CHAPTER SIXTEEN

Syntax and Language Typology

We have now dealt with phonology (the sound system of language) and have looked briefly at morphology and word formation, both of which are concerned with the internal structures of words.

In this chapter we will be concerned with two topics. One is how words are put together to form sentences, which is called *syntax*, and the other is with how the languages of the world differ in the devices they use to communicate certain grammatical information. The latter topic is especially interesting for us, since German and English use entirely different devices and are therefore not merely different languages but different **types** of languages as well.

Syntax is, as just stated, concerned with how words are put together to form sentences. In the previous chapter you were given an inadequate but intuitively accurate definition of word. It is infinitely more difficult to give a definition of sentence. We all do know what a sentence is, however. We know, for instance, that of the following pair, 1) is a sentence of English and 2) is not, even though they might communicate roughly the same information:

　　1) The dog bit the man and hurt him.
　　2) Dog bite man ouch.

1) conforms to the syntactic patterns, or rules, of English. 2) does not. Each language has its own rules concerning how words are to be ordered with respect to one another, what sorts of helping words are introduced etc. I will use the following definition of sentence: "An utterance in a language which will be accepted by a native speaker of that language as a well-formed sentence of that language is a sentence of that language." This is, of course, obviously circular and therefore inadequate, but intuitively we should all know what this means. It accepts as a sentence the first paragraph of this chapter. It excludes possible utterances like "Boy run", "sky blue" and "although he is ill" and also excludes utterances like the second paragraph of this chapter, which consists of more than one sentence. A sentence, then, is a minimal unit of some sort, just as a phoneme and a morpheme are minimal units.

Sentences are composed of words. Words are composed of morphemes. Morphemes have meaning. Therefore sentences, and words, have meaning. One obvious question is how we arrive at the meaning of a sentence. One possible answer which comes quickly to mind is that we add up the meanings of all of the morphemes in the sentence. The sum of the meanings of the individual morphemes is the meaning of the sentence. Thus if a sentence contains 6 morphemes,

the meanings of which we will symbolise with *a, b, c, d, e, f* respectively, we could express the meaning of the sentence formally as something like 3):

3) $\Sigma(a + b + c + d + e + f)$

i.e. the meaning of the sentence would be the sum (Σ) of the meanings of the individual morphemes. Thus a sentence like 4),

4) The dog bit the man

consists of the morphemes {THE}, {DOG}, {BITE}, {-D} (i.e. past tense), {THE} and {MAN}. We will use for the purposes of this problem *italic* spellings of the ordinary English words to indicate the *meanings* of those words and use *-ed* to indicate the meaning of {-D}. The meaning of 4) could then be represented as:

4') $\Sigma(the + dog + bite + -ed + the + man)$

At first glance this seems reasonable. However, since addition is commutative, i.e. a + b = b + a, the same formula would also give the meaning of 5), which contains exactly the same words as 4) but which is a different sentence with a different meaning.

5) The man bit the dog.

There is obviously more to the meaning of a sentence, and hence of any well-formed utterance, than the meanings of the morphemes and words in the sentence. The meaning of a sentence also must include information about, in these cases, the performer of the action and the victim of the action. 4) and 5), despite containing the same words, and therefore the same morphemes, differ in that in 4) a dog performs the act of biting on a man, whereas in 5) the roles are reversed. The terms **Agent** and **Patient** are often used to indicate the performer and the victim of an action respectively. Traditionally the terms **Subject** and **Direct Object** are used. The meanings of 4) and 5) obviously have to include the meanings of the morphemes, but in addition they include an identification of the grammatical roles played by *man* and *dog* in each.

The notions **Agent** and **Patient** are semantic notions. Semantics is the study of meaning. **Subject and Direct Object**, on the other hand, are syntactic notions, sometimes called *grammatical relations*. In 4) and 5) the Agent is the syntactic subject and the Patient is the syntactic direct object, but this is not always the case. In 6) the syntactic subject is the semantic *Patient*, and the Agent is neither the subject nor the direct object.

