

But not always. Nouns and adjectives can function as part of a predicate, as in *Dana is a phonologist* and *Chris was sick*. Even though they are part of the predicate, they are still formally distinct from verbs; they do not take **tense suffixes** like verbs do, i.e. *\*Dana phonologists* or *\*Chris sicked*. The **copula** *be*, a kind of verb, carries these verbal inflections. Contrast this with the situation in Lakhota, in which nouns and adjective-like words do bear verbal inflections when functioning as predicates, in this instance **agreement in number** with the subject.

- (1.5) a. Wičháša ki hená lowá-pi. Lakhota  
           man the those sing-PL  
           ‘Those men are singing.’<sup>1</sup>  
 a’. Lakhóta ki hená lowá-pi.  
       Sioux the those sing-PL  
       ‘Those Siouxs (Indians) are singing.’  
 b. Wičháša ki hená lakhóta-pi.  
       man the those Sioux-PL  
       ‘Those men are Siouxs (Indians).’  
 b’. Lakhóta ki hená wičháša-pi.  
       Sioux the those man-PL  
       ‘Those Siouxs (Indians) are men.’  
 c. Wičháša ki hená khúža-pi.  
       man the those sick-PL  
       ‘Those men are sick.’

<sup>1</sup> In most examples from languages other than English, there will be an interlinear **gloss** with a translation for each meaningful element in the sentence directly under it as well as a free translation into English in the third line. In the interlinear gloss, the translation will be lined up directly under the element being translated. Complex words will be broken up into their meaningful parts (see section 1.3 below) separated by hyphens, and the translation for each part will be joined to the translations for the other parts by hyphens and placed below the whole word. Thus in (1.5a), for example, *wičháša* means ‘man’, *ki* means ‘the’ and *hená* means ‘those’; the last word, *lowápi*, is broken up into two parts, *lowá* and *pi*, which are linked by a hyphen, and each part is translated (*lowá* means ‘sing’ and *pi* means ‘plural subject’), with the translations linked by a corresponding hyphen and placed below the Lakhota word. If an element requires a translation involving more than one English word, the words will be joined by a ‘.’, e.g. ‘was.washed’ in (2.4a). Finally, grammatical notions like tense and number are glossed using abbreviations which are listed at the beginning of the book.

*An introduction to syntax*

Nouns in Lakhota do not normally carry any indication of number; the only way to tell that the NP containing *wičháša* ‘man’ is **plural** in (1.5a) is by means of the plural demonstrative *hená* ‘those’ (cf. *hé* ‘that’). In particular, the plural suffix *-pi* is impossible on the noun *wičháša* in (1.5a); based on (1.5a), (1.5a’), one could conclude that it occurs only on verbs. But this would be incorrect, as the sentences in (1.5b, c) show. Nouns like *wičháša* ‘man’ do take *-pi* when they function as a predicate, rather than as an argument. Hence nouns in Lakhota seem to function readily as predicates, something their English counterparts do not do. Adjective-like words also function directly as predicates, as (c) illustrates; there is no copular element analogous to English *be* in either of these sentence types. Verbs and adjective-like words can also serve as arguments in Lakhota, as in (1.6).

- (1.6) a. *Hokšíla ki hená čhéya-pi.* Lakhota  
 boy the those cry-PL  
 ‘Those boys are crying.’  
 b. *Čhéya ki hená hokšíla-pi.*  
 cry the those boy-PL  
 ‘The ones crying are boys.’  
 c. *Khúža ki hená wičháša-pi.* (cf. (1.5c))  
 sick the those man-PL  
 ‘The sick ones are men.’

The verb *čhéya* ‘cry’ serves as the predicate in (1.6a) and the argument in (1.6b). Note that in the English translation the verb *cry* cannot simply function as the subject; it must, rather, occur in a complex expression *the ones crying*. In Lakhota, by contrast, *hokšíla* ‘boy’ and *čhéya* ‘cry’ simply exchange positions in the sentence without any formal modification. The same is true of the noun *wičháša* ‘man’ and the adjective-like word *khúža* ‘sick’ in (1.6c). Thus, the expected correlations between noun and argument, verb and predicate and adjective and modifier are not as strong in Lakhota as they are in English.

An even more striking example of this lack of correlation between lexical class and syntactic function can be seen in Nootka, a Wakashan language spoken on Vancouver Island in British Columbia, Canada (Swadesh 1939).

- (1.7) a. *Wafa:k-ma qo:ʔas-ʔi.* Nootka  
 go-3sgPRES man-the  
 ‘The man is going.’  
 a’. *Qo:ʔas-ma wafa:k-ʔi.*  
 man-3sgPRES go-the  
 ‘The one going is a man.’  
 aʹʹ. *Qo:ʔas-ma.*  
 man-3sgPRES  
 ‘He is a man.’  
 b. *ʔi:h-ma qo:ʔas-ʔi.*  
 large-3sgPRES man-the  
 ‘The man is large.’  
 b’. *Qo:ʔas-ma ʔi:h-ʔi.*  
 man-3sgPRES big-the  
 ‘The large one is a man.’