1
Goals

1.1 Overview

The aim of this chapter is to give you some idea of the goals of linguistic theory, and to introduce you to some simple concepts which will be used throughout the rest of the book. Among the notions which will be explained in this chapter are terms such as *theory of language*, *grammar of a language*, *particular/universal grammar*, *competence*, *performance*, *grammaticality*, *linguistic intuition*, *rule-governed creativity*, *generate*, *observational/descriptive/explanatory adequacy*, *constraint*, *markedness*, and *innateness*.

1.2 Grammatical competence

Linguistics is the study of *Language*. But why should we be interested in the phenomenon of *Language*? Chomsky gives an avowedly *mentalist* answer to this question. For him, the most fundamental reason for studying language is that language is a mirror of the mind – i.e. by detailed study of language, we might hope to reach a better understanding of how the human mind produces and processes language. As Chomsky remarks (*Language and Mind* (1972a), p. 103):

> There are a number of questions which might lead one to undertake a study of language. Personally, I am primarily intrigued by the possibility of learning something, from the study of language, that will bring to light inherent properties of the human mind.

But what aspects of language should be the focus of our study? Chomsky's answer is that there are three inter-related theories which any detailed study of language ultimately seeks to develop, namely:

(1) (i) Theory of Language Structure
(ii) Theory of Language Acquisition
(iii) Theory of Language Use

The Theory of Language Structure will concern itself with what are the defin-

ing structural properties of natural (i.e. human) languages; the Theory of Language Acquisition with the question of how children acquire their native language(s); and the Theory of Language Use with the question of how linguistic and nonlinguistic knowledge interact in speech comprehension and production. Of the three, the task (i) of developing a Theory of Language Structure is logically prior to the other two, since only if we first know what 'language' is can we develop theories about how it is acquired and used. It is perhaps not surprising, therefore, that most of Chomsky's work has been devoted to the attempt to develop a Theory of Language Structure.

But what exactly is it that such a theory seeks to characterise? The answer is that any adequate Theory of Language Structure must provide answers to questions such as the following:

- What is language?
- What is it that you know when you know a language?
- What are the essential defining characteristics of natural languages which differentiate them from, for example, artificial languages like those used in Mathematics or Computing, or from animal communication systems?
- Do languages differ from each other in unpredictable ways, or do they all share certain common, universal properties?

But how do we attempt to develop a Theory of Language Structure which will answer such questions? The first step is to formulate detailed descriptions (known technically as *grammars*) of particular languages (e.g. English): this is the study of *Particular Grammar*. So, for example, we might try and devise a grammar of English, a grammar of French, a grammar of Swedish, a grammar of Swahili, a grammar of Chinese . . . and so on and so forth. A grammar of a particular language will take the familiar form of a set of rules or principles which tell you how to 'speak' and 'understand' the language: more precisely, a grammar will comprise a set of rules or principles which specify how to form, pronounce, and interpret Phrases and Sentences in the language concerned. The word *grammar* in this technical sense has a much broader sense than that familiar from school textbooks, since it covers not only Morphology (i.e. the internal structure of words) and Syntax (i.e. how words are combined together to form phrases and sentences), but also Phonology (i.e. pronunciation) and some aspects of Semantics (i.e. meaning) as well. When we have compiled detailed grammars of a number of different languages, the second step in our quest for a Theory of Language Structure is to abstract from particular grammars common, universal properties that they all share: this is the study of *Universal Grammar* – i.e. the search for linguistic universals.