6) The man was bitten by the dog.

Not all languages have the same ways of indicating subjects and objects. English uses word order. In English a noun phrase (including pronouns) which comes before the verb is the subject and a noun phrase which comes immediately after a verb is the direct object. In active sentences like 4) and 5) subjects tend to be Agents and direct objects tend to be Patients. Thus we are able, upon encountering 4) or 5), to identify the subject and the direct object by their positions in the sentence and to assign to them the semantic roles Agent and Patient respectively. In passive sentences like 6), on the other hand, subjects tend to be Patients and Agents, if present at all, tend to be preceded by "by". We thus identify the subject by its position and then assign it the Patient role. It is therefore vital to the understanding of a sentence that we identify the subject. Then, depending on the type of sentence it is, which we must also be able to identify, we can assign the subject the

correct semantic role. Only then can we deduce the meaning of the sentence. Therefore the meaning of 4) must include not only 4') but also a statement to the effect that "dog" is the Agent and "man" the Patient. The meaning of 5) could be summed up by 4') plus a statement to the effect that "man" is the Agent and "dog" the Patient. And 4') plus a statement that "man" is the Patient and "dog" the Agent would account for the meaning of 6). Thus the meanings of 4) and 6) would be analysed as essentially the same which, in fact, they are.

German does not use word order to identify subjects and objects. Rather German uses grammatical case. Particular grammatical cases are identified with particular grammatical relations. The Nominative case tends to be identified with grammatical subjects and the Accusative case tends to be identified with direct objects. The Dative case is identified with indirect objects. The Nominative, Accusative and Dative cases have other uses as well, but their major function is to identify subjects, direct objects and indirect objects respectively. Thus word order plays little if any role in identification in German, whereas it is vital in English, and grammatical case plays no role in identification of grammatical relation in English. In fact, English nouns do not display grammatical case at all, and the few vestiges of grammatical case in the pronouns are better described as positional variants than as actual case forms. 4) could be translated into German as 7):

> 7) Der Hund biss den Mann

which has the same word order as 4), but it could also be translated as 8), which has the same word order as 5).

> 8) Den Mann biss der Hund.

And 6), where the Patient comes before the Agent, could be translated as either 9) or 10):

> 9) Der Mann wurde von dem Hund gebissen.
> 10) Von dem Hund wurde der Mann gebissen.

In 10) the Agent precedes the Patient, which is the order of 7), and in 9) the Patient precedes the Agent, which is the order of 8). Yet all four sentences have the same meaning. And 11), 12), 13) and 14) all have the same meaning, which corresponds to the meaning of 5).

> 11) Der Mann biss den Hund.
> 12) Den Hund biss der Mann.
> 13) Der Hund wurde von dem Mann gebissen.
> 14) Von dem Mann wurde der Hund gebissen.

Strictly speaking, of course, 11) - 14) do not all have *exactly* the same meaning. By "have the same meaning", I mean that the semantic roles of "Mann" and "Hund" are identical in all four, i.e. the **biter** and the **bitten** are identical. Since German does not depend on word order to identify grammatical relations like subject and object, German can utilise differences in word order for other purposes, for emphasis, for instance. The difference between 7) and 8) is largely a difference in emphasis, a difference between what is already given and what is then said about it. Old, or given, information tends to come early in the sentence and new information comes later.

William G. Moulton, whose seminal book *The Sounds of English and German* has often been referred to in these volumes, has written that he was once engaged in a discussion in German with a German colleague about the works of the English poet Robert Browning. At some point his

German colleague said: "Browning hat Rilke übersetzt". Moulton initially understood the sentence to mean that Robert Browning had translated the works of the German poet Rainer Maria Rilke into English. He then immediately recalled that Browning was dead before Rilke was born and therefore reinterpreted the sentence to mean what his colleague had intended it to mean - that Rilke had translated Browning's works into German. The sentence as it stands is ambiguous. There is absolutely no clue in the sentence to tell the hearer or reader whether Browning or Rilke is the subject. In this instance Moulton had to rely on his knowledge of the literary histories of both English and German to enable him to assign the correct interpretation to the sentence. Because Browning was the object of the discussion (the so-called *Topic* or *Theme*), Moulton's colleague put him first in the sentence. The fact that Rilke was the one who translated the works of Browning is the new information in the sentence (the *Comment* or *Rheme*) and comes later. German word order tends to be Topic-Comment. Thus even when there is no overt case marker to distinguish a nominative from an accusative, German does not "fall back on" word order. German has never used word order as a means of identifying grammatical relations. In a situation like Moulton's one has to rely on one's knowledge of the situation under discussion and of the real world to get the intended interpretation. Some philosophers like to talk about "possible worlds". One possible world would be one in which, for instance, everything is exactly the same as in this one except that Rilke lived before Browning. In *that* world the sentence "Browning hat Rilke übersetzt" would mean exactly what Moulton interpreted it to mean originally in *this* world.

It is customary to refer to languages like English, which use word order to identify grammatical relations, as **word order languages** and to refer to languages like German, which use grammatical case to identify grammatical relations, as **case languages**. There are many case languages in the world. Indeed, English was once a case language. Modern German has four grammatical cases. One thousand years ago German had five. Russian has six. Estonian has fourteen. And grammatical cases are not always distributed the same way. For instance, related European languages like German, Old English, Latin and Russian put all grammatical subjects in the case which is called **Nominative** and put direct objects in the case which is called **Accusative**. In Basque, spoken in Spain and totally unrelated to other European languages, **subjects** of **intransitive** sentences (a transitive sentence has a direct object, an intransitive sentence does not) and **direct objects** of **transitive** sentences are put in the same case, which is called the **Absolutive** case, and subjects of transitive sentences appear in the **Ergative** case. Thus there are categories within the case language type. German is a Nominative-Accusative language, Basque is an Ergative-Absolutive language. Australian aboriginal languages are typically Ergative-Absolutive where nouns are concerned and Nominative-Accusative where pronouns are concerned.

Typing languages as word order or as case languages is just one way of categorising languages. Another way of categorising languages has to do with the relationships between morphemes and words. This way of typing languages goes back to Friedrich von Schlegel in the nineteenth century. According to this typology, languages in which words tend to consist of only one morpheme are called *isolating* languages. Chinese and Vietnamese are the cases usually given. Languages in which words typically contain more than one morpheme fall roughly into two sub-types, *agglutinating* languages and *inflecting* languages.

Turkish is the language usually cited as an example of an agglutinating language. In Turkish the noun *ev* means "house". The plural suffix for this noun is *-ler*. *-im* means "my", and *-den* means "out of". In the following list these morphemes are transparent in every instance.

ev	"house"	evim	"my house"
evler	"houses"	evlerim	"my houses"
evden	"out of the house"	evimden	"out of my house"
evlerden	"out of the houses"	evlerimden	"out of my houses"

In each word each of the morphemes appears in exactly the same form and is therefore instantly recognisable. To *agglutinate* means to join together by adhesion, as with glue. The morphemes are "glued" together in an unaltered and unalterable form in a specific order.

In inflecting languages it is often much more difficult to isolate morphemes than in agglutinating languages. Latin is the example usually cited, but German will do as well. Let us look at the declension of the definite article, a paradigm familiar to you.