Consider first the study of Particular Grammar. What exactly is it that a grammar of a particular language sets out to describe? Chomsky gives an essentially *mentalist* answer to this question: for him, a grammar is a *model* (= systematic description) of those linguistic abilities of native speakers of a language which enable them to speak and understand their language fluently. These linguistic abilities, Chomsky terms the *competence* of the native speaker. Thus, a *grammar* of a language is a model of the linguistic competence of the fluent native speaker of the language. *Competence* (the fluent native speaker's knowledge of the language) is contrasted by Chomsky with *performance* (what people actually say or understand by what someone else says on a given occasion). Competence is 'the speaker–hearer's knowledge of his language', while Performance is 'the actual use of language in concrete situations' (Chomsky, *Aspects* (1965), p. 4). Very often, performance is an imperfect reflection of competence: for example, the fact that people make occasional slips of the tongue in everyday speech does not mean that they don't know their native language, or don't have fluency (i.e. competence) in it. Slips of the tongue and like phenomena are – for Chomsky – *performance errors*, attributable to a variety of performance factors like tiredness, boredom, drunkenness, drugs, external distractions, and so forth. Linguistics is – for Chomsky – primarily concerned with competence, since a Theory of Competence will be a subpart of an eventual Theory of Performance: that is, you have to understand what a native speaker knows about his language before you can study the effects of tiredness, drunkenness, etc. on this knowledge. Thus, what we mean by saying that a grammar is a model of the native speaker's competence is that a grammar tells us what we need to know in order to be fluent in a language.

Chomsky distinguishes two types of competence: (i) *grammatical competence*, and (ii) *pragmatic competence* (see e.g. Chomsky, *Essays* (1977a), p. 40). The former belongs to the Theory of Language Structure, and the latter to the Theory of Language Use. Pragmatics is concerned with the role played by nonlinguistic information such as background knowledge and personal beliefs in our use of sentences. To take one of Chomsky's own examples (from *Essays* (1977a), p. 40), suppose I have a friend who says to me 'Today was a disaster.' If I know (by way of background information) that he was giving a special lecture today, then on the basis of this background knowledge I infer that he probably means that his lecture went down very badly. It is the native speaker's *pragmatic competence* which enables him to bring into play nonlinguistic information in the interpretation of sentences. By contrast, in the case of a sentence such as:

(2) He thinks that John is wrong

it is the native speaker's *grammatical competence* (his knowledge of the grammar of his language) which tells him that *he* cannot be interpreted as referring to the same person as *John* in a sentence like (2). Since, as we noted earlier, Chomsky has devoted himself primarily to the study of language structure rather than language use, he has focussed almost exclusively on the task of attempting to characterise *grammatical* rather than *pragmatic* competence.

The native speaker's grammatical competence is reflected in two types of *intuition* which speakers have about their native language(s) – (i) intuitions about sentence *well-formedness*, and (ii) intuitions about sentence *structure*. The word *intuition* is used here in a technical sense which has become standardised in Linguistics: by saying that a native speaker has *intuitions* about the well-formedness and structure of sentences, all we are saying is that he has the ability to make *judgments* about whether a given sentence is well-formed or not, and about whether it has a particular structure or not. The term *well-formed* is also a standard technical term in the linguistic literature: for the time being, you can think of it as meaning 'OK' – but a little later, we'll try and define *well-formedness* a little more precisely.

These *intuitions* about sentences span four different aspects of language – namely *Phonology* (= the study of sounds and sound systems), *Morphology* (= the study of how morphemes (grammatical units smaller than the word) are combined together into words), *Syntax* (= the study of how words are combined together to form sentences), and *Semantics* (= the study of meaning). Hence, we can say that native speakers have phonological, morphological, syntactic, and semantic competence, and that this competence is reflected in their intuitions about the phonological, morphological, syntactic, and semantic well-formedness and structure of sentences in their native language(s). We'll look briefly at each of these different aspects of competence in turn.

Let's begin by illustrating typical intuitions reflecting a native speaker's *phonological* competence. All native speakers of English would agree that (3) (a) below is phonologically well-formed in respect of its stress pattern (i.e. it's OK to pronounce the sentence with primary stress on the capitalised syllables), whereas (3) (b) is phonologically ill-formed in respect of its stress pattern (i.e. it isn't OK to pronounce the sentence with primary stress on the capitalised syllables):

(3) (a) **THIS** is a gra**MMA**tical **SEN**tence
(b) This is **A** grammati**CAL** sen**TENCE**

So, we all have intuitions about possible and impossible stress patterns in sentences. Moreover, we all have strong *phonotactic* intuitions – i.e. intuitions

about what are possible and impossible sound sequences among native words in English. For instance, we'd probably all agree that *blick* is a possible, but non-occurring English word, whereas **bnick* by contrast is not a possible native English word (an asterisk in front of a word, phrase, or sentence indicates that it is ill-formed in some way): such a word could only occur in English as a foreign borrowing. Phonological competence is also reflected in intuitions about phonological structure: any English speaker intuitively feels, for example, that the sequence 'black bird' can either be a single phonological word (**BLACK**bird, with primary stress on *black* = a species of bird, like thrush, robin, etc.), or two independent phonological words (**BLACK BIRD** or black **BIRD** = bird which is black, as opposed to 'white bird', 'yellow bird', etc.).