	Masculine	**Neuter**	**Feminine**	**Plural**
Nominative	der	das	die	die
Accusative	den	das	die	die
Dative	dem	dem	der	den
Genitive	des	des	der	der

The only thing common to all of these forms is /d-/. We can (tentatively) associate the meaning "definite article" with this. It is not at all clear how we should analyse the rest of the forms. The following paradigm gives the forms in phonemic transcription with /d-/ divided from the rest of the form by a hyphen:

	Masculine	**Neuter**	**Feminine**	**Plural**
Nominative	/d-er	d-as	d-i	d-i/
Accusative	/d-en	d-as	d-i	d-i/
Dative	/d-em	d-em	d-er	d-en/
Genitive	/d-εs	d-εs	d-er	d-er/

Beginning with the first form, we can say that /-er/ represents nominative singular masculine, but it is not clear whether these are three separate morphemes or whether these are even the correct morphemes. Nor is it clear which morpheme should be associated with which bit of phonemic material. In fact, since there are only two phonemes but possibly three morphemes, it is impossible to associate any morpheme with any phoneme, even arbitrarily. The forms spelled *die* consist of only two phonemes. /d-/ is definite article, but what is /-i/? It is at least feminine singular plus either nominative or accusative. Either that or it is plural nominative or plural accusative. But it is only one phoneme. It is therefore impossible to assign any of these morphemes, if they are indeed morphemes, to anything phonological.

How should the form *dem* be analysed? /d-/ is "definite article"; /-em/ could be specifically dative singular masculine or specifically dative singular neuter, or it might be dative singular non-feminine.

In the last chapter we looked at one verb inflection. Let us now look at another fragment. /-st/ seems to have the constant meaning "2nd person singular". "2nd person plural" is always

indicated by /-t/ or /-ət/. /-t/ and /-ət/ also sometimes mean "3rd person singular", of course, but for the moment we will concentrate on the 2nd person. Comparing /-st/ and /-t/ we find one constant piece of common phonological material, /t/. The common meaning of "2nd person singular" and "2nd person plural" is "2nd person". Following the principles laid out in the previous chapters we could therefore (tentatively) assign the meaning "2nd person" to /t/. Then we could assign the meaning "singular" to /s/. But what will we then assign "plural" to? /t/ is already taken up, and there isn't anything else there. Perhaps we should decide that it is not necessary to have a separate plural morpheme - it would be sufficient to have a singular morpheme, and in its absence we have, by default, the plural. That seems to work for this example, but it falls down when we come to the first person. In the present tense the usual first person singular verb inflection is /-ə/. In the plural it is /-ən/. Here we could assign the meaning "1st person" to /ə/ and the meaning "plural" to /n/. But we have nothing to assign the meaning "singular" to.

This is a typical situation in inflecting languages, and the usual pedagogical approach with inflecting languages is to give the student a paradigm like the one above and say "Here it is - learn it." With inflecting languages we often have to content ourselves with saying that there are certain stretches of phonological material, call them **morphs**, which *simultaneously* indicate more than one morpheme. Thus /-st/ is simultaneously {2ND PERSON} and {SINGULAR}. In the definite article paradigm /-er/ is simultaneously *either* {NOMINATIVE}, {SINGULAR}, {MASCULINE} *or* {DATIVE}, {SINGULAR}, {FEMININE} *or* {GENITIVE}, {SINGULAR}, {FEMININE} *or* {GENITIVE}, {PLURAL}. Let me emphasise that linguists are not admitting defeat by doing this. This is simply what the language forces us to do. In fact, the first step in a morphological analysis of a text is to find the recurrent partials and to call them *morphs*. Once that has been done the process of deciding whether the morphs represent individual morphemes or simultaneously represent more than one morpheme can begin. If we are lucky enough to be dealing with an agglutinating language we will have relatively few morphs which simultaneously represent more than one morpheme. If we are dealing with an inflecting language we will have relatively more.

No language is entirely isolating or entirely agglutinating or entirely inflecting. German has certain categories of words which are largely isolating, i.e. one word = one morpheme. The prepositions are an example, as are the conjunctions. In compounding and derivation German tends towards agglutination, with clear-cut boundaries between elements of words. And even within the inflected categories certain inflections are easily assigned to one morpheme whereas others cannot be. *Inflecting*, *agglutinating* and *isolating* are matters of degree. English is more isolating than German, but English has some inflections and it agglutinates in compound words. German is more inflecting than English but has some isolates and some agglutinations.

Some nineteenth century scholars saw in these different types of languages and in the mixtures of types within individual languages an evolutionary process. Complicated inflections like those in the German definite article or the Latin noun were seen as developments from an earlier, primitive stage in which languages were agglutinating. Over time some sort of linguistic evolutionary law had determined that those separate elements should begin to melt together, to coalesce, into more complicated, less transparent elements. Since the scholars who held this view were overwhelmingly German, and since German (and Latin, their scholarly language) were like this, these languages were seen as the ultimates in linguistic evolution, and languages like Turkish were seen as still at a primitive stage. Complicated inflections were seen as a sign of "strength" and simple inflections, including particularly agglutinations, as a sign of "weakness". This is the origin of the terms "strong" and "weak" for irregular and regular verbs respectively. Verbs like *machen*, which form their past tense by adding /-t-/ and then a predictable person and number ending and which form the past participle by "sticking on" (agglutinating) at the beginning of the stem a /gə-/ and at the end a /-t/, were seen as being "weak", meaning like forms in "weak"

languages such as Turkish, whereas verbs like *binden*, which change the vowel of the stem in the past tense and change it again in the past participle, were felt to be showing "strength". This is, of course, absurd, as is the notion that languages evolve naturally to an inflecting type like German or Latin. Another absurd notion is that technologically more advanced people speak more advanced languages than so-called primitive people, leading to the notion that some languages are "better" than others, or that some languages are "simple" and others are "difficult", the degree of ease or difficulty involved in learning the language supposedly having something to do with the intelligence of the native-speaking population, a proposition with which most practicing linguists would disagree. These notions are all part and parcel of **Romanticism**, the social and esthetic movement of the 19th and late 18th centuries which emphasised a love for the past, for the far-away (often leading its practitioners to "make up the facts" about distant places and peoples, including their languages), and for the strange, the beautiful, the wild and the irregular. Ideas of strong and weak languages are also bound up with notions of Teutonic superiority, also an integral part of German Romanticism which was to resurface in the 20th century under Adolf Hitler. Unfortunately the German grammatical tradition has retained these terms, which were adopted by the British grammatical tradition, from which they passed to us. We appear to be stuck with them. I doubt that any of the readers of this work will ever have seen "strength" in vowel change and "weakness" in merely adding endings. Nevertheless we should be aware that these terms have suspect origins. They are harmless enough as long as they are used only as labels for particular classes of verbs, nouns and adjective endings and as long as it is recognised that they have nothing whatsoever to do with genuine strength or weakness. Similarly the terms *masculine*, *feminine* and *neuter* are simply terms for three noun classes, those which appear with *der*, *die* and *das* respectively in the nominative singular. There is nothing inherently "masculine" about masculine nouns, nothing inherently "feminine" about feminine nouns and nothing inherently "neutral with regard to sex" about neuter nouns, certain pseudo-Freudian attempts to show the opposite notwithstanding.

The great American linguist Edward Sapir summed up the evolutionary hypothesis of linguistic development as follows: "A linguist who insists on talking about the Latin type ... as though it were necessarily the high-water mark of linguistic development is like the zoologist that sees in the organic world a huge conspiracy to evolve the race-horse or the Jersey cow." (Quoted in Lyons 1969)

* * * * * * * * * *

In summary, then, German is a **case** language, specifically a **Nominative-Accusative** language. English is a **word-order** language. German is primarily an **inflecting** language; English is primarily an **isolating** language, although it has a few inflections. Both languages are **mixtures** of isolating, agglutinating and inflecting languages. The proportions of each are different in the two languages. English and German are therefore different *types* of languages on at least two counts. This is particularly interesting because English is genetically closely related to German. The Angles and Saxons, after all, went to England from what is now northern Germany. There are still ethnic groups in north Germany who call themselves Angles and Saxons.

In the next chapter we will begin a discussion of the history of the German language, a study of its origins and linguistic development. This will include its relationship to English and to most of the other languages of Europe and many of the languages of Asia.