In much the same way, *morphological competence* is reflected in the native speaker's intuitions about morphological well-formedness and structure. For example, native speakers of English know that *van* and *can* have the respective plural forms *vans* and *cans*, but that the plural of *man* is *men* and not **mans*. Likewise, native English speakers know that *fold* and *scold* have the respective past tense forms *folded* and *scolded*, but that the past tense form of *hold* is not **holded*, but rather *held*. In the same way, anyone fluent in English knows that the Verbs *approve* and *refuse* have corresponding Nouns *approval* and *refusal*, but that the Noun counterparts of *prove* and *amuse* are not **proval* and **amusal*, but rather *proof* and *amusement*. In addition, native speakers also have intuitions about morphological structure: for example, English speakers intuitively feel that words like *overload*, *overplay*, and *overwork* are structured out of two independent morphemes, a prefix *over* (meaning 'excessively') and a stem *load–play–work*, whereas by contrast *overture* does not comprise the two morphemes *over* and *ture*.

The native speaker's *semantic* competence is reflected in intuitions about semantic well-formedness and structure. For example, any native speaker of English would agree that (4) (a) below is semantically well-formed, but that (4) (b) is semantically ill-formed (i.e. 'odd' in some way, by virtue of its meaning, so that it 'doesn't make sense'):

(4) (a) I thought that Mary was ill, but it turned out that she wasn't
(b) !I realised that Mary was ill, but it turned out that she wasn't

(In this book, we use ! in front of a sentence to show it is 'anomalous' (i.e. semantically or pragmatically 'odd'); generally speaking, we follow the standard practice of using an asterisk * in front of a sentence to indicate that it is *syntactically* ill-formed, though occasionally we extend the use of the asterisk to indicate that a sentence is simply ill-formed in some way, without specifying

in what way(s)). A second type of semantic intuition which native speakers have about their language concerns semantic structure and semantic relations. To take an example from Chomsky (*Knowledge* (1986), p. 8), any native speaker of English knows that *them* can be interpreted as being coreferential to (i.e. referring to the same set of individuals as) *the men* in (5) (a) below, but not in (5) (b):

(5) (a) I wonder who *the men* expected to see *them*
(b) *The men* expected to see *them*

Hence, intuitions about coreference relations in sentences are part of the set of intuitions we have about semantic relations in and between sentences.

Having looked briefly at how phonological, morphological, and semantic competence is reflected in intuitions about well-formedness and structure, let's now turn to examine the nature of *syntactic* competence. Here, too, we find that competence is reflected in two types of intuition: intuitions about syntactic well-formedness, and intuitions about syntactic structure. To say that a native speaker has intuitions about syntactic well-formedness in his language is to say that he is able to judge whether such-and-such a sequence of words is a grammatical sentence in his language or not. For example, any native speaker of English would intuitively recognise (leaving aside for the moment differences of style or dialect) that all the examples in (6) below are grammatical (i.e. syntactically well-formed) sentences in English:

(6) (a) I gave back the car to him
(b) I gave the car back to him
(c) I gave him back the car
(d) I gave him the car back

but that the following are ungrammatical as sentences of English:

(7) (a) *I gave the car to him back
(b) *I gave back him the car

(Recall that an asterisk in front of a sentence means that it is ill-formed in some way (usually, *syntactically ill-formed*, i.e. *ungrammatical*); by convention, any sentence which does not have an asterisk in front of it is assumed to be well-formed; note that asterisks go at the *beginning*, not the *end* of sentences!)

But what does it mean to say that native speakers have intuitions about the *syntactic structure* of sentences in their language? All this means is that native speakers have 'gut feelings' about which words in a sentence 'go with' or 'modify' which other words. For example, in the case of a sentence such as:

(8) Some people can be very selfish

we'd all agree that *some* 'goes with' *people* (and not, for example, *very*), so that the sequence [*some people*] forms a Phrase of some sort. Likewise, we'd all agree that *very* 'modifies' *selfish* (and not, for example, *be*), so that the sequence [*very selfish*] forms a rather different kind of Phrase. And, if you think about it carefully, you'd probably agree that in a sentence such as:

(9) You could not go to her party

we could take the word *not* to 'go with' either *could* or *go*, as becomes clearer if we look at the two 'variants' of (9) given in (10) below:

(10) (a) What you *could not* do is go to her party
(b) What you could do is *not go* to her party

Thus, native speakers have an intuitive knowledge of the syntactic relations between the words in sentences in their language; in other words, they intuitively know how words are combined together to form Phrases, and Phrases are combined together to form sentences. We should note, however, that native speakers' intuitions about syntactic structure are often much less sharp (or reliable) than their intuitions about syntactic well-formedness (grammaticality).

1.3 Basic concepts and fundamental misconceptions

Before we go any further, it is useful to clarify some of the basic concepts which we have introduced so far, and to try and eradicate some of the fundamental misconceptions which people sometimes have about these concepts. One basic concept which often gives rise to unfortunate misunderstanding is that of *grammaticality*. In this connection, we should point out that it is important not to confuse the descriptive notion *grammatical* with the corresponding prescriptive notion *correct*. For example, there are many varieties of English in which sentences like:

(11) Mine is bigger than what yours is

are perfectly grammatical, and for speakers of these dialects such sentences are syntactically well-formed. But at the same time, sentences like (11) are of a type stigmatised as 'incorrect' or 'bad grammar' by a certain self-styled sociocultural elite (= pedants!). This poses an apparent dilemma for the linguist: should he *describe* what people actually say, or should he attempt to *prescribe* what he or others think they ought to say? In other words, should Linguistics be *descriptive* or *prescriptive*?

In actual fact, it is hard to see how anyone could defend the *prescriptive* approach. In any other field of enquiry, it would be seen as patently absurd. What would we say of the social anthropologist who, instead of describing the way a given society is, sets about prescribing the way he thinks it ought to be? (We'd probably suggest he ought to give up Anthropology and take up Politics!) And what would we think of the scientist who, regretting the unfortunate tendency for the moon to orbit round the earth, instead proposes an alternative model in which the earth orbits round the moon, simply because he thinks things *ought* to be that way? No-one these days would take any such enterprise seriously; and the same is true of Linguistics. Modern Linguistics is purely *descriptive*, not prescriptive. Hence, we describe sentences as 'grammatical/ungrammatical', or 'well-formed/ill-formed'; we do not use prescriptive terminology such as 'correct/incorrect' (yes, that is a *prescriptive* statement!)

A related issue which often gives rise to muddled thinking concerns the relative importance of written language on the one hand, and spoken language on the other. There is a tendency for the uninformed to think that written language is a 'purer', 'more correct' form of language than spoken language. As Palmer (1983, p. 27) notes, 'All too often people tend to think of the spoken language as a rather poor version of the written language.' This has been counteracted by the opposite trend in work in Descriptive Linguistics in the last fifty years, which has concerned itself primarily with spoken language, and generally regarded the written language as being of secondary importance. Why should this be? There are a variety of reasons why descriptive linguists might suppose that spoken language is a more fertile and rewarding field of study than written language.

For one thing, spoken language tends to be less subjected to prescriptive pressures than written language, and hence is a less artificial medium of communication (written language is often a kind of 'censored' version of spoken language). Secondly, spoken language is a more spontaneous form of communication than written language: you can spend half an hour composing a sentence in a letter, but not in a conversation! Thirdly, spoken language shows a much greater range of variation between individuals than written language (and hence is more interesting if we want to study how language varies according to social class, etc.). Fourthly, the spoken language generally contains more linguistic information than the written language: for example, all spoken sentences have an associated intonation contour which is not properly represented in the written language. Fifthly, spoken language is far more frequently used as a medium of communication than written language: we talk to friends, etc. all day long, but we don't spend all day writing letters to people! For these

and many other reasons, many linguists would argue that Linguistics should concern itself primarily with spoken rather than written language.

There is a sense, however, in which the whole debate about whether we should concentrate on the written or the spoken language is a non-issue. For, after all, what we are seeking to describe is the native speaker's *competence*, i.e. his linguistic knowledge: this is a *mental* property, which commonly has two physical manifestations (in the form of written or spoken language), but which might have alternative physical instantiations (e.g. in sign language, etc.). What we are interested in is the abstract grammatical competence underlying the physical realisations of that competence; to the extent that both written and spoken forms of language (provided we can be sure they are free from prescriptive influences, errors induced by performance factors, etc.) reflect this abstract competence, then both provide us with useful clues to the nature of *competence*. Thus, the linguist doesn't study only the spoken language, or only the written language: rather, he studies *language* (in all its physical instantiations).

A more serious problem which arises with Chomsky's conception of a Grammar as a model of the linguistic intuitions of the fluent native speaker of a given language concerns what to do about disagreements among native speakers about the well-formedness or structure of particular sentences. One of the abstractions that Chomsky makes in studying language is to assume that speech communities are homogeneous: i.e. to assume that all native speakers of a given language will have essentially the same well-formedness and structural intuitions. Chomsky himself remarks (*Aspects* (1965), p. 3) that: 'Linguistic theory is concerned primarily with an ideal speaker–listener, in a completely homogeneous speech community.' But the problem with this abstraction is that it is plainly not the case that speech communities are homogeneous: all native speakers have to some extent their own individual way of speaking (or *idiolect*), and there are, of course, larger linguistic groupings within society. Speakers with a common geographical background may share a common *dialect*, while speakers from a common social background may share a common *sociolect*. We can illustrate the linguistic differences within a given speech community in terms of the examples in (12) below: each of these sentences would probably be accepted as well-formed by only a certain percentage of English speakers (hence the use of the % prefix):

(12) (a) %Your car wants mending
(b) %That's to do tomorrow
(c) %I gave it her
(d) %Can the both of us come?

(e) %There's a man sells vegetables in the village
(f) %It were me what told her

The obvious question is what the linguist is to do in such cases. The broad answer is that in general the problem of linguistic variation within a speech community is one which is more appropriately dealt with in a partially separate discipline (*Sociolinguistics*), and since it is not a problem which is essentially *syntactic* in nature, it is not the kind of problem which ought to be the primary focus of attention in the attempt to develop an adequate theory of *Syntax*. For practical purposes, most linguists describing a language of which they are native speakers rely on their own intuitions, and thus the grammar they devise is essentially a grammar of their own idiolect, which they assume is representative of the language as a whole.

An even more tricky problem which arises with the notion of *well-formedness* concerns the attempt to identify in what way(s) a given ill-formed sentence is 'odd'. Let's first draw a distinction between sentences which are 'pragmatically anomalous' in some way, and those which are 'linguistically ill-formed'. While the distinction may be clear enough in principle, it is often very hard in practice to decide which side of the dividing line a given sentence falls. For example, what is the status of sentences such as the following, (taken from Lakoff (1971), p. 332):

(13) (a) My uncle realises that I'm a lousy cook
(b) My cat realises that I'm a lousy cook
(c) My goldfish realises that I'm a lousy cook
(d) My pet amoeba realises that I'm a lousy cook
(e) My frying pan realises that I'm a lousy cook
(f) My sincerity realises that I'm a lousy cook
(g) My birth realises that I'm a lousy cook

Intuitively, most people would regard (13) (a) as perfectly well-formed, (13) (b) as slightly less natural, (13) (c) as a bit eccentric, (13) (d) as implausible, (13) (e) as just plain daft, and (13) (f) and (g) as absolutely inconceivable. But what precisely is the nature of the oddity of the more unusual sentences? The answer is that the oddity seems to be largely *pragmatic* (i.e. nonlinguistic) in nature. Thus, whether or not you find expressions like *My goldfish thinks that* . . . well-formed depends on whether or not you believe that goldfish do (or might) possess powers of thought; a sentence like (13) (c) presupposes that goldfish are capable of thought, and a person who rejects sentences like (13) (c) is in effect rejecting the implied proposition that goldfish can think. Why should he reject such a proposition? Presumably because